IL DIRITTO DELL'INFORMAZIONE E DELL'INFORMATICA
ISSN 1593-5795
Anno XXXVIII Fasc. 2 - 2022

Giusella Finocchiaro

LA PROPOSTA DI REGOLAMENTO SULL'INTELLIGENZA ARTIFICIALE: IL MODELLO EUROPEO BASATO SULLA GESTIONE DEL RISCHIO

Estratto

GIUFFRÈ
GIUFFRÈ FRANCIS LEFEBVRE

GIUSELLA FINOCCHIARO *

LA PROPOSTA DI REGOLAMENTO SULL'INTELLIGENZA ARTIFICIALE: IL MODELLO EUROPEO BASATO SULLA GESTIONE DEL RISCHIO

1. Il contesto geopolitico di riferimento.

Una riflessione sulla proposta di Regolamento europeo sull'intelligenza artificiale [1] non può prescindere da alcune considerazioni, necessariamente sintetiche, sullo scenario geopolitico attuale. Il mercato delle nuove tecnologie e, specificamente, dell'intelligenza artificiale è senza dubbio un mercato globale, privo di barriere geografiche. L'azione normativa, invece, è attualmente caratterizzata dai confini degli Stati nazionali o delle istituzioni sovranazionali, come l'Unione europea. D'altronde questo contrasto, fra la natura eminentemente nazionale del diritto, e la dimensione globale del fenomeno da disciplinare, che oggi appare in tutta la sua evidenza nello sforzo di regolare le conseguenze giuridiche delle applicazioni di intelligenza artificiale, ha caratterizzato fin dall'origine il problema delle fonti del diritto su Internet [2].

Oggi il mercato appare sostanzialmente diviso in tre aree di influenza: quella europea, quella statunitense e quella cinese.

* Il presente contributo riproduce il testo, con aggiunta di note, della Relazione pronunciata in occasione del Convegno "La via europea per l'intelligenza artificiale", tenutosi presso l'Università Ca' Foscari di Venezia nei giorni 25-26 novembre 2021, i cui Atti saranno pubblicati a cura del Dipartimento di Economia.

[1] Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, 21 aprile 2021, COM(2021) 206 *final*, disponibile *on line* al seguente *link*: *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206* (consultato il 28 febbraio 2022).

[2] Si consenta di rinviare al mio *Lex mercatoria e commercio elettronico. Il diritto applicabile ai contratti conclusi su Internet*, in *Contr. e impr.*, 2001, pp. 571-610, ove amplia bibliografia.

Il modello adottato in Europa è quello regolatorio: si intende non soltanto normare e disciplinare i nuovi fenomeni, le nuove tecnologie e i nuovi beni, ma anche fare sì che il modello europeo divenga un riferimento globale e possa essere adottato nelle altre regioni geopolitiche (il cosiddetto "effetto Bruxelles") [3]. All'interno di questo modello si salvaguardano non soltanto i diritti fondamentali [4], ma anche i "valori" europei, e quest'ultimo termine è più volte citato nell'ambito della proposta in commento, a sottolineare che il modello proposto non è solo normativo, ma culturale. Si vuole rendere evidente che non si tratta soltanto di regole giuridiche, ma anche della cultura che quelle regole esprimono.

Il modello adottato negli Stati Uniti, con le necessarie semplificazioni contenute in questa sintesi, è un modello auto-regolatorio e basato sull'antitrust.

Quello cinese, invece, appare un modello dirigistico e basato sul capitalismo di Stato. Certamente la Cina si caratterizza per essere sempre più attiva anche nella produzione di norme: nell'ambito della protezione dei dati personali, basti ricordare la *Personal Information Protection Law* (PIPL) in vigore dal 1° novembre 2021 [5], la *Data Security Law* (DSL) in vigore dal 1° settembre 2021 [6] e la *Cybersecurity Law* (CSL) in vigore dal 1° giugno 2021 [7]. E sotto il profilo strategico, la recente creazione della *Shanghai Data Exchange* (SDE), la borsa di Shanghai per lo scambio dei dati, persegue anche l'obiettivo di creare lo "Shanghai Model" per la compravendita di dati. Il "modello Shanghai" ha l'ambizione di risolvere i problemi che oggi rendono difficile la circolazione dei dati e di proporsi come modello globale di riferimento per eliminare i rischi dell'incertezza giuridica.

Dunque, come sempre, la proposta regolatoria persegue anche obiettivi di natura geopolitica, cercando di estendere l'ambito di applicazione del Regolamento. L'art. 2, infatti, con tecnica analoga a quella utilizzata dall'art. 3 del Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 "relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali

[3] Sul tema, compiutamente Bradford, *The Brussels Effect: How the European Union Rules the World*, New York, 2020.

[4] Sono espressamente richiamati i seguenti diritti sanciti dalla Carta dei diritti fondamentali dell'Unione europea: il diritto alla dignità umana (art. 1), al rispetto della vita privata e alla protezione dei dati di carattere personale (art. 7 e 8), alla non discriminazione (art. 21) e alla parità tra donne e uomini (art. 23).

[5] *Personal Information Protection Law of the People's Republic of China*, 20 agosto 2021.

[6] *Data Security Law of the People's Republic of China*, 10 giugno 2021.

[7] *Cybersecurity Law of the People's Republic of China*, 6 novembre 2016.

dati" [8], dispone che il Regolamento si applichi ai fornitori che immettono sul mercato o mettono in servizio sistemi di intelligenza artificiale (di seguito in sigla anche "IA") nell'Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un Paese terzo, nonché agli utenti dei sistemi di IA situati nell'Unione e ai fornitori e agli utenti di sistemi di IA situati in un Paese terzo, ove l'*output* prodotto dal sistema sia utilizzato nell'Unione [9].

2. Il rischio della retorica e la necessità di un nuovo modello.

Come è noto, normare sulla tecnologia è molto difficile. Il principio della neutralità tecnologica, ormai affermato in ambito internazionale, costituisce l'esito di un complesso dibattito. Secondo tale principio, che si è affermato nell'elaborazione dell'UNCITRAL (*United Nations Commission on International Trade Law*) [10], il diritto deve rimanere neutro rispetto alla tecnologia, astenendosi dall'individuare la soluzione tecnica che consenta di

[8] Dispone, infatti, l'art. 3 del Regolamento (UE) 2016/679: "Il presente regolamento si applica al trattamento dei dati personali effettuato nell'ambito delle attività di uno stabilimento da parte di un titolare del trattamento o di un responsabile del trattamento, indipendentemente dal fatto che il trattamento sia effettuato o meno nell'Unione. Il presente regolamento si applica al trattamento dei dati personali di interessati che si trovano nell'Unione, effettuato da un titolare del trattamento o da un responsabile del trattamento che non è stabilito nell'Unione, quando le attività di trattamento riguardano: a) l'offerta di beni o la prestazione di servizi ai suddetti interessati nell'Unione, indipendentemente dall'obbligatorietà di un pagamento dell'interessato; oppure b) il monitoraggio del loro comportamento nella misura in cui tale comportamento ha luogo all'interno dell'Unione. Sull'ambito di applicazione territoriale del Regolamento (UE) 2016/679, Catanzariti, *Art. 3*, in D'Orazio-Finocchiaro-Pollicino-Resta (a cura di), *Codice della* privacy *e* data protection, Milano, 2021, pp. 143-153; De Hert-Czerniawski, *Expanding the European data protection scope beyond territory: Article 3 of the General Data Protection Regulation in its wider context*, in *International Data Privacy Law*, 2016, pp. 230-243; Finocchiaro, *Il quadro d'insieme sul Regolamento europeo sulla protezione dei dati personali* e Spangaro, *L'ambito di applicazione materiale della disciplina del Regolamento Europeo 679/2016*, in Finocchiaro (diretto da), *La protezione dei dati personali in Italia. Regolamento UE n. 2016/679 e d.lgs. 10 agosto 2018, n. 101*, Bologna, 2019, rispettivamente pp. 1-25 e pp. 27-62; Reccia, *Art. 3*, in Riccio-Scorza-Belisario (a cura di), *GDPR e normativa* privacy. *Commentario*, Milano, 2018, pp. 18-24.

[9] In questo senso si esprimono chiaramente anche il considerando n. 10 e il considerando n. 11. Si legge, infatti, nel considerando n. 10: "Al fine di garantire condizioni di parità e una protezione efficace dei diritti e delle libertà delle persone in tutta l'Unione, è opportuno che le regole stabilite dal presente regolamento si applichino ai fornitori di sistemi di IA in modo non discriminatorio, a prescindere dal fatto che siano stabiliti nell'Unione o in un paese terzo, e agli utenti dei sistemi di IA stabiliti nell'Unione". Recita, invece, il considerando n. 11: "Alla luce della loro natura di sistemi digitali, è opportuno che determinati sistemi di IA rientrino nell'ambito di applicazione del presente regolamento anche quando non sono immessi sul mercato, né messi in servizio, né utilizzati nell'Unione (…) Al fine di impedire l'elusione del presente regolamento e di garantire una protezione efficace delle persone fisiche che si trovano nell'Unione, è opportuno che il presente regolamento si applichi anche ai fornitori e agli utenti di sistemi di IA stabiliti in un paese terzo, nella misura in cui l'output prodotto da tali sistemi è utilizzato nell'Unione (…)".

[10] Il principio di neutralità tecnologica veniva affermato dall'UNCITRAL già nel *Model Law on Electronic Commerce* del 1996. Sul punto si consenta di richiamare il mio *Il ruolo dell'UNCITRAL nello sviluppo della disciplina sul commercio elettronico*,

implementare i principi giuridici affermati. In altri termini, la norma giuridica non dovrebbe riferirsi a un livello di sicurezza predeterminato o a una tecnologia specifica, bensì dovrebbe limitarsi a dettare lo scopo da raggiungere senza indicare le modalità tecniche per il suo perseguimento [11].

Nel normare sull'intelligenza artificiale, occorre tenere in considerazione altre due criticità, che possono condizionare l'azione del legislatore. La prima è quella costituita dalla paura: la paura che l'applicazione di intelligenza artificiale possa assumere decisioni autonomamente e, possa, ispirandosi anche alla letteratura fantascientifica, rivolgersi contro gli esseri umani e operare in modi non previsti. Queste paure trovano la loro origine nella difficile prevedibilità degli esiti dei processi di *machine learning*. Anche di questo condizionamento, come di quello della retorica, bisogna essere consapevoli perché la paura porta, in questo caso, all'ansia del controllo. Il legislatore dominato, o quanto meno condizionato, dalla paura ha l'esigenza di controllare tutto, dettando norme estremamente dettagliate, fotografando l'esistente, senza consentire così alle regole di potersi evolvere col tempo, scendendo nel dettaglio, abbandonando la neutralità tecnologica che, a livello internazionale, invece è il principio fondamentale a cui fare riferimento. Non si può lasciare che le suggestioni indotte dalle grandi opere artistiche sul tema, citate anche nella Risoluzione del Parlamento europeo del 2017 che riportava le leggi della robotica di Asimov [12], così come la cinematografia e altri sviluppi artistici, alimentino paure infondate e condizionanti. Ciò potrebbe sfociare nel bisogno di controllo assoluto e quindi generare il rischio di normare in dettaglio ciò che in realtà ancora non si

nonché Castellani, *I testi dell'UNCITRAL in materia di diritto del commercio elettronico* e Ratti, *La Convenzione sull'uso delle comunicazioni elettroniche: le principali disposizioni*, in Finocchiaro-Delfini (a cura di), *Diritto dell'informatica*, Milano, 2014, rispettivamente, pp. 63-70, pp. 43-62 e pp. 71-107.

[11] Il principio di neutralità tecnologica è stato adottato dal legislatore europeo, ad esempio, nel Regolamento (UE) n. 910/2014, il c.d. Reg. e-IDAS, allorché ha introdotto delle definizioni di firma elettronica e firma elettronica avanzata non riferite a specifiche tecnologie. Non può dirsi, invece, "tecnologicamente neutra" la nozione di firma elettronica qualificata, collegata a livelli di sicurezza predeterminati.

[12] V. Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2018/C 252/25). In particolare, si legge nella Risoluzione: "le leggi di Asimov devono essere considerate come rivolte ai progettisti, ai fabbricanti e agli utilizzatori di robot, compresi i robot con capacità di autonomia e di autoapprendimento integrate, dal momento che tali leggi non possono essere convertite in codice macchina". Veniva successivamente riportato in nota il testo delle leggi di Asimov, che, come è noto, recitano: "(0) Un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno. (1) Un robot non può recar danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno. (2) Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Prima Legge. (3) Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Prima o con la Seconda Legge".

conosce del tutto e che ha bisogno di un modello normativo e principi fermi, di una cornice chiara entro cui svilupparsi, e non di regole minuziose e fardelli burocratici. Non si deve cadere nell'errore di volere disciplinare il processo decisionale nel particolare.

La seconda criticità è quella della retorica: indulgere in essa e finire col rendere le applicazioni di intelligenza artificiale soggetto giuridico senza che questa operazione sia funzionale ad un nuovo modello normativo. Il rischio è che l'uso generico del termine "intelligenza" implicitamente induca ad assumere che c'è un "soggetto intelligente" e non applicazioni tecnologiche che fanno cose che, se fatte da umani, sarebbero considerate intelligenti, per riprendere la lezione di Turing [13].

Si può certamente elaborare un modello giuridico di responsabilità basato su una sintesi linguistica o sulla metafora [14], così come è accaduto per il concetto di personalità giuridica, anche per la soggettività giuridica nelle applicazioni di intelligenza artificiale, ma sempre che questo sia idoneo a meglio normare il fenomeno [15]. Occorre che questa sia una scelta precisa e non una mera conseguenza di una sottintesa e presupposta soggettività, portato della retorica [16] e che il nuovo modello giuridico sia definito [17].

Ad esempio, con riguardo al problema del risarcimento del

[13] Alan Turing, considerato il padre fondatore della scienza informatica e dell'intelligenza artificiale, affermava: "The idea behind digital computers may be explained by saying that these machines are intended to carry out any operations which could be done by a human computer" nel suo *Computing Machinery and Intelligence*, in *Mind*, New Series, 1950, p. 436.

[14] Sull'utilizzo della metafora come tecnica giuridica e sui limiti di questa tecnica, si rinvia alle belle pagine di Galgano, *Le insidie del linguaggio giuridico. Saggio sulle metafore nel diritto*, Bologna, 2010.

[15] Sul tema del riconoscimento della personalità giuridica, Pagallo, *The Laws of Robots*, Springer, 2013; Sartor, *Cognitive automata and the law: electronic contracting and the intentionally of software agents*, in *Artificial Intelligence Law*, 2009, pp. 253-290; Teubner, *Rights of Non-Humans? Electronic Agents and Animals as New Actors in Politics and Law*, in *Journal of Law and Society*, 2016, pp. 497-521. Teubner approfondisce il tema nel recente volume *Soggetti giuridici digitali? Sullo status privatistico degli agenti software*, Napoli, 2019. In senso contrario, invece, Bertolini-Aiello, *Robot companions: A legal and ethical analysis*, in *Information Society*, 2018, pp. 130-140; Bertolini, *Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules*, in *Law, Innovation and Technology*, 1, 2013 pp. 214-247; Coppini, *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in *Politica del diritto*, 2018, pp. 713-739; Toffoletto, *IoT e intelligenza artificiale: le nuove frontiere della responsabilità civile (e del risarcimento)*, note a margine del convegno "Intelligenza artificiale e primi profili applicativi: Giustizia, IoT e Lavoratori" (Aula Magna del Palazzo di Giustizia di Milano), 17 aprile 2018. Cfr. anche il mio contributo, *La conclusione del contratto telematico mediante i* "software agents"*: un falso problema giuridico? Brevi considerazioni*, in *Contr. e impr.*, 2, 2002, pp. 500-509.

[16] Ho sviluppato più ampiamente questi argomenti nel mio *Intelligenza artificiale e responsabilità*, in *Contr. e impr.*, 2, 2020, pp. 713-731.

[17] Il rapporto tra responsabilità civile e intelligenza artificiale è esaminato approfonditamente all'interno della sezione monografica della rivista *Giurisprudenza italiana* dedicata al tema "Intelligenza artificiale e responsabilità", a cura di Ruffolo e di Gabrielli. In particolare, si vedano i contributi di Costanza, *L'intelligenza artificiale*

danno cagionato dall'applicazione di intelligenza artificiale occorre considerare che essa non disporrebbe di un patrimonio con il quale risarcire il danno. Si potrebbe costituire un patrimonio da riservare all'applicazione di intelligenza artificiale, proprio allo scopo di consentire il risarcimento del danno [18]. Tuttavia, se si vuole preservare un patrimonio a questo scopo, non è necessario costruire il complesso edificio della soggettività dell'applicazione. Si può comunque costituire un patrimonio riservato al risarcimento di queste tipologie di danni.

A differenza di quanto accadde per la persona giuridica [19], in questo caso sembra che l'attribuzione della soggettività giuridica alle applicazioni di IA sia una costruzione che aumenti la com-

e gli stilemi della responsabilità civile, pp. 1686-1689, che individua la responsabilità in capo al soggetto più vicino al fatto lesivo causato dall'applicazione di intelligenza artificiale, toccando un tema ampiamente approfondito anche da Franzoni, *La "vicinanza della prova", quindi…*, in *Contr. e impr.*, 2016, p. 360 ss.; di Gambini, *Algoritmi e sicurezza*, p. 1726-1740, che affronta il tema della sicurezza nel settore dell'intelligenza artificiale esaminando le soluzioni già esistenti e adottate nell'ambito dei servizi della società dell'informazione e dei trattamenti automatizzati di dati personali; di Ruffolo, *Intelligenza artificiale*, machine learning *e responsabilità da algoritmo*, pp. 1689-1704, che vaglia le possibilità di individuazione e qualificazione della responsabilità da intelligenza artificiale alla luce del Codice civile, esaminando altresì le ipotesi di riconoscimento di personalità giuridica. Con particolare riferimento al tema della responsabilità civile delle c.d. *autonomous car*, si rinvia a Calabresi-Al Mureden, Driverless Cars. *Intelligenza artificiale e futuro della mobilità*, Bologna, 2021 e Al Mureden, Autonomous cars *e responsabilità civile tra disciplina vigente e prospettive* de iure condendo, in *Contr. e impr.*, 3, 2019, pp. 895-924.

Approfondisce il tema della responsabilità derivante dalle applicazioni di sistemi di intelligenza artificiale anche il volume curato da Pizzetti, *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, 2018 e in particolare il contributo di Bassini-Liguori-Pollicino, *Sistemi di Intelligenza Artificiale, responsabilità e* accountability. *Verso nuovi paradigmi?*, pp. 333-371, nel quale gli autori offrono una panoramica dei diversi orientamenti espressi da dottrina e giurisprudenza in relazione al tema della responsabilità civile e penale derivante da eventuali danni causati da sistemi di intelligenza artificiale; Massolo, *Responsabilità civile e IA*, pp. 373-382, valuta invece l'idoneità del quadro giuridico nazionale ed europeo a regolare i rapporti civili nell'era dell'intelligenza artificiale.

Per una rassegna delle problematiche sollevate, v. anche il volume curato da De Franceschi e Schulze, *Digital Revolution. New challenges for Law*, C.H.Beck-Nomos, 2019, e in particolare i contributi di Mazzini, *A System of Governance for Artificial Intelligence through the Lens of Emerging Intersections between AI and EU Law*, pp. 245-298, e di Mezzanotte, *Risk Allocation and Liability Regimes in the IoT*, pp. 169-189; Rachum-Twaig, *Whose Robot Is It Anyway?: Liability for Artificial-Intelligence-Based Robots*, in *University of Illinois Law Review*, 2020; Weber-Staiger, *New Liability Patterns in the Digital Era*, in Synodinou-Jougleux-Markou-Prastitou (a cura di), *EU Internet Law*, Springer, 2017, pp. 197-214.

[18] Cfr. Pagallo, *Robotrust and Legal Responsibility*, in *Know Techn Pol*, 2010, pp. 367-379; Sartor, *Gli agenti* software*: nuovi soggetti del ciberdiritto?*, in *Contr. e impr.*, 2, 2002, pp. 465-499.

[19] Cfr. Galgano, *cit.*, il quale attribuisce alla metafora il ruolo di semplificazione linguistica della complessità dei rapporti giuridici e, citando a sua volta D'Alessandro, *Persone giuridiche e analisi del linguaggio*, Padova, 1989, p. 70 ss., afferma: "la persona giuridica è, pertanto, un'entità 'rilevabile solo sulla scena giuridica verbalizzata', mentre 'sulla scena effettuale o esistenziale non v'è che un soggetto di diritto: l'uomo'. Ma, individuato l'ambito entro il quale il concetto di persona giuridica assolve la propria funzione, non si può fare a meno di apprezzare il valore di questo concetto, che nessuna parafrasi — 'per quanto complessa e ingegnosa' — potrebbe sostituire; a esso si deve guardare, anzi, come a una "ammirevole creazione originale del linguaggio giuridico" (p. 49).

plessità giuridica piuttosto che diminuirla [20]. Soprattutto l'attribuzione della soggettività all'applicazione di intelligenza artificiale non risolve il problema più complesso: quello di individuazione dei criteri di allocazione della responsabilità, che sembra costituire il vero nodo della questione. È necessario, infatti, elaborare un nuovo modello di responsabilità che dovrebbe completamente affrancarsi da condizionamenti di natura soggettiva [21] e che sia piuttosto un modello di allocazione del rischio [22].

3. Il modello europeo.

La proposta di Regolamento europeo sull'intelligenza artificiale si inserisce, dunque, all'interno di un disegno strategico, articolato fondamentalmente in quattro ambiti: quello della protezione dei dati personali, con il già menzionato Regolamento (UE) 2016/679; quello dei servizi digitali e del mercato digitale, con il *Digital Services Act* [23] e il *Digital Markets Act* [24]; quello che riguarda l'identità digitale, con la revisione del Regolamento e-IDAS del

[20] Cfr. Di Giovanni, *Attività contrattuale e intelligenza artificiale*, in *Giur. It.*, 2019, pp. 1677-1686.

[21] Pure presenti nell'elaborazione della Risoluzione del Parlamento europeo del febbraio 2017, che, come osserva Ruffolo, *Per i fondamenti di un diritto alla robotica* self-learning*; dalla* machinery *produttiva all'auto* driverless*: verso una "responsabilità da algoritmo"?*, in Ruffolo (a cura di), *Intelligenza artificiale e responsabilità*, Milano, 2017, pp. 11-12: "pare non affrancarsi del tutto da una certa centralità dell'elemento soggettivo colposo come criterio di attribuzione della responsabilità per lesioni cagionate da robot *self-learning*". Nella Risoluzione, infatti, al punto 59, lett. *f)*, si invita la Commissione europea a valutare "l'istituzione di uno status giuridico specifico per i robot nel lungo termine, di modo che almeno i robot autonomi più sofisticati possano essere considerati come persone elettroniche responsabili di risarcire qualsiasi danno da loro causato, nonché eventualmente il riconoscimento della personalità elettronica dei robot che prendono decisioni autonome o che interagiscono in modo indipendente con terzi".

[22] Ho sviluppato più ampiamente queste considerazioni nel mio *Intelligenza artificiale e responsabilità*, op. cit. Certamente un principio che può rivelarsi di grande utilità è quello basato sull'*accountability*. Sul punto v. anche Comandé, *Intelligenza artificiale e responsabilità tra* liability *e* accountability, in *Analisi giuridica dell'economia*, 1, 2019, pp. 169-188; Costanza, *cit.*, p. 1689 e il mio contributo, *L'*accountability *nel Regolamento europeo*, in Barba-Pagliantini (a cura di), *Commentario del Codice Civile delle persone*, Torino, 2019. In favore di un regime di "strict liability", v. Vladeck, *Machines without Principals: Liability Rules and Artificial Intelligence*, in *Wash. L. Rev.*, 2014, pp. 117-150, il quale tuttavia propende per la creazione della soggettività pura delle applicazioni autonome. Per approfondimenti sulla responsabilità dell'organizzazione nel suo complesso, con riguardo al compimento di atti informatici, si consenta di rinviare al mio volume, *I contratti informatici*, in Galgano (diretto da), *Tratt. dir. comm. e dir. pubbl. econ.*, XXII, Padova, 1997, in particolare p. 187 ss. Già Calabresi parlava di "*enterprise liability*" nel suo *Some Thoughts on Risk Distribution and the Law of Torts*, in *Yale Law Journal*, 1961, pp. 449-553.

[23] Proposta di Regolamento del Parlamento europeo e del Consiglio relativo a un mercato unico dei servizi digitali (legge sui servizi digitali) e che modifica la direttiva 2000/31/CE, 15 dicembre 2020, COM(2020) 825 *final*, disponibile *on line* al seguente *link: https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:52020PC0825* (consultato il 28 febbraio 2022).

[24] Proposta di Regolamento del Parlamento europeo e del Consiglio relativo a mercati equi e contendibili nel settore digitale (legge sui mercati digitali), 15 dicembre 2020, COM(2020) 842 *final*, disponibile *on line* al seguente *link*: *https://eur-lex.europa*

2014 [25] e infine il modello in commento che riguarda, appunto, l'intelligenza artificiale.

L'obiettivo strategico è quello di assicurare la costituzione di un mercato unico digitale europeo, e consentire all'Unione europea "di essere un leader mondiale nello sviluppo di un'intelligenza artificiale sicura, affidabile ed etica" [26] anche attraverso l'elaborazione di un modello normativo. Gli obiettivi specifici sono i seguenti:

"— assicurare che i sistemi di IA immessi sul mercato dell'Unione e utilizzati siano sicuri e rispettino la normativa vigente in materia di diritti fondamentali e i valori dell'Unione;

— assicurare la certezza del diritto per facilitare gli investimenti e l'innovazione nell'intelligenza artificiale;

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

— facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure e affidabili nonché prevenire la frammentazione del mercato" [27].

La proposta di Regolamento in materia di intelligenza artificiale è una proposta molto articolata che consta di ben 85 articoli [28].

È stato scelto un approccio orizzontale, che investe tutti i settori. Secondo la Commissione, si è scelta una metodologia che prevede un quadro normativo soltanto per i sistemi di IA ad alto rischio, con la possibilità per tutti i fornitori di sistemi di IA non ad alto rischio di seguire un codice di condotta. "I requisiti riguarderanno i dati, la documentazione e la tracciabilità, la

.eu/legal-content/it/TXT/?uri=COM:2020:842:FIN (consultato il 28 febbraio 2022).

[25] Proposta di Regolamento del Parlamento europeo e del Consiglio che modifica il regolamento (UE) n. 910/2014 per quanto riguarda l'istituzione di un quadro per un'identità digitale europea, 3 giugno 2021, COM(2021) 281 *final*, disponibile *on line* al seguente *link*: *https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:52021PC0281* (consultato il 28 febbraio 2022).

[26] Così la Relazione della Commissione europea che accompagna la proposta di Regolamento, p. 2.

[27] Così la Relazione della Commissione europea che accompagna la proposta di Regolamento, p. 3.

[28] Sulla proposta di Regolamento, cfr. Alpa, *Quale modello normativo europeo per l'intelligenza artificiale?*, in *Contr. e impr.*, 2021, p. 1003-10026; Ebers, S*tandandizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act*, in Di Matteo-Cannarsa-Poncibò (a cura di), *Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, in corso di pubblicazione, Cambridge, 2022; Floridi, *The European Legislation on AI: a Brief Analysis of its Philosophical Approach*, in *Philos. Technol*, 2021, p. 215-222; MacCarthy-Propp, *Machines learn that Brussels writes the rules: The EU's new AI regulation. Editor's Note*, in *Brookings.edu*, 2021; Mcfadden-Jones-Taylor-Osborn, *Harmonising Artificial Intelligence; The role of standards in the EU AI Regulation Harmonising Artificial Intelligence*, Oxford Commission on AI & Good, Oxford Information Labs, 2021; Mökander-Axente-Casolari-Floridi, *Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation*, in *Minds and Machines*, 2021; Towsend, *Decoding the Proposed European Union Artificial Intelligence Act*, in *Insights*, 2021; Veale-Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, in *Computer Law Review International*, 2021, pp. 97-112.

fornitura di informazioni e la trasparenza, la sorveglianza umana nonché la robustezza e la precisione e saranno obbligatori per i sistemi di IA ad alto rischio. Le imprese che introducessero codici di condotta per altri sistemi di IA lo farebbero su base volontaria" [29].

Dunque, la proposta disciplina l'immissione nel mercato, la messa a disposizione, la messa in servizio e l'uso dei sistemi di intelligenza artificiale, come recita l'art. 1, in generale, e non in un settore specifico.

La definizione di intelligenza artificiale, coerentemente con quest'approccio, è estremamente generale [30].

Il modello adottato dalla Commissione è un modello basato sul rischio, come si afferma nella stessa relazione alla proposta di Regolamento, differenziando tra gli usi dell'IA che creano rispettivamente un rischio inaccettabile, un rischio alto e un rischio basso o minimo.

Il legislatore prevede che ci siano applicazioni di intelligenza artificiale che debbano essere vietate e le elenca all'art. 5 del Regolamento [31]. Per esempio, vieta la messa in servizio di sistemi che utilizzino tecniche subliminali, nonché la messa in servizio di sistemi che sfruttino le vulnerabilità di un gruppo di persone.

In particolare, con riguardo a queste pratiche di intelligenza

[29] Così la Relazione della Commissione europea che accompagna la proposta di Regolamento, p. 11.

[30] L'art. 3 definisce sistema di intelligenza artificiale: "un *software* sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie di obiettivi definiti dall'uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono". L'all. I così li elenca:

"*a*) Approcci di apprendimento automatico, compresi l'apprendimento supervisionato, l'apprendimento non supervisionato e l'apprendimento per rinforzo, con utilizzo di un'ampia gamma di metodi, tra cui l'apprendimento profondo (*deep learning*);

b) approcci basati sulla logica e approcci basati sulla conoscenza, compresi la rappresentazione della conoscenza, la programmazione induttiva (logica), le basi di conoscenze, i motori inferenziali e deduttivi, il ragionamento (simbolico) e i sistemi esperti;

c) approcci statistici, stima bayesiana, metodi di ricerca e ottimizzazione".

[31] Recita, in particolare, l'art. 5, 1° comma: "Sono vietate le pratiche di intelligenza artificiale seguenti:

a) l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che utilizza tecniche subliminali che agiscono senza che una persona ne sia consapevole al fine di distorcerne materialmente il comportamento in un modo che provochi o possa provocare a tale persona o a un'altra persona un danno fisico o psicologico;

b) l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che sfrutta le vulnerabilità di uno specifico gruppo di persone, dovute all'età o alla disabilità fisica o mentale, al fine di distorcere materialmente il comportamento di una persona che appartiene a tale gruppo in un modo che provochi o possa provocare a tale persona o a un'altra persona un danno fisico o psicologico;

c) l'immissione sul mercato, la messa in servizio o l'uso di sistemi di IA da parte delle autorità pubbliche o per loro conto ai fini della valutazione o della classificazione dell'affidabilità delle persone fisiche per un determinato periodo di tempo sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste, in cui il punteggio sociale così ottenuto comporti il verificarsi di uno o di entrambi i seguenti scenari: i) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi

artificiale, il legislatore europeo inserisce il conseguimento dell'obiettivo come fine caratterizzante l'applicazione vietata: essa deve essere volta a conseguire determinati obiettivi, quali distorcere materialmente il comportamento o provocare un danno fisico o psicologico [32].

È vietata altresì l'immissione nel mercato e la messa in servizio di sistemi di punteggio sociale, lesivi della dignità umana [33].

È infine vietato anche il ricorso a sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico, particolarmente invasivo dei diritti e delle libertà delle persone interessate, a fini di attività di contrasto [34].

La criticità sta naturalmente nelle eccezioni previste dal legislatore, ad esempio per i casi in cui l'uso dei sistemi di identificazione biometrica remota sia strettamente necessario per "la ricerca mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi" o per "la prevenzione di una minaccia specifica,

di persone fisiche in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti; ii) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi di persone fisiche che sia ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla sua gravità;

d) l'uso di sistemi di identificazione biometrica remota 'in tempo reale' in spazi accessibili al pubblico a fini di attività di contrasto, a meno che e nella misura in cui tale uso sia strettamente necessario per uno dei seguenti obiettivi: *i*) la ricerca mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi; *ii*) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di un attacco terroristico; *iii*) il rilevamento, la localizzazione, l'identificazione o l'azione penale nei confronti di un autore o un sospettato di un reato di cui all'articolo 2, paragrafo 2, della decisione quadro 2002/584/GAI del Consiglio, punibile nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno tre anni, come stabilito dalla legge di tale Stato membro".

[32] Così l'art. 5, 1° comma, lett. *a*) e *b*). Così il considerando n. 16: "Si tratta di azioni compiute con l'intento di distorcere materialmente il comportamento di una persona, in un modo che provoca o può provocare un danno a tale persona o a un'altra".

[33] Chiarisce il considerando n. 17: "I sistemi di IA che forniscono un punteggio sociale delle persone fisiche per finalità generali delle autorità pubbliche o di loro rappresentanti possono portare a risultati discriminatori e all'esclusione di determinati gruppi. Possono inoltre ledere il diritto alla dignità e alla non discriminazione e i valori di uguaglianza e giustizia. Tali sistemi di IA valutano o classificano l'affidabilità delle persone fisiche sulla base del loro comportamento sociale in molteplici contesti o di caratteristiche personali o della personalità note o previste. Il punteggio sociale ottenuto da tali sistemi di IA può determinare un trattamento pregiudizievole o sfavorevole di persone fisiche o di interi gruppi in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti, o a un trattamento pregiudizievole che risulta ingiustificato o sproporzionato rispetto alla gravità del loro comportamento sociale. È pertanto opportuno vietare tali sistemi di IA".

[34] Così il considerando n. 18: "L'uso di sistemi di IA di identificazione biometrica remota 'in tempo reale' delle persone fisiche in spazi accessibili al pubblico a fini di attività di contrasto è ritenuto particolarmente invasivo dei diritti e delle libertà delle persone interessate, nella misura in cui potrebbe avere ripercussioni sulla vita privata di un'ampia fetta della popolazione, farla sentire costantemente sotto sorveglianza e scoraggiare in maniera indiretta l'esercizio della libertà di riunione e di altri diritti fondamentali. L'immediatezza dell'impatto e le limitate opportunità di eseguire ulteriori controlli o apportare correzioni in relazione all'uso di tali sistemi che operano 'in tempo reale' comportano inoltre un aumento dei rischi per quanto concerne i diritti e le libertà delle persone oggetto di attività di contrasto".

sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di un attacco terroristico" [35]. Su questo il dibattito è aperto, dal momento che le eccezioni, se troppo ampie, possono finire con il vanificare quanto vietato in altre disposizioni.

La proposta di Regolamento europeo, dall'art. 6 in poi, classifica i sistemi di intelligenza artificiale ad alto rischio. Prevede che se i sistemi di intelligenza artificiale sono sistemi ad alto rischio, allora il soggetto che li produce e che li immette nel mercato deve adempiere a una serie di prescrizioni e procedere ad una valutazione della conformità *ex ante*.

Sono sistemi classificati ad alto rischio: i sistemi destinati ad essere utilizzati come componenti di sicurezza di prodotti soggetti a valutazione di conformità *ex ante* da parte di terzi; altri sistemi di IA indipendenti, che presentano implicazioni in relazione ai diritti fondamentali, elencati all'allegato III del Regolamento [36].

Gli obblighi e i requisiti che un sistema di intelligenza artificiale ad alto rischio deve rispettare sono dettati dall'art. 8 e seguenti.

Secondo l'art. 8, nel garantire conformità a tali obblighi e requisiti, si deve tenere conto di due elementi: la finalità prevista

[35] Così l'art. 5, 1° comma, lett. *d*).

[36] Così la Relazione della Commissione europea che accompagna la proposta di Regolamento, p. 14. Sono classificati ad alto rischio, ad esempio, i sistemi di IA destinati a essere utilizzati per l'identificazione biometrica remota "in tempo reale" e "a posteriori" delle persone fisiche e quelli destinati a essere utilizzati come componenti di sicurezza nella gestione del traffico stradale e nella fornitura di acqua, gas, riscaldamento ed elettricità. Ancora, l'allegato III fa riferimento ai sistemi di IA destinati a essere utilizzati al fine di determinare l'accesso o l'assegnazione agli istituti di istruzione e formazione professionale, ai sistemi di IA destinati a essere utilizzati per l'assunzione o la selezione di persone fisiche e all'IA destinata a essere utilizzata per adottare decisioni in materia di promozione e cessazione dei rapporti contrattuali di lavoro, per l'assegnazione dei compiti, nonché per il monitoraggio e la valutazione delle prestazioni e del comportamento delle persone nell'ambito di tali rapporti di lavoro. Sono classificati ad alto rischio anche i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche o per conto di autorità pubbliche per valutare l'ammissibilità delle persone fisiche alle prestazioni e ai servizi di assistenza pubblica, i sistemi di IA destinati a essere utilizzati per valutare l'affidabilità creditizia delle persone fisiche, a eccezione dei sistemi di IA messi in servizio per uso proprio da fornitori di piccole dimensioni, e i sistemi di IA destinati a essere utilizzati per inviare servizi di emergenza di primo soccorso o per stabilire priorità in merito all'invio di tali servizi. Sono poi presi in considerazione i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per varie finalità, tra cui determinare il rischio di reato o recidiva in relazione a una persona fisica o il rischio per potenziali vittime di reati, rilevare lo stato emotivo di una persona fisica, valutare l'affidabilità degli elementi probatori nel corso delle indagini o del perseguimento di reati o individuare i "*deep fake*", ossia, come si dirà, i sistemi di IA che generano o manipolano immagini o contenuti audio o video in modo che persone, oggetti, luoghi o altre entità o eventi esistenti potrebbero apparire falsamente autentici o veritieri. Sono inclusi nell'elenco dei sistemi di IA ad alto rischio anche quelli destinati a essere utilizzati dalle autorità pubbliche competenti per rilevare lo stato emotivo di una persona fisica, per valutare un rischio posto da una persona fisica che intende entrare o è entrata nel territorio di uno Stato membro, per verificare l'autenticità dei documenti di viaggio delle persone fisiche, per l'esame delle domande di asilo, di visto e di permesso di soggiorno e dei relativi reclami. Infine, l'allegato III fa ri-

del sistema di IA ad alto rischio e il sistema di gestione dei rischi, oggetto di una specifica disciplina dettata dall'art. 9.

In primo luogo è necessario istituire, attuare, documentare e mantenere un sistema di gestione dei rischi, ossia un processo iterativo continuo, eseguito nel corso dell'intero ciclo di vita del sistema di IA ad alto rischio, che comprenda l'identificazione e l'analisi dei rischi noti e prevedibili associati al sistema, la stima e la valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio sia usato conformemente alla propria finalità e in condizioni di uso improprio ragionevolmente prevedibile, la valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato, nonché l'adozione di adeguate misure di gestione dei rischi [37].

Successivamente, sono individuati alcuni criteri di qualità per i *set* di dati impiegati a fini di addestramento, convalida e prova nell'ambito dei sistemi di IA ad alto rischio [38]. In particolare, tali *set* di dati devono essere sottoposti a specifiche pratiche di *governance* e gestione ed essere pertinenti, rappresentativi, esenti da errori, completi nonché statisticamente appropriati [39].

Ancora, è richiesto che, prima dell'immissione sul mercato o

ferimento ai sistemi di IA destinati ad assistere un'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto e nell'applicazione della legge a una serie concreta di fatti.

[37] Così l'art. 9, 1° e 2° comma. I commi dal 4° al 7° introducono alcune specificazioni con riferimento alle misure di gestione dei rischi, che devono essere individuate tenendo conto dello stato dell'arte generalmente riconosciuto e devono essere tali che qualsiasi rischio residuo associato a ciascun pericolo nonché il rischio residuo complessivo dei sistemi di IA ad alto rischio siano considerati accettabili, a condizione che il sistema di IA ad alto rischio sia usato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile. I rischi residui dovranno essere comunicati all'utente. In ogni caso, in virtù del 5° comma, al fine di individuare le misure di gestione dei rischi più appropriate, i sistemi di IA ad alto rischio sono sottoposti a prove che garantiscano che i sistemi funzionino in modo coerente per la finalità prevista e siano conformi ai requisiti individuati dal Regolamento. Tali prove devono essere effettuate, a seconda dei casi, in un qualsiasi momento dell'intero processo di sviluppo e, in ogni caso, prima dell'immissione sul mercato o della messa in servizio, sulla base di metriche e soglie probabilistiche definite in via preliminare e adeguate alla finalità perseguita dal sistema di IA ad alto rischio.

[38] Secondo le definizioni dell'art. 3, costituiscono "dati di addestramento" i dati "utilizzati per addestrare un sistema di IA adattandone i parametri che può apprendere, compresi i pesi di una rete neurale"; sono "dati di convalida" i dati "utilizzati per fornire una valutazione del sistema di IA addestrato e per metterne a punto, tra l'altro, i parametri che non può apprendere e il processo di apprendimento, al fine di evitare l'eccessivo adattamento ai dati di addestramento (*overfitting*)"; sono "dati di prova" i dati "utilizzati per fornire una valutazione indipendente del sistema di IA addestrato e convalidato al fine di confermarne le prestazioni attese prima della sua immissione sul mercato o messa in servizio".

[39] Così l'art. 10. In particolare, ai sensi del 2° comma, le pratiche di *governance* e gestione dei dati devono riguardare, in particolare: "*a*) le scelte progettuali pertinenti; *b*) la raccolta dei dati; *c*) le operazioni di trattamento pertinenti ai fini della preparazione dei dati, quali annotazione, etichettatura, pulizia, arricchimento e aggregazione; *d*) la formulazione di ipotesi pertinenti, in particolare per quanto riguarda le informazioni che si presume che i dati misurino e rappresentino; *e*) una valutazione preliminare della disponibilità, della quantità e dell'adeguatezza dei set di dati necessari; *f*)

della messa in servizio del sistema di IA ad alto rischio, ne sia redatta la relativa documentazione tecnica, deputata a dimostrare che il sistema di IA ad alto rischio sia conforme ai requisiti stabiliti dal Regolamento e a fornire alle autorità e agli organismi competenti tutte le informazioni necessarie per valutare tale conformità [40].

Inoltre, i sistemi di IA ad alto rischio devono essere progettati e sviluppati in modo da garantire, tramite la registrazione automatica degli eventi e per tutto il ciclo di vita, la tracciabilità del proprio funzionamento [41], che deve risultare sufficientemente trasparente da consentire agli utenti di interpretarne l'*output* e utilizzarlo adeguatamente. In quest'ottica, i sistemi di IA ad alto rischio devono essere accompagnati da istruzioni per l'uso, in un formato digitale o non digitale, che comprendano informazioni concise, complete, corrette, chiare e pertinenti, accessibili e comprensibili per gli utenti [42].

I sistemi di IA ad alto rischio devono poi essere progettati e sviluppati con strumenti di interfaccia uomo-macchina che ne consentano un'efficace supervisione da parte delle persone fisiche, volta a prevenire o ridurre al minimo i rischi per la salute, la

un esame atto a valutare le possibili distorsioni; *g*) l'individuazione di eventuali lacune o carenze nei dati e il modo in cui tali lacune e carenze possono essere colmate". Inoltre, in base al 4° comma dell'art. 10, "i *set* di dati di addestramento, convalida e prova tengono conto, nella misura necessaria per la finalità prevista, delle caratteristiche o degli elementi particolari dello specifico contesto geografico, comportamentale o funzionale all'interno del quale il sistema di IA ad alto rischio è destinato a essere usato".

[40] Così l'art. 11, che rimanda all'allegato IV della proposta per l'individuazione del contenuto minimo della documentazione tecnica.

[41] Così l'art. 12, che specifica come le capacità di registrazione debbano consentire, in particolare, di monitorare il funzionamento del sistema di IA ad alto rischio per quanto riguarda il verificarsi di situazioni che possono far sì che il sistema presenti un rischio per la salute, la sicurezza o la tutela dei diritti fondamentali delle persone o comporti una modifica sostanziale, e devono agevolare il monitoraggio successivo all'immissione sul mercato. Inoltre, la disposizione in esame individua i dati necessariamente oggetto di registrazione con riferimento ai sistemi di IA per l'identificazione biometrica remota "in tempo reale" e "a posteriori" delle persone fisiche, tra cui: la registrazione del periodo di ciascun utilizzo del sistema; la banca dati di riferimento con cui il sistema ha verificato i dati di *input*; i dati di *input* per i quali la ricerca ha portato a una corrispondenza; l'identificativo delle persone fisiche che partecipano alla verifica dei risultati.

[42] Così l'art. 13 che, tra le informazioni da fornire all'utente, include: l'identità e i dati di contatto del fornitore e, ove applicabile, del suo rappresentante autorizzato; le caratteristiche, le capacità e i limiti delle prestazioni del sistema di IA ad alto rischio, tra cui la finalità prevista; il livello di accuratezza, robustezza e cibersicurezza rispetto al quale il sistema di IA ad alto rischio è stato sottoposto a prova e convalidato e che ci si può attendere; qualsiasi circostanza nota e prevedibile che possa avere un impatto sul livello atteso di accuratezza, robustezza e cibersicurezza; qualsiasi circostanza nota o prevedibile, connessa all'uso del sistema di IA ad alto rischio in conformità alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile, che possa comportare rischi per la salute e la sicurezza o per i diritti fondamentali; le sue prestazioni per quanto riguarda le persone o i gruppi di persone sui quali il sistema è destinato a essere utilizzato; ove opportuno, le specifiche per i dati di *input* o qualsiasi altra informazione pertinente in termini di *set* di dati di addestramento, convalida e prova, tenendo conto della finalità prevista del sistema di IA.

sicurezza o i diritti fondamentali [43] e, infine, devono essere progettati e sviluppati in modo tale da conseguire, alla luce della propria finalità, un adeguato livello di accuratezza, robustezza e cibersicurezza, che perduri per l'intero ciclo di vita del sistema [44].

Ulteriori obblighi e requisiti per i sistemi di IA ad alto rischio derivano dalle previsioni che stabiliscono specifici adempimenti in capo ai soggetti coinvolti nella filiera di creazione e di utilizzo di tali sistemi.

Si tratta del *provider* del sistema di IA, definito come la persona fisica o giuridica, l'autorità pubblica, l'agenzia o altro organismo che sviluppa o fa sviluppare un sistema di intelligenza artificiale al fine di immetterlo sul mercato con il proprio nome o marchio, a titolo oneroso o gratuito [45]; dell'utente, vale a dire qualsiasi persona fisica o giuridica, autorità pubblica, agenzia o altro ente che utilizza, per scopi professionali e non personali, un sistema di intelligenza artificiale sotto la propria autorità [46]; del c.d. *importer*, ossia qualsiasi persona fisica o giuridica stabilita nell'Unione

[43] Così l'art. 14, in base al quale la sorveglianza umana è garantita almeno mediante una delle seguenti misure: *a*) misure individuate e integrate nel sistema di IA ad alto rischio dal fornitore prima della sua immissione sul mercato o messa in servizio, ove tecnicamente possibile; *b*) misure individuate dal fornitore prima dell'immissione sul mercato o della messa in servizio del sistema di IA ad alto rischio, adatte ad essere attuate dall'utente. Ai sensi del 4° comma dell'art. 14, queste misure, a seconda delle circostanze, devono consentire a chi sia incaricato della sorveglianza del sistema IA ad alto rischio, tra l'altro, di: *a*) comprenderne appieno le capacità e i limiti ed essere in grado di monitorarne debitamente il funzionamento, in modo che i segnali di anomalie, disfunzioni e prestazioni inattese possano essere individuati e affrontati quanto prima; *b*) restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'*output* prodotto dal sistema, c.d. "distorsione dell'automazione", in particolare per i sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche; *c*) essere in grado di interpretare correttamente l'*output* del sistema di IA ad alto rischio, tenendo conto in particolare delle caratteristiche del sistema e degli strumenti e dei metodi di interpretazione disponibili.

[44] Così l'art. 15, in base al quale i sistemi di IA ad alto rischio devono essere anche resilienti per quanto riguarda errori, guasti o incongruenze che possono verificarsi all'interno del sistema o nell'ambiente in cui esso opera, in particolare a causa della loro interazione con persone fisiche o altri sistemi Ancora, secondo l'art. 15, 4° comma, i sistemi di IA ad alto rischio devono essere resilienti ai tentativi di terzi non autorizzati di modificarne l'uso o le prestazioni sfruttando le vulnerabilità del sistema. Infine, si prevede che le soluzioni tecniche volte a garantire la cibersicurezza siano adeguate alle circostanze e ai rischi pertinenti, mentre le soluzioni tecniche finalizzate ad affrontare le vulnerabilità specifiche dell'IA devono includere, ove opportuno, misure volte a prevenire e controllare gli attacchi che cercano di manipolare il *set* di dati di addestramento (c.d. "avvelenamento dei dati"), gli *input* progettati in modo da far sì che il modello commetta un errore (c.d. "esempi antagonistici") o i difetti del modello. L'art. 15 chiude il capo 2 della proposta, espressamente dedicato ai requisiti per i sistemi di IA ad alto rischio.

[45] Così l'art. 3, 1° comma, n. 2. Accanto alla figura del *provider*, l'art. 3, n. 5 individua quella del rappresentante autorizzato, ossia qualsiasi persona fisica o giuridica stabilita nell'Unione che riceva un mandato scritto dal *provider* al fine di adempiere ed eseguire per suo conto gli obblighi e le procedure stabiliti nel regolamento. Secondo l'art. 25, prima di mettere a disposizione i propri sistemi sul mercato dell'Unione, qualora non possa essere identificato un importatore, i *provider* stabiliti al di fuori dell'Unione devono nominare, mediante mandato scritto, un rappresentante autorizzato stabilito nell'Unione.

[46] Così l'art. 3, 1° comma, n. 4.

che immette sul mercato o mette in servizio un sistema di intelligenza artificiale che reca il nome o il marchio di una persona fisica o giuridica stabilita al di fuori dell'Unione [47]; del *distributor*, qualsiasi persona fisica o giuridica nella catena di approvvigionamento, diversa dal fornitore o dall'importatore, che mette a disposizione un sistema di intelligenza artificiale sul mercato dell'Unione senza modificarne le proprietà [48].

Nel quadro delineato dalla proposta, il *provider* di un sistema di IA ad alto rischio è innanzitutto chiamato a garantire che il sistema sia conforme ai requisiti sopra descritti [49], nonché, in caso di difformità, ad adottare immediatamente le necessarie misure correttive, richiamarlo o ritirarlo dal mercato, a seconda dei casi, informando i *distributor* e, ove presente, l'*importer* [50]. La conformità del sistema di IA ad alto rischio al Regolamento dovrà essere altresì attestata mediante una procedura di valutazione, che il *provider* garantisce venga svolta prima della sua immissione sul mercato o della sua messa in servizio. Se la procedura di valutazione della conformità ha esito positivo, il *provider* deve redigere una dichiarazione di conformità e apporre la marcatura CE [51]. In ogni caso, il *provider* dovrà dimostrare la conformità del sistema di IA ad alto rischio ai requisiti del Regolamento ove richiesto da un'autorità nazionale competente [52]. Inoltre, istituisce e documenta un sistema di monitoraggio successivo all'immissione sul mercato che, proporzionato alla natura delle tecnologie e ai rischi posti dal sistema di IA, gli consenta di valutarne il costante perdurare della conformità [53].

Ancora, il *provider* è tenuto ad istituire un sistema di gestione della qualità che, tramite *policy*, procedure e istruzioni scritte, garantisca la conformità del sistema di IA al Regolamento [54]; a

[47] Così l'art. 3, 1° comma, n. 6.
[48] Così l'art. 3, 1° comma, n. 7.
[49] Così l'art. 16, 1° comma, lett. *a*).
[50] Così l'art. 16, 1° comma, lett. *g*) e l'art. 21.
[51] Così l'art. 16, 1° comma, lett. *e*) e *i*) e l'art. 19. Dispongono in merito alla valutazione di conformità, ai certificati, alla dichiarazione di conformità e alla marcatura CE il capo V del titolo terzo del Regolamento e, in particolare, gli artt. 43, 48 e 49.
[52] Così l'art. 16, 1° comma, lett. *j*) e l'art. 23, in base al quale, su richiesta motivata di un'autorità nazionale competente, i *provider* forniscono anche l'accesso ai *log* generati automaticamente dai sistemi di IA ad alto rischio che si trovino sotto il proprio controllo.
[53] Così l'art. 61, in base al quale "il sistema di monitoraggio successivo all'immissione sul mercato raccoglie, documenta e analizza attivamente e sistematicamente i dati pertinenti forniti dagli utenti o raccolti tramite altre fonti sulle prestazioni dei sistemi di IA ad alto rischio per tutta la durata del loro ciclo di vita". Tale sistema di monitoraggio si basa su un piano di monitoraggio, incluso nella documentazione tecnica di cui ogni sistema di IA ad alto rischio deve essere dotato.
[54] Così l'art. 26, 1° comma, lett. *b*) e l'art. 17, in base al quale il sistema di gestione della qualità deve essere documentato in modo sistematico e includere almeno i seguenti elementi: *a*) una strategia per la conformità normativa, compresa la conformità alle procedure di valutazione della conformità e alle procedure per la gestione delle modifiche dei sistemi di IA ad alto rischio; *b*) le tecniche, le procedure e gli interventi sistematici da utilizzare per la progettazione, il controllo della progettazione e la

redigere la documentazione tecnica del sistema [55]; a conservare i *log* generati automaticamente che siano sotto il proprio controllo [56] e, prima che il sistema di IA ad alto rischio sia immesso sul mercato o messo in servizio, a registrarlo nella apposita banca dati dell'Unione europea, ove si tratti di un sistema ad alto rischio indipendente [57]. Il *provider* ha altresì l'obbligo di informare immediatamente le autorità nazionali competenti degli Stati membri in cui ha messo a disposizione il sistema di IA ad alto rischio e, ove applicabile, l'organismo che ha rilasciato un certificato per il sistema, qualora gli sia noto che il sistema presenti un rischio per la salute, la sicurezza o la tutela dei diritti fondamentali delle persone [58], e di segnalare qualsiasi incidente grave o malfunzionamento alle autorità di vigilanza del mercato degli Stati membri in cui tali incidenti o violazioni si siano verificati [59].

Quanto agli *importer*, è previsto che, prima di immettere sul mercato un sistema di IA ad alto rischio, garantiscano che il *provider* abbia eseguito l'appropriata procedura di valutazione della conformità e redatto la relativa documentazione tecnica. Gli

verifica della progettazione del sistema di IA ad alto rischio; *c*) le tecniche, le procedure e gli interventi sistematici da utilizzare per lo sviluppo e per il controllo e la garanzia della qualità del sistema di IA ad alto rischio; *d*) le procedure di esame, prova e convalida da effettuare prima, durante e dopo lo sviluppo del sistema di IA ad alto rischio e la frequenza con cui devono essere effettuate; *e*) le specifiche tecniche, comprese le norme, da applicare e, qualora le pertinenti norme armonizzate non siano applicate integralmente, i mezzi da usare per garantire che il sistema di IA ad alto rischio sia conforme ai requisiti stabiliti dal regolamento; *f*) i sistemi e le procedure per la gestione dei dati, compresa la raccolta, l'analisi, l'etichettatura, l'archiviazione, la filtrazione, l'estrazione, l'aggregazione, la conservazione dei dati e qualsiasi altra operazione riguardante i dati effettuata prima e ai fini dell'immissione sul mercato o della messa in servizio di sistemi di IA ad alto rischio; *g*) il sistema di gestione dei rischi; *h*) la predisposizione, l'attuazione e la manutenzione di un sistema di monitoraggio successivo all'immissione sul mercato; *i*) le procedure relative alla segnalazione di incidenti gravi e di malfunzionamenti; *j*) la gestione della comunicazione con le autorità nazionali competenti, le autorità competenti, comprese quelle settoriali, che forniscono o sostengono l'accesso ai dati, gli organismi notificati, altri operatori, clienti o altre parti interessate; *k*) i sistemi e le procedure per la conservazione delle registrazioni e di tutte le informazioni e la documentazione pertinenti; *l*) la gestione delle risorse, comprese le misure relative alla sicurezza dell'approvvigionamento; *m*) un quadro di responsabilità che definisca le responsabilità della dirigenza e di altro personale per quanto riguarda tutti gli aspetti elencati nel regolamento. In virtù del 2° comma dell'art. 17, l'attuazione di questi aspetti è proporzionata alle dimensioni dell'organizzazione del fornitore.

[55] Così l'art. 18. La documentazione tecnica cui si fa riferimento è quella di cui all'art. 11, sopra menzionata.

[56] Così l'art. 20, che, al 1° comma, recita: "I fornitori di sistemi di IA ad alto rischio conservano i *log* generati automaticamente dai loro sistemi di IA ad alto rischio, nella misura in cui tali *log* sono sotto il loro controllo in virtù di un accordo contrattuale con l'utente o in forza di legge. I *log* sono conservati per un periodo adeguato alla luce della finalità prevista del sistema di IA ad alto rischio e degli obblighi giuridici applicabili a norma del diritto dell'Unione o nazionale".

[57] Così l'art. 16, 1° comma, lett. *f*) e l'art. 51. Reca, invece, la disciplina della banca dati dell'Unione europea l'art. 61, in base al quale la Commissione europea, in collaborazione con gli Stati membri, istituisce e mantiene una banca dati contenente informazioni relative ai sistemi di IA ad alto rischio indipendenti. Tali informazioni, individuate all'allegato VIII del regolamento, sono accessibili al pubblico.

[58] Così l'art. 22.

[59] Così l'art. 62.

importer garantiscono anche che il sistema rechi la necessaria marcatura di conformità e sia accompagnato dalla documentazione e dalle istruzioni per l'uso necessarie. Inoltre, ove ritenga o abbia motivo di ritenere che un sistema di IA ad alto rischio non sia conforme al Regolamento, l'*importer* è tenuto a non immetterlo sul mercato fino a quando non sia stato reso conforme e, qualora il sistema presenti un rischio per la salute, la sicurezza o la tutela dei diritti fondamentali delle persone, deve informarne il *provider* e le autorità di vigilanza del mercato [60].

Obblighi similari ricadono, a cascata, sul *distributor*, che prima di mettere a disposizione sul mercato un sistema di IA ad alto rischio, non solo verifica che il sistema rechi la marcatura CE, la documentazione tecnica e le istruzioni per l'uso, ma, in generale, che il *provider* e l'*importer* abbiano rispettato gli obblighi del Regolamento [61].

Si giunge così agli obblighi degli utenti, tenuti ad utilizzare i sistemi di IA ad alto rischio conformemente alle istruzioni per l'uso, a monitorare il funzionamento del sistema e ad informare il *provider* o il *distributor*, nonché a sospendere l'uso del sistema, qualora abbiano motivo di ritenere che questo presenti un rischio per la salute, la sicurezza o la tutela dei diritti fondamentali delle persone o nel caso in cui individuino un incidente grave o un malfunzionamento [62].

A chiusura di questa breve disamina, giova rilevare come, in

[60] Così l'art. 26, 1° e 2° comma. In virtù del 3° e del 4° comma, l'*importer* deve altresì indicare il proprio nome, la propria denominazione commerciale registrata o il proprio marchio registrato e l'indirizzo al quale può essere contattato sul sistema di IA ad alto rischio oppure, ove ciò non sia possibile, sull'imballaggio o in un documento di accompagnamento, e deve garantire che, fintantoché un sistema di IA ad alto rischio sia sotto la propria responsabilità, le condizioni di stoccaggio o di trasporto, ove applicabili, non pregiudichino la conformità ai requisiti previsti dal regolamento. Infine, il 5° comma istituisce anche in capo a tali soggetti un dovere di cooperazione con le autorità nazionali competenti, a cui, su richiesta motivata, l'*importer* fornisce tutte le informazioni e la documentazione necessarie per dimostrare la conformità del sistema di IA ad alto rischio ai requisiti del regolamento, compreso l'accesso ai *log* generati automaticamente che si trovino sotto il suo controllo.

[61] Così l'art. 27, in base al quale anche il *distributor*, al pari dell'*importer*, qualora ritenga o abbia motivo di ritenere che un sistema di IA ad alto rischio non sia conforme ai requisiti del regolamento, non lo mette a disposizione sul mercato fino a quando non sia stato reso conforme e, qualora il sistema presenti un rischio per la salute, la sicurezza o la tutela dei diritti fondamentali delle persone, ne informa il *provider* o l'*importer* e le autorità competenti. Inoltre, il *distributor*, al pari del *provider*, è chiamato ad attuare misure correttive in caso di difformità del sistema di IA rispetto ai requisiti del regolamento. Infine, anche il *distributor* è destinatario di un dovere di collaborazione con le autorità nazionali ed è chiamato a garantire che fintantoché un sistema di IA ad alto rischio sia sotto la propria responsabilità, le condizioni di stoccaggio o di trasporto non pregiudichino la conformità del sistema ai requisiti del regolamento.

[62] Così l'art. 29, in base al quale l'utente è anche tenuto a garantire che i dati di *input* siano pertinenti rispetto alla finalità prevista del sistema di IA ad alto rischio, ove eserciti su di questi un controllo, e a conservare i *log* generati automaticamente dai sistemi di IA ad alto rischio, nella misura in cui tali *log* siano sotto il suo controllo.

virtù dell'art. 28, qualsiasi *distributor*, *importer*, utente o altro terzo è considerato un *provider* ai fini del Regolamento, ed è soggetto ai relativi obblighi, se immette sul mercato o mette in servizio un sistema di IA ad alto rischio con il proprio nome o marchio; se modifica la finalità prevista di un sistema di IA ad alto rischio già immesso sul mercato o messo in servizio e, infine, se apporta una modifica sostanziale al sistema di IA ad alto rischio. Ove ricorrano gli ultimi due casi, il *provider* che ha inizialmente immesso sul mercato o messo in servizio il sistema di IA ad alto rischio, ai fini del Regolamento, non è più considerato tale.

Anche per i sistemi di IA diversi da quelli ad alto rischio sono previsti alcuni obblighi, che si sostanziano nei requisiti minimi di trasparenza stabiliti per determinati sistemi [63]. È il caso, ad esempio, dei sistemi di IA destinati a interagire con le persone fisiche, che i *provider* garantiscono siano progettati e sviluppati in modo che le persone fisiche siano informate di stare interagendo con un sistema di IA, a meno che non risulti evidente dalle circostanze e dal contesto di utilizzo. Ancora, è il caso dei sistemi di riconoscimento delle emozioni o di categorizzazione biometrica, per i quali è prescritto agli utenti di informare delle loro modalità di funzionamento le persone fisiche che vi siano esposte, o dei c.d. "deep fake", sistemi che generano o manipolano immagini o contenuti audio o video che assomigliano notevolmente a persone, oggetti, luoghi o altre entità o eventi esistenti e che potrebbero apparire falsamente autentici o veritieri, per i quali è previsto che gli utenti rendano noto che il contenuto è stato generato o manipolato artificialmente.

4. Conclusioni.

La proposta di Regolamento europeo sull'intelligenza artificiale non rappresenta il primo caso in cui il legislatore europeo costruisce un modello basato fondamentalmente sulla gestione del rischio. Il caso più recente e più importante è quello costituito dal Regolamento europeo relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati. Ma in quest'ultimo il sistema di gestione del rischio è accompagnato dal principio di *accountability*, cioè dal principio introdotto proprio da quel Regolamento secondo il quale il titolare del trattamento deve adottare le misure adeguate ad attuare i principi e le disposizioni del Regolamento conformemente alle caratteristiche specifiche del trattamento e comprovare di avere svolto questa attività [64]. Questo principio presuppone che

[63] Così l'art. 52.

[64] Sul principio di *accountability*, si

il titolare del trattamento sia il soggetto nella migliore posizione per gestire e valutare il rischio.

Nella proposta di Regolamento europeo sull'intelligenza artificiale l'approccio è diverso. È il legislatore che decide quali sono i sistemi ad alto rischio e come il rischio che essi procurano deve essere affrontato.

Se si considera che la definizione di applicazioni di intelligenza artificiale è assai ampia, cioè "un *software* sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata serie di obiettivi definiti dall'uomo, generare *output* quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono", appare critico che le applicazioni di intelligenza artificiale siano rigidamente classificate: qualunque sviluppo è destinato ad essere inserito nella griglia definita dal legislatore.

La problematicità sta nel pericolo che le applicazioni di intelligenza artificiale, anche future, siano normate con la prospettiva dell'attuale presente e che quindi il sistema non sia sufficientemente dinamico per seguire poi i successivi sviluppi dell'intelligenza artificiale.

Per gestire questo rischio, nella proposta di Regolamento europeo si pongono in capo ai soggetti che immettono nel mercato le applicazioni di intelligenza artificiale una serie di obblighi, riguardanti certificazioni, notifiche, produzione di documentazione e marcature, che sono obblighi anche di natura amministrativa. Un approccio di questo tipo presenta alcune criticità. La prima è la mancanza di neutralità tecnologica, come si è già illustrato. L'altra criticità è quella di prevedere, a fronte di una generalissima definizione di intelligenza artificiale, in capo ad ogni tipo di imprenditore i medesimi obblighi, a prescindere dalle dimensioni dell'impresa e della criticità delle attività. Un modello di questo genere ha all'origine il grande difetto di trattare tutte le applicazioni di intelligenza artificiale in maniera sostanzialmente omogenea, mentre le applicazioni di intelligenza artificiale possono essere diversissime tra loro e declinate in maniera differente.

Questo errore è stato già commesso in altri contesti, con riguardo proprio alla disciplina in materia di trattamento dei dati

consenta di rinviare ai miei contributi *Il quadro d'insieme sul Regolamento europeo*, op. cit.; *L'*accountability *nel Regolamento europeo*, op. cit., e *Il principio di* accountability, in *Giur. It.*, 2019, pp. 2778-2783. Cfr. anche Lucchini Guastalla, Privacy *e* Data Protection*: principi generali*, in Tosi (a cura di), *Privacy digitale. Riservatezza e protezione dei dati personali tra GDPR e nuovo Codice* Privacy, Milano, 2019, pp. 58-92; Id., *Il nuovo regolamento europeo sul trattamento dei dati personali: i principi ispiratori*, in *Contr. e impr.*, 2018, pp. 106-125; Malgieri, *Art. 5*, in *Codice della* privacy *e* data protection, op. cit., pp. 89-190; Mantelero, *Responsabilità e rischio nel Reg. UE 2016/679*, in *Nuove leggi civ.*, 2017, pp. 144-164; Resta, *Art. 5* e Siano, *Art. 24*, in *GDPR e normativa* privacy. *Commentario*, op. cit., rispettivamente pp. 61-63 e pp. 236-245.

personali, che poi, infatti, in tempi più recenti, è stata ripensata con il principio dell'*accountability* per consentire di modulare le misure da adottare in ragione delle caratteristiche proprie del caso specifico.

Gli obblighi dettati dal legislatore europeo producono effetti differenti a seconda dei soggetti nei confronti dei quali si rivolgono. Le società di grandi dimensioni presumibilmente non avranno problemi a gestire oneri di documentazione, certificazione, marcatura e quant'altro. Le piccole imprese, e in particolare le *start-up*, invece, vedranno oneri economici molto pesanti e rilevanti a seguito di questi obblighi previsti dal legislatore europeo. Inevitabilmente gli oneri e i costi della tutela saranno diversi a seconda del soggetto nei cui confronti saranno indirizzati. Quindi si profila un forte rischio per le piccole imprese, per le *start-up* e per l'attività dei ricercatori che caratterizzano la realtà italiana.

La proposta di Regolamento europeo ha, da un lato, il pregio di essere un primo modello organico in materia. Il modello, basato sulla gestione del rischio, sembra però poco dinamico e troppo oneroso dal punto di vista economico. Oltre a ciò, la normativa europea rischia ancora una volta di isolare l'Europa. È certamente fondamentale su questi temi e nel mercato del digitale proporre e affermare un modello, come l'Europa ha fatto in tanti altri casi, ma altrettanto importante è lavorare sulla cooperazione internazionale, cioè cercare di costruire delle modalità di dialogo, di cooperazione — anche normativa — con le altre regioni del mondo, perché costruire un modello isolato significa isolare l'Europa. Riprendendo una metafora ormai nota, se vogliamo costruire delle fortezze, dobbiamo poi ricordarci di costruire anche i ponti che ci consentano di collegarle ad altri sistemi.

Abstract

The article provides a critical appraisal of the EU Artificial Intelligence Act. After a detailed analysis of its most relevant provisions, it points out what are, according to the Author, the most debatable aspects, i.e. the rigidity of certain key definitions, the ensuing difficulty to keep up with the inevitable technological developments, the lack of distinguishing between the size of the entities using AI programmes, the difficult compatibility of the risk-based approach in the context of international cooperation.