Capitolo terzo

INTELLIGENZA ARTIFICIALE E REGOLAZIONE

1. *La «volontà di potenza» dell'IA e la capacità di presa del diritto*

Se in passato nel dibattito pubblico la capacità di sviluppo dell'IA è stata a volte sopravvalutata, all'inverso dal 2012 in avanti il fenomeno cui si assiste è di segno opposto, vale a dire che appare strisciante la percezione della tendenza a sottovalutare l'intelligenza artificiale, il suo potenziale sviluppo ed i rischi connessi a tale tecnologia[1]. Ciò sta avvenendo soprattutto in concomitanza con l'implementazione degli algoritmi di apprendimento automatico. Benché siano indubbi i vantaggi in termini di efficienza che l'intelligenza artificiale è in grado di assicurare, non possono essere però sminuiti gli elementi critici[2], a cominciare dall'opacità dei

Questo capitolo è di Marco Macchia e Antonella Mascolo.

[1] In un'intervista rilasciata nel 2017 Stephen Hawking ha messo in guardia sui possibili scenari di sviluppo dell'IA con queste parole: «Success in creating effective AI, could be the biggest event in the history of our civilization. Or the worst. We just don't know. So, we cannot know if we will be infinitely helped by AI, or ignored by it and side-lined or conceivably destroyed by it. Unless we learn how to prepare for, and avoid the potential risks, AI could be the worst event in the history of our civilization. It brings dangers, like powerful autonomous weapons, or new ways for the few to oppress the many. It could bring great disruption to our country».

[2] Su questo tema la letteratura è oramai molto ampia. Senza pretesa di completezza cfr. M. Ebers, *Regulating AI and Robotics: Ethical and Legal Challenges*, in M. Ebers e S. Navas (a cura di), *Algorithms and Law*, Cambridge, 2020; S. Sassi, *Gli algoritmi nelle decisioni pubbliche tra trasparenza e responsabilità*, in «Analisi Giuridica dell'Economia», 2019, n. 1, pp. 109 ss.; B. Lepri, N. Oliver, E. Letouzé, A. Pentland e P. Vinck, *Fair, Transparent and Accountable Algorithmic Decision-making Processes*, in «Philosophy and Technology», 2018, n. 31; B.D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter e L. Floridi, *The Ethics of Algorithms: Mapping the*

processi decisionali, dalla difficoltà di esercitare un controllo «umano» sino alla possibilità di errori e discriminazioni (i cd. *biases*).

Ne consegue che lo sviluppo di tali strumenti, tenuto conto delle implicazioni sociali, economiche e biologiche connesse all'utilizzo dell'intelligenza artificiale, non sembra poter essere affidato alle sole forze del mercato. Spetta al diritto il compito di regolare tale tecnologia[3], così da orientarne virtuosamente il funzionamento e declinarla al perseguimento delle finalità di interesse pubblico sottese all'azione di legislatori e regolatori[4]. Duplice è lo sforzo che il diritto

Debate, in «Big Data & Society», 2016, n. 3; M. Ananny, *Toward an Ethics of Algorithms: Convening, Observation, Probability and Timeless*, in «Science, Technology & Human Values», 25, 2015, n. 1.

[3] Sul tema della regolazione dell'IA si vedano, tra i contributi italiani più recenti, C. Casonato e B. Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, in «BioLaw Journal – Rivista di BioDiritto», 2021, n. 3; G. Di Rosa, *Quali regole per i sistemi automatizzati intelligenti?*, in «Rivista di diritto civile», 2021, n. 5, pp. 823 ss.; F. Rodi, *Gli interventi dell'Unione europea in materia di intelligenza artificiale e robotica: problemi e prospettive*, in G. Alpa (a cura di), *Diritto e intelligenza artificiale*, Pisa, 2020, pp. 187-210; L. Parona, *Prospettive europee e internazionali di regolazione dell'intelligenza artificiale tra principi etici, soft law e self regulation*, in «Rivista della regolazione dei mercati», 2020, n. 1, pp. 70 ss.; A. Adinolfi, *L'Unione europea dinanzi allo sviluppo dell'intelligenza artificiale: la costruzione di uno schema di regolamentazione europeo tra mercato unico digitale e tutela dei diritti fondamentali*, in S. Dorigo (a cura di), *Il ragionamento giuridico nell'era dell'intelligenza artificiale*, Pisa, 2020, pp. 13 ss.; A. Amidei, *La governance dell'intelligenza artificiale: profili e prospettive di diritto dell'Unione europea*, in U. Ruffolo (a cura di), *Intelligenza artificiale. Il diritto, i diritti, l'etica*, Torino, 2020, pp. 571 ss.; A. Celotto, *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in «Analisi Giuridica dell'Economia», 2019, n. 1, pp. 47 ss.; G. Resta, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in «Politica del diritto», 2019, n. 2, pp. 199 ss.

[4] Secondo A. Soro, *Democrazia e potere dei dati*, Milano, 2019, p. 189, alla volontà di potenza intrinseca alla tecnica deve essere opposto «da parte del diritto e dell'etica – e verosimilmente da entrambe queste scienze sociali – un limite a tutela della centralità della persona e delle sue libertà. E la stessa politica non deve abdicare, in questo contesto, al suo ruolo essenziale di individuazione degli obiettivi da perseguire per massimizzare l'utilità sociale, nel rispetto della dignità della perso-

è chiamato a compiere: da un lato, si rende necessaria una puntuale verifica circa le norme dell'ordinamento positivo e la loro perdurante abilità a riflettere alcuni principi ispiratori che ne giustificano la collocazione all'interno del sistema delle fonti; dall'altro, occorre una regolazione proattiva *ad hoc*, in grado di direzionare lo sviluppo dei processi in atto e correggere quelle distorsioni capaci di riflettersi sul godimento dei diritti e delle libertà individuali.

È chiaro che il rapporto tra algoritmi e regolazione può essere duplice. Per un verso, i primi possono essere uno efficiente strumento per migliorare la regolazione. Difatti,

> algorithms can be harnessed to play a role not only in the implementation of regulatory schemes, technical or discretionary, but also in their evaluation and eventually in formation process of alternative schemes. The development of the predictive algorithms may be useful in assessing not only a particular case, but the more general relationship between regulatory means and ends. It may shed light on what measure is likely to work, and under what conditions. It may also inform the policy makers with respect to the probable cost-benefit analysis of achieving certain policy goals. Such algorithms may be conceptualised as «policy algorithms», since the problem they are designed to solve is the overall risk allocation in a given socio-economic field, or the adequacy (likelihood) of a certain regulatory scheme as applied to achieve its goals, compared to (tested) alternatives. Obviously, such algorithms can also be designed so that they «learn» and adapt, as they analyse policy decisions at the aggregate level, to detect those with greater probabilities of achieving a desired goal (and lower probability for achieving unintended negative consequences)[5].

na, salvaguardando principi-cardine come quelli di libertà, eguaglianza, solidarietà e pluralismo».

[5] Così in A. Reichman e G. Sartor, *Algorithms and Regulation*, in H.-W. Micklitz, O. Pollicino, A. Reichman, A. Simoncini, G. Sartor e G. De Gregorio (a cura di), *Constitutional Challenges in the Algorithmic Society*, Cambridge, 2022, p. 154.

Per un altro verso, all'opposto, gli algoritmi sono loro stessi l'oggetto dell'intervento regolatorio. E in questa seconda accezione sono in questo scritto indagati.

Le iniziative susseguitesi fino ad oggi nei vari ambiti[6] – e in modo particolare in quello giuridico regolatorio – documentano una frammentarietà nella declinazione di linee guida, tendenzialmente settoriali, nonché la carenza di interventi di carattere organico sotto forma di *hard law*. Comune a tali documenti è la generale preoccupazione di assicurare che lo sviluppo dei moderni sistemi di intelligenza artificiale e delle tecnologie algoritmiche si svolga in armonia con le esigenze di tutela dei diritti e delle libertà individuali. Accanto alla diffusa percezione circa gli innumerevoli vantaggi che l'implementazione di queste tecnologie può recare entro una varietà di settori, si assiste infatti alla ferma consapevolezza in ordine alla necessità di conservare una funzionalizzazione dei loro utilizzi alla salvaguardia della dignità umana e dei diritti.

Benché tali documenti abbiano il merito di fornire spunti per orientare il dibattito e – almeno potenzialmente – le scelte di *policy*, essi non sono in grado di assicurare la vera *governance* che il settore richiede. Risulta, invece, indispensabile assicurare un quadro giuridico di riferimento, che, adottando un approccio «di sistema», si dimostri capace di superare tanto gli steccati settoriali dei singoli campi di ricerca che i confini geografici degli Stati.

Dal punto di vista del contenuto normativo, considerato l'elevata complessità tecnica del fenomeno, è necessario che la regolazione giuridica sia svolta in chiave interdisciplinare, con la collaborazione tecnica di programmatori ed ingegneri informatici e con l'apporto di studiosi di scienze sociali, «altrimenti il regolatore solo giuridico rischia di porre norme

[6] Solo negli ultimi tre anni sono stati stilati oltre 70 documenti che hanno ad oggetto l'individuazione dei principi etici che debbono guidare l'intelligenza artificiale, redatti da aziende (Google, Ibm, ecc.), istituzioni sovranazionali (l'OCSE, l'Unione europea) e organizzazioni accademiche (la Dichiarazione di Montreal, Future of Life Institute, AI4People).

del tutto avulse dal contesto operativo e come tali inutili o quanto meno poco performative»[7].

Sul secondo aspetto, proprio perché i fenomeni in atto non hanno dimensione di carattere nazionale, occorre che la regolazione sia anch'essa – per quanto possibile – di carattere sovranazionale[8]. In questa direzione muovono i recenti sforzi dell'Unione europea, per ora solo a livello programmatico, convergenti verso la creazione di un quadro regolatorio comune a livello europeo per la creazione di un'intelligenza artificiale «degna di fiducia».

Ogni tentativo di regolazione dell'IA deve, in ogni caso, fare i conti con l'intrinseca fluidità dell'intelligenza artificiale, una tecnologia difficile da regolare sia perché «ancor più di altre tecnologie innovative, è caratterizzata da incessanti sviluppi che rendono rapidamente obsoleta qualsiasi disciplina volta a regolarla» e sia perché essa «si contraddistingue per una forte dose di autonomia e imprevedibilità di funzionamento, la quale, accompagnata all'inspiegabilità dei processi interni (fenomeno della *black box*) può rappresentare una potenziale fonte di rischi, non calcolabili *ex ante*»[9].

[7] Celotto, *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, cit.; in termini analoghi, M. Giuliano, *Regolare l'infosfera*, in «Contratto e impresa», 2021, n. 3, p. 885 osserva anche come «il giurista, senza la piena ed effettiva conoscenza delle funzioni che il codice abilita, raggiunge solo in minima parte le potenzialità che ha a disposizione, immaginando di risolvere solo il problema della sua documentazione, senza rendersi conto invece come in gioco sia la sua stessa capacità cognitiva».

[8] *Ibidem*, p. 886, rileva come «con Internet è nato un nuovo territorio, mondiale, senza confini, senza territorio, strumento di crescita individuale e sociale, in cui le relazioni umane si svolgono al di fuori di qualsiasi tentativo di programmazione e di regolamentazione. Non tutte le attività che si svolgono nella rete sono regolabili con la legge di un determinato Stato. La sua essenza transnazionale rende inefficace qualsiasi normativa statale».

[9] Casonato e Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, cit.

In ogni caso, propedeutico ad ogni tentativo di regolazione è la definizione[10] dell'IA poiché ciò incide sulla effettiva capacità di «cattura» e di regolamentazione della tecnologia[11]. Il termine IA è, infatti, suscettibile di abbracciare tecniche ed applicativi anche molto diversi tra loro, il cui utilizzo può comportare risultati e fattori di rischio per i soggetti coinvolti altrettanto diversificati. Mentre, infatti, negli algoritmi *model based* risulta relativamente agevole ricostruire il «ragionamento» algoritmico (nonché i correlativi profili di eventuale responsabilità), nei sistemi dotati di *machine learning*[12] lo stesso sviluppatore incontra un ostacolo, anche

[10] R. Girasa, *Artificial Intelligence as a Disruptive Technology. Economic Transformation and Government Regulation*, Cham, 2020 individua cinque definizioni di IA tra le più accreditate in letteratura: *i*) IA come abilità di un computer digitale di svolgere compiti comunemente associati a soggetti intelligenti; *ii*) IA come studio e programmazione di agenti intelligenti; *iii*) IA come teoria e sviluppo di sistemi in grado svolgere compiti che normalmente richiedono un'intelligenza umana; *iv*) IA come intelligenza artificiale in opposizione all'intelligenza naturale propria degli uomini e di altri animali; *v*) IA come sistema in grado di svolgere compiti al ricorrere di circostanze impreviste ed imprevedibili senza una supervisione umana oppure capaci di imparare dall'esperienza e migliorare le proprie performance. Sul tema della difficoltà di definizione dell'IA cfr. tra gli altri S. Russell e P. Norvig, *Artificial Intelligence: A Modern Approach*, Upper Saddle River, 2020, p. 17.

[11] Come osservato da M. Favaretto, E. De Clercq, O. Schneble e B.S. Elger, *What Is Your Definition of Big Data? Researchers' Understanding of the Phenomenon of the Decade*, in «PlosOne», 2020, con riferimento ai *big data*, ma le considerazioni sono trasponibili anche alle tecnologie di IA, «As long as definitions are unclear, laws, regulations and guidelines that are bound to govern Big Data research in these two fields of research are unlikely to be effective, especially if researchers are unaware of the regulatory framework or refrain from defining their research as Big Data research out of fear for regulatory restrictions».

[12] Per M. Van Otterlo, *A Machine Learning View on Profiling*, in M. Hildebrandt e K. de Vries, *Privacy, Due Process and the Computational Turn - Philosophers of Law Meet Philosophers of Technology*, London, 2016, secondo cui «machine learning is any methodology and set of techniques that can employ data to come up with novel pattern and knowledge and generate models that can be used for effective predictions about the data». Sul tema cfr. anche A. Mackenzie, *The Production of Prediction: What Does Machine Learning Want?*, in «European Journal of Cultural Studies», 18, 2015, n. 4-5. Per una panoramica aggiornata dell'attuale livello di sviluppo dell'IA nei diversi campi di impiego

insormontabile, nella imprevedibilità dei processi decisionali che ne rende impossibile la supervisione e la correzione[13].

Su questo tema è recentemente intervenuto il Consiglio di Stato[14], adottando, in controtendenza rispetto all'approccio delle istituzioni europee, un'accezione notevolmente restrittiva del termine. Chiamato a giudicare circa l'esatta perimetrazione tecnica della nozione di «algoritmo di trattamento» nell'ambito di una gara per la fornitura di *pacemaker*, il Consiglio di Stato ha approfondito la nozione di algoritmo evidenziando le differenze rispetto alla stessa nozione di intelligenza artificiale. L'algoritmo si risolve in una «sequenza finita di istruzioni, ben definite e non ambigue, così da poter essere eseguite meccanicamente e tali da produrre un determinato risultato», sia pure ineludibilmente collegata – nel contesto applicativo di sistemi tecnologici – al concetto di automazione, ossia a «sistemi di azione e controllo idonei a ridurre l'intervento umano». Diversa è invece, per i giudici amministrativi, la nozione di intelligenza artificiale, con tale termine dovendosi intendere i soli applicativi

> in cui l'algoritmo contempla meccanismi di *machine learning* e crea un sistema che non si limita solo ad applicare le regole software e i parametri preimpostati (come fa invece l'algoritmo «tradizionale») ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni, secondo un processo di apprendimento automatico[15].

cfr. Girasa, *Artificial Intelligence as a Disruptive Technology. Economic Transformation and Government Regulation*, cit.

[13] B. Marchetti, *The Algorithmic Administrative Decision and the Human in the Loop*, in «BioLaw Journal – Rivista di BioDiritto», 2021, n. 2. Come spiegato da Andy Rubin, co-creatore del sistema Android, quando gli ingegneri scrutano in una profonda rete neurale, quello che vedono è «un oceano di matematica»: un insieme massiccio e multistrato di problemi di calcolo che – derivando costantemente relazione tra miliardi di punti dati – genera ipotesi sul mondo.

[14] Sentenza n. 7891 del 25/11/2021.

[15] Par. 9.1.

Com'è evidente, tale nozione di IA lascia fuori un numero significativo di applicativi, dai motori deduttivi e inferenziali, ai sistemi esperti sino agli altri approcci basati sulla conoscenza. Anche al di là del merito tecnico della definizione, il rischio è che un elevato numero di applicativi finiscano per sfuggire ai tentativi di regolazione. Diametralmente opposta è la nozione fatta propria, come si vedrà, dalle istituzioni dell'Unione europea, che hanno invece dilatato al massimo la nozione di intelligenza artificiale, ricomprendendovi sia sistemi di IA «forte» che applicativi di IA «debole», ossia quei sistemi informatici capaci di prestazioni normalmente attribuite all'intelligenza umana, pur senza assumere alcuna analogia tra le menti e i sistemi informatici.

2. *La «via europea» allo sviluppo dell'intelligenza artificiale*

Negli ultimi anni le istituzioni europee hanno mostrato una crescente attenzione al tema dell'intelligenza artificiale, in particolare enfatizzando la necessità di una *governance* integrata a livello europeo[16] – anche a fronte della «corsa all'IA» ingaggiata da USA e Cina[17] – per implementarne lo

[16] Nel documento *Progetto di relazione sull'intelligenza artificiale in un'era digitale* redatto dalla Commissione speciale sull'intelligenza artificiale in un'era digitale (2020/2266[INI]), pubblicato il 2/11/2021, viene rimarcato come, rispetto a USA, Cina e Giappone, l'Europa sia tutt'ora notevolmente indietro, con il rischio che «i valori europei siano sostituiti a livello globale, che le nostre imprese siano emarginate e che il nostro tenore di vita sia drasticamente ridotto», a tal punto che «se l'UE non agirà rapidamente e con coraggio, finirà per diventare una colonia digitale di Cina, Stati Uniti e di altri Stati, rischiando di perdere la propria stabilità politica, la sicurezza sociale e le libertà individuali». Entrambi i Paesi hanno, inoltre, il vantaggio di un mercato unico unificato, una maggiore flessibilità nella *governance* digitale e un più forte impegno politico finalizzato a confermare la propria leadership nell'IA.

[17] L'intento della Cina, reso esplicito nel Piano di sviluppo dell'IA del 2017, è quello di imporsi quale leader mondiale dell'IA entro il 2030.

sviluppo in armonia con il quadro dei principi e dei diritti fondamentali condivisi dagli Stati europei.

Nel febbraio 2017, il Parlamento europeo ha approvato una risoluzione «con raccomandazioni alla Commissione per le norme di diritto civile sulla robotica»[18]. La Risoluzione invita gli Stati membri a disciplinare in maniera omogenea gli aspetti civilistici della robotica e suggerisce la creazione di un'Agenzia europea per la robotica e l'IA composta da regolatori ed esperti esterni in grado di assicurare «le competenze tecniche, etiche e normative necessarie per sostenere gli attori pubblici interessati, sia a livello dell'Unione che degli Stati membri, nel loro sforzo di garantire una risposta tempestiva, etica e ben informata alle nuove opportunità e alle nuove sfide». La Risoluzione auspica una serie di cautele etiche, per assicurare che lo sviluppo e l'utilizzo dei robot avvenga in condizioni tali da preservare la dignità, l'autonomia e l'autodeterminazione degli individui, la tutela della privacy e che si presti attenzione «alla possibilità che nasca un attaccamento emotivo tra gli uomini e i robot, in particolare per i gruppi vulnerabili (bambini, anziani, disabili), per attenuare gli impatti emotivi e fisici»[19]. Inoltre, gli Stati membri sono invitati ad affrontare la questione dei robot innanzitutto dal punto di vista della responsabilità, creando un sistema di registrazione dei robot, una assicurazione obbligatoria e un fondo di garanzia per i danni causato da robot non assicurati nonché forme di responsabilità oggettiva[20].

[18] Risoluzione del Parlamento europeo del 16/2/2017, recante *Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*.

[19] Punto 3 della Risoluzione.

[20] La Risoluzione giunge ad auspicare anche «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» (punto 59).

A distanza di un anno, con le due comunicazioni del 25/4/2018[21] e del 7/12/2018[22], la Commissione europea ha delineato tre pilastri fondamentali su cui deve basarsi lo sviluppo di un'«IA *made in Europe*»: *i*) l'aumento degli investimenti pubblici e privati nel campo dell'intelligenza artificiale; *ii*) la capacità di prepararsi per tempo ai cambiamenti socioeconomici; *iii*) la garanzia di un quadro etico e giuridico adeguato a rafforzare i valori europei.

Per la Commissione, la via europea all'intelligenza artificiale – realizzabile anche grazie alle eccellenze scientifiche ed industriali presenti nel contesto europeo[23] – passa per lo sviluppo di un'IA «etica, sicura e all'avanguardia», posta al servizio dell'uomo e delle sfide cruciali del XXI secolo, «dalla cura delle malattie alla lotta contro i cambiamenti climatici e alla previsione delle catastrofi naturali, dall'aumento della sicurezza dei trasporti alla lotta alla criminalità al miglioramento della cybersicurezza»[24]. Essenziale a tale scopo, secondo la Commissione europea, è l'adozione di un approccio coordinato a livello dell'Unione, capace di sfruttare al massimo le opportunità offerte dall'IA e affrontare le nuove sfide, senza però rinunciare ai valori fondanti dell'Unione, quali la dignità umana e la tutela della privacy.

A tali primi documenti programmatici ha fatto seguito la pubblicazione, nel mese di febbraio 2020, del *Libro bianco sull'intelligenza artificiale*[25], con lo scopo di fissare gli obiettivi ed i traguardi per la realizzazione di una so-

[21] Comunicazione COM(2018) 237 del 25/4/2018 (*Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni – L'intelligenza artificiale per l'Europa*).

[22] Comunicazione COM(2018) 795 del 7/12/2018 (*Piano coordinato sull'intelligenza artificiale*).

[23] La Commissione ravvisa l'opportunità di una mobilitazione totale, di mobilitare anche soggetti privati, quali imprese, organizzazioni di consumatori, sindacati ed altri rappresentanti della società civile, nell'ottica di creare una piattaforma multilaterale (l'*Alleanza europea per l'IA*) che si occupi, con un approccio integrato, di tutti gli aspetti connessi allo sviluppo dell'IA europea.

[24] Comunicazione COM(2018) 795 del 7/12/2018.

[25] Documento COM(2020) 65 del 19/2/2020.

cietà europea in cui soluzioni digitali, tecnologie affidabili, persone e imprese possano convivere in modo sinergico e sicuro, in un ambiente aperto, democratico, nell'ambito di un'economia dinamica e sostenibile.

La capacità dell'Europa di riuscire ad imporsi come leader nel settore delle tecnologie digitali e dell'intelligenza artificiale dipende – secondo la Commissione europea – dalla definizione di un solido quadro normativo europeo e dalla capacità di mettere a punto un modello coerente di *governance*, in grado di scongiurare il rischio di approcci individuali, isolate fughe in avanti o soluzioni frammentarie da parte dei singoli Stati.

La sfida sta, dunque, nella capacità di trovare un giusto equilibrio tra la creazione di un ecosistema in cui l'intelligenza artificiale possa prosperare – garantendo all'Europa di diventare un leader mondiale nella tecnologia – e la protezione della società dai rischi che tale tecnologia può portare con sé. La posta in gioco della sfida sono diritti fondamentali, quali il diritto alla privacy, alla dignità umana, alla libertà di espressione e alla non discriminazione, ma anche il futuro economico dell'Europa.

Le azioni che la Commissione si propone di avviare per raggiungere questo obiettivo si muovono lungo tre direttrici principali, vale a dire: la necessità di un approccio sovranazionale, uno sforzo regolatorio proattivo e, in ultimo, un approccio *risk-based.*

Per quanto concerne il primo profilo, la Commissione rimarca la necessità di intensificare la collaborazione con gli Stati membri, adottando un «solido e coordinato approccio europeo», soprattutto in settori chiave quali la ricerca, gli investimenti, l'adozione da parte del mercato, le competenze e i talenti, i dati e la cooperazione internazionale[26]. Comple-

[26] In particolare, nel piano della ricerca viene ravvisata la necessità di concentrare gli sforzi della comunità della ricerca e dell'innovazione verso la creazione di un centro per la ricerca e l'innovazione a livello europeo, superando l'attuale panorama frammentato di centri di competenza, nessuno dei quali è in grado di raggiungere le dimensioni necessarie per competere con i principali istituti a livello mondiale. Secondo il *Libro bianco*, «i centri e le reti dovrebbero concentrarsi nei settori in

mentare a tale misura è l'implementazione dello sviluppo di competenze necessarie per lavorare nel settore dell'IA ed il miglioramento del livello attuale delle competenze della forza lavoro, al fine di prepararla all'uso delle tecnologie basate sull'IA.

Carattere sovranazionale dovrebbe avere – secondo il *Libro bianco* – anche la *governance* in materia di IA. Sul punto, la Commissione parla di cooperazione tra autorità nazionali competenti all'interno di un quadro di riferimento europeo, quale forum per lo scambio di informazioni e *best practice*, per identificare le tendenze emergenti e fornire pareri anche al fine di facilitare l'attuazione comune delle normative. Tali misure debbono però iscriversi entro un contesto normativo comune a livello dell'Unione, volto ad assicurare la promozione delle tecnologie di IA nel contesto dei diritti e delle garanzie fondamentali a tutela dei cittadini europei. Secondo la Commissione, la regolazione delle nuove tecnologie non può essere affidata esclusivamente alle autorità nazionali, anche perché l'adozione di normative non coordinate e difformi ha effetti negativi per il mercato comune, essendo destinata a tradursi in barriere agli scambi.

In stretta connessione con il tema della regolazione comune è anche la questione relativa alla certezza del diritto: la Commissione ritiene che il quadro normativo vigente debba essere migliorato per affrontare i fattori di rischio specificatamente connessi con l'intelligenza artificiale.

Segnatamente, le caratteristiche principali dell'IA rendono difficile garantire la corretta applicazione e il rispetto della normativa nazionale e dell'UE. A causa di opacità dell'IA è difficile individuare e dimostrare eventuali violazioni delle disposizioni normative (comprese quelle che tutelano i diritti

cui l'Europa ha il potenziale per diventare leader a livello mondiale, come l'industria, la salute, i trasporti, la finanza, le catene del valore agroalimentari, l'energia/l'ambiente, il settore forestale, l'osservazione della Terra e lo spazio. In tutti questi settori si sta svolgendo la corsa alla leadership mondiale e l'Europa offre notevoli potenzialità, conoscenze e competenze».

fondamentali), attribuire la responsabilità e soddisfare le condizioni per chiedere un risarcimento.

Per assicurare la trasparente applicazione dell'IA, secondo la Commissione non è tanto necessario adottare nuove regole, ma preliminarmente verificare la capacità di tenuta delle regole già esistenti. La valutazione se adottare nuove normative a livello sovranazionale deve seguire un approccio *risk-based*, assicurando una regolamentazione su misura per le applicazioni di intelligenza artificiale considerate ad «alto rischio» sulla base dei due criteri cumulativi indicati dalla Commissione.

3. *Per un'intelligenza artificiale antropocentrica*

Nel giugno del 2018 la Commissione europea ha incaricato un gruppo di esperti ad alto livello sull'intelligenza artificiale di redigere un documento contenente gli orientamenti etici per l'IA[27]. Dopo quasi un anno di lavoro, il gruppo di lavoro ha pubblicato il documento *Orientamenti etici per un'IA affidabile*[28], al dichiarato scopo di fornire – al legislatore europeo e agli Stati nazionali – un quadro di riferimento per la realizzazione di un'intelligenza artificiale robusta e affidabile.

Il principio guida che ha orientato la redazione del documento è che l'intelligenza artificiale non rappresenti un fine in sé, quanto piuttosto «un mezzo promettente per aumentare la prosperità umana, migliorando così il benessere individuale e sociale e il bene comune nonché favorendo progresso e innovazione»[29].

L'approccio europeo dovrebbe, quindi, essere nel senso di realizzare un'intelligenza artificiale antropocentrica, in

[27] Si veda Comunicazione COM(2018) 237 del 25/4/2018 (*Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni – L'intelligenza artificiale per l'Europa*).

[28] Il documento è stato pubblicato in data 8/4/2019.

[29] Par. 9.

cui le tecnologie dell'intelligenza artificiale siano poste al servizio dell'umanità e del bene comune, con l'obiettivo di migliorare il benessere e la libertà degli esseri umani.

Per consolidare l'alleanza tra uomo e macchina nel conseguimento del comune obiettivo di implementare il benessere individuale e collettivo, il gruppo di lavoro ritiene indispensabile che il cemento sia costituito dalla fiducia che gli uomini devono poter riporre nelle tecnologie dell'IA. L'obiettivo, quindi, deve essere quello di mettere a punto un'intelligenza artificiale che sia «degna di fiducia» da parte dell'uomo[30].

Affinché sia affidabile è necessario che l'intelligenza artificiale possieda tre componenti, le quali devono essere sempre presenti durante l'intero ciclo di vita del sistema, ossia: la legalità[31], nel senso che l'IA deve ottemperare a tutte le leggi ed ai regolamenti; l'eticità[32], intesa come adesione dell'IA a principi etici condivisi in sede europea; e la

[30] La fiducia nello sviluppo, nella distribuzione e nell'utilizzo di sistemi di IA non deve riguardare soltanto le proprietà intrinseche della tecnologia, ma anche le qualità dei sistemi sociotecnici che comportano applicazioni dell'IA.

[31] Come precisato al par. 22, «i sistemi di IA non operano in un mondo senza leggi. A livello europeo, nazionale e internazionale un corpus normativo giuridicamente vincolante è già in vigore o è pertinente per lo sviluppo, la distribuzione e l'utilizzo dei sistemi di IA. Le fonti giuridiche pertinenti sono, a titolo esemplificativo, il diritto primario dell'UE (i trattati dell'Unione europea e la sua Carta dei diritti fondamentali), il diritto derivato dell'UE (ad esempio, il regolamento generale sulla protezione dei dati, le direttive antidiscriminazione, la direttiva macchine, la direttiva sulla responsabilità dei prodotti, il regolamento sulla libera circolazione dei dati non personali, il diritto dei consumatori e le direttive in materia di salute e sicurezza sul lavoro), ma anche i trattati ONU sui diritti umani e le convenzioni del Consiglio d'Europa (come la Convenzione europea dei diritti dell'uomo) e numerose leggi degli Stati membri dell'UE. Oltre alle norme applicabili orizzontalmente, esistono varie norme specifiche per settore applicabili a particolari applicazioni di IA (ad esempio il regolamento sui dispositivi medici nel settore sanitario)».

[32] Come chiarito al par. 26, per ottenere un'IA affidabile «non è sufficiente il rispetto della legge, che è solo una delle tre componenti. Il diritto non è sempre al passo con gli sviluppi tecnologici, e a volte non lo è nemmeno con le norme etiche o semplicemente non è adatto ad affrontare determinate questioni. Affinché i sistemi di IA siano affidabili,

robustezza[33] dal punto di vista tecnico, al fine di assicurare che l'IA non provochi, anche con le migliori intenzioni, danni non intenzionali.

Uno degli aspetti di maggiore interesse è rappresentato proprio dalla centralità dell'etica nello sviluppo dell'IA *made in Europe*, non solo in funzione di protezione di individui e gruppi, ma anche nell'ottica di «migliorare la prosperità individuale, e il benessere collettivo generando agiatezza, creando valore e massimizzando la ricchezza»[34].

A partire dal quadro dei diritti inviolabili previsti dal diritto internazionale in materia di diritti umani, dai trattati UE e della Carta dei diritti fondamentali dell'Unione europea, il documento individua quattro imperativi etici che gli operatori pubblici ed i privati dovrebbero impegnarsi ad assicurare nello sviluppo di dispositivi di intelligenza artificiale.

Il primo imperativo etico è il principio di rispetto dell'autonomia umana: gli esseri umani che interagiscono con i sistemi di intelligenza artificiale devono poter mantenere la propria piena ed effettiva autodeterminazione e devono poter essere partecipi del processo democratico.

In questo senso, i sistemi di intelligenza artificiale non devono essere progettati in modo da subordinare, costringere, ingannare, manipolare, condizionare o aggregare in modo

essi dovrebbero quindi essere anche etici garantendo la compatibilità con le norme etiche».

[33] Viene precisato al par. 27 del documento che «anche qualora il fine etico sia garantito, gli individui e la società devono comunque essere sicuri che i sistemi di IA non causeranno alcun danno involontario. Tali sistemi dovrebbero funzionare in modo sicuro e affidabile e dovrebbero essere previste misure di salvaguardia per prevenire qualsiasi effetto negativo indesiderato. È quindi importante garantire che i sistemi di IA siano robusti. Tale componente è necessaria sia da un punto di vista tecnico (garantendo la robustezza tecnica del sistema in un dato contesto, ad esempio il settore di applicazione o la fase del ciclo di vita), sia da un punto di vista sociale (tenendo in debita considerazione il contesto e l'ambiente in cui il sistema opera). L'eticità e la robustezza dell'IA sono quindi componenti strettamente correlate che si integrano a vicenda. I principi enunciati nel capitolo I e i requisiti tratti da essi nel capitolo II riguardano entrambe le componenti».

[34] Par. 34.

ingiustificato gli esseri umani. Al contrario, tali sistemi devono essere progettati in modo tale da «aumentare, integrare e potenziare le abilità cognitive, sociali e culturali umane. La distribuzione delle funzioni tra esseri umani e sistemi di IA dovrebbe seguire i principi di progettazione antropocentrica e lasciare ampie opportunità di scelta all'essere umano»[35].

Il secondo principio cardine per la costruzione di un'intelligenza artificiale affidabile è individuato nel principio di prevenzione dei danni: i sistemi di IA non devono causare danni né aggravarli e neppure influenzare negativamente gli esseri umani. Pertanto, non solo i sistemi devono essere tecnicamente robusti e gli ambienti in cui essi operano devono essere sicuri e protetti, ma si rende anche necessario prestare attenzione alle situazioni in cui i sistemi di IA possono causare o aggravare gli effetti negativi dovuti ad asimmetrie di potere o di informazione (ad esempio tra datori di lavoro e dipendenti, imprese e consumatori o governi e cittadini).

Il terzo imperativo etico individuato dal documento consiste nel principio di equità, da intendersi sia in chiave sostanziale che procedurale. La dimensione sostanziale del principio implica un impegno, da parte dell'UE e degli Stati membri, a garantire una distribuzione giusta ed equa di costi e benefici, assicurando che gli individui ed i gruppi siano liberi da distorsioni inique, discriminazioni e stigmatizzazioni. Ciò vuol dire che i sistemi di IA dovrebbero essere progettati in modo tale da promuovere le pari opportunità in termini di accesso all'istruzione, ai beni, ai servizi e alla tecnologia. Assicurando inoltre che sia rispettato il principio di proporzionalità tra mezzi e fine da perseguire[36]. Nella sua dimensione procedurale, invece, il principio di equità implica che le decisioni elaborate attraverso sistemi di IA possano essere impugnate. Per dare effettività a tale principio, si rende necessario che l'organismo responsabile della decisione sia identificabile ed i processi decisionali siano compiutamente spiegabili.

[35] Par. 13.
[36] Par. 52.

Infine, l'ultimo imperativo etico è il principio di spiegabilità, il quale implica che i processi decisionali siano trasparenti e che «la capacità e lo scopo dei sistemi di IA devono essere comunicati apertamente e le decisioni, per quanto possibile, devono essere spiegate a coloro che ne sono direttamente e indirettamente interessati»[37], allo scopo di rendere il processo decisionale – e, soprattutto, la decisione finale – compiutamente intellegibile ai suoi destinatari.

Dal punto di vista pratico, per assicurare la tenuta dei quattro imperativi etici fondamentali il documento individua sette requisiti concreti, che i sistemi di IA dovrebbero assicurare: l'intervento e la sorveglianza umani[38]; la robustezza tecnica e la sicurezza della struttura[39], compresa la resilienza ad attacchi esterni; la riservatezza e la *governance* dei dati[40]; la trasparenza[41], inclusi il rispetto della riservatezza, la qualità e l'integrità dei dati e l'accesso a essi; il rispetto della diversità, la non discriminazione e l'equità[42],

[37] Par. 53.

[38] Tale requisito è strettamente connesso con l'imperativo etico di autodeterminazione umana, implicando che «gli utenti dovrebbero essere in grado di adottare decisioni autonome e informate in merito ai sistemi di IA» (par. 64) e che sia garantita la sorveglianza umana sull'intero ciclo di vita del sistema di IA.

[39] Il requisito della robustezza tecnica è strettamente connesso al principio di prevenzione dei danni. Per garantire la robustezza tecnica è necessario che i sistemi siano sviluppati con un approccio di prevenzione dei rischi e in maniera tale che si comportino in maniera attendibile secondo le previsioni, riducendo al minimo il rischio di danni non intenzionali e prevenendo danni inaccettabili.

[40] Il requisito di riservatezza e *governance* dei dati si pone in correlazione stretta con il principio di prevenzione dei danni ed implica un'adeguata *governance* dei dati sotto il profilo della qualità e dell'integrità dei dati utilizzati, la loro pertinenza rispetto al settore in cui i sistemi di IA sono distribuiti e la capacità di trattare i dati in modo da assicurare la riservatezza.

[41] Tale requisito è strettamente connesso con il principio di esplicabilità, e comprende la tracciabilità del set di dati e dei processi tecnici, la spiegabilità nonché il diritto degli utenti di essere a conoscenza del fatto che essi stiano interagendo con un sistema di IA.

[42] Questo requisito è strettamente connesso al principio di equità. Per ottenere un'IA affidabile, occorre che l'inclusione e la diversità siano permesse durante l'intero ciclo di vita del sistema. Oltre al fatto

incluse la prevenzione di distorsioni inique, l'accessibilità e la progettazione universale e la partecipazione dei portatori di interessi; il benessere sociale ed ambientale[43], inclusi la sostenibilità e il rispetto ambientale, l'impatto sociale, la società e la democrazia; l'*accountability*[44], inclusi la verificabilità, la riduzione degli effetti negativi e la loro segnalazione, i compromessi e i ricorsi.

Tali requisiti – pur variamente modulabili a seconda del campo specifico di applicazione e del relativo contesto – dovrebbero essere assicurati durante l'intero ciclo di vita di un sistema di IA. Benché privo di carattere vincolante, il documento ha un'indiscussa valenza simbolica e programmatica in ordine al modello di sviluppo dell'IA *made in Europe*: un'intelligenza artificiale, legale, etica e robusta.

4. *Verso una regolamentazione dell'IA «made in Europe»*

Muovendo dal *Libro bianco sull'intelligenza artificiale. Un approccio europeo all'eccellenza e alla fiducia*, la Commissione ha inaugurato in questi mesi una grande stagione costituente, ponendo le basi per un significativo sforzo regolatorio funzionale all'aumento degli investimenti e dell'innovazione nello spazio europeo. Oltre al *Digital Services Act* e al *Digital Market Act*, è stata elaborata una articolata Proposta

che tutti i portatori di interessi influenzati dall'IA devono essere presi in considerazione e coinvolti nel corso del processo, tale principio comporta anche la necessità di garantire la parità di trattamento e la parità di accesso attraverso processi di progettazione inclusivi.

[43] Il requisito si pone in relazione con i principi di equità e di prevenzione dei danni, ed implica la sostenibilità – ecologica e sociale – dei sistemi di IA.

[44] Questo requisito integra quelli enunciati sopra ed è strettamente connesso con il principio di equità. Per conseguire tale requisito occorre mettere in atto meccanismi che garantiscano l'*accountability* dei sistemi di IA e dei loro risultati, sia prima che dopo la loro attuazione. Tale requisito implica la verificabilità degli algoritmi, la riduzione al minimo degli effetti negativi e la relativa segnalazione nonché la possibilità di ricorso in caso di effetti negativi o ingiusti ricollegantesi al processo automatizzato ed ai suoi esiti.

di Regolamento dell'IA[45]. Come riferito nel Memorandum di accompagnamento «l'obiettivo principale della proposta è garantire il corretto funzionamento del mercato interno stabilendo norme armonizzate, in particolare sullo sviluppo, l'immissione sul mercato dell'Unione e l'uso di prodotti e servizi che utilizzano tecnologie di IA o forniti come IA autonoma di sistemi». La filosofia di una digitalizzazione centrata sull'uomo è condivisa anche da altre organizzazioni internazionali, come l'indagine del Consiglio d'Europa sul *legal framework for AI* sta a dimostrare.

Sviluppando queste tematiche, il 21/4/2021 la Commissione europea ha pubblicato una Proposta di Regolamento sull'approccio europeo all'intelligenza artificiale[46], che intende proporsi come il primo quadro giuridico europeo sull'IA[47]. Fondata sui comuni valori europei del *Bill of Rights*, la proposta bilancia benefici e rischi per la salute dei consumatori e per i diritti fondamentali. Si tratta di un approccio regolamentare che nasce sul campo, calato dall'alto, e non scaturisce dal procedimento di assemblea. Al centro è posta la misurazione del rischio – un sistema *risk-based*, appunto – secondo uno specifico livello di classificazione. L'IA va trattata come una merce pericolosa, è necessaria per la società digitale, ma non è aliena da pericoli. Perché il rischio sia normativamente accettabile deve essere calcolato secondo un modello di mediazione giuridica improntato alla

[45] Per un commento a tale Proposta di Regolamento si vedano, tra gli altri, Casonato e Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, cit.; G. Proietti, *Intelligenza artificiale: una prima analisi della proposta di regolamento europeo*, in «Diritto bancario», 2021; Di Rosa, *Quali regole per i sistemi automatizzati?*, cit., p. 823.

[46] COM(2021) 206 final.

[47] Come osservato da A. Pajno, *Introduzione allo studio della proposta della Commissione europea di Regolamento sull'Intelligenza artificiale*, in *Introduzione al seminario Astrid-LED su Proposta della Commissione europea di Regolamento sull'Intelligenza Artificiale*, Roma, 7/6/2021, tale proposta «si pone come tentativo di regolazione dell'impiego dell'IA in un unico quadro giuridico: si registra il tentativo di passare da una logica frammentaria e settoriale, riguardante i singoli usi o singole tecnologie ad una prospettiva di regolazione generale».

tollerabilità: se l'algoritmo, ad esempio, serve ad attribuire credito sociale è reputato ad alto rischio ed è dunque vietato, altrimenti il rischio può essere minimizzato mediante alcune restrizioni spesso legate ad obblighi di trasparenza.

Sebbene siano numerosi gli aspetti su cui valga la pena soffermarsi attentamente, in questa sede pare opportuno concentrarsi su alcuni profili. Innanzitutto, la proposta della Commissione di regolazione del *machine learning* rappresenta un rilevante esperimento globale, in cui la centralità del «pacchetto IA» è funzionale a trasformare l'Europa in un *hub* globale per un'intelligenza artificiale affidabile e «umana». Nel rispetto di questa prospettiva, infatti, la proposta è accompagnata da un piano coordinato con gli Stati membri volto a rafforzare contestualmente gli investimenti e l'innovazione nel settore in tutta Europa. Trova così conferma l'intento della Commissione di presentare un progetto di grande portata per regolare all'interno di un unico quadro giuridico ciò che finora è stato considerato dai più ingovernabile. Sfida ambiziosa che solleva un inevitabile steccato rispetto agli altri, gli ordinamenti statunitense e cinese che, al contrario, hanno finora optato per un approccio di *laissez faire*.

Da una logica di intervento pubblico frammentario e settoriale, avente ad oggetto singoli usi o specifiche tecnologie, si passa dunque ad un intervento regolatorio di portata generale. Proprio perché così imponente, la prospettiva non può che essere *de jure condendo*, giacché saranno necessari almeno due o tre anni prima che la proposta diventi legislazione dal momento che il negoziato è ancora agli inizi. Ci troviamo in sostanza all'inizio di un lungo percorso in cui nessuno dei *player* globali intende compiere passi falsi.

Il Regolamento nasce dalla ravvisata necessità di un quadro giuridico dell'Unione che stabilisca norme armonizzate sull'intelligenza artificiale per promuovere lo sviluppo, l'uso e la diffusione dell'intelligenza artificiale nel mercato interno, al contempo assicurando un elevato livello di protezione degli interessi pubblici, come la salute e la sicurezza e tutela dei diritti fondamentali, riconosciuti e protetti dal diritto dell'Unione. A parte i molti usi benefici dell'intelligenza

artificiale, quella tecnologia può anche essere usata in modo improprio e fornire strumenti nuovi e potenti per pratiche manipolative, di sfruttamento e di controllo sociale. Tali pratiche sono particolarmente dannose e dovrebbero essere vietate perché contraddicono i valori dell'Unione di rispetto della dignità umana, libertà, uguaglianza, democrazia e Stato di diritto e diritti fondamentali dell'Unione, compreso il diritto alla non discriminazione, i diritti del bambino. L'obiettivo dichiarato dell'UE è quello di porsi come leader mondiale nello sviluppo di un'intelligenza artificiale sicura, affidabile ed etica.

Appare coerente con tale obiettivo di fondo anche il fatto che la regolazione sia affidata ad un regolamento, un atto «di portata europea e di applicazione diretta, volta a evitare la frammentazione del mercato interno e l'incertezza giuridica che deriverebbe per tutti da una regolamentazione esclusivamente nazionale»[48], pur lasciando, come si vedrà, ampio spazio a diversi livelli di azione da parte degli Stati membri.

Il secondo profilo da evidenziare sin d'ora riguarda l'impiego della tecnica della marcatura, un sistema già ampiamente sperimentato nel settore commerciale con cui si vorrebbe regolare lasciando il controllo a monte ai privati per riservare a valle all'autorità pubblica unicamente controlli a campione. La marcatura CE serve, difatti, ad indicare la conformità di un prodotto ai requisiti imposti dalla normativa dell'Unione europea. Per comprenderne il funzionamento di base basti dire che l'apposizione della marcatura a un sistema di IA ad alto rischio richiede il rispetto da parte del fornitore dei seguenti passaggi procedurali: *a*) determinare se il sistema di *machine learning* messo sul mercato è classificato ad alto rischio in base alle disposizioni regolamentari; *b*) garantire che la progettazione, lo sviluppo e il sistema di gestione della qualità siano conformi alla normativa; *c*) attivare una procedura di valutazione della conformità volta a valutare e dimostrare la conformità del sistema secondo un modello di autocertificazione e di *auditing* privato; *d*) apporre la marca-

[48] *Ibidem*.

tura al sistema firmando una dichiarazione di conformità; *e*) immettere il software sul mercato. La funzione del marchio CE è quella di tutelare interessi pubblici, come la salute e la sicurezza degli utilizzatori dei prodotti, appartenenti ad una determinata tipologia, assicurando che essi siano conformi a tutte le disposizioni sovranazionali che prevedono il loro utilizzo, così che la marcatura CE non funge da marchio di qualità o d'origine, ma costituisce un puro marchio amministrativo, che segnala che il prodotto marcato può circolare liberamente nel mercato unico europeo.

L'esteso ambito di applicazione del Regolamento riflette questi obiettivi. Il Regolamento si applica all'immissione sul mercato, alla messa in servizio e all'uso di «sistemi di IA», con tale termine intendendosi qualsiasi «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» (art. 3, n. 1). Sono, quindi, ricompresi nella nozione di intelligenza artificiale non solo i sistemi di apprendimento automatico, ma anche gli approcci basati sulla logica e sulla conoscenza nonché gli approcci statistici, compresi i metodi di ricerca e ottimizzazione[49].

A fronte della complessità e dell'eterogeneità dei sistemi inclusi nello spettro applicativo della proposta, che potrebbero anche indebolire la capacità di presa, la regolamentazione si basa su un approccio *risk-based*, differenziando la regolamentazione dei sistemi sulla base non delle relative caratteristiche intrinseche bensì dei relativi

[49] Ai sensi dell'Allegato I sono considerati «tecniche e approcci di intelligenza artificiale»: *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.

profili di rischio ricollegantesi all'immissione nel circuito di utilizzo, con l'obiettivo di garantire adeguati standard di tutela e, soprattutto, assicurare la predeterminazione dei diversi livelli di responsabilità. Da un lato, pertanto, la previa immissione sul mercato di un sistema di intelligenza artificiale a rischio elevato implica la sottoposizione a una serie di controlli preordinati a garantire la sicurezza (attraverso una valutazione di conformità); dall'altro, il tema di un'appropriata progettazione e del consequenziale sviluppo del sistema di intelligenza artificiale risulta correlato alla possibilità di un'effettiva ed efficace sorveglianza da parte di persone fisiche allorquando lo stesso è utilizzato[50].

4.1. *I sistemi vietati*

Sono vietati in termini assoluti i sistemi che mirano a manipolare in base a tecniche subliminali la condotta delle persone oppure fanno leva sulle vulnerabilità di alcuni soggetti al fine di condizionarne la condotta e provocare un danno fisico o psicologico all'utente o ad un'altra persona (art. 5, lett. *a* e *b*).

Risultano, invece, proibiti soltanto in linea di principio i sistemi di IA utilizzati «da parte di autorità pubbliche o per loro conto» per stabilire l'affidabilità delle persone in base alla loro condotta sociale o alle caratteristiche personali[51] (art. 5, lett. *c*). Tali applicazioni sono proibite solo se esse

[50] Rileva Di Rosa, *Quali regole per i sistemi automatizzati?*, cit., che tale modello sembra riproporre la disciplina in tema di danno da prodotto difettoso, richiamandosi cioè il generale principio secondo cui i prodotti che non soddisfano le norme di sicurezza obbligatorie sono considerati difettosi (indipendentemente dalla colpa del produttore), non escludendo la possibile apertura a regimi di responsabilità oggettiva.

[51] Ad esempio, è il caso dei sistemi di *credit scoring*. Sui sistemi di *credit scoring* cfr. G. Sciascia, *Reputazione e potere: il social scoring tra distopia e realtà*, in «Giornale di diritto amministrativo», 2021, n. 3, pp. 317 ss.; G. Biferali, *Big Data e valutazione del merito creditizio per l'accesso al peer to peer lending*, in «Diritto dell'informazione e dell'informatica», 3, 2018, n. 1, pp. 487 ss.; C. Alvisi, *I trattamenti nel settore bancario, finanziario e assicurativo*, in L. Califano e C. Colapietro (a cura di), *In-*

determinano un trattamento pregiudizievole o, comunque, sfavorevole in un contesto scollegato a quello in cui i dati sono stati generati oppure ad un trattamento pregiudizievole che sia ingiustificato o sproporzionato rispetto alla condotta sociale e alla sua gravità.

Viene altresì impedito (lett. *d*) l'uso di sistemi di identificazione biometrica[52] in modalità *real time* in spazi aperti al pubblico per finalità di polizia[53], a meno che non siano strettamente necessari per la ricerca mirata di potenziali vittime di azioni criminose, come bambini scomparsi, per la

novazione tecnologica e valore della persona. Il diritto alla protezione dei dati personali nel Regolamento UE 2016/679, Napoli, 2017.

[52] Il Reg. n. 2016/679 (GDPR) definisce come «dati biometrici» quelli ottenuti «da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica che ne consentono o confermano l'identificazione univoca» (art. 4, comma 14, Reg. n. 2016/679), riconoscendo ad essi il rango – ed il conseguente regime di tutela – di dati sensibili. Ai sensi dell'art. 9, par. 1, del GDPR «È vietato trattare dati personali che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, o l'appartenenza sindacale, nonché trattare dati genetici, dati biometrici intesi a identificare in modo univoco una persona fisica, dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona». Per un inquadramento generale del tema cfr. I. Berle, *Face Recognition Technology. Compulsory Visibility and Its Impact on Privacy and the Confidentiality of Personal Identifiable Images*, Cham, 2020; L.L. Morrison, *Portraits of Automated Facial Recognition: On Machinic Ways of Seeing the Face*, Bielefeld, 2019; K.A. Gates, *Our Biometric Future. Facial Recognition Technology and the Culture of Surveillance*, New York, 2011; L. Introna e H. Nissenbaum, *Facial Recognition Technology: A Survey of Policy and Implementation Issues*, in Lancaster University Management School Working Paper, 2010. Per una panoramica globale sul contesto di utilizzo nel settore pubblico si veda la ricerca redatta per l'organizzazione German Marshall Fund of the United States da R. Richardson, *Facial Recognition in the Public Sector: The Policy Landscape*, pubblicata nel mese di febbraio 2021. Con specifico riferimento al contesto di impiego da parte di autorità pubbliche: A. Mascolo, *Riconoscimento facciale e autorità pubbliche*, in «Giornale di diritto amministrativo», 2021, n. 3, pp. 462 ss.

[53] Secondo la ricerca *At least 11 police forces use face recognition in the EU, AlgorithmWatch reveals*, redatta nel 2019 da Algorithm Watch, dei venticinque Stati membri dell'Unione europea passati in rassegna, almeno dieci hanno una forza di polizia che utilizza il riconoscimento del volto e altri otto prevedono di introdurlo nei prossimi anni.

prevenzione di un pericolo specifico, sostanziale e imminente alla vita o alla sicurezza di una persona o di un attacco terroristico o, infine, per la individuazione, localizzazione o incriminazione di un soggetto sospetto di reati previsti dall'art. 2(2) della decisione quadro del Consiglio 2002/584 per i quali lo Stato membro interessato preveda una pena detentiva pari o superiore a tre anni.

Ad ogni modo, anche nei casi in cui l'utilizzo di sistemi di identificazione biometrica in *real time* in spazi accessibili al pubblico è consentito, l'autorità pubblica deve graduarne l'utilizzo tenendo conto della «gravità, la probabilità e l'entità del danno causato dal mancato uso del sistema» nonché delle relative possibili implicazioni per i diritti e le libertà delle persone coinvolte[54].

Tuttavia, come è stato evidenziato[55], anche in virtù dell'utilizzo di concetti indeterminati e suscettibili di un margine di interpretazione, la disposizione lascia un notevole margine di manovra agli Stati membri, con il rischio di applicazioni difformi della disciplina all'interno del territorio unionale ed un conseguente altrettanto differenziato regime di tutela per i soggetti coinvolti nell'attività di trattamento automatizzato.

4.2. *I sistemi ad alto rischio*

La parte preponderante della disciplina regolamentare è dedicata ai sistemi di IA classificati come «ad alto rischio». Sono considerati ad alto rischio i dispositivi che soddisfino

[54] Per quanto riguarda il paragrafo 1, lett. *d*, e il paragrafo 2, ogni singolo uso di un sistema di identificazione biometrica remota «in tempo reale» in spazi accessibili al pubblico a fini di attività di contrasto è subordinato a un'autorizzazione preventiva rilasciata da un'autorità giudiziaria o da un'autorità amministrativa indipendente dello Stato membro in cui deve avvenire l'uso, rilasciata su richiesta motivata e in conformità alle regole dettagliate del diritto nazionale di cui al paragrafo 4. Tuttavia, in una situazione di urgenza debitamente giustificata, è possibile iniziare a usare il sistema senza autorizzazione e richiedere l'autorizzazione solo durante o dopo l'uso.

[55] Casonato e Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, cit.

entrambe le condizioni[56] fissate dall'art. 6 della Proposta di Regolamento nonché quelli specificatamente individuati all'allegato III. Tra di essi rientrano, in particolare, i sistemi di IA destinati a essere utilizzati per l'identificazione biometrica remota – sia in tempo reale che «a posteriori» delle persone fisiche; i sistemi di IA destinati a essere utilizzati come componenti di sicurezza nella gestione del traffico stradale e nella fornitura di acqua, gas, riscaldamento ed elettricità; i sistemi utilizzati per la gestione dell'istruzione e della formazione professionale nonché per la gestione dei lavoratori e per l'accesso al lavoro autonomo; i sistemi utilizzati per l'accesso a prestazioni e servizi pubblici e a servizi privati essenziali e fruizione degli stessi.

In relazione a tali dispositivi, la normativa europea impone che il *provider* adotti, documentandolo ed aggiornandolo, un sistema di gestione dei rischi che comprende: l'identificazione e l'analisi dei rischi noti e/o prevedibili associati all'utilizzo del sistema di IA; la stima e la valutazione dei rischi che possono insorgere anche nel caso di uso improprio ragionevolmente prevedibile del sistema; la valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivamente all'immissione sul mercato; l'adozione di adeguate misure di gestione dei rischi. In particolare, tali ultime misure devono essere tali da assicurare che «qualsiasi rischio residuo associato a ciascun pericolo nonché il rischio residuo complessivo dei sistemi di IA ad alto rischio sono considerati accettabili» e comprendono l'eliminazione o la riduzione dei rischi attraverso un'adeguata progettazione o fabbricazione, l'attuazione di misure di mitigazione in relazione a quei fattori di rischio

[56] Le condizioni fissate dall'art. 6 della Proposta di Regolamento sono le seguenti*: a)* il sistema di IA è destinato a essere utilizzato come componente di sicurezza di un prodotto, o è esso stesso un prodotto, disciplinato dalla normativa di armonizzazione dell'Unione elencata nell'allegato II; *b)* il prodotto, il cui componente di sicurezza è il sistema di IA, o il sistema di IA stesso in quanto prodotto è soggetto a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio di tale prodotto ai sensi della normativa di armonizzazione dell'Unione elencata nell'allegato II.

che non sono suscettibili di essere eliminati del tutto e la fornitura di informazioni adeguate agli utenti.

La Proposta di Regolamento specifica i requisiti minimi affinché i sistemi ad alto rischio siano ammessi all'interno dell'Unione europea. Il primo requisito si riferisce alla qualità e ad un sistema di *governance* dei dati, prevedendosi l'impiego di set di dati pertinenti, completi ed esenti da errori[57] (art. 10). Il secondo requisito concerne la documentazione tecnica, redatta in modo da dimostrare che il sistema sia conforme ai requisiti richiesti dal Regolamento (art. 11). Essenziale è anche l'accuratezza, la robustezza e la sicurezza dei sistemi di IA, da assicurare sin dalla fase di progettazione di sviluppo (art. 15) anche al fine di rendere tali sistemi resilienti rispetto ad errori, guasti ed incongruenze anche dovute all'interazione con persone fisiche.

In linea di continuità con il *Libro bianco* sull'IA, rilievo centrale assume la garanzia di un livello adeguato di trasparenza[58], ritenuta necessaria per assicurare la relativa

[57] Ai sensi del par. 2 dell'art. 10 i set di dati di addestramento, convalida e prova sono soggetti ad adeguate pratiche di *governance* e gestione dei dati. Tali pratiche riguardano 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*) 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.

[58] Sul tema della trasparenza algoritmica si richiamano, tra gli scritti recenti maggiormente significativi, J. Kemper e D. Kolkman, *Transparent to Whom? No Algorithmic Accountability without a Critical Audience*, in «Information, Communication & Society», 2019; A. Datta, S. Sen e Y. Zick, *Algorithmic Transaprency via Quantitative Input Influence*, in T. Cerquitelli, D. Quercia e F. Pasquale, *Transparent Data Mining for Big and Small Data*, Cham, 2017; M. Ananny e K. Crawford, *Seeing without Knowing: Limitations of Transparency Ideal and Its Application to Algorithmic Accountability*, in «New Media and Society», 2016, pp. 1-17; N. Diakopoulos, *Accountability in Algorithmic Decision Making*, in «Communications of the ACM», 59, 2016, n. 2; F. Pasquale, *The Black Box Society. The Secret Algorithms That Control Money and Informa-*

accountability e rendere possibile l'eventuale contestazione delle determinazioni assunte dall'IA. La Proposta di Regolamento richiede, a tale scopo, che ogni sistema di IA ad alto rischio sia disegnato e sviluppato in modo da tale «da assicurare che il loro funzionamento sia sufficientemente trasparente da consentire agli utenti di interpretare l'output del sistema e utilizzarlo adeguatamente» (art. 13). In tal modo, la Proposta di Regolamento recepisce un'accezione di trasparenza intesa come *smart disclosure*: ciò che rileva non è tanto l'accessibilità al codice sorgente[59], neppure menzionato, bensì l'effettiva comprensibilità della logica e dei meccanismi di funzionamento dell'algoritmo[60].

tion, Cambridge-London, 2016; J. Burrell, *How the Machine «Thinks»: Understanding Opacity in Machine Learning Algorithms*, in «Big Data & Society», 2016, n. 1; Ananny, *Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness*, cit.

[59] La giurisprudenza italiana è da tempo orientata nel senso che il diritto di accesso vada esteso anche al codice sorgente, ritenendo irrilevanti eventuali opposte esigenze di tutela del copyright (TAR Lazio, Roma, sez. III-*bis*, 1/7/2020, n. 7526; Cons. Stato, sez. V, 13/12/2019, n. 8472). Aderendo ad una tesi che ritenesse preclusivo l'accesso ai codici sorgente, che del programma informatico costituiscono la scrittura in linguaggio informatico delle attività digitalizzate, rileva la sentenza TAR Lazio n. 7526/2020, «si finirebbe per legittimare l'oscuramento di atti che incidono su rilevanti porzioni di attività amministrativa afferenti alla gestione di pubblici concorsi, con evidente vulnus al principio di trasparenza. Si produrrebbe, in sostanza, una insostenibile situazione di "doppio binario" dove nei concorsi gestiti con l'ausilio di strumenti informatici la regola della trasparenza avrebbe una portata ridotta rispetto alle procedure concorsuali tradizionali nelle quali, peraltro, molte delle attività concorsuali sono lasciate alla sola autonomia e responsabilità dei candidati, per cui non si pongono affatto problemi di accessibilità a previ atti che, come nel caso di specie, possano incidere sugli effetti di tali attività».

[60] La prescelta impostazione si pone in linea di continuità con il Regolamento n. 679/2016. Gli artt. 13, comma 2, lett. *f* e 14, comma 2, lett. *g*, stabiliscono che il titolare del trattamento è tenuto fornire ai soggetti interessati informazioni in ordine all'«esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'art. 22, paragrafi 1 e 4 e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l'importanza e le conseguenze previste di tale trattamento per l'interessato». Per una ricostruzione del dibattito sull'esistenza, o meno, di un diritto alla spiegazione si vedano tra i principali contributi

Sempre nell'ottica di assicurare un approccio antropocentrico all'IA, va letto l'art. 14 della Proposta di Regolamento, il quale impone che i sistemi di IA ad alto rischio siano progettati e sviluppati «in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui il sistema di IA è in uso», al precipuo fine di prevenire o ridurre al minimo i rischi per la sicurezza o i diritti fondamentali che possano emergere nel contesto dell'utilizzo del sistema di IA. In linea di ideale continuità con l'art. 22 del Regolamento n. 679/2016[61], la Proposta di

S. Wachter, B. Mittelstadt e L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, cit.; I. Mendoza e L.A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, in «Eu Internet Law: Regulation And Enforcement», 2017; L. Edwards e M. Veale, *Slave to the Algorithm? Why a «Right to Explanation» is Probably Not the Remedy You Are Looking For*, in «Duke Law & Technology Review», 2017, n. 16; M. Brkan, *Do Algorithms Rule the World? Algorithmic Decision-making and Data Protection in the Framework of the GDPR and Beyond*, in «International Journal of Law and Information Technology», 2018, n. 27; G. Malgieri e G. Comandé, *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection*, in «International Data Privacy Law», 7, 2017; A. Selbst e J. Powles, *Meaningful Information and the Right to Explanation*, in «International Data Privacy Law», 7, 2017, n. 4; B. Goodman e S. Flaxman, *EU Regulations on Algorithmic Decision-Making and a «Right to Explanation»*, in «AI Magazine», 2017 (accessibile al link https://arxiv.org/abs/1606.08813).

[61] L'art. 22 del GDPR ha attribuito al soggetto interessato il «diritto di non essere sottoposto ad una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona». La portata di tale disposizione risulta, però, dall'ampio catalogo di eccezioni previste dall'art. 22, par. 2, e dall'art. 23 del Regolamento. Resta, tuttora, non chiarito se la norma abbia l'effetto di introdurre un diritto in capo al soggetto interessato di non essere sottoposto ad un trattamento completamente automatizzato ovvero un divieto generalizzato in capo al titolare del trattamento – salvo il ricorrere di una delle ipotesi di cui all'art. 22, par. 2, lett *a)* e *c)* – all'utilizzo di procedure completamente automatizzate. L'ambiguità lessicale è di non poco conto poiché le due interpretazioni offrono due protezioni completamente diverse: in un caso vi è un divieto generale all'utilizzo di procedure automatizzate e nell'altro un diritto, azionabile dal soggetto se e allorquando riceva dal titolare del trattamento la notifica dell'avvio di un procedimento. Su tale dibattito cfr. Mendoza e Bygrave, *The Right Not to Be Subject*

Regolamento ha inteso consacrare un generale obbligo di una supervisione umana[62], esercitata da un soggetto dotato delle competenze idonee a comprendere appieno le capacità e i limiti del sistema di IA ad alto rischio 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 tempestivamente (art. 14, lett. *a*).

4.3. *Il sistema di «governance»*

Il titolo VI della Proposta di Regolamento è dedicata al sistema di *governance* dell'IA, che si caratterizza per l'adozione di un approccio di cooperazione tra istituzioni dell'Unione e Stati membri.

A livello dell'Unione, viene istituito un Comitato europeo per l'intelligenza artificiale (art. 56), con il compito di

to Automated Decisions Based on Profiling, cit.; D. Sancho, *Automated Decision-Making under Article 22 GDPR: Towards a More Substantial Regime for Solely Automated Decision-Making*, in Ebers e Navas (a cura di), *Algorithms and Law*, cit.

[62] Ai sensi dell'art. 14, par. 4, le misure devono essere in grado di consentire le seguenti azioni, a seconda delle circostanze, alle persone alle quali è affidata la sorveglianza umana: *a*) comprendere appieno le capacità e i limiti del sistema di IA ad alto rischio 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 da un sistema di IA ad alto rischio («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; *d*) essere in grado di decidere, in qualsiasi situazione particolare, di non usare il sistema di IA ad alto rischio o altrimenti di ignorare, annullare o ribaltare l'output del sistema di IA ad alto rischio; *e*) essere in grado di intervenire sul funzionamento del sistema di IA ad alto rischio o di interrompere il sistema mediante un pulsante di «arresto» o una procedura analoga.

fornire consulenza ed assistenza alla Commissione europea[63], di promuovere la cooperazione tra la Commissione e le autorità nazionali e di garantire un'uniforme applicazione del Regolamento, anche attraverso la formulazione di pareri, raccomandazioni e contributi scritti. Com'è evidente dalla disamina dei compiti affidati, il Comitato non si pone in una posizione sovraordinata rispetto alle autorità degli Stati membri. Piuttosto esso si pone come struttura di raccordo, come emerge anche dalla sua stessa composizione, del Comitato, al cui tavolo, affiancano la presidenza della Commissione, siedono i vertici delle amministrazioni nazionali di vigilanza ed il Garante europeo per la protezione dei dati personali.

L'applicazione e l'implementazione del Regolamento è affidata, invece, alle autorità degli Stati membri, «organizzate e gestite in modo che sia salvaguardata l'obiettività e l'imparzialità dei loro compiti e attività». Viene così in rilievo, come è stato osservato[64], un modello di amministrazione comunitaria indiretta, alla stregua del quale «la Commissione, pur lasciando agli Stati la decisione in ordine all'assetto istituzionale ottimale, incide in vario modo sulle scelte organizzative interne».

[63] Ai sensi del par. 2 dell'art. 56 il Comitato fornisce consulenza e assistenza alla Commissione europea al fine di: *a*) contribuire all'efficace cooperazione delle autorità nazionali di controllo e della Commissione per quanto riguarda le materie disciplinate dal Regolamento; *b*) coordinare e contribuire agli orientamenti e all'analisi della Commissione, delle autorità nazionali di controllo e di altre autorità competenti sulle questioni emergenti nel mercato interno in relazione alle materie disciplinate dal Regolamento; *c*) assistere le autorità nazionali di controllo e la Commissione nel garantire l'applicazione uniforme del Regolamento.

[64] Casonato e Marchetti, *Prime osservazioni sulla proposta di regolamento dell'Unione europea in materia di intelligenza artificiale*, cit. che osservano come tale modello di *governance* è interessante sotto un duplice ordine di profili: da un lato in quanto esso dimostra «la crescente tendenza del diritto dell'Unione a conformare i sistemi amministrativi nazionali anche in relazione ai profili, organizzativi e procedurali, che storicamente erano considerati terreno riservato all'autonomia procedurale degli Stati»; e dall'altro perché «conferma la necessità, sempre più avvertita, di competenze ad abilità trasversali e interdisciplinari all'interno delle pubbliche amministrazioni in ragione della complessità tecnica delle funzioni da esercitare».

5. *Considerazioni conclusive*

Anche nella Proposta del nuovo regolamento emerge la permanenza di una fondamentale tensione – che ha sino ad ora caratterizzato l'approccio delle istituzioni europee all'IA in tutti i documenti sopracitati – tra gli obiettivi economici e la tutela dei diritti fondamentali da parte delle istituzioni europee. Per un verso, infatti, l'Europa intende imporsi come leader nel settore delle tecnologie digitali e dell'intelligenza artificiale; per l'altro, è presente nell'approccio delle istituzioni europee l'obiettivo di assicurare che lo sviluppo dell'IA non comprima il godimento dei diritti e delle libertà fondamentali riconosciute ai cittadini europei. La vera sfida che l'Europa è chiamata ad affrontare è quella di essere in grado di raggiungere contemporaneamente ambedue gli ordini di obiettivi.

Un profilo da segnalare riguarda l'assetto di *governance*. Le norme sono stabili, poiché il Regolamento è dotato di diretta applicabilità ed efficacia diretta, e sono in grado di superare le incertezze e i mosaici di norme nazionali, ma non sono prescrittive fino in fondo[65]. Per garantire che i sistemi di gestione dei rischi siano resi più flessibili, agili e adattivi per far fronte alla velocità di innovazione e trasformazione in atto rispetto agli approcci e ai modi di pensare tradizionali, le regole europee sono basate su principi, ma le stesse possono essere integrate mediante atti delegati, potendo così mutare nel tempo come in un processo dinamico. Inoltre, anche la scelta dell'approccio basato sul rischio appare

[65] Nel documento *Progetto di relazione sull'intelligenza artificiale in un'era digitale* redatto dalla Commissione speciale sull'intelligenza artificiale in un'era digitale (2020/2266[INI]), pubblicato il 2/11/2021, viene rimarcato come il forte ritardo dell'UE rispetto a Cina e USA dipenda soprattutto dalla frammentazione normativa a livello dell'Unione («le incongruenze che caratterizzano il diritto dell'UE, le contraddizioni tra il diritto dell'UE e quello nazionale, le diverse interpretazioni giuridiche e la carente applicazione a livello di Stati membri») nonché dalla persistente incertezza giuridica («in alcuni settori mancano standard e norme comuni, mentre altri sono danneggiati da un'eccessiva regolamentazione o dalla presenza di proposte legislative rimaste pendenti per molto tempo senza essere adottate»).

particolarmente congeniale avuto riguardo della complessità e del carattere cangiante dei sistemi di IA, rispetto alla cui evoluzione un sistema regolamentare rigido difficilmente riuscirebbe a stare al passo.

A fronte di ciò, l'attuazione amministrativa è lasciata alla combinazione e alla cooperazione tra livello europeo e autorità nazionali, senza abbandonare il principio di sussidiarietà. Il *Board* per l'intelligenza artificiale è chiamato a difendere un approccio comune, condividendo *best practices*, sviluppando pratiche amministrative uniformi, fornendo opinioni, linee guida o raccomandazioni, nonché standard armonizzati e specifiche tecniche in tema di IA. Mentre alle autorità nazionali, responsabili dell'applicazione e dell'attuazione del Regolamento a livello statale, è lasciato il compito di supervisione del sistema e di vigilanza sugli organismi di valutazione della conformità.

È indubbio che con la Proposta di Regolamento sull'IA si compia un salto innovativo avente un forte impatto sulle Big Tech. Si prende atto che la tutela da accordare ai diritti fondamentali, che possono essere pregiudicati dal processo di digitalizzazione, non può essere garantita da un sistema che si autoregola e che non assicuri l'utilità sociale. L'IA deve essere affidabile, ingenerare fiducia e sicurezza nei consumatori, ed essere incentrata sulla tutela della persona secondo il concetto kantiano dell'uomo come fine e non come mezzo. Sintetizzato così, l'approccio europeo è emblematico perché fondato su un'architettura robusta, elegante e chiara con cui si intende attirare investimenti senza sospendere la tutela dei valori e dei diritti fondamentali. Che bilancia l'interesse alla ricerca con la salvaguardia della *rule of law*, in funzione della certezza e della prevedibilità per gli operatori.

D'altronde, l'approntamento di uno specifico modello regolatorio serve ad avviare politiche di gestione e organizzazione amministrativa che possano fungere da solido presupposto alle procedure informatizzate garantendo l'interoperabilità. Non è sufficiente abilitare con una norma legittimante le autorità a fare uso di questa tecnologia, poiché è necessario fornire piuttosto un quadro regolatorio di promozione e implementazione di tale sviluppo tecnologico e

di standardizzazione dei dati, con regole precise sui livelli di accesso alle informazioni, assicurando l'adempimento degli obblighi imposti dal GDPR a norma del quale vi deve essere una totale trasparenza sul trattamento dei dati personali degli utenti[66]. Solo con una regolamentazione chiara e certa, posta primariamente a tutela dei diritti fondamentali della persona e basata sui principi della trasparenza e della responsabilità, è auspicabile un vasto impiego di tale strumento innovativo.

[66] In questa logica è ampiamente condivisa la raccomandazione del MISE, *Proposte per la Strategia italiana*, cit., circa «l'attivazione di una struttura unitaria, con le modalità individuate dal Governo, per una Governance nazionale per le tecnologie innovative con lo scopo di *i*) contribuire a definire in termini coordinati politiche e interventi concreti, anche mediante consultazioni pubbliche, nel campo delle nuove tecnologie, nel rispetto del principio della neutralità tecnologica; *ii*) coordinare gli investimenti in un disegno di interventi unitario e sinergico; *iii*) attivare e favorire, se necessario, la collaborazione tra istituzioni, comitati e uffici pubblici con competenze e compiti in materia, anche segnalando l'opportunità di regolamentazioni congiunte delle Autorità di settore; *iv*) dialogare con istituzioni e uffici competenti dell'UE e degli altri Paesi; *v*) monitorare l'impatto degli interventi realizzati; *vi*) esprimere pareri nelle fasi di formazione della normativa primaria e secondaria e nelle fasi di recepimento della normativa europea nell'ordinamento italiano».