PROPOSTA DI REGOLAMENTO DELLA COMMISSIONE EUROPEA DEL 21 APRILE 2021 SULL'INTELLIGENZA ARTIFICIALE CON PARTICOLARE RIFERIMENTO ALLE IA AD ALTO RISCHIO.

Giovanna Marchianò*

Abstract

Nel presente scritto si vuole esaminare, sia pure nella consapevolezza di un maggior approfondimento, la proposta di regolamento della Commissione sull'intelligenza artificiale, con particolare riferimento all'IA ad alto rischio. La proposta presenta due aspetti sicuramente positivi: prima di tutto l'aver tentato di dare una cornice giuridica all'IA e, di conseguenza, un'armonizzazione del mercato in merito alle stesse. Non risulta invece soddisfacente nella regolamentazione delle IA e oscura appare la parte delle IA ad alto rischio compreso l'allegato a cui si dovrebbe far riferimento. La proposta di regolamento richiede, ad avviso di chi scrive, una revisione anche sulla tecnica utilizzata nella redazione della futura norma.

The purpose of this paper is to examine, albeit not in depth, the Commission's proposal for a regulation on artifical intelligence, with particular reference to high-risk AI. The proposal has two very positive aspects: firstly, the attempt to give the IA a legal framework and, consequently, market harmonization in relation to them. On the other hand, it is not satisfactory in the regulation of Ics and the higt-risk part of Ics, including the annex, to which reference should be made, is obscured. In my opinion, the proposal for a regulation also requires a review of the technique used in drafting the future standard.

SOMMARIO: 1. Perimetro definitorio di I.A. nella proposta di regolamento presentata dalla Commissione; **- 2.** L'obiettivo dell'armonizzazione del mercato interno in merito all'IA, basato sul "rischio"; **- 3.** Le IA vietate e/o regolamentate; **- 4.** Le IA ad alto rischio, *ex* artt. 6 e 7 della proposta di regolamento; **- 5.** Il sistema di *governance* nell'IA ad alto rischio; **- 6.** Possibili conclusioni.

1-. Perimetro definitorio di I.A. nella proposta di regolamento presentata dalla Commissione

In data 21 aprile 2021 è stata presentata una proposta della Commissione europea in tema d'intelligenza artificiale (I.A.) nella convinzione, peraltro fondata, che tale tema sia tra le priorità dell'agenda politica nel futuro dell'Unione. L'assunto trova conferma nell'intenso programma di interventi sul tema dell'IA che il Parlamento europeo e la Commissione hanno adottato in questi ultimi anni[1]. La proposta, infatti, deve essere inserita nel più ampio disegno europeo che, col *Digital Service Act* (DSA), è intervenuta sulla responsabilità delle piattaforme e con il *Digital Market Act* (DMA), ha puntato a proteggere le piccole imprese dalla predominanza dei giganti del *web*.

Ora, la Commissione europea ha elaborato un ambizioso progetto con l'obiettivo non solo di regolamentare ma anche di promuovere l'uso dell'intelligenza artificiale in Europa nel rispetto delle libertà individuali e delle disposizioni contenute nella Carta europea dei diritti fondamentali. E' evidente il tentativo dell'Unione di recuperare terreno rispetto ad altri Paesi come gli Stati Uniti e la Cina.

La prima osservazione attiene al dover prendere atto che, nella proposta della Commissione, di fatto si è adottato il perimetro definitorio d'intelligenza artificiale, così come contenuto nella comunicazione della Commissione al Parlamento europeo, al Consiglio e al Comitato economico e sociale europeo e al Comitato delle Regioni ovvero per IA, l'Europa fa riferimento a *quei sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni con un certo grado di autonomia per raggiungere specifici obiettivi.* Si tratta di *software* che agiscono nel mondo virtuale o dispositivi *hardware* quali i *robot* avanzati, l'auto a guida autonoma, i droni o le applicazioni dell'*internet*. Si è di fronte a sistemi complessi capaci di migliorare sé stessi con progressione esponenziale il ché già da tempo ha lanciato un allarme sociale per quanto attiene gli interessi e le posizioni della collettività, onde evitare che tali IA non siano oggetto né di abuso né di cattiva utilizzazione con conseguente

1 * Professore associato di Diritto pubblico, *Alma Mater Studiorum* - Università di Bologna.

1. Il 16 febbraio 2017 il Parlamento ha approvato una risoluzione di raccomandazione alla Commissione avente a oggetto norme di diritto civile sulla robotica, *(2015/2013 (INL)* A questa ha fatto seguito la successiva comunicazione del 25 aprile 2018 dal titolo *Intelligenza artificiale per l'Europa*. Nel maggio 2019 è stato pubblicato un *report* di un gruppo di esperti i cui risultati sono stati ripresi nella relazione del febbraio 2020, della stessa Commissione, *Sulle implicazioni dell'intelligenza artificiale dell'internet e della robotica in materia di sicurezza e responsabilità* (*Commissione europea COM (2020) 64 final*), 16 febbraio 2020. Il 20 gennaio 2021, nuovamente è stata emanata una risoluzione in tema di armi letali nella quale si chiede alla Commissione l'adozione di una strategia volta a proibire i sistemi d'arma se non sono soggetti al controllo umano.

incidenza negativa sui diritti e sulle libertà fondamentali garantiti dai principi europei.

Molteplici, e non sempre coincidenti sono le definizioni, peraltro di tipo descrittivo d'IA talché ci si è riferiti a «*the science of making machine do things that would require intelligence if done by men*, o a un *cross-disciplinary approach to understanding, modeling, and replicating intelligence and cognitive processes by invoking various computational, mathematical, logical, mechanical, and even biological principles and devices*»[2].

Il vero è che ormai non si può far riferimento a una singola tecnologia, quanto piuttosto a una pluralità e a un insieme di modalità tecniche che, sia pure a livelli diversi e in differenti modi, hanno in comune il fatto di porre in essere un comportamento "intelligente" in numerosi scenari[3].

Nella proposta di regolamento non vi è una definizione se non implicita d'IA, facendosi riferimento a un *software* sviluppato con una o più tecniche e approcci tra quelle elencate nell'allegato I, destinati a raggiungere determinati obiettivi definiti dall'uomo generando risultati quali previsioni, raccomandazioni o decisioni che influenzano gli ambienti con i quali interagiscono. Al fine di circoscrivere i sistemi d'IA a cui si applica la proposta in esame, occorre quindi far riferimento all'allegato I ove si legge:

«*(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;*

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

2 La prima definizione d'IA fa riferimento a L. COPLAND, *Artificial Intelligence: Philosophical Instroduction*, New Jersey, 1993, 1; la seconda definizione fa riferimento a K. FRANKISH, W. M. RAMSEY (EDS), *The Cambridge Handbook of Artificial Intelligence*, Cambridge, 2014, 7. Riprende tali definizioni C. CASONATO, *Intelligenza artificiale e diritto costituzionale: prime considerazioni*, in *Diritto pubblico comparato ed europeo*, Il Mulino, Fascicolo speciale, maggio 2019.

3 Quanto alla ricerca del perimetro volto a delineare il fenomeno dell'IA, occorre far riferimento al gruppo di esperiti istituito nel 2018 dalla Commissione europea *AI HLEG*, che ha elaborato una definizione ampia dell'IA secondo cui, *i sistemi d'intelligenza artificiale sono sistemi di software (e verosimilmente anche di hardware) progettati da esseri umani che, dato un obiettivo complesso, agiscono all'interno di una dimensione fisica o digitale percependo il loro insieme attraverso l'acquisizione di dati, interpretando i dati raccolti sia essi strutturali o non strutturali, ragionando sulle conoscenze o elaborando informazioni derivanti da questi dati e selezionando, tra tutte le azioni possibili, le migliori per il raggiungimento dell'obiettivo indicato.* AI-HLEG, *High-Level Expert Group on Artifical Intelligence*, *A definition of AI: Main capabilities and Scientific Diciplines*, 2019. Il 25 aprile 2018 la Commissione europea ha istituito ben tre gruppi di esperti: il primo gruppo per mettere a fuoco i principi etici d'IA, ed è quello a cui fa riferimento il testo; il secondo per studiare le eventuali modifiche della direttiva comunitaria del 1985 in materia di responsabilità da prodotto difettoso; il terzo gruppo sul tema della responsabilità giuridica *extra* contrattuale per le tecnologie emergenti. V. U. PAGALLO, *Etica e dritto dell'Intelligenza Artificiale nella governance del digitale: il Middle-out Approach, op. cit.*, p. 29

(c) Statistical approaches, Bayesian estimation, search and optimization methods».

Sostanzialmente trattasi di sistemi d'IA complessi capaci di un autoapprendimento automatico e di un alto livello di automazione[4]. E' noto che tali sistemi possono applicare diversi metodi di ragionamento pratico sui dati disponibili nonché differenti modi di apprendimento[5]. In genere nei sistemi basati sull'apprendimento automatico è possibile distinguere due componenti ovvero l'algoritmo addestratore e l'algoritmo addestrato: il secondo realizza il compito affidato al sistema, il primo modifica il secondo in modo che questo svolga meglio quel compito. In questo quadro si distinguono vari metodi di apprendimento: - apprendimento supervisionato; - apprendimento per rinforzo; - apprendimento non supervisionato. L'apprendimento supervisionato che a oggi è il metodo più utilizzato e che viene richiamato nella proposta di regolamento, prevede l'insegnamento alla macchina, vale a dire un insieme di dati di addestramento (*training set*), in prospettiva al compito che la stessa deve compiere. All'apprendimento supervisionato si accosta il c.d. "apprendimento per rinforzo", il sistema apprende dai propri risultati o delle altrui azioni; in genere si fa riferimento a sistemi che, attraverso l'autoapprendimento, migliorano le proprie strategie "di gioco" sulla base di risultati cui tali strategie conducono. Nell'addestramento non supervisionato, il sistema apprende senza ricevere indicazioni o istruzioni dall'esterno. Tecniche per l'addestramento non supervisionato sono utilizzate per il *cloustering* cioè per raggruppare, all'interno di un sistema, un insieme di elementi che hanno profili tra loro connessi, quali ad es. documenti relativi al medesimo argomento.

Fatte tali premesse, ai fini di mera rammentazione, occorre tener presente che, nel Libro Bianco[6] dal titolo *Un approccio europeo all'eccellenza e alla fiducia*, la Commissione

4 G. MARCHIANÒ, *Intelligenza artificiale: IA specifiche e l'amministrazione provvedimentale – IA generali e i servizi pubblici*, in www.federalismi.it 21 aprile 2021. Nel modello dell'apprendimento automatico i sistemi d'IA hanno una capacità d'incamerare grandi masse di dati (*machine learning*): in questo caso l'uomo fornisce un *software* da applicare ai dati cui la macchina ha accesso senza alcun altro intervento. Nei sistemi basati sull'apprendimento automatico è possibile distinguere due componenti funzionali: l'algoritmo addestratore e l'algoritmo addestrato. Il secondo realizza il compito affidato al sistema, il primo modifica il secondo in modo che questo svolga meglio quel compito. Pertanto, come osservava TURING, già alla fine degli anni '50, una macchina capace di apprendere realizza gli obiettivi assegnati senza che l'uomo abbia indicato ad essa come procedere anzi, senza che lo stesso abbia consapevolezza di ciò che accade all'interno della macchina.

5 D. HAREL, Y. FELDMAN, *Algorithmics: The Spirit of Computing*, Boston, 2004; D.U. GALETTA, J. G. CORVALAN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in www.federalismi.it, 3/2019: «*I sistemi di Intelligenza artificiale utilizzano computer, algoritmi e varie tecniche per elaborare informazioni e risolvere problemi o prendere decisioni che in precedenza potevano essere prese solo dall'intelligenza naturale. In fin dei conti – per semplificare al massimo – proprio come il cervello estrae, seleziona ed organizza le informazioni disponibili per prendere decisioni, l'Intelligenza artificiale fa lo stesso, ma con altri metodi e un'altra velocità*»; S. RUSSEL, P. NORVIG, *Artificial Intelligenze. A Modern Approch*, *Upple Saddle River*, 2016, i quali a p. 693, osservano che un agente impara da se stesso e migliora le proprie prestazioni rispetto ai compiti futuri dopo aver compiuto osservazioni sul mondo.

6 Libro Bianco, *Bruxelles, 19.2.2020 COM (2020) 65 final.*

aveva già sottolineato la necessità di dar luogo ad un quadro normativo della materia orientando gli investimenti in tema d'IA ad un duplice obiettivo, da un lato promuovere l'adozione dell'IA, dall'altro affrontare e/o limitare i rischi associati all'utilizzo delle stesse. La Commissione, sotto questo profilo, nulla dice di nuovo, poiché questa prospettiva era già stata sottolineata sia dagli organismi comunitari che dalla stessa Commissione nonché dalla dottrina. Si era da tempo messo in evidenza che, soprattutto per quanto riguarda la c.d. intelligenza artificiale capace di autoapprendimento, nella sua applicazione veniva a modificare gli assetti economici, politici e sociali; tali sistemi «…*si accompagnano tuttavia a gravi rischi tra cui la disoccupazione, le disuguaglianze, le esclusioni sociali*» e, la stessa dottrina suggeriva che, per fronteggiare questi rischi, senza limitarne la ricerca e i benefici sull'economia dell'IA, si dovesse approntare una normativa in materia. La Commissione non sottovaluta il fatto che, attraverso l'utilizzo dell'IA, si avrà uno sviluppo delle imprese in particolare nei servizi e nei settori in cui l'Europa già gode di una situazione particolarmente favorevole (macchinari, trasporti, *cyber* sicurezza, agricoltura economica verde e circolare, assistenza sanitaria e settori ad alto valore aggiunto come la moda e il turismo) ma emerge la rilevanza d'imporre un quadro etico per un'IA "*incentrata sull'uomo*"[7]. E' in questa prospettiva che si muove l'Europa e che trova conforto nella dottrina in materia; la proposta di regolamento ha fatto propri tali principi per perseguirne gli obiettivi. E' stata lanciata dalla Commissione una *survey* ove gli Stati membri vengono sollecitati insieme alle altre istituzioni europee, ai portatori d'interesse compresa l'industria, ai ricercatori e a tutte le parti interessate, al fine di contribuire al futuro processo decisionale della Commissione nel settore in esame.

Tuttavia, accanto agli elementi negativi sopra paventati (disoccupazione, disuguaglianza ed esclusione sociale) come sottolineato dal Libro Bianco è evidente che molti sono gli elementi positivi derivanti dall'utilizzo di tali sistemi. L'Europa ha come obiettivo quello di sviluppare un sistema d'IA che consenta alla società e all'economia, nel loro complesso, di godere a pieno dei benefici apportati dalla tecnologia, basti pensare che i cittadini potranno usufruire di nuovi vantaggi come il miglioramento dell'assistenza sanitaria, dei sistemi di trasporto nonché dei servizi pubblici venendo a beneficiare ad es. della riduzione dei costi di fornitura di servizi quali l'istruzione, l'energia e la gestione dei rifiuti, accentuando altresì la sostenibilità

7 G. SARTOR E F. LAGIOIA, *Le decisioni algoritmiche tra etica e diritto, op. cit.,* p. 66, ove si rileva che i rischi legati all'emergere di una "intelligenza generale artificiale" non devono essere sottovalutati, diversamente l'"intelligenza artificiale specifica" sta già trasformando gli assetti economici, politici e sociali. A tali opportunità si accompagnano tuttavia gravi rischi, tra cui la disoccupazione, le disuguaglianze, le esclusioni sociali. Per fronteggiare questi rischi, senza limitare la ricerca e gli usi benefici dell'IA, sono state adottate numerose iniziative volte a prospettare un quadro etico e normativo per un'IA "incentrata sull'uomo".

dei prodotti e dotando le forze dell'ordine di strumenti "appropriati" per garantire la sicurezza dei cittadini: tutto ciò però a condizione che non vengano stravolti il rispetto dei loro diritti e delle loro libertà fondamentali.

Sotto quest'ultimo profilo, l'impatto dei sistemi d'IA deve essere considerato non solo in una prospettiva individuale ma anche dal punto di vista della società nel suo complesso: infatti l'uso dei sistemi d'IA può svolgere un ruolo significativo per uno sviluppo sostenibile, rinforzare il processo democratico e, addirittura, il sostegno dei diritti sociali, contrariamente a quanto sottolineato dalla dottrina più risalente.

2. L'obiettivo dell'armonizzazione del mercato interno in merito all'IA, basato sul "rischio".

Nella relazione che accompagna la proposta di regolamento sull'IA, la Commissione rileva che, mentre il Libro Bianco è volto a definire le opzioni politiche su come raggiungere il duplice obiettivo di promuovere l'adozione dell'IA e, contestualmente, affrontare i rischi associati a determinati usi delle nuove tecnologie, la proposta mira invece a creare un quadro giuridico per una "*IA affidabile*" preservando i valori e i diritti fondamentali dell'Unione e della Carta europea; così, per creare un'IA affidabile, occorre che essa sia rispettosa della dignità umana, delle libertà individuali, dell'uguaglianza, della non discriminazione e della solidarietà.

L'obiettivo principale della proposta, sostenuta dal principio di sussidiarietà, è quello di assicurare un corretto funzionamento del mercato unico digitale per lo sviluppo, l'immissione e l'utilizzo di prodotti e servizi che si avvalgono della tecnologia dell'IA. Stante la diversità delle legislazioni nei singoli Paesi, la Commissione rileva che, data l'ampia circolazione transfrontaliera di tali "prodotti" e "servizi", è necessaria l'armonizzazione attraverso la legislazione comunitaria in quanto, norme nazionali divergenti, verrebbero a ostacolare la garanzia di sicurezza e di protezione dei diritti fondamentali e dei valori dell'Unione.

A ben vedere, anche sotto questo profilo, la proposta di regolamento non dice nulla di nuovo ponendosi sulla scia di quanto previsto nel Libro Bianco ove già si sottolineava l'opportunità di un approccio comune europeo all'IA al fine di raggiungere dimensioni sufficienti ad evitare la frammentazione del mercato stesso. L'introduzione d'iniziative nazionali rischia infatti di compromettere la certezza del diritto, d'indebolire la fiducia dei cittadini e di ostacolare l'emergere di un'industria europea dinamica. Si auspicava e si auspica, quindi, la realizzazione di un quadro strategico normativo che stabilisca misure per allineare gli sforzi a livello europeo,

nazionale e regionale anche tramite un partenariato tra settore pubblico e privato. L'obiettivo di tale quadro è mobilitare risorse per consentire un "*ecosistema d'eccellenza*" nonché un "*ecosistema di fiducia*" garantendo la tutela dei diritti fondamentali soprattutto per le strutture d'IA ad alto rischio operanti nell'Unione: la Commissione infatti sostiene con forza un approccio basato sulla direzione «*Creare fiducia nell'intelligenza artificiale antropocentrica*».

All'art. 2 della proposta in esame si sottolinea che, al fine di garantire un'efficace tutela dei diritti di libertà degli individui in tutta Europa, il regolamento dovrebbe applicarsi ai fornitori di sistemi d'IA indipendentemente dal fatto che essi operino nell'Unione o in un Paese terzo nonché a tutti coloro che utilizzano tali sistemi nel mercato interno.

La proposta di regolamento non trova applicazione nei sistemi d'IA utilizzati esclusivamente per scopi militari così come non si applica alle autorità pubbliche dei Paesi terzi né alle organizzazioni internazionali a meno ché non ci si trovi di fronte a fattispecie relative a casi di cooperazione giudiziaria tra Paesi terzi e l'Unione.

Infine si osservi come, la proposta di regolamento, non pregiudica l'applicazione delle disposizioni sulla responsabilità dei prestatori intermedi di servizi di cui al capo II, sez. IV, direttiva 2000/31/CE del Parlamento europeo e del Consiglio, anzi, la proposta pare avere l'obiettivo d'integrare i dati del Regolamento n. 679/2016 sulla protezione dei dati[8] nonché la Direttiva 680/2016/UE, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali[9].

All'art. 2 della proposta di regolamento, si esplicitano gli obiettivi che s'intendono raggiungere:

- garantire che i sistemi d'IA immessi nel mercato europeo siano sicuri e rispettosi dei diritti fondamentali e dei valori fondanti dell'Unione;
- garantire la certezza del diritto per facilitare gli investimenti e l'innovazione nell'IA; migliorare la *governance* e l'effettiva applicazione della legislazione esistente sui diritti fondamentali e i requisiti di sicurezza applicabili ai sistemi d'IA;

8 Regolamento generale sulla protezione dei dati – Regolamento n. 679/2016 del Parlamento europeo e del Consiglio, 27 aprile 2016. Aggiornato alle rettifiche pubblicate sulla Gazzetta Ufficiale dell'Unione europea 127 del 23 maggio 2018.

9 Direttiva 2016/680/UE del Parlamento europeo e del Consiglio del 27 aprile 2016 relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio.

- facilitare lo sviluppo di un mercato unico per l'IA disciplinato dal punto di vista giuridico in modo che sia "*sicura e degna di fiducia nell'applicazione*" e prevenirne la frammentazione delle discipline dei singoli Stati.

E' evidente come la Commissione abbia un obiettivo molto ambizioso: quello di fornire "*un quadro giuridico solido e flessibile*". In verità, nella proposta non è dato rinvenire un quadro giuridico "*solido*" anche se, data la complessità della materia e il suo continuo divenire, risulta sicuramente difficile cogliere tale obiettivo. Nella relazione di accompagnamento alla stessa si sottolinea che la proposta sarebbe completa e a prova di futuro nelle sue scelte fondamentali, in particolare sui "*principi*" comunitari che tutti i meccanismi d'IA devono rispettare, venendosi a prevedere un sistema normativo proporzionato incentrato su una regolamentazione basata sul "rischio" il ché vuol dire un approccio al tema che non crei restrizioni del mercato; tuttavia non va dimenticato, al di là di tali dichiarazioni, che le disposizioni in materia si troveranno a disciplinare fattispecie concrete che possono mettere a rischio i sopra citati principi. Si è invece d'accordo sul fatto che la disciplina normativa deve prevedere meccanismi flessibili al fine di consentire l'adattamento della regola giuridica rispetto all'evoluzione tecnologica, tratto tipico di tale settore. Sotto quest'ultimo profilo appare interessante l'art. 84 che prevede una valutazione e un riesame dell'Allegato III (sistemi d'IA ad alto rischio) una volta all'anno dopo l'entrata in vigore del regolamento; nel medesimo articolo si prevede che ogni quattro anni la Commissione valuti l'impatto e l'efficacia dei codici di condotta per favorire l'applicazione dei requisiti stabiliti nell'art. 3, capo 2 ed eventualmente altri requisiti aggiuntivi per l'IA ad alto rischio. E' interessante il fatto che, ogni quattro anni successivi all'entrata in vigore del regolamento, la Commissione presenti una relazione sulla valutazione e il riesame dello stesso al Parlamento europeo e al Consiglio così come la Commissione può apporre modifiche al presente regolamento, tenendo conto degli sviluppi tecnologici che, nel frattempo, sono venuti in essere. Per quanto scritto in modo oscuro sicuramente l'art. 84 cerca di garantire la "*flessibilità*" della normativa in parola. Flessibilità che la s'intravede anche su un diverso piano, ad es. nell'art. 4 ove si prevede che, alla Commissione, sia conferito il potere di adottare atti delegati volti a modificare l'elenco delle tecnologie contenute nell'allegato I, con l'obiettivo di conformare la normativa allo sviluppo dei sistemi d'IA. Il ricorso agli atti delegati, infatti, a modifica degli allegati, lo si rinviene varie volte nella proposta di regolamento sicché effettivamente, come si legge nel Libro Bianco e nella parte esplicativa della proposta, la disciplina in questione può definirsi *flessibile* in quanto, al continuo modificarsi di tali sistemi verrebbe a modificarsi anche l'oggetto della disciplina normativa.

Un solido quadro normativo europeo per l'IA affidabile, sarebbe l'unico strumento in grado di garantire condizioni di parità tra tutti i cittadini europei, rafforzando al contempo la competitività e lo sviluppo nell'Unione dell'IA, come si legge esplicitamente nella relazione di accompagnamento: *Only common action at Union level can also protect the Union's digital sovereignty and leverage its tools and regulatory powers to shape global rules and standards.* In effetti, la prima osservazione da farsi sulla proposta di regolamento è che, l'elemento di maggiore interesse contenuto nella stessa è rinvenibile nella volontà della Commissione di creare un'armonizzazione delle legislazioni degli Stati membri; anche nei Paesi ove non ci sia una legge *ad hoc* sull'IA, il tentativo più o meno riuscito, come si vedrà nel prosieguo delle presenti riflessioni, è sicuramente positivo almeno nello sforzo di creare un'architettura giuridica basata su un approccio proporzionale evitando di non utilizzare le nuove tecnologie[10]. Coerentemente, l'obiettivo porta a quanto si legge nell'art. 1, ove si stabiliscono i punti cardini sui quali si fonda la proposta in questione:

«*(a) harmonised rules for the placing on the market, the putting into service and the use of artificial intelligence systems ('AI systems') in the Union; prohibitions of certain artificial intelligence practices;*

(b) specific requirements for high-risk AI systems and obligations for operators of such systems;

(c) harmonised transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content;

(d) rules on market monitoring and surveillance».

10 Nei Paesi Bassi è in discussione un disegno di legge volto a prevedere che, decisioni automatiche, siano consentite solo a determinate condizioni e siano basate sulla profilazione automatica solo quando la legge lo consente: *https://marliesvaneck.wordpress.com/2017/01/31/algorithms-in-public-administration/* Nel diritto francese, la legge del 7 novembre 2016 ha introdotto nel codice dei rapporti tra il pubblico e l'amministrazione (Crpa) due serie di disposizioni in materia di *governance* algoritmica. La prima, oggetto dell'articolo L. 311-3-1, stabilisce che ogni «*decisione individuale, presa sulla base di un'elaborazione algoritmica, richiede una menzione esplicita che informi l'interessato. Le norme che definiscono questo trattamento e le caratteristiche principali della sua attuazione devono essere comunicate dall'amministrazione alla persona interessata se ne fa domanda*». Questo testo è completato dall'articolo R. 311-3-1-2, secondo il quale in caso di richiesta di accesso, l'amministrazione deve fornire in forma intelligibile: 1. il grado e le modalità di contributo dell'elaborazione algoritmica al processo decisionale; 2. i dati trattati e le loro fonti; 3. i parametri di trattamento applicati alla situazione della persona interessata e, se applicabile, la loro ponderazione; 4. le operazioni eseguite dal trattamento. Successivamente è stato introdotto l'articolo L. 312-1-3 del Crpa ove si prevede che «*le amministrazioni […] ad eccezione delle persone giuridiche il cui numero di agenti o dipendenti è al di sotto di una soglia stabilita per decreto, pubblicano online le regole che definiscono i principali trattamenti algoritmici utilizzati nell'esecuzione dei loro compiti quando su di essi basano delle decisioni individuali*». J. B. AUBY, *Il diritto amministrativo di fronte alle sfide digitali*, in *Istituzioni del federalismo*, 3/2019, p. 626. Sul rapporto tra il procedimento amministrativo italiano e l'algoritmo, si rinvia, tra gli altri, a G. MARCHIANÒ, *La legalità algoritmica nella giurisprudenza amministrativa,* in *Il diritto pubblico dell'economia*, 3/2020.

In questo caso l'armonizzazione del mercato viene assunta in senso negativo ovvero il divieto di utilizzare talune pratiche d'IA nonché di far riferimento a requisiti specifici per i sistemi d'IA ad alto rischio, con conseguenti obblighi per gli operatori di tali sistemi.

Tuttavia, l'obiettivo di creare una "*solida base giuridica*" in materia non sembra essere stato colto dalla Commissione in questa sede: il sommario rinvio ai principi dell'ordinamento comunitario e alle tutele dei diritti dei cittadini, appare quanto meno troppo generico e generale il ché già pone il problema del rapporto tra le nuove tecnologie e i principi dell'ordinamento comunitario in tema di diritti sociali, pubblici e privati. E' vero che in questi anni la giurisprudenza della Corte di Giustizia e la Corte europea dei diritti dell'uomo (CEDU), hanno forgiato la materia attraverso le loro decisioni ma, desta qualche perplessità, un rinvio così generico che dovrebbe tra l'altro rappresentare il limite all'utilizzo di taluni sistemi; cosicché, nell'applicazione alle singole fattispecie della disciplina contenuta nella proposta, tale limite potrebbe non arginare la decisione d'immettere nel mercato europeo, talune IA, soprattutto ove queste abbiano ripercussioni positive sull'economia travolgendo così, i principi a cui la proposta di regolamento fa riferimento.

3. Le IA vietate e/o regolamentate.

Più specifica è la disciplina dell'IA vietate, *ex* art. 5, sulla base del fatto che la loro utilizzazione potrebbe violare i valori fondanti dell'Unione, relative a determinati usi e sistemi d'IA, quale l'identificazione biometrica remota.

La *ratio* della disciplina contenuta nell'articolo in esame è rinvenibile nel fatto che, tali tipi d'IA, si pongono al di fuori dell'impianto della proposta che è quello di coniugare il mercato europeo in materia di tecnologie digitali con i diritti e i principi garantiti dall'Unione ai suoi cittadini. Sostanzialmente l'IA vietate attengono all'utilizzo di tecniche capaci di distorcere il comportamento di una persona o di sfruttarne la vulnerabilità a causa dell'età, della disabilità fisica o mentale, in modo da causare danni fisici o psicologici. Parimenti, al punto (c) dell'art. 5, si vieta l'uso dei sistemi d'IA volto ad una valutazione o alla classificazione dell'affidabilità delle persone fisiche, basandosi sul comportamento sociale o su caratteristiche personali dell'interessato, con la conseguenza di creare trattamenti pregiudizievoli verso talune persone o trattamenti di favore verso altre. In particolare è vietato, come si legge al punto (d) dell'art. 5, l'uso dei sistemi d'identificazione biometrica (riprendendo la definizione contenuta nell'art. 3.18 del Regolamento n. 1725/2018 ovvero «*i dati*

personali 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, quali l'immagine facciale o i dati dattiloscopici)»[11], nel caso in cui tali dati siano assunti a distanza, in spazi pubblici o accessibili a tutti a meno ché, tale misura, non sia specificamente mirata alla ricerca di vittime potenziali della criminalità, tra cui: i bambini scomparsi; la prevenzione di una minaccia specifica sostanziale imminente della sicurezza fisica delle persone, di un attacco terroristico o l'individuazione, la localizzazione, l'identificazione di un soggetto sospettato di reato per una pena detentiva di durata di almeno tre anni, secondo la legge di ciascuno Stato membro. Tuttavia al punto 3, è previsto che l'identificazione biometrica a distanza sia ammissibile, previa autorizzazione preventiva rilasciata o da un giudice o da un'autorità amministrativa indipendente dello Stato membro, su richiesta motivata del soggetto agente. Quest'ultimo dovrebbe operare un bilanciamento a fronte della natura della situazione e della gravità del danno causato dalla mancanza dell'uso di tale sistema, che potrà in questi casi essere invece utilizzato; pur tuttavia è evidente la difficoltà d'inquadrare a quali fattispecie si faccia riferimento stante la genericità della disposizione in oggetto.

In una situazione d'"*urgenza giustificata*", (concetto peraltro non chiaro in quanto la situazione d'urgenza viene solo successivamente giustificata dal soggetto agente non essendo rinvenibile da un punto di vista giuridico una situazione d'urgenza giustificata *tout court*), l'autorità giudiziaria o amministrativa competente dovrebbe essere in grado di fondare il rilascio di tale autorizzazione in base a fatti oggettivi con indicazioni chiare e proporzionate al conseguimento di uno di quegli obiettivi previsti nel par. 1, lett. d) dello stesso art. 5. Se si è di fronte ad una fattispecie "*d'urgenza giustificata*" riesce difficile coniugare la richiesta del soggetto agente con una motivazione derivante non solo dai fatti oggettivi ma anche da una valutazione di bilanciamento tra i vari fattori in gioco! Si pensi a un attacco terroristico, caso in cui l'applicazione di queste disposizioni sembra più ipotetica che non reale.

Il vero è che, il riconoscimento facciale non è proibito *tout court*, come sembrerebbe dalla prima parte dell'articolo in questione ma è disciplinato in relazione a singole fattispecie che appaiono talmente complesse da lasciare perplessi, circa un corretto utilizzo di autorizzazioni preventive che siano idonee a giustificare il ricorso a tale sistema. Forse vi rientra il caso del bambino scomparso, dove vi è un lasso di tempo tra la scomparsa e la ricerca dello stesso e, tale lasso di tempo, consentirebbe un'autorizzazione sulla falsa riga richiesta dalla proposta di regolamento ma, negli

11 Per definire la nozione di dati biometrici occorre far riferimento al Regolamento n. 679/2016, art. 3, e del Regolamento n. 1725/2018 del Parlamento e del Consiglio nonché all'art. 3 della Direttiva 2016/680/UE del Parlamento europeo e del Consiglio.

altri casi, diventa difficile tant'è che sembra rinviarsi questa procedura allo Stato membro. L'articolo, infatti, si chiude affermando che uno Stato membro può decidere di autorizzare l'uso di sistemi d'identificazione biometrica nei casi previsti nel par. 1, lett. (d), 2 e 3. Lo Stato membro stabilisce, nell'ambito della propria normativa, le modalità di applicazione necessarie per la richiesta, l'emissione e l'esercizio delle autorizzazioni di cui si è detto precedentemente. Dette disposizioni dovrebbero inoltre precisare per quali reati sia consentito l'uso dell'identificazione biometrica e le autorità che possono essere autorizzate all'applicazione di tali sistemi.

4. Le IA ad alto rischio, *ex* artt. 6 e 7 della proposta di regolamento.

Il criterio sul quale si basa la disciplina proposta è "*il rischio*", da ciò la suddivisione dell'IA secondo una graduazione dello stesso: dagli usi d'IA ad alto rischio a quelli che presentano un rischio più limitato, per i quali si prevedono particolari obblighi di trasparenza, infine ai casi di rischi minimi sottoposti al controllo delle autorità nazionali[12]. L'art. 6, dal titolo Regole di classificazione per i sistemi d'IA ad alto rischio, sostanzialmente rinvia all'Allegato II e all'Allegato III, cioè non vi è una spiegazione dettagliata, si legge solo che taluni sistemi d'IA utilizzati come componente di un prodotto e il prodotto stesso devono essere sottoposti ad una valutazione di conformità in quanto rientranti, secondo gli Allegati II e III, tra le IA ad alto rischio. La norma non brilla certo di chiarezza e solo

12 D. U. GALETTA E J. G. CORVALAN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, *op. cit.*, pp. 10 ss., ove si fa riferimento a tre diversi livelli di automazione che a loro volta possono essere aggiunti e combinati con livelli di innovazione che facilitano e semplificano l'integrazione con l'agente pubblico e/o col cittadino: «*A) Primo livello. Automazione completa. In questo caso, gli algoritmi collegano automaticamente i dati e le informazioni con i documenti, tramite l'uso di sistemi di Intelligenza Artificiale basati su regole, meglio conosciuti come "sistemi esperti" … B) Secondo livello. Automazione e intervento umano ridotto. In molti casi è invece necessario (ed inevitabile) che un operatore interagisca con un sistema automatizzato, al fine di completare la creazione di un documento. Questo per diverse ragioni: può succedere, infatti, che parti di un documento richiedano aggiornamenti costanti e che non ci sia modo di automatizzare questa attività; oppure, a volte la circostanza che un operatore umano possa intervenire ponendo domande, o interagendo con il sistema, risulta più conveniente in vista delle altre fasi del procedimento (…) sebbene comporti ritardi di secondi (o minuti) nell'esecuzione delle operazioni automatizzate. C) Terzo Livello. Automazione più predizione. L'apprendimento automatico è una tecnica di Intelligenza Artificiale che, in una spiegazione molto elementare e rudimentale, può essere concettualizzata in modo seguente: uno o più algoritmi rilevano molti dati al fine di stabilire dei modelli, che vengono poi tradotti in previsioni, sulla base di alcuni criteri statistici (…) In tutti e tre i casi di automazione testé descritti, tuttavia la possibilità di far ricorso a diversi gradi di innovazione menzionati passa per il tramite di una governance dei dati … che sono il prodotto di alcuni passaggi strutturali di flussi o "alberi decisionali". Questi alberi decisionali sono diagrammi di costruzioni logiche, basate su regole, che servono a rappresentare e classificare una serie di condizioni che si verificano successivamente per la risoluzione di un problema*».

l'Allegato III riesce a dare un quadro di quali siano, ad oggi, le IA ad alto rischio! Stante la rilevanza delle regole di classificazione dei sistemi ad alto rischio sarebbe stato opportuno o un rinvio direttamente agli Allegati o una diversa formulazione della norma di apertura ma così è difficile comprendere cosa la Commissione abbia voluto aggiungere alla materia con l'art. 6.

Positiva è invece la disposizione contenuta nell'art. 9 della proposta di regolamento ove si delinea il sistema di gestione del rischio come processo in continuo divenire che richiede un sistematico aggiornamento; a tal fine si prevede che, per l'IA ad alto rischio, le strutture siano sottoposte preventivamente a *test* al fine di evitare possibili errori, almeno quelli più prevedibili; inoltre, per i sistemi ad alto rischio, si prevede, *ex* art. 14, la sorveglianza umana che può sempre intervenire qualora vi siano delle *bias* rispetto l'obiettivo a cui deve tendere l'IA.

4.a In verità nonostante l'adozione di tale criterio, non sempre nella proposta di regolamento è dato intravedere un'organica distinzione tra IA ad alto rischio e IA a rischio minimo: ciò lo si evince soprattutto negli allegati che accompagnano gli enunciati dell'art. 6, rispetto all'art. 7 che dovrebbero riguardare i sistemi d'IA ad alto rischio. Nell'art. 7, al comma 2, così testualmente si legge: «*When assessing for the purposes of paragraph 1 whether an AI system poses a risk of harm to the health and safety or a risk of adverse impact on fundamental rights that is equivalent to or greater than the risk of harm posed by the high-risk AI systems already referred to in Annex III, the Commission shall take into account the following criteria: (a) the intended purpose of the AI system; (b) the extent to which an AI system has been used or is likely to be used; (c) the extent to which the use of an AI system has already caused harm to the health and safety or adverse impact on the fundamental rights or has given rise to significant concerns in relation to the materialisation of such harm or adverse impact, as demonstrated by reports or documented allegations submitted to national competent authorities; (d) the potential extent of such harm or such adverse impact, in particular in terms of its intensity and its ability to affect a plurality of persons; (e) the extent to which potentially harmed or adversely impacted persons are dependent on the outcome produced with an AI system, in particular because for practical or legal reasons it is not reasonably possible to opt-out from that outcome; (f) the extent to which potentially harmed or adversely impacted persons are in a vulnerable position in relation to the user of an AI system, in particular due to an imbalance of power, knowledge, economic or social circumstances, or age; (g) the extent to which the outcome produced with an AI system is easily reversible, whereby outcomes having an impact on the health or safety of persons shall not be considered as easily reversible; (h) the extent to which existing Union legislation provides for: (i) effective measures of redress in relation to the risks posed by an AI system,*

with the exclusion of claims for damages; (ii) effective measures to prevent or substantially minimise those risks».

Molte delle fattispecie richiamate nel valutare il rischio, dipendono in verità da come l'IA è stata utilizzata o è probabile che venga utilizzata ma, se questo è il criterio, nulla ha a ché vedere con quanto stabilito, ad es., al punto 2, (c) il quale fa riferimento a sistemi d'IA che hanno già causato un danno e dei quali ovviamente dovrebbe essere vietato l'utilizzo[13]. A ben vedere la fattispecie in esame dovrebbe rientrare tra le pratiche d'IA vietate, *ex* art. 5, a meno ché non si sia provveduto ad una modifica e ad una nuova valutazione di conformità. Parimenti, le fattispecie contenute nei punti (d) - (e), appaiono non coordinate con la *ratio* della proposta, in quanto l'oggetto attiene all'entità del danno causato dall'IA sulle persone ma, la valutazione del danno, è cosa ben diversa dai criteri enunciati dalla stessa proposta per individuare le fattispecie d'IA ad alto rischio nel primo comma dell'art. 6: il danno, infatti, potrebbe derivare anche da IA a basso rischio. Anche gli altri punti attengono alle misure del danno o alle misure necessarie per ridurre sostanzialmente il rischio ma, nulla aggiungono al criterio che la Commissione ha voluto indicare nello stabilire se, un sistema d'IA, sia ascrivibile tra quelli ad alto rischio o meno. Ciò che manca è un accorpamento delle fattispecie secondo una logica ben precisa così come un accorpamento e un approfondimento normativo a secondo del profilo che s'intende disciplinare. A leggere l'art. 7, comma 2, si ha la sensazione che siano state messe insieme le fattispecie più disparate, senza un criterio logico-giuridico che possa fornire un quadro chiaro volto a spiegare quali siano i criteri che la Commissione utilizza per definire un'IA ad alto rischio. Ciò è tanto più grave se si considera che le fattispecie esaminate possono anche ampliarsi o diminuire, via via che il dato tecnico ovvero le IA, si sviluppano in più direzioni. A questo punto la flessibilità di cui parla in apertura la proposta di regolamento, deve per forza avere dei gangli certi perché solo così è possibile che la fattispecie giuridica insegua, con coerenza, la fattispecie tecnica.

La disciplina dell'IA ad alto rischio dovrebbe peraltro innescarsi sulle disposizioni in materia già adottate dall'Unione implicando l'opportunità, su un piano giuridico, che la proposta di regolamento introduca "anelli di congiunzione" o rinvii, rispetto alle normative precedenti in materia. Nelle disposizioni transitorie sembra invece che la proposta voglia bypassare la normativa pregressa[14], ciò lo si evince dal fatto che,

13 Si legge testualmente: «*(c) the extent to which the use of an AI system has already caused harm to the health and safety or adverse impact on the fundamental rights or has given rise to significant concerns in relation to the materialisation of such harm or adverse impact, as demonstrated by reports or documented allegations submitted to national competent authorities*».

14 (a) Regolamento (CE) n. 300/2008 del Parlamento europeo e del Consiglio, 11 marzo 2008, che istituisce norme comuni per la sicurezza dell'aviazione civile e che abroga il regolamento n. 2320/2002;

nelle disposizioni finali, si prevede la modifica di alcune norme già emanate dall'Unione[15], attraverso la previsione di commi aggiuntivi volti all'applicazione della futura normativa.

4.b Molto ambiguo risulta quanto disposto dall'art. 6 in merito al perimetro dei sistemi d'IA ad alto rischio: da una prima lettura di tale articolo, si desume che trattasi di casi in cui l'IA contengono elementi che potrebbero presentare fattori di rischio che vengono meglio specificati nell'Allegato III. In verità in tale allegato si accorpano IA capaci di auto apprendimento con IA molto più semplici, come gli algoritmi di *clustering* o per meglio dire, IA specifiche e IA generali[16]. Gli algoritmi

(b) Regolamento n. 167/2013 del Parlamento europeo e del Consiglio, 5 febbraio 2013, relativo all'omologazione e alla vigilanza del mercato dei veicoli agricoli e forestali;
(c) Regolamento n. 168/2013 del Parlamento europeo e del Consiglio, 15 gennaio 2013, relativo all'omologazione e alla vigilanza del mercato dei veicoli a motore a due o tre ruote e dei quadricicli;
(d) Direttiva 2014/90/UE del Parlamento europeo e del Consiglio, 23 luglio 2014, sull'equipaggiamento marittimo e che abroga la Direttiva 96/98/CE del Consiglio;
(e) Direttiva 2016/797/UE del Parlamento europeo e del Consiglio, 11 maggio 2016, relativa all'interoperabilità del sistema ferroviario dell'Unione europea (rifusione);
(f) Regolamento n. 858/2018 del Parlamento europeo e del Consiglio, 30 maggio 2018, relativo all'omologazione e alla vigilanza del mercato dei veicoli a motore e dei loro rimorchi nonché dei sistemi, dei componenti e delle entità tecniche indipendenti destinati a tali veicoli che modifica i Regolamenti n. 715/2007, n. 595/2009 e abroga la Direttiva 2007/46/CE;
(g) Regolamento n. 1139/2018 del Parlamento europeo e del Consiglio, 4 luglio 2018, recante norme comuni nel settore dell'aviazione civile, che istituisce un'Agenzia dell'Unione europea per la sicurezza aerea e che modifica i Regolamenti n. 2111/2005, n. 1008/2008, n. 996/2010, n. 376/2014 e le Direttive 2014/30/UE e 2014/53/UE del Parlamento europeo e del Consiglio e abroga i Regolamenti n. 552/2004 e n. 216/2008 del Parlamento europeo e del Consiglio e il Regolamento n. 3922/91 del Consiglio;
(h) Regolamento n. 2144/2019 del Parlamento europeo e del Consiglio, 27 novembre 2019, relativo ai requisiti di omologazione dei veicoli a motore e dei loro rimorchi nonché di sistemi, componenti ed entità tecniche destinati a tali veicoli, per quanto riguarda la loro sicurezza generale e la protezione degli occupanti dei veicoli e degli altri utenti vulnerabili della strada che modifica il Regolamento n. 2018/858 del Parlamento europeo e del Consiglio e abroga i Regolamenti n. 78/2009, n. 79/2009 e n. 661/2009 del Parlamento europeo e del Consiglio e i Regolamenti n. 631/2009, n. 406/2010, n. 672/2010, n. 1003/2010, n. 1005/2010, n. 1008/2010, n. 1009/2010, n. 19/2011, n. 109/2011, n. 458/2011, n. 65/2012, n. 130/2012, n. 347/2012, n. 351/2012, n. 1230/2012 e n. 166/2015 della Commissione.

15 Così si prevede che si debba aggiungere all'art. 4, par. 3 del Regolamento n. 300/2008, un nuovo comma in merito alle procedure per l'approvazione dell'uso delle apparecchiature di sicurezza relative ai sistemi d'IA. L'art. 76 invece, modifica il Regolamento n. 167/2013, rilevando che qualora si adottino atti delegati ai sensi del primo comma riguardanti i sistemi d'IA, questi devono avere i requisiti di cui al Titolo III e si deve tener conto del Capo 2 della proposta di regolamento. Parimenti all'art. 77 si prevede una modifica del Regolamento n. 168/2013 in tema di atti delegati e, sempre sugli atti delegati, l'art. 78 modifica la Direttiva 2014/90/UE; l'art. 79 modifica la Direttiva 2016/797/UE, sempre sugli atti delegati e di esecuzione che devono ottemperare ai requisiti contenuti nella proposta; l'art. 80 modifica il Regolamento n. 858/2018, sempre sugli atti delegati e l'art. 81 modifica il Regolamento n. 138/2018; l'art. 82 modifica il Regolamento n. 2144/2019 sempre in tema di atti di esecuzione che devono essere rispettosi di quanto previsto nella proposta di regolamento.

16 Come noto, il *clustering* consiste in un insieme di metodi per raggruppare oggetti in classi omogenee. Un *cluster* è un insieme di oggetti che presentano tra loro delle similarità, ma che, per contro, presentano dissimilarità con oggetti in altri *cluster*. L'*input* di un algoritmo di *clustering* è costituito da un campione di elementi, mentre l'*output* è dato da un certo numero di *cluster* in cui gli elementi del campione sono suddivisi in base a una misura di similarità. Gli algoritmi di *clustering* forniscono come *output* anche la descrizione delle

informatici possono essere semplici o complessi ma ovviamente non tutti gli algoritmi riguardano l'IA anche se, tutti i sistemi d'IA, presuppongono l'uso di algoritmi. Ora, partendo da tale assunto, i casi contenuti nell'Allegato III risultano non seguire una linea precisa che contenga la necessaria distinzione tra le IA ad alto rischio rispetto agli algoritmi di *clustering*. Se come detto la proposta intende adottare il criterio del rischio, base per una graduazione delle IA, non è dato comprendere il perché negli allegati vengano messi insieme strumenti che presentano aspetti diversi uniti dal solo fatto dell'utilizzo di macchine, venendosi quasi a smentire i presupposti sui quali si basa la proposta.

Si deve sempre aver chiaro il dato che i sistemi in questione non sono perfetti e che il solo fatto dell'utilizzo di sistemi digitali non implica *tout court* che tali sistemi siano scevri da rischi, errori o neutrali. Basti pensare che ad es. gli algoritmi sono venuti a diventare i protagonisti di un "*mash-washing*" attraverso i numeri tuttavia, anche utilizzando tali sistemi, si possono operare scelte dal preciso impatto sociale che si ammantano di neutralità e di perfezione solo perché assunte dalle macchine. Al contrario tutto dipende dall'operatore che li ha realizzati e dalle modalità del loro utilizzo; per gli algoritmi che non fanno parte di quei sistemi di auto apprendimento non si possono certo imputare agli stessi, errori o distorsioni nel risultato ma, errori e distorsioni, si possono invece imputare al soggetto che ha realizzato l'algoritmo.

E' ormai assodato che l'impiego di tali strumenti comporti in realtà una serie di scelte tutt'altro che neutre, anche negli algoritmi più semplici i criteri in base ai quali i dati sono raccolti, selezionati, ordinati e messi insieme, sono frutto di precise operazioni interpretative, compiute a monte da un soggetto fisico ma messe in atto dal programmatore del *software* con conseguente formulazione di giudizio.

Fatte queste premesse di carattere generale, al primo punto nell'Allegato III si fa riferimento all'identificazione e classificazione biometrica delle persone fisiche: sistemi d'IA che vengono utilizzati per la biometria remota in tempo reale comprensivi dell'identificazione delle persone fisiche. Queste sono sicuramente dell'IA a cui può essere attribuita l'etichetta ad alto rischio, parimenti nel caso d'IA – secondo punto dell'allegato – relative alla gestione delle infrastrutture critiche, quale il traffico stradale, la fornitura di acqua, gas, riscaldamento ed elettricità. Il terzo punto riguarda l'istruzione e la formazione del personale, in particolare al punto (a) si parla di sistemi d'IA destinati a essere utilizzati per l'accesso o l'assegnazione di persone fisiche agli istituti d'istruzione e formazione professionale. In realtà più che d'IA, in questo caso, parrebbe trattarsi di algoritmi *cloustering* ovvero sistemi non di auto apprendimento ma di organizzazione di dati. In genere viene fatto l'esempio di

caratteristiche di ciascun *cluster*, il che è fondamentale per poi prendere decisioni strategiche sulle azioni da compiere verso tali gruppi (*marketin*g mirato, promozioni *ad hoc*, creazione di nuovi prodotti e servizi).

quanto è successo in Belgio, dove il Governo centrale ha deciso d'implementare il sistema basato su tecniche d'intelligenza artificiale, al fine di assegnare l'istituto scolastico più opportuno per ogni studente, il ché ha portato però a creare classi non eterogenee, come politicamente si voleva, collocando gli studenti più talentuosi tutti in una stessa classe. Questo è dovuto al fatto che, dove vi siano scelte di natura politica discrezionale delle due l'una: o si riesce a trasferire queste scelte nell'algoritmo, casomai soppesando taluni fattori rispetto ad altri, o è chiaro che l'algoritmo opera mettendo insieme i casi secondo una logica di dati omogenei; il punto è che bisogna vedere cosa il *software* ha previsto. Lo stesso ragionamento vale per i sistemi di decisione automatizzata in merito alla promozione o alla bocciatura dello studente. Anche in questo caso esistono già dei precedenti come quello della scuola inglese, dove, al termine dello scorso anno scolastico, i docenti, non avendo sufficienti dati per capire se promuovere o meno i propri studenti, causa pandemia Covid-19, hanno deciso di affidarsi a un algoritmo. Purtroppo tale scelta è risultata del tutto forviante a causa di una *bias* del sistema che ha dato come risultato il fatto che, tutti gli alunni che appartenevano a famiglie benestanti venivano promossi mentre quelli bocciati appartenevano a famiglie in difficoltà. E' vero che ci sono state distorsioni nell'atto prodotto dall'algoritmo, ma qui tutto dipende da come viene strutturato l'algoritmo ed è quello che, sia pure nelle diversità delle fattispecie, è avvenuto in Italia nel famoso caso di assegnazione delle cattedre a professori di scuole medie il ché ha portato, peraltro, a un dibattito giurisprudenziale che ha fatto sì che il giudice chiedesse di rendere conoscibile l'algoritmo e di chiarire, ai soggetti interessati, il perché di determinate opzioni operate dall'algoritmo stesso. La giurisprudenza, che si è snodata tra TAR e Consiglio di Stato[17], ha riconosciuto non solo il diritto dell'interessato ad accedere all'algoritmo, ma l'amministrazione è stata obbligata a dare puntuali motivazioni e spiegazioni sull'*iter* logico seguito.

In questi casi l'attività richiesta alla macchina è più "elementare" poiché fa riferimento a sequenze d'istruzioni predefinite dall'uomo in modo univoco per raggiungere l'obiettivo che ci si è prefissati. Generalmente si tratta di una procedura seriale standardizzata utilizzata per l'elaborazione d'ingenti quantità di dati ed è sempre l'uomo o l'operatore tecnico che precisa i punti rilevanti sui quali si deve basare la scelta, ecco perché per questi tipi di algoritmi si parla di algoritmi *cloustering*, in contrapposizione all'IA ad alto rischio; il ritrovarsi tali algoritmi nell'Allegato III che dovrebbe elencare i sistemi d'IA ad alto rischio è indice di quella mancanza di organicità già prima segnalata nell'elencazione di tali algoritmi.

17 Sentenza TAR Lazio, sez. III *bis*, Roma, n. 3742 del 21.03.2017, alla quale hanno fatto seguito la sentenza TAR Lazio, Roma, sez. III, *bis*, n. 6606 del 27.05.2019 nonché il Consiglio di Stato, sez. IV, 8.4.2019, n. 2270.

Coerente al solo titolo dell'Allegato III, è invece l'ipotesi contenuta nel punto 2, (b), che fa riferimento all'IA destinate ad essere utilizzate per prendere decisioni in merito alle promozioni, al monitoraggio delle prestazioni lavorative, ai comportamenti delle persone in ambito lavorativo. Se in queste fattispecie si può parlare d'IA ad alto rischio tuttavia, queste non possono intervenire sui rapporti contrattuali lavorativi, altrimenti si verserebbe nell'ipotesi dell'art. 5 ovvero una valutazione da parte di un'IA vietata in quanto vertente sul comportamento di un soggetto, le sue espressioni, il suo modo di rapportarsi sul luogo di lavoro ecc.

Il punto 5 sempre dell'Allegato III, fa riferimento all'accesso e fruizione dei servizi privati essenziali e di pubblici servizi e prestazioni dove al punto (a) all'IA viene affidato il compito di valutare la fruibilità dei soggetti privati all'assistenza pubblica alle prestazioni di servizio; nel punto (b) si parla di sistemi d'IA destinati a essere utilizzati per valutare l'affidabilità creditizia delle persone, cosa che se vogliamo già viene fatta dalle banche; certo è che se tale valutazione viene lasciata ad una macchina delle due l'uno: o l'algoritmo è talmente minuzioso da riuscire per ogni singolo utente a stabilirne l'affidabilità creditizia (a questo punto trattasi di un algoritmo *ad personam*), o l'IA si basa su forme di auto apprendimento per cui ad es. sa che al di sotto di una certa soglia il soggetto che non dimostri di essere solvibile non può ottenere il credito ma, in questo caso, vi è poi un grosso margine di discrezionalità affidata alla macchina. Al punto 5, (c) dell'Allegato III, si parla di sistemi d'IA destinati a essere utilizzati per stabilire le priorità nel caso di contestuale richiesta di servizio di pronto intervento da parte dell'assistenza medica o dei vigili del fuoco: la questione diventa delicata perché neanche l'uomo può stabilire questa scelta tanto meno la macchina o l'IA! Vi è una determinante valutazione di opportunità nello stabilire una priorità tra interventi urgenti, ci si chiede: si possono creare sistemi che ci diano, con un margine di tranquillità, una risposta? Certamente sì nel futuro ma oggi non è dato sapere.

Più interessanti appaiono invece i sistemi d'IA destinati ad assistere le autorità pubbliche nei permessi di soggiorno, nei reclami relativi all'ammissibilità delle persone fisiche che presentano la domanda ecc. L'interesse nasce dal fatto che, in questo caso l'IA lavorano insieme alle autorità pubbliche, anzi assistono le autorità pubbliche; nell'Allegato III si fa questo esempio ma, la combinazione tra uomo e IA è sicuramente vincente nel settore sanitario, nel settore farmaceutico, nel settore agricolo cioè, invece di lasciare tutto alla macchina, la macchina diventa uno strumento di ausilio nell'attività umana.

Infine al punto 6, sempre dell'Allegato III, si parla di sistemi d'IA destinati ad essere utilizzati dalle autorità in base a valutazioni del rischio cui incorre una

potenziale vittima rispetto ad una persona che si sia resa responsabile di reati o che sia recidiva per quella tipologia di reato: in questo caso si affida all'IA la facoltà di derogare alla disciplina contenuta nell'art. 5 della proposta di regolamento. E' noto il caso in cui i cittadini di colore erano indicati come i più probabili soggetti recidivi rispetto agli altri cittadini ed è evidente che qui vi è un errore a monte dell'IA o di colui che ha programmato la macchina. Anche l'affidamento all'IA della gestione delle migrazioni e del controllo delle frontiere, che rientra sempre tra le IA ad alto rischio, suscita una qualche perplessità. La materia dell'immigrazione cui fa riferimento l'Allegato III al punto 7, di fatto è coperta per intero, basti pensare al punto (d) in merito all'esame delle domande d'asilo politico, di visto e di permesso di soggiorno, nel quale tuttavia la macchina assiste le autorità pubbliche competenti per l'esame delle suddette domande.

5. Il sistema di *governance* nell'IA ad alto rischio.

Non è possibile in questa sede, valutare o commentare le singole norme in merito al sistema di *governance* pertanto si è operata la scelta di richiamare taluni aspetti che, sostanzialmente, appaiono più interessanti.

In questa prospettiva rileva quanto disposto nell'articolo 10 il quale prevede che le IA che utilizzino tecniche di autoapprendimento devono soddisfare i criteri di qualità già indicati negli articoli precedenti nonché un *set* d'informazioni quali, la convalida e le pratiche di gestione. Tali pratiche riguardano in particolare le scelte di progettazione, la raccolta di dati, una valutazione preventiva della disponibilità e dell'idoneità della serie di dati che risultano necessari, un esame in vista di possibili distorsioni e, se rintracciate, eventuali lacune o carenze nei dati. Nell'individuazione e correzione dell'IA ad alto rischio, ai fini del controllo preventivo, occorre far riferimento al Regolamento n. 679/2016 nonché all'art. 10 della Direttiva 2016/280/UE e al Regolamento n. 1725/2018, per garantire i diritti di libertà delle persone fisiche comprese le nuove tecniche di riutilizzo e l'uso delle misure di sicurezza a tutela della *privacy*. L'art. 10 chiude al comma 6, rilevando che le pratiche in materia di *governance* e gestione dei dati, che si applicano ai sistemi d'IA ad alto rischio, siano specificamente previste nella presente proposta di regolamento, secondo quanto si legge nel paragrafo 2. Accanto a questo, sempre per l'IA ad alto rischio, è previsto che l'IA forniscano una documentazione tecnica e le necessarie informazioni alle autorità nazionali competenti e agli organismi notificati per

consentire a tali soggetti di valutare la conformità del sistema ai requisiti di cui all'Allegato IV.

Più interessante, sempre per i sistemi d'IA ad alto rischio, appare quanto disposto dall'art. 12, ove si prevede che, tali sistemi devono consentire la registrazione automatica degli eventi durante il funzionamento dell'IA, al fine di promuovere un livello di tracciabilità dei sistemi d'IA e, contestualmente, monitorare il funzionamento degli stessi. Il sistema di monitoraggio dovrebbe assicurare la tracciabilità del funzionamento dell'IA, ciò rileva, non solo in via generale ma, in particolare, al verificarsi di situazioni che comportino modifiche sostanziali allo stesso sistema. Si osservi come l'art. 16 preveda che i fornitori dei sistemi d'IA ad alto rischio non possano immettere sul mercato l'IA prima che siano espletate le procedure di valutazione relative alla loro conformità e agli obblighi di registrazione di cui all'art. 51.

Come si legge nella relazione che accompagna la proposta in esame, nell'ipotesi in cui vi sia una modifica, il sistema d'IA è sottoposto ad una nuova valutazione; solo a seguito di tutti questi passaggi l'IA dovrebbe ottenere una marcatura CE ma, a ben vedere, non è dato comprendere che tipo di marcatura un'IA possa ottenere non trattandosi di un oggetto o di un bene specifico al quale si appone la marcatura CE!

Interessante infine, appare quanto disposto nel Titolo III, Capitolo IV della proposta di regolamento che istituisce gli "organi notificati" ovvero quegli organi o organismi ai quali deve essere notificato il sistema che s'intende introdurre (secondo un'interpretazione logica) prima del loro utilizzo sul mercato; in tal senso si prevede che ciascuno Stato membro designi un'istituzione o un'autorità di notifica responsabile delle procedure necessarie per la valutazione e il monitoraggio soprattutto dei sistemi ad alto rischio. Di conseguenza gli Stati membri dovrebbero designare un organo nazionale di accreditamento secondo quanto stabilisce il Regolamento n. 765/2008[18]: trattasi di autorità indipendenti[19] che vagliano la legalità dell'IA ad alto rischio secondo una procedura molto dettagliata. La Commissione, *ex* art. 35, assegna un numero d'identificazione all'organismo notificato e mette a disposizione l'elenco di tali organismi compresi i numeri d'identificazione attribuiti, il ché dovrebbe soddisfare (il condizionale è d'obbligo) un livello minimo di trasparenza, almeno sotto questo profilo. Ove possano nascere dei dubbi sull'operato dell'organismo notificato, la Commissione interviene per controllare che tale

18 Regolamento n. 765/2008 del Parlamento europeo e del Consiglio del 9 luglio 2008 che pone norme in materia di accreditamento e vigilanza del mercato per quanto riguarda la commercializzazione dei prodotti e che abroga il Regolamento n. 339/93.

19 L'art. 33 della presente proposta di regolamento, prevede infatti che gli organismi notificati siano indipendenti, non solo dal fornitore di sistemi d'IA, ma da qualsiasi altro operatore avente un interesse economico nei sistemi d'IA ad alto rischio.

organismo soddisfi i requisiti richiesti dalla presente proposta di normativa. In questo modo nei sistemi d'IA ad alto rischio si viene a creare una sorveglianza del mercato in conformità a quanto stabilito nel Regolamento n. 1020/2019 che dovrebbe trovare integrale applicazione.

6. Possibili conclusioni.

In definitiva, la proposta di regolamento appare positiva in merito allo sforzo di disegnare un quadro normativo nella materia in questione, così come appare positivo l'aver assunto "il rischio" quale criterio discriminatorio tra le varie IA. Pur tuttavia, si auspicano delle modifiche proprio in relazione alle regole di classificazione dei sistemi e in particolare dei sistemi ad alto rischio. Appare invece più coerente la tenuta dei registri, la trasparenza e la fornitura delle informazioni agli utenti ma, in generale, la redazione del testo richiederà una profonda revisione non tanto concettuale o quanto meno non solo concettuale ma anche volta a chiarire meglio taluni passaggi rilevanti che risultano, ad una prima lettura, quanto meno oscuri e di difficile applicazione pratica.