

NADIA ZORZI GALGANO

IL REGOLAMENTO UE 2024/1689 DEL 13 GIUGNO 2024 SUL C.D. ALTO RISCHIO INACCETTABILE: LE PRATICHE INERENTI SISTEMI DI INTELLIGENZA ARTIFICIALE VIETATE DAL LEGISLATORE EUROPEO

SOMMARIO: 1. Premessa introduttiva. – 2. Le pratiche tassative di IA vietate dalla legge: il primo gruppo di pratiche vietate e la possibilità di valutarle unitariamente. Le pratiche di IA manipolative della persona fisica; le pratiche di IA che sfruttano “in altro modo” il comportamento di una persona fisica o di uno specifico gruppo di persone fisiche per la loro vulnerabilità; le pratiche di IA di *social scoring*. – 3. Un ulteriore gruppo di pratiche di IA vietate: le pratiche di sistemi di IA vietate per effettuare valutazioni o previsioni del rischio che una persona fisica commetta un reato; le pratiche di sistemi di IA vietate che creano o ampliano banche dati di riconoscimento facciale mediante *scraping* non mirato; le pratiche di sistemi di IA vietate per il loro inferire le emozioni di una persona fisica nell’ambito del luogo di lavoro o degli istituti di istruzione; le pratiche di sistemi di IA vietate volte alla categorizzazione biometrica per trarne deduzioni o inferirne logicità in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale e orientamento sessuale. – 4. L’ultimo gruppo di pratiche di sistemi di IA vietate dal legislatore europeo concernente l’uso di sistemi di identificazione biometrica remota “in tempo reale” in spazi accessibili al pubblico per finalità di attività di contrasto. – 5. Qualche spunto di conclusione sulle pratiche di sistemi di IA vietate.

1. – Il complesso Regolamento Europeo su sistemi di intelligenza artificiale è composto non solo dalla disciplina confluita nel corpo normativo del Regolamento stesso, costituito da 113 disposizioni normative, ma anche da ben XIII allegati. Quasi un terzo di queste⁽¹⁾, si focalizzano

⁽¹⁾ Senza voler contare gli allegati qualcuno dei quali sarebbe comunque rilevante in proposito, così, a titolo indicativo, l’allegato III, sui «Sistemi di IA ad alto rischio di cui all’articolo 6, paragrafo 2», ma soprattutto l’allegato II, rubricato «Elenco dei reati di cui all’articolo 5, paragrafo 1, primo comma, lettera b), punto iii)» che più in specifico dà luogo ad una lista tassativa anche per il legislatore nazionale di reati per cui, purché strettamente necessario e sottoposta al principio di proporzionalità, la pratica di IA di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico per attività di contrasto è consentita, ossia viene sottratta al divieto di legge. E mi permetto di rinviare per una prima prospettiva d’insieme al mio contributo in corso di pubblicazione su *Contratto e impresa/Europa*.

proprio sul c.d. “alto rischio” dei sistemi di IA, ovvero tale da risultare idoneo a mettere in pericolo salute, sicurezza e diritti fondamentali su cui può venire ad incidere un sistema di IA, ossia idoneo alla probabilità di un danno grave a tutto il pregnante sistema valoriale di cui alla disciplina stessa, e che, per espresso riconoscimento del legislatore europeo, in proposito, quest’ultima è preordinata a proteggere.

L’entrata a pieno regime del Regolamento è demandata a fine 2027 ⁽²⁾ e, al pari del resto al GDPR ⁽³⁾, risulta scaglionata nella decorrenza di inizio efficacia, a seconda delle parti in cui risulta suddiviso l’AI Act ⁽⁴⁾

⁽²⁾ Ossia a quasi 3 anni dall’emanazione del Regolamento.

⁽³⁾ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati). In dottrina, v. *Persona e mercato dei dati. Riflessioni sul GDPR*, a cura di Zorzi Galgano, Milano, 2019; *I dati personali nel diritto europeo*, a cura di Cuffaro, D’Orazio e Ricciuto, Torino, 2019; *La protezione dei dati personali in Italia. Regolamento UE 2016/679 e d.lgs. 10 agosto 2018, n. 101*, a cura di Finocchiaro, Bologna, 2019; *Innovazione tecnologica e persona. Il diritto alla protezione dei dati personali nel regolamento UE 2016/679*, a cura di Califano e Colapietro, Napoli, 2018; nonché il volume collettaneo curato da FINOCCHIARO, *Il nuovo Regolamento europeo sulla privacy e sulla protezione dei dati personali*, Bologna, 2017; CUFFARO, DI CIOMMO, D’ORAZIO, MANTELERO e GAMBINI, *Trattamento dei dati personali e Regolamento UE n. 2016/679*, e.book speciale di Corriere giuridico; BERNARDI e MESSINA, *Privacy e Regolamento Europeo*, Milano, 2017; DE FRANCESCHI, *La circolazione dei dati personali tra privacy e contratto*, Napoli, 2017; BISTOLFI, BOLOGNINI e PELINO, *Il Regolamento Privacy europeo. Commentario alla nuova disciplina sulla protezione dei dati personali*, Milano, 2016; SICA, D’ANTONIO e RICCIO (a cura di), *La nuova disciplina europea della privacy*, Milano, 2016; PIZZETTI, *Privacy e il diritto europeo alla protezione dei dati personali*, tomo I (*Dalla dir. 95/46 al nuovo regolamento europeo*) e tomo II (*Il regolamento europeo 2016/679*), Torino, 2016; BUSIA, LIGUORI e POLLICINO (a cura di), *Le nuove frontiere della privacy nelle tecnologie digitali: bilanci e prospettive*, Roma, 2016. Tra i contributi, THOBANI, *La libertà del consenso al trattamento dei dati personali e lo sfruttamento economico dei diritti della personalità*, in *Eur. dir. priv.*, 2016, p. 513 ss.; STANZIONE, *Il Regolamento europeo sulla privacy: origini e ambito di applicazione*, in *Eur. dir. priv.*, 2016, p. 1249 ss.; SPINA, *Alla ricerca di un modello di regolazione per l’economia dei dati. Commento al Regolamento (Ue) 2016/679*, in *Riv. regolaz. merc.*, 2016, p. 143 ss.; PIRAINO, *Il Regolamento generale sulla protezione dei dati ed i diritti dell’interessato*, in *NGCC*, 2017, p. 369 ss.; GUASTALLA, *Il nuovo Regolamento europeo sul trattamento dei dati personali: i principi ispiratori*, in *questa rivista*, 2018, 1, p. 106 ss.; RESTA e ZENO ZENCHOVICH, *Volontà e consenso nella fruizione dei servizi in rete*, in *Riv. trim. dir. e proc. civ.*, 2018, p. 411 ss.; CUFFARO, *Il diritto europeo sul trattamento dei dati personali*, in *questa rivista*, 2018, 3, p. 1098 ss.; BRAVO, *Lo “scambio di dati personali” nei contratti di fornitura di servizi digitali e il consenso dell’interessato tra autorizzazione e contratto*, in *questa rivista*, 2019, 1, p. 34 ss.

⁽⁴⁾ Così, ex art. 113, il Regolamento sui sistemi di IA è entrato in vigore il 1° agosto 2024, ossia nel tempo previsto di regola per la decorrenza iniziale dei Regolamenti dell’Unione Europea, con entrata in vigore il 20° giorno successivo alla pubblicazione nella *Gazzetta ufficiale dell’Unione Europea*. Tuttavia, a) i capi I e II si applicano a decorrere dal 2 febbraio 2025, inclusa la normativa di cui al tema qui oggetto di attenzione; b) il capo III,

e, per questo, in definitiva, sono stabiliti tempi ancora lunghi. È chiaro che per una valutazione più ponderata della disciplina dovremo attendere di verificarne il funzionamento a pieno regime e ciò implica che solo allora ritorneremo in punto. Tuttavia, le pratiche di IA qui in oggetto sono entrate come già detto da poco in vigore, in questo febbraio 2025, e perciò giustificano questa prima analisi critica.

Come rilevato, per più di un terzo delle sue disposizioni (senza contare gli allegati a riguardo) il modello di disciplina dell'AIR appare finalizzato a fronteggiare il c.d. alto rischio, promuovendo così un accesso al mercato dell'IA ed a garantirne un democratico e partecipato sviluppo con prevalente protezione di salute, sicurezza ed altri diritti fondamentali. L'alto rischio viene graduato. Se risulta "inaccettabile", le corrispondenti pratiche di IA sono vietate *ex art. 5*. Se invece l'alto rischio non è "inaccettabile", il mercato dell'IA, in base alla disciplina dell'AIA, lo accetta e lo fronteggia attraverso le specifiche regole previste. Anche in questo caso, come detto, in parte in uno specifico allegato⁽⁵⁾, si prevedono già pratiche di IA ad alto rischio delineate dal legislatore europeo e va precisato che l'allegato può nel corso del tempo essere soggetto a ripetute e rapide modifiche ad opera della Commissione UE, rassicurando così su uno sviluppo delle lista di pratiche ad alto rischio non inaccettabile destinato a rimanere sempre in linea nel corso del tempo, anche tenendo in ogni caso conto dello sviluppo tecnologico dei sistemi di IA.

Per altra parte, tuttavia, a norma dell'art. 6, par. 1, viene demandato alle figure soggettive competenti delineate dall'AIA, il controllo sui requi-

sezione IV, il capo V, il capo VII, il capo XII e l'art. 78 si applicano dal 2 agosto 2025 ad eccezione dell'art. 101; c) l'art. 6, paragrafo 1) e i corrispondenti obblighi di cui al presente Regolamento si applicano a decorrere dal 2 agosto 2027. In altri termini, il primo rinvio di entrata in vigore concerne il capo I, ossia le disposizioni generali e il capo II, ossia le pratiche di IA vietate qui in discussione; il secondo rinvio, ad agosto 2025, concerne il capo III sui «Sistemi di IA ad alto rischio», sezione 4, che è dedicata alla «Autorità di notifica e organismi notificati», che è la disciplina riguardante le Autorità di cui ogni legislatore nazionale ne deve designare almeno una, quale autorità responsabile della predisposizione e dell'esecuzione delle procedure previste per la valutazione, la designazione e la notifica degli organismi di valutazione della conformità, e per il loro monitoraggio *ex art. 28* che si conclude prevedendo: «(...) nell'ambito della collaborazione tra le Autorità di notifica di tutti gli Stati membri»: un arduo compito da allestire per i legislatori nazionali entro agosto di questo 2025. Infine, il terzo e ultimo rinvio al 2 agosto 2027 riguarda l'entrata in vigore dell'art. 6, par. 1, che concerne i requisiti richiesti per l'alto rischio dei sistemi di IA e, soprattutto, la valutazione di conformità, anche quest'ultima, procedura in parte demandata alle regole armonizzate previste dal legislatore nazionale.

⁽⁵⁾ L'allegato III, rubricato infatti «Sistemi di IA ad alto rischio», di cui all'art. 6, par. 2.

siti di alto rischio, attraverso il funzionamento delle procedure di conformità e altre regole all'uopo rilevanti, in questo modo rendendo responsabili in punto gli operatori economici del settore incentrato sull'IA. Certo è che, laddove è il legislatore a delineare pratiche di IA ad alto rischio, sempre in caso di alto rischio inaccettabile *ex art. 5*, in parte invece nel caso di alto rischio "non inaccettabile" di cui all'allegato III, ne risulta facilitato il corrispondente compito in concreto, perché è sufficiente in proposito che la pratica di IA posta in essere corrisponda ad una di quelle presenti nell'elenco stabilito dal legislatore. Con garanzia della certezza del diritto e con compito, si ribadisce, estremamente semplice e rapido.

Tutto ciò premesso, al fine di un primo inquadramento del problema correlato all'alto rischio dei sistemi di IA, in questo scritto, ci si intende soffermare più specificamente sulle singole pratiche di IA di alto rischio "inaccettabile" e per questo vietate dalla legge (*ex art. 5*). A titolo di premessa, occorre innanzi tutto coordinarne la disciplina con altre disposizioni del Regolamento, innanzi tutto quelle di cui all'art. 2 che delimita l'ambito di applicazione e che non tocca, a dire il vero, regole qui oggetto di interesse. Occorre invece tener presente la norma di chiusura in tema di pratiche di IA vietate, di cui al par. 8 dell'art. 5, secondo cui «il presente Regolamento lascia impregiudicati i divieti che si applicano qualora una pratica di IA violi altre disposizioni del diritto dell'Unione»⁽⁶⁾.

2. – Chiarisce al riguardo del tema in oggetto il *considerando* 28): «l'IA presenta, accanto a molti utilizzi benefici, la possibilità di essere impiegata impropriamente e di fornire strumenti nuovi e potenti per pratiche di manipolazione, sfruttamento e controllo sociale. Tali pratiche sono particolarmente dannose e abusive e dovrebbero essere vietate poiché sono contrarie ai valori dell'Unione relativi al rispetto della dignità umana, alla libertà, alla uguaglianza, alla democrazia ed allo Stato di diritto e ai diritti fondamentali sanciti dalla Carta, compresi il diritto alla non discriminazione, alla protezione dei dati, alla vita privata e i diritti dei minori». E si continua nel lungo *considerando* 29) con una generale illustrazione della ragione giustificatrice che ha indotto il legislatore europeo a stabilire il divieto di pratiche di IA di alto rischio inaccettabile, ossia per impedire un significativo danno alle persone o anche a gruppi particolarmente vulnerabili di persone, a una loro possibile manipolazione da parte dei sistemi di IA. E il danno, si specifica da parte del *considerando* stesso, non occorre

⁽⁶⁾ E seguiranno alcuni esempi, *infra*, nel testo.

che sia voluto, ossia con accertamento di un elemento soggettivo, essendo sufficiente che esso derivi da «pratiche in sé manipolative o di sfruttamento consentito (*rectius*: reso possibile) (7) dall'IA».

Un punto importante, toccato sempre dal *considerando* 29), dove, in realtà, vengono anche indicate dal legislatore le ragioni di giustificazione del divieto delle prime specifiche pratiche di sistemi di IA interdette, concerne la c.d. “complementarità” fra la disciplina che in riferimento ai sistemi di IA vieta le pratiche ad alto rischio “inaccettabile” e la normativa europea in tema di pratiche scorrette del 2005 (8). In merito a come si dispieghi una tale complementarità può essere precisato che nell'eventualità di una violazione del divieto diretto a impedire di porre in essere una delle pratiche vietate *ex art. 5*, che risulti al contempo anche pratica scorretta, vuoi in quanto ingannevole, vuoi in quanto aggressiva, in base alla disciplina della direttiva 2005/29/CE, all'operatore economico interessato potrebbe essere comminata la doppia sanzione, da un lato quella molto elevata prevista dall'AIA, a protezione degli interessi specifici di cui alla presente normativa, ossia interessi correlati al sistema valoriale protetto all'apice delle fonti (Carta di Nizza, nella versione di Strasburgo, e Costituzione) a tutela di salute, sicurezza e (altri) diritti fondamentali, e d'altro lato, la sanzione comminata al riguardo dalla direttiva 2005/29/CE, a tutela degli interessi economici dei consumatori lì in oggetto. È anche il caso di puntualizzare in proposito che i soggetti protetti dalle rispettive discipline non sono, a rigore, i medesimi sempre sovrapponibili perché mentre l'AIA considera, in proposito, in modo esplicito *ex art. 5*, in alcune delle fattispecie previste, una tutela rapportata alle persone, così, ad esempio, in par. 1, lettera *a*) e lettera *b*), nella disciplina sulle pratiche scorrette, al centro della tutela ci sono i consumatori che sono sempre persone fisiche ma che compiono atti rientranti nella disciplina stessa per fini privati (e ciò significa che in alcuni casi marginali potrebbe non verificarsi piena coincidenza). Il che, del resto, è pacifico se si pensa che l'AIA è, come più volte detto, posto con riferimento all'alto rischio, a protezione di valori all'apice delle fonti e ciò naturalmente impedisce di lasciar fuori dalla tutela prevista qualsiasi persona fisica, perché salute, sicurezza e diritti fondamentali sono propri di ciascuno, senza nessuna possibilità di esclusione, come invece può accadere nella tutela consumeristica dove le persone fisiche, che compiono atti per fini che attengono alla loro attività

(7) Parentesi mia.

(8) E in punto mi permetto di rinviare al mio scritto in corso di pubblicazione su *Contratto e impresa/Europa*.

imprenditoriale o professionale, sono escluse dalla corrispondente protezione legislativa. Il caso di una doppia tutela ricorre, a titolo esemplificativo, qualora risulti posta in essere una pratica di c.d. “*deep fake*” correlata a sistemi di IA ⁽⁹⁾, nel caso in cui questa si presti ad integrare una tipologia delle pratiche di sistemi di IA vietate dall’art. 5 soprattutto *sub* profilo di cui alla lettera *a*). D’altra parte, la disciplina qui in considerazione incide, a sua volta, anche in altro modo sulla precedente inerente le pratiche scorrette. In quella è stabilita una c.d. *black list* di pratiche in ogni caso scorrette e vietate quindi dal legislatore e comprensiva anche del caso in cui una delle pratiche scorrette ivi delineate venisse realizzata attraverso un sistema di IA. In altri termini, è chiaro che, se una delle pratiche lì indicate venisse da oggi posta in essere da un sistema di IA, essa ricadrebbe nell’ambito della tutela di quella e sottoposta alla corrispondente sanzione da parte dell’Autorità all’uopo competente, ossia l’Autorità Garante della Concorrenza e del Mercato.

Lo stesso *considerando* n. 29), infine, puntualizza che i divieti di pratiche manipolative o di sfruttamento, di cui all’AIR, non pregiudicano le pratiche lecite consentite dal diritto europeo o nazionale nel contesto di un trattamento medico, né parallelamente esse siano atte a pregiudicare le pratiche lecite ammesse dal diritto nella pubblicità (come la pubblicità iperbolica) ⁽¹⁰⁾. Nelle Linee guida della Commissione europea sulle pratiche vietate in materia di intelligenza artificiale ⁽¹¹⁾, pubblicate il 4 febbraio 2025, che rappresentano uno strumento di *soft law* rivolto prevalentemente agli operatori economici di settore nei sistemi di IA, figura, in proposito, l’esempio di un sistema di intelligenza artificiale che simula ed elabora conversazioni umane (un *chatbot*), destinato dal fornitore a sostenere e guidare gli utenti nel mantenimento di uno stile di vita sano e a fornire consigli su misura per esercizi psicologici e fisici ⁽¹²⁾.

L’art. 5, par. 1, comma 1 stabilisce dunque un elenco da intendersi come tassativo di pratiche di IA vietate, comprensivo di otto tipi di prati-

⁽⁹⁾ Si veda il caso del *deep fake*, inteso quale «immagine o un contenuto audio o video generato o manipolato dall’IA che assomiglia a persone, oggetti, luoghi, entità o eventi esistenti e che apparirebbe falsamente autentico o veritiero a una persona» (art. 3, n. 60), su cui v. anche art. 50 e *considerando* n. 134).

⁽¹⁰⁾ Si segnala che anche i successivi *considerando*, da 30) a 44) incluso, riguardano casi specifici di pratiche di IA vietate dall’art. 5.

⁽¹¹⁾ V. EU Commission, *Approval of the content of the draft Communication from the Commission – Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*, Brussels, 4.2.2025 C(2025) 884 *final*, disponibili al momento solo in lingua inglese.

⁽¹²⁾ EU Commission, *op. cit.*, p. 27-28.

che di IA interdette (dalla lettera *a*) alla lettera *b*)). A titolo di premessa, può essere segnalato che nelle fattispecie previste dal legislatore europeo si utilizzano i concetti di immissione sul mercato, messa in servizio o uso, che sono concetti tecnici, ben precisati nel loro contenuto precettivo, perché corrispondenti a specifiche definizioni normative, di cui all'art. 3 *sub* n. 9) e *sub* n. 11) ⁽¹³⁾, dove quest'ultimo ricomprende anche l'uso.

La prima pratica di IA vietata *ex* art. 5, par. 1, comma 1, lettera *a*) si riferisce al divieto di immissione sul mercato, messa in servizio od uso di sistemi di IA che impieghino tecniche subliminali o altre tecniche manipolative o ingannevoli (ma anche le ingannevoli sono manipolative) aventi, per scopo o per effetto, la materiale distorsione del comportamento di una persona o di un gruppo di persone, pregiudicando la loro capacità di prendere una decisione informata, ed inducendole a prendere così una decisione che non avrebbero altrimenti preso, in modo da provocare a tale persona ma anche ad altra persona o ad un gruppo di persone un danno significativo. Le tecniche subliminali sono tecniche che agiscono sotto la soglia di consapevolezza dei soggetti e proprio per questo sono tecniche manipolative.

Il problema era emerso già alla fine degli anni Settanta del secolo scorso in riferimento alla pubblicità subliminale dove, al di là del messaggio visibile, si impiegavano immagini o suoni non percettibili per pochi millisecondi con l'obiettivo di influenzare i consumatori in maniera inconsapevole. Per questo la nostra legge in tema di pubblicità ha sempre vietato la pubblicità subliminale ⁽¹⁴⁾. La pubblicità subliminale è in definitiva vietata perché pubblicità contraria al principio cardine in tema di pubblicità, la quale deve essere trasparente, vale a dire una pubblicità sempre riconoscibile come tale ⁽¹⁵⁾.

⁽¹³⁾ L'art. 3 *sub* n. 9) definisce l'«immissione sul mercato» come la prima messa a disposizione di un sistema di IA o di un modello di IA per finalità generali sul mercato dell'Unione, mentre l'art. 3 *sub* n. 11) definisce la «messa in servizio» come la fornitura di un sistema di IA direttamente al *deployer* per il primo uso o per uso proprio nell'Unione per la finalità prevista.

⁽¹⁴⁾ Dapprima nel previgente art. 4, comma 3, d. lgs. n. 74 del 1992, in attuazione della direttiva 84/450/CEE in materia di pubblicità ingannevole, ed ora, nell'art. 5, co. 3, della disciplina di cui al d. lgs. n. 145 del 2007, che tutela sempre in attuazione di direttiva europea, in particolare dell'articolo 14 della direttiva 2005/29/CE che modifica la direttiva 84/450/CEE sulla pubblicità ingannevole, i professionisti dalla pubblicità ingannevole e dalla pubblicità comparativa illecita. Cfr., in proposito, SPANGARO, *Minori e mass media. Vecchi e nuovi strumenti di tutela*, Milano, 2011, in particolare capitolo III.

⁽¹⁵⁾ Su pratiche in generale manipolative ed ingannevoli vedi anche la disciplina in attuazione della direttiva 2005/29/CE sulle pratiche scorrette, confluita oggi nel Codice del Consumo, *sub* artt. 18 ss.

Stabilisce a riguardo delle tecniche subliminali o altre tecniche manipolative il *considerando* 29): «Le tecniche di manipolazione basate sull'IA possono essere utilizzate per persuadere le persone ad adottare comportamenti indesiderati o per indurle con l'inganno a prendere decisioni in modo da sovvertirne e pregiudicarne l'autonomia, il processo decisionale e la libera scelta. L'immissione sul mercato, la messa in servizio o l'uso di determinati sistemi di IA con l'obiettivo o l'effetto di distorcere materialmente il comportamento umano, con il rischio di causare danni significativi, in particolare aventi effetti negativi sufficientemente importanti sulla salute fisica, psicologica o sugli interessi finanziari, sono particolarmente pericolosi e dovrebbero (*rectius*: devono) ⁽¹⁶⁾ pertanto essere vietati. Tali sistemi di IA impiegano componenti subliminali quali stimoli audio, grafici e visivi che le persone non sono in grado di percepire poiché tali stimoli vanno al di là della percezione umana o altre tecniche manipolative o ingannevoli che sovvertono o pregiudicano l'autonomia, il processo decisionale o la libera scelta di una persona senza che sia consapevole di tali tecniche o, se ne è consapevole, senza che sia in grado di controllarle o resistervi o possa evitare l'inganno (...)» ⁽¹⁷⁾.

Nelle Linee guida sono indicate diverse ipotesi di tecniche subliminali, ad esempio per addestrare il sistema di IA, riconducibili all'art. 5, lett. a), quali i messaggi subliminali visivi ⁽¹⁸⁾, uditivi ⁽¹⁹⁾, o tattili ⁽²⁰⁾, le immagini incorporate ⁽²¹⁾ o le tecniche di manipolazione temporale ⁽²²⁾.

⁽¹⁶⁾ Parentesi mia.

⁽¹⁷⁾ L'eventuale avvertenza che, ad esempio, nel video si possono impiegare tecniche subliminali, ora, nel sistema di IA, non impedisce l'inganno.

⁽¹⁸⁾ V. EU Commission, *op. cit.*, p. 20, che rinvia a un sistema di intelligenza artificiale che può mostrare o incorporare immagini o testi che lampeggiano brevemente durante la riproduzione di un video, tecnicamente visibili, ma che lampeggiano troppo velocemente perché la mente cosciente possa registrarli, pur essendo in grado di influenzare atteggiamenti o comportamenti.

⁽¹⁹⁾ Si fa riferimento a un sistema di intelligenza artificiale che può diffondere suoni o messaggi verbali a basso volume o mascherati da altri suoni influenzando l'ascoltatore senza che ne sia consapevole. Questi suoni sono ancora tecnicamente all'interno della gamma dell'udito ma non vengono notati coscientemente dall'ascoltatore a causa della loro sottigliezza o del mascheramento da parte di altri suoni (v. EU Commission, *op. ult. loc. cit.*).

⁽²⁰⁾ Si rinvia a un sistema di intelligenza artificiale che può stimolare sottili sensazioni fisiche percepite inconsciamente in grado di influenzare gli stati emotivi o il comportamento (v. EU Commission, *op. ult. loc. cit.*).

⁽²¹⁾ Si tratta di un sistema di intelligenza artificiale che può nascondere immagini all'interno di altri contenuti visivi che non vengono percepiti coscientemente ma che possono comunque essere elaborati dal cervello e influenzare il comportamento (v. EU Commission, *op. ult. loc. cit.*).

⁽²²⁾ È il caso di un sistema di intelligenza artificiale che alteri la percezione del tempo

Malgrado siano sorti molti dubbi sull'efficacia diretta in tema di pubblicità subliminale, o più in generale sulle tecniche subliminali: se mando per pochi millisecondi nel messaggio promozionale, sotto la soglia di consapevolezza, ad esempio, l'immagine di un tipo di panino Mac Donald, non ottengo il risultato di indurre potenziali consumatori che assistono al messaggio a ordinare proprio quel panino, ma stimolo o posso stimolare anche solo in alcuni un senso di fame, così che il soggetto interessato consuma, ad esempio, patatine fritte, che ha più immediatamente alla sua portata. Certo è che comunque, al di là dell'efficacia messa in dubbio dagli esperimenti, si tratta di tecniche manipolative del comportamento dei soggetti, proprio perché agenti sotto la loro soglia di consapevolezza. Ed in questo senso, insieme a tutte le altre possibili tecniche manipolative, appare più che giustificato il divieto delle corrispondenti pratiche di IA stabilito in proposito dal legislatore europeo. Non è, in ogni caso, una minima soglia quantitativa ad impedire il potenziale effetto ingannevole.

Nelle Linee guida figura, tra gli altri, l'esempio di una tecnica di IA ingannevole e quindi vietata, consistente nell'impiego di un *chatbot* che impersona, simulandone la voce, un amico o un parente di una persona reale per estorcergli denaro⁽²³⁾. Per contro, precisa la Commissione che un sistema di IA generativa che presenti incidentalmente informazioni false o fuorvianti potrebbe non ricadere nel divieto *ex art. 5, par. 1, comma 1, lett. a), AIA*, tenuto conto dei limiti e dello stato dell'arte in tema di intelligenza artificiale generativa. Ciò, in particolare, può avvenire se il fornitore del sistema ha informato adeguatamente gli utenti sui limiti del sistema stesso e vi ha integrato le opportune salvaguardie, al fine di ridurre al minimo tali risultati falsificatori o fuorvianti. In ogni caso, il sistema non deve essere destinato a contesti sensibili, quali salute, istruzione, elezioni, in cui risulta probabile che si verifichino gravi conseguenze dannose, né deve essere utilizzato in simili contesti⁽²⁴⁾.

Va poi precisato quanto segue. Nella disciplina del 2005 ripetutamente citata, con riferimento alle pratiche scorrette, che possono dar luogo, vuoi a pratiche ingannevoli, quali sono anche le pratiche di IA qui in considerazione, vuoi a pratiche aggressive, si operava riferimento al concetto non solo di consumatore ma anche di consumatore vulnerabile, che poteva essere tale, sia in riferimento a singoli individui, sia in riferimento a gruppi

nelle interazioni con l'utente influenzandone il comportamento e causando impazienza e dipendenza (v. EU Commission, *op. ult. loc. cit.*).

⁽²³⁾ V. EU Commission, *op. cit.*, p. 23.

⁽²⁴⁾ *Ibidem*, p. 24.

determinati di persone (ad esempio come bambini ed adolescenti per le pratiche ingannevoli in pubblicità). Va però precisato che il concetto di vulnerabilità lì connesso a singoli o a gruppi di consumatori è, in realtà, un concetto ontologicamente più ampio, perché si riferisce sempre a vulnerabilità (o fragilità) che possono essere proprie degli esseri umani (non solo degli umani in quanto consumatori).

Nel caso di questa disciplina che per le pratiche vietate di IA considera oggetto di protezione salute, sicurezza e diritti fondamentali, che sono diritti inviolabili propri di ciascuno, è pacifico, come già accennato, che il concetto di vulnerabilità viene riferito a persone o gruppi determinati di persone. Del resto, già anteriormente alla disciplina normativa qui in oggetto, il legislatore europeo aveva emanato una specifica direttiva, la 2019/882 sui requisiti di accessibilità dei prodotti e dei servizi in cui, con riferimento espresso alla vulnerabilità, operava rinvio, sotto il profilo soggettivo a “persone” o a “gruppi di persone”. Così, a titolo indicativo, il *considerando* 5) di quella normativa, dove si parlava di «dispositivi assistivi per particolari gruppi di persone con disabilità (...)»; e ancora, l'art. 3 n. 1) nella definizione di «persone con disabilità», ove si stabiliva: «Coloro che hanno minorazioni fisiche, mentali, intellettuali, o sensoriali a lungo termine che, in interazione con varie barriere, possono impedire la loro piena ed effettiva partecipazione nella società su una base di eguaglianza con gli altri»; e nel n. 36) della stessa disposizione, riferito alla definizione di «tecnologia assistiva», si affermava: «qualsiasi elemento, parte di apparecchiatura, servizio o sistema di prodotti, compresi i *software*, utilizzato per accrescere, mantenere, sostituire o migliorare le capacità funzionali delle persone con disabilità oppure per alleviarne o compensare minorazioni, limitazioni dell'attività, o restrizioni della partecipazione», in cui oggi potrebbero annoverarsi, perché con effetti benefici, anche sistemi di IA.

Che comunque la vulnerabilità sia propria di persone fisiche o di gruppi determinati di persone fisiche (e non solo di consumatori) se ne ha riprova anche con riferimento alle discipline preesistenti da cui, al di fuori dei sistemi di IA, è stata mutuata questa tipologia di pratiche vietate (per sistemi di IA), così, ad esempio, in tema di pubblicità o pratiche pubblicitarie, ove la categoria più immediata di consumatori, *rectius*, persone vulnerabili, era costituita da “bambini ed adolescenti” ossia da minori, soggetti beneficiari di una protezione rafforzata di fronte alla disciplina sulla pubblicità ingannevole in attuazione prima di una direttiva europea di armonizzazione minima⁽²⁵⁾ e poi modificata ma mantenuta a

(25) Si veda nota precedente n. 14.

seguito della direttiva del 2005 ripetutamente citata in tema di pratiche commerciali scorrette. A titolo esemplificativo, si era anche precisato, al riguardo, che per il più severo controllo correlato alla vulnerabilità di bambini ed adolescenti fosse ritenuto sufficiente che “la pubblicità fosse suscettibile di raggiungere bambini ed adolescenti (...), non anche che fosse ad essi precipuamente diretta”, come nel caso di messaggi che promuovevano ricambi per moto e motocicli, dove pur trattandosi di campagna pubblicitaria ad ampio spettro su numerosissime riviste, parte di queste erano certamente in grado di raggiungere se non bambini, adolescenti. Veniva allora stabilito che le modalità enfatiche di presentazione del messaggio sembravano idonee a fare maggior presa proprio sui minori, in quanto soggetti dotati di filtri di valutazione meno sviluppati, per vagliare l’attendibilità delle affermazioni pubblicitarie e pertanto si è ritenuto che i messaggi oggetto di allora valutazione fossero da interdire⁽²⁶⁾. Al riguardo, si trattava di valutare una pubblicità presentata anche attraverso riviste specialistiche per moto, ad ampia diffusione, avente per oggetto componenti tecniche per scooter che consentivano di ottenere prestazioni migliori. La pubblicità in realtà indicava con un’avvertenza scritta che si trattava di componenti utilizzabili solo in contesti agonistici o privati, ma non su moto circolanti su strade pubbliche, ma l’indicazione, rispetto all’insieme, non veniva ritenuta sufficiente per neutralizzare il forte impatto grafico visivo della pubblicità stessa e quindi come tale ritenuta idonea a minacciare (anche) la sicurezza di adolescenti in quanto minori⁽²⁷⁾.

Rispetto allo sviluppo che ha interessato sempre il tema delle pratiche di pubblicità che si è registrato nel nostro Paese nel corso del tempo, con riferimento alla categoria di bambini ed adolescenti meritevoli di una normativa rafforzata di fronte alla pubblicità, è particolarmente interessante il cambiamento che ha investito il mutamento della stessa disciplina normativa fra l’epoca in cui vigeva il regime successivo alla prima direttiva contro la pubblicità ingannevole che era come detto una direttiva di armonizzazione minima⁽²⁸⁾, e il regime successivo all’emanazione della direttiva 2005/29/CE sulle pratiche commerciali scorrette, già ripetutamente ricordata, che è di armonizzazione massima. Nella prima e anteriore, si indicavano tre tipologie di una tutela rafforzata di bambini ed adolescenti

⁽²⁶⁾ Così, caso *Maxi Speed Control*, PI/4115, Provv. n. 12, 355, in *Boll.* 32/2003.

⁽²⁷⁾ Sotto il profilo dell’impatto anche e soprattutto visivo della pubblicità stessa, e quindi come tale ritenuta idonea a minacciare (anche) la sicurezza di adolescenti in quanto minori, cfr. anche caso *Glo Hyper X2*, PS/12524, Provv. n. 31053, in *Boll.* 7/2024.

⁽²⁸⁾ Vedi *supra*, nota 14.

contro la pubblicità ingannevole. La prima (già esaminata in riferimento all'esempio riportato) considerava la pubblicità che, essendo in grado di raggiungere bambini ed adolescenti, potesse minacciare direttamente o indirettamente la loro sicurezza⁽²⁹⁾; la seconda considerava la pubblicità, sempre suscettibile di essere ricevuta da bambini ed adolescenti, che abusasse della loro naturale credulità e mancanza di esperienza⁽³⁰⁾. La terza tipologia di tutela rafforzata in pubblicità per bambini ed adolescenti considerava quella pubblicità che, impiegando bambini ed adolescenti in via strumentale, si rivolgesse agli adulti, sfruttando abusivamente i loro sentimenti per i più giovani. È questa una categoria che nell'evoluzione normativa del nostro sistema è scomparsa dalla normativa successiva alla dir. sulle pratiche scorrette tipizzate dal legislatore nazionale dell'attuazione. Tuttavia, in proposito, ha stabilito l'Autorità di controllo in materia: la norma che vietava esplicitamente la pubblicità che abusi dei naturali sentimenti degli adulti per i più giovani non era intesa, a bene vedere, a tutelare direttamente i minori vulnerabili per la loro età, quanto piuttosto ad assicurare che gli operatori pubblicitari non *condizionassero indebitamente* il comportamento economico di consumatori adulti, sfruttando in qualsiasi modo la loro disposizione d'animo nei confronti di bambini ed adolescenti (pratiche da ricomprendersi quindi, oggi, fra le pratiche aggressive verso adulti, piuttosto che in quelle ingannevoli verso minori). Spiegato meglio, queste strategie di consumo potranno eventualmente, alla luce delle circostanze del caso di specie, essere qualificate (non già come pratiche ingannevoli di bambini ed adolescenti ma piuttosto) come pratiche aggressive per adulti vulnerabili o per gruppi di adulti vulnerabili, per le loro particolari condizioni (ad esempio nonni, genitori, ecc., dove dunque appare più appropriato riportare questa categoria di consumatori vulnerabili alla lettera *b*), anche con riferimento oggi a pratiche di IA vietate)⁽³¹⁾.

In altri termini, bambini ed adolescenti sono soggetti sempre vulnerabili ed infatti come tali privi in ogni ordinamento di capacità di agire (ed anche i minori emancipati che sono adolescenti hanno capacità di agire limitata e per atti di straordinaria amministrazione non possono, ad esem-

⁽²⁹⁾ Sempre a titolo indicativo, per tutti, caso *Gilera Runner*, PI/1420, Provv. n. 5146, in *Boll.* 32-33/1997, in cui la potenzialità lesiva della sicurezza del minore veniva ravvisata in una pubblicità relativa ad un ciclomotore e diretta ad esaltare il mito della velocità.

⁽³⁰⁾ Stesso caso *Maxi Speed Control*, già indicato in nota precedente.

⁽³¹⁾ V. caso *Treccani Reti*, PS/12138, Provv. n. 30257 del 2022, in *Boll.* 30/2022, in cui l'acquisto del prodotto viene prospettato «come un sicuro investimento anche per gli eredi».

pio, esprimere autonomo valido consenso, essendo invece richiesto in proposito anche il consenso del curatore). Orbene, bambini ed adolescenti sono sempre vulnerabili perché “strutturalmente” tali, non avendo ancora raggiunto la piena maturità.

Ma, sempre con riferimento alla categoria costituita da persone vulnerabili o a gruppi di persone vulnerabili, traendo spunto dall’esperienza pregressa in tema di pubblicità ingannevole (e quindi qui da oggi di pratiche di IA ingannevoli e per questo vietate), si può pensare ad altre categorie soggettive (o di determinati gruppi di persone) per le loro specifiche condizioni (malattia, o disagio fisico o psichico o altre condizioni di vulnerabilità, come la mancanza di adeguata cultura, o altre condizioni socio-economiche rilevanti). In proposito avevo tuttavia chiarito che, mentre per i minori si tratta di vulnerabilità di tutti e di ciascuno per tutti gli appartenenti allo specifico gruppo dei minori, per gli appartenenti alle altre categorie di gruppi di persone vulnerabili, si tratta di vulnerabilità che può venire in gioco, solo a seconda delle circostanze del caso. Così, traendo spunto dall’esperienza pregressa in tema di pratiche pubblicitarie ingannevoli, appaiono meritevoli di protezione, in quanto gruppo vulnerabile, gli obesi di fronte (allora ma anche oggi) a pratiche pubblicitarie per prodotti dietetici, mentre gli stessi non sono affatto soggetti vulnerabili o appartenenti ad uno specifico gruppo di persone vulnerabile, ad esempio di fronte a pratiche pubblicitarie riferibili a televisori. Così, a titolo indicativo, con riferimento al controllo di cui all’Autorità competente in tema di pratiche pubblicitarie correlabili all’esperienza pregressa, quando si trattava di pratiche pubblicitarie relative a prodotti dietetici, o trattamenti dimagranti in cui venivano omesse informazioni rilevanti quali il fatto che il prodotto dietetico non assicura di per sé il risultato se non accompagnato anche da altro, ad esempio, da esercizio fisico come da passeggiate ogni giorno, ecc. In un caso, ad esempio, si trattava di pubblicità diffusa via Internet, in cui si promuovevano trattamenti dimagranti offerti dal professionista nei suoi centri specializzati dove si impiegavano macchinari di elettrostimolazione. Più in particolare, l’ingannevolezza si è correlata al fatto che i messaggi in questione non indicassero le controindicazioni al trattamento dimagrante oggetto di pubblicità, che venivano illustrate solo nelle condizioni generali di contratto del servizio presentate all’atto di conclusione del contratto stesso, ma non in quanto essenziali per l’invito all’acquisto, esplicitate anche nel primo aggancio del consumatore, e quindi in pubblicità. Si è dichiarato dall’Autorità adibita al controllo in materia: «(...) In un settore come quello dei trattamenti estetici dimagranti, in cui l’asimmetria informativa a vantaggio del professionista impone una

particolare attenzione in ordine alla completezza informativa delle comunicazioni ai consumatori (...) anche in considerazione delle condizioni di debolezza psicologica (*rectius*: di vulnerabilità) in cui potrebbero trovarsi i destinatari dei messaggi, soggetti che spesso versano in forte sovrappeso, (...) impone (come detto) un maggior grado di completezza all'informazione fornita in pubblicità anche al fine di non porre al rischio la salute dei consumatori». Si è allora concluso: «Il suddetto messaggio pubblicitario risulta ingannevole sotto il profilo omissivo nella misura in cui tace i rischi e le controindicazioni connessi a trattamenti di elettrostimolazione, potendo indurre il consumatore a trascurare le normali regole di prudenza a tutela della propria salute e sicurezza»⁽³²⁾.

Ancora, l'Autorità adibita al controllo ha argomentato l'ingannevolezza di messaggi pubblicitari aventi per oggetto talismani, carte che predicono il futuro, oggetti magici, operando riferimento a destinatari vulnerabili in ragione delle loro condizioni socioculturali particolarmente disagiate. Si trattava, nella specie, di un messaggio pubblicitario che promuoveva la vendita di un talismano. L'Autorità osservava che destinataria della campagna non poteva intendersi la generalità dei consumatori, quanto piuttosto quelle persone o gruppi determinati di persone che, per ragioni socioculturali, erano portate a credere all'esistenza di fenomeni caratterizzati per loro natura da "intrinseca indimostrabilità"⁽³³⁾. In realtà, sotto il

⁽³²⁾ Caso *Centri Dimagranti Sobrino*, PS/5445, Prov. n. 21260, in *Boll.* 24/2010; ma cfr. anche caso *Ariix-Newage*, PS/12323, Prov. n. 30423 del 2022. Cfr. anche caso *Clean Forte*, PS/12440, Prov. n. 30703 del 2023, in *Boll.* 27/2023 e caso *Nutriscore-Vivil*, PS/12187, Prov. n. 30317, in *Boll.* 5/2023. Vedi pure caso *T-Shock 31-Timodellaclub*, PS/9388, Prov. n. 30325 del 2022, relativamente a una pratica commerciale scorretta in violazione degli artt. 20, 21, comma 1, lettera *b*), e 22 del Codice del Consumo, nell'ambito della produzione e commercializzazione di cosmetici e attrezzature nel settore dei trattamenti estetici e dimagranti, consistente nell'aver diffuso a mezzo stampa e attraverso siti Internet messaggi ingannevoli e omissivi circa le proprietà, i risultati ottenibili e l'efficacia del trattamento T-Shock31.

⁽³³⁾ Così, fra i tanti casi, *Benaglio Dorian Magia*, PI/880, Prov. n. 4222, in *Boll.* 36/37 96; 2. Cfr. anche il caso *Sensitiva Adelia Felice*, PS/2681, Prov. n. 1979 del 2009. In questo caso, la pratica è stata ritenuta ingannevole, in quanto in violazione degli articoli 20 e 23, comma 1, lettera *s*), del Codice del Consumo, consistente in affermazioni relative a prodotti che, contrariamente al vero, sarebbero in grado di curare malattie, disfunzioni o malformazioni, che nel caso concreto faceva leva sulla debolezza e disperazione di alcune persone che possano trovarsi in gravi difficoltà per problemi di salute, lavoro, denaro, prospettandosi in realtà anche profili di aggressività. Come risulta dal provvedimento dell'Autorità, infatti, la professionista non si limitava ad una generica esaltazione dei propri poteri di "Astrologa – Telepatica e Carismatica" come ella stessa si definiva e delle sue eventuali capacità, ma appariva particolarmente suscettibile di attrarre quelle categorie di consumatori propense ad un approccio fideistico ed irrazionale verso l'attività di sensitive e astrologiche, facendo leva sulla debolezza, fragilità psicologica e disperazione di alcune

profilo dei sistemi di IA vietati, come vedremo *infra*, è più opportuno riportare questa pratica entro la lett. *b*), perché si sfrutta la vulnerabilità di singoli o di gruppi vulnerabili, approfittando della loro eccessiva credulità.

Il secondo tipo di pratiche di IA vietate è, *ex art. 5 par. 1, comma 1, lett. b*): l'immissione sul mercato, messa in servizio o uso, di un sistema di IA che sfrutta la vulnerabilità di una persona fisica, o di uno specifico gruppo di persone, per l'età (bambini e adolescenti, ossia minori), per disabilità, o per una specifica situazione sociale o economica (anziani in stato di malattia, solitudine, o povertà) con lo scopo o l'effetto di distorcere materialmente il comportamento di tale persona, o di una persona che appartiene a tale gruppo, in modo che possa provocare a lei o ad altra persona un danno significativo. Si tratta, in definitiva, di pratiche interdette, perché sfruttano la vulnerabilità di persone fisiche o di gruppi di persone vulnerabili al fine e per gli effetti già ricordati. Anche qui, al di fuori delle pratiche di IA, si tratta di tecniche già vietate dalla legge, ad esempio dalla disciplina di matrice europea sulle pratiche scorrette (in particolare, le pratiche scorrette in quanto aggressive), dove, fra il resto, veniva previsto, come già riportato, anche il concetto di consumatore vulnerabile. In quest'ultimo caso, rinviando a condizioni quali l'infermità mentale o fisica, l'età, o ingenuità (o altre condizioni considerate in riferimento specifico alle pratiche commerciali scorrette *sub art. 5*, nonché *sub artt. 8 e 9*, di cui alla dir. 2005/29/CE).

In passato, non avevo mancato di osservare, che anche il codice civile, per quanto con diversa prospettiva di tutela rispetto a quella della disciplina delle pratiche scorrette, opera rinvio all'interno della violenza morale per la tutela del contraente, ad una pluralità di categorie di vittime distinte in ragione "dell'età, del sesso e della condizione delle persone", *ex art. 1435*, così, come nelle pratiche aggressive dovute a molestie, coercizione o indebito condizionamento si opera rinvio ad analoghe categorie di persone vulnerabili o di gruppi vulnerabili di soggetti⁽³⁴⁾. In entrambi i casi, tuttavia, le condizioni soggettive delle "persone" fisiche vittime delle pratiche

persone che potessero trovarsi in gravi difficoltà per i più vari problemi, in particolare di salute.

⁽³⁴⁾ La differenza di prospettiva attiene al fatto che mentre la citata disciplina europea interviene sull'*attività* dell'operatore economico suscettibile di sfruttare la volontà del consumatore vulnerabile, o di un determinato gruppo di consumatori vulnerabili e può anche prescindere dalla già avvenuta conclusione di un atto di acquisto, mirando ad impedirne la conclusione, il codice civile ha riguardo all'*atto* posto in essere dal contraente debole messo in condizione di costrizione a contrarre, e mira all'invalidazione dell'atto.

di sfruttamento del loro comportamento sono sostanzialmente le medesime. È proprio con riferimento a questa tipologia di pratiche vietate nei sistemi di IA che, nel *considerando* n. 29), si parla di «sfruttamento in altro modo», rispetto a quanto stabilito nella precedente lettera a) della vulnerabilità di una persona o di uno specifico gruppo di persone dovuto «all'età, a disabilità (ed ecco che proprio in proposito si menziona la già citata direttiva 2019/882) o a una specifica situazione sociale, o economica, che potrebbe rendere tali persone, o il corrispondente specifico gruppo di appartenenza, più vulnerabili allo sfruttamento, come le persone che vivono in condizioni di povertà, e le minoranze etniche o religiose».

La vulnerabilità di persone o di gruppi specifici di persone, con riferimento a quanto in passato già oggetto di controllo al di fuori dei sistemi di IA, è stata ravvisata, a titolo esemplificativo, innanzi tutto in bambini ed adolescenti per la loro età in rapporto alla pubblicità, che abusasse della loro naturale credulità e mancanza di esperienza. In alcuni casi si trattava di pratiche pubblicitarie dirette a bambini ma il cui oggetto appariva destinato invece ad un pubblico adulto. Il fine consisteva nello sfruttare il comportamento del minore per la sua credulità o mancanza di esperienza affinché facesse indebita pressione sugli adulti (genitori o nonni) per l'acquisto del prodotto interessato. Così, ad esempio, una pubblicità che comunicava l'organizzazione di una festa a cui venivano invitati espressamente a partecipare gratuitamente i bambini accompagnati dai genitori, ma si trattava in realtà di un'iniziativa promozionale diretta alla vendita di un'enciclopedia ⁽³⁵⁾.

La terza tipologia di tutela rafforzata in pubblicità per bambini e adolescenti considerava quella pubblicità che, impiegando bambini ed adolescenti in via strumentale, si rivolgesse agli adulti sfruttando abusivamente i loro sentimenti per i più giovani. È questa una categoria che nell'evoluzione normativa del nostro sistema è scomparsa dalla normativa successiva alla dir. sulle pratiche scorrette, tipizzate dal legislatore nazionale dell'attuazione. Tuttavia, in proposito e come già visto, l'Autorità di controllo in materia ha stabilito che la pratica non si configuri quale pratica ingannevole di bambini ed adolescenti, ma piuttosto come pratica aggressiva per adulti vulnerabili o per gruppi di adulti vulnerabili per le loro particolari condizioni, ad esempio nonni per un loro (eccessivo) attaccamento alla generazione dei più giovani.

⁽³⁵⁾ Caso *didattica*, PI/7417, Provv. n. 2771, in *Boll.* 4/1995. Cfr. anche *Treccani Reti*, PS/12138, Provv. n. 30257 del 2022, in *Boll.* 30/2022 sopra richiamato.

In riferimento alla disciplina sulle pratiche scorrette, sono state considerati consumatori vulnerabili o appartenenti a gruppi vulnerabili, soggetti in ragione delle loro particolari condizioni (ad esempio di disagio fisico o psichico o culturale o socioeconomico). Così, ad esempio, iniziative promozionali dirette alla vendita di talismani o scaccia malocchio, o la lettura di carte a pagamento, che sfruttano il comportamento di persone vulnerabili colpite da un evento luttuoso (vero e proprio come la morte di un familiare, ma anche solo un accadimento luttuoso sentimentalmente, come la perdita di un amore), portandole “irragionevolmente” a credere a vincite facili e sicure, o a risoluzione di problemi di varia natura, ecc. attribuendo all’oggetto messo in vendita, o al comportamento dell’operatore economico “caratteristiche ed effetti palesemente inverosimili” che non trovavano riscontro nella realtà ma solo nell’eccessiva credulità degli interessati di cui, in definitiva, si sfruttava indebitamente il comportamento, portandoli ad acquistare il prodotto o servizio interessato⁽³⁶⁾. E quello che qui c’è da notare è che in certe circostanze, come nell’accadimento luttuoso prima citato, ciascuno quindi anche uno di noi può perdere l’equilibrio e divenire vulnerabile.

La Commissione, nel dare molti esempi di ognuna delle pratiche vietate, di cui all’art. 5, *sub* lettere *a*) e *b*)⁽³⁷⁾, pone tuttavia indicazioni interessanti concernenti le pratiche che non ricadono nel divieto.

Si segnala, tra gli altri, l’esempio di un *chatbot* terapeutico che utilizza tecniche subliminali per indirizzare gli utenti verso uno stile di vita più sano e per abbandonare le cattive abitudini come il fumo. Come chiarito nelle Linee guida, anche se gli utenti che seguono i consigli e la terapia subliminale sperimentano un certo disagio fisico e stress psicologico a causa dello sforzo compiuto per smettere di fumare, il *chatbot* abilitato all’intelligenza artificiale non può essere considerato in grado di causare danni significativi. Tale disagio temporaneo è, infatti, inevitabile e compensato dai benefici a lungo termine per la salute degli utenti. Non ci sono tentativi nascosti di influenzare il processo decisionale al di là della promozione di abitudini sane⁽³⁸⁾.

Qui a dire il vero c’è un impiego di tecniche subliminali, ma non c’è manipolazione, perché, in primo luogo, la persona è consapevole dell’uti-

⁽³⁶⁾ Caso *Mago Vito lo Cascio*, PS/717, Prov. 21174, in *Boll.* 22/2010.

⁽³⁷⁾ Perché evidentemente la posizione è la stessa: i soggetti vulnerabili non vengono in gioco solo nella lettera *b*), dove essi sono sempre vulnerabili, ma anche nella lettera *a*) come rilevato nel testo.

⁽³⁸⁾ V. EU Commission, *op. cit.*, p. 46.

lizzo delle tecniche subliminali e, anzi, utilizza il sistema di IA in oggetto per il fine per cui è preordinato perfettamente consapevole di tutto, e, d'altra parte, si tratta comunque di una tecnica impiegata per scopo terapeutico. L'unica questione rimane non l'ammissibilità, ma una ammissibilità sotto sorveglianza medica che, di solito, nel caso degli scopi terapeutici o altrimenti ammessi si richiede. Del resto, è indubbio, in parallelo, che lo stesso scopo potrebbe essere raggiunto attraverso un sistema di IA che utilizza l'ipnosi. Ma anche in questo caso, rimane pacifico che il tutto deve avvenire sotto il controllo medico.

Il terzo tipo di pratiche di IA vietate, consiste, *ex art.* 5, par. 1, comma 1, lettera c), nella: immissione sul mercato, messa in servizio o uso di sistemi di IA per la valutazione o la classificazione delle persone fisiche o di gruppi di persone fisiche per un determinato periodo di tempo sulla base del loro comportamento sociale o di caratteristiche personali o della personalità, conosciute (ma anche inferite o previste), in cui il punteggio sociale negativo così ottenuto comporti il verificarsi di uno o di entrambi gli scenari seguenti: *i*) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche, o di gruppi di persone, in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti; *ii*) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di gruppi di persone che sia ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla sua gravità.

A proposito del caso *sub i*), le Linee guida riportano, tra gli altri, l'esempio dell'utilizzo, da parte di un'autorità fiscale, di uno strumento predittivo di IA su tutte le dichiarazioni dei redditi dei contribuenti di un Paese, al fine di selezionare le dichiarazioni dei redditi da sottoporre ad un'ispezione più approfondita. Lo strumento di IA ipotizzato rientrerebbe nel divieto, giacché utilizzerebbe variabili rilevanti, come il reddito annuale, il patrimonio, i dati sui familiari dei beneficiari, ma anche dati non correlati, come le abitudini sociali dei contribuenti o le connessioni a Internet, per individuare specifiche persone da sottoporre a ispezione. Per contro, la Commissione esclude dal divieto il caso esemplificativo di un'agenzia di assistenza sociale che utilizzi un sistema di IA per stimare la probabilità di frode da parte dei beneficiari di assegni familiari, basandosi su dati strettamente rilevanti per l'assegnazione delle prestazioni e raccolti legalmente⁽³⁹⁾.

⁽³⁹⁾ V. EU Commission, *op. cit.*, p. 57.

Con riferimento al caso *sub ii*), nelle Linee guida viene riportato, tra gli altri, l'esempio di un comune che utilizzi un sistema di IA per assegnare un punteggio all'affidabilità dei residenti sulla base di molteplici dati relativi al loro comportamento sociale in una varietà di contesti. Il punteggio generato per i residenti considerati "meno degni di fiducia" verrebbe utilizzato per l'inserimento in una lista nera, che comporterebbe la revoca di benefici pubblici o altre gravi misure punitive e, in ogni caso, una maggiore sorveglianza. In questo esempio, tra i fattori oggetto di valutazione, vi sarebbero anche l'insufficiente volontariato e/o piccoli comportamenti scorretti, come la mancata restituzione dei libri alla biblioteca in tempo, l'abbandono dei rifiuti per strada al di fuori del giorno di raccolta e il ritardo nel pagamento delle tasse locali. Tale esempio ricade nel divieto, perché, come è chiaro, consiste in un trattamento del tutto ingiustificato e sproporzionato rispetto al singolo comportamento sociale preso in considerazione⁽⁴⁰⁾. Diversamente, viene escluso dall'ambito di applicazione del divieto il caso esemplificativo di un sistema di intelligenza artificiale che utilizzi i dati raccolti nei campi profughi (ad esempio la conformità comportamentale) per le decisioni relative al reinsediamento o all'occupazione, dal momento che tali dati sarebbero rilevanti ai fini della valutazione; ciò, sempre a condizione che siano rispettate le procedure previste dal diritto dell'Unione in materia di migrazione e che, quindi, il trattamento risulti giustificato e proporzionato⁽⁴¹⁾.

Si tratta, in definitiva, di pratiche di IA che impieghino il c.d. *social scoring* o punteggio sociale. Appare opportuno chiarire soprattutto alle giovani generazioni, al riguardo, che anche il modello sociale italiano, fino allo scoppio della seconda guerra mondiale era un modello sociale piuttosto "ingessato" che premiava con punteggio sociale positivo o toglieva punteggio sociale in negativo per determinate caratteristiche di (mancata) omologazione sociale (ad esempio, più semplicemente, un vestire standardizzato per passeggiare in centro città, dove non poteva mancare nei giorni festivi il cappello in qualsiasi stagione e non erano ammessi jeans o tute da lavoro), con stigmatizzazione invece di chi non si omologava al modello stesso; allo stato attuale, il principio di libertà proprio di tutte le società democratiche ha di fatto finito via via per togliere peso al c.d. punteggio sociale, che nel suo profilo in negativo, qui adeguatamente considerato dalla legge, può costituire fattore (e nella storia, purtroppo non mancano

⁽⁴⁰⁾ V. *op. ult. loc. cit.*, p. 58.

⁽⁴¹⁾ *Ibidem*, p. 63.

gli esempi), di gravi discriminazioni per individui o per appartenenti a determinati gruppi di persone anche eventualmente vulnerabili (ad esempio, per la loro appartenenza etnica, o per caratteristiche personali o sociali), punteggio sociale che di fatto può essere ancora presente, invece, sia in società meno evolute della nostra attuale sia in determinati gruppi sociali, o in sette, o più in generale, in società anche politiche chiuse, non improntate a libertà o non democratiche. Si afferma così al riguardo nel *considerando* 31): «I sistemi di IA che permettono ad attori pubblici o privati di attribuire un punteggio sociale alle persone fisiche possono portare a risultati discriminatori e all'esclusione di determinati gruppi (...). Tali sistemi di IA valutano o classificano le persone o gruppi di persone sulla base di vari punti di dati riguardanti il loro comportamento sociale in molteplici contesti o di caratteristiche personali o della personalità note, inferite o previste nell'arco di determinati periodi di tempo. Il punteggio sociale ottenuto da tali sistemi di IA può determinare un trattamento pregiudizievole o sfavorevole di persone fisiche o di interi gruppi in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti, o a un trattamento pregiudizievole che risulta ingiustificato o sproporzionato rispetto alla gravità del loro comportamento sociale. I sistemi di IA che comportano tali pratiche inaccettabili di punteggio aventi risultati pregiudizievoli o sfavorevoli dovrebbero (*rectius*: devono) ⁽⁴²⁾ pertanto essere vietati (...)».

In riferimento al punteggio sociale occorre puntualizzarne un aspetto. Esso a rigore può manifestarsi in due profili fra loro antitetici. Può essere punteggio sociale in negativo come quello chiaramente evidenziato *sub* lett. c), par. 1, comma 1, art. 5, oggetto di divieto, per le ragioni già evidenziate. Ma può essere anche punteggio sociale positivo, ove si danno punti (o premi o riconoscimenti o misure a favore) verso coloro che si uniformano ad un determinato positivo modello sociale promosso di regola dallo Stato. A titolo esemplificativo, per quanto riguarda il nostro Paese, si può pensare alle misure di sostegno per le famiglie numerose al fine di combattere la crescita 0. In questo caso non c'è ragione per negarne l'ammissibilità anche da parte di sistemi di IA per il perseguimento di un dato modello che lo stato nazionale nel quadro dei valori dell'Unione europea voglia promuovere e perseguire. Anche le Linee guida riportano un esempio di *social scoring* positivo rapportato ad un determinato modello sociale. Si tratta dell'ipotesi di un sistema di intelligenza artificiale

(42) Parentesi mia.

applicato ad una piattaforma di acquisti *online* che offra privilegi agli utenti con una solida storia di acquisti e un basso tasso di restituzione dei prodotti, come una procedura di richiesta di restituzione più rapida ovvero la possibilità di accedere ai rimborsi senza restituzione. Come chiarito dalla Commissione europea, tale esempio non ricadrebbe nel divieto *ex art. 5, par. 1, comma 1, lettera c)*, dal momento che i vantaggi sarebbe giustificati e proporzionati per premiare un comportamento positivo e gli altri utenti continuerebbero comunque ad avere accesso alla procedura di restituzione *standard* ⁽⁴³⁾.

Le pratiche di IA, finora considerate, anche per tutto quanto illustrato in proposito, presentano un tratto comune. Corrispondono a pratiche al di fuori dei sistemi di IA già note al pubblico e, come detto, anche in passato poste in essere che, dando luogo comunque a tipologie di pratiche in contrasto con valori fondamentali protetti all'apice delle fonti del diritto Europeo (Trattati o Carta di Nizza nella versione di Strasburgo), ma anche contrari ad alcuni nostri valori costituzionalmente protetti, erano già state vietate dal diritto europeo e/o dal nostro diritto interno. Così, solo a titolo indicativo, la pubblicità subliminale costituisce esempio classico di una pubblicità che agisce sotto la soglia di consapevolezza e per questo manipolativa dell'autonomia decisionale delle persone fisiche e per questo già vietata. Le pratiche di IA che sfruttano il comportamento umano ⁽⁴⁴⁾ sono gravemente pregiudizievoli per la dignità umana e foriere di discriminazioni in violazione del principio di uguaglianza, in particolare verso persone vulnerabili o gruppi di persone vulnerabili che meritano invece protezione al fine di garantire loro il pieno esercizio dei diritti inviolabili nel rispetto di una autentica partecipazione ed inclusione sociale. Il *social scoring* negativo è vietato entro i limiti stabiliti dalla legge europea, perché è atto a dar luogo a gravi discriminazioni. In altri termini, si guarda ai sistemi di IA che, classificando le persone in base al comportamento sociale o a caratteristiche personali, possono comportare pratiche lesive del diritto alla dignità e alla non discriminazione e dei valori di uguaglianza e giustizia.

È pacifico allora che il presente Regolamento vieta tali pratiche di IA, corrispondenti alle lettere *a) b) c)* del par. 1, comma 1, dell'art. 5, perché si tratta di pratiche che, se impiegate all'interno dei sistemi di IA, violano valori fondamentali protetti proprio all'interno dei sistemi di IA, ed in

⁽⁴³⁾ V. EU Commission, *op. cit.*, p. 63.

⁽⁴⁴⁾ Già considerate *sub* lettera *b)*.

particolare i due valori più preganti e specifici al riguardo menzionati dall'AIA, come il valore della «affidabilità» e, parallelamente, il valore del voler costituire una «tecnologia antropocentrica». Sotto il primo profilo, si tratta di sistemi di IA, vietati dalla disciplina dell'AIR, perché si tratta di sistemi di IA sviluppati e impiegati in modo in ogni caso opaco e non trasparente, e tali da minare l'affidabilità del mercato e degli umani sui sistemi di IA corrispondenti⁽⁴⁵⁾. Sotto il secondo profilo, rileva il valore dell'antropocentrismo dei sistemi di IA⁽⁴⁶⁾, perché le indicate pratiche di IA non evidenziano certo il "(...) fine ultimo di migliorare il benessere degli esseri umani" ⁽⁴⁷⁾, ma, tutto all'opposto, sono pratiche lesive degli umani e che possono arrecare danni significativi agli umani stessi (ed ai loro diritti fondamentali), quindi certamente in palese e significativo contrasto con il loro benessere.

3. – Le prossime pratiche di IA vietate presentano anch'esse un tratto comune. Appartengono a categorie di pratiche di IA che implicano o possono implicare, perché correlate a trattamenti massivi di dati relativi a persone fisiche, un'ingiustificata invasione della loro sfera personale e della dignità umana non improntata al principio di proporzionalità e anche per questo fortemente lesiva. Proprio perché pratiche di IA correlate al trattamento di dati di carattere personale, non possono comunque violare i principi in proposito espressi dalla Carta dei diritti e delle libertà fondamentali di cui *sub* art. 8, né, di regola, quelli del GDPR del 2016. Nelle disposizioni qui includibili figura l'espressione "per tale finalità specifica" riferita alla messa in servizio di tali pratiche. Più in particolare, si tratta di pratiche di IA che vanno dalla lettera *d*) alla lettera *g*).

Va qui precisato che nell'articolo 5 (dalla lettera *d*)) si fa riferimento, con riguardo alla "messa in servizio", alla "finalità specifica", laddove nelle definizioni dui all'art. 3 n. 12) il riferimento è alla "finalità prevista". La mancanza di identica espressione figura anche nella versione inglese e

⁽⁴⁵⁾ I principi in tema di IA, al fine di garantire un'IA affidabile e antropocentrica, sono stati originariamente elaborati all'interno dell'*High-Level Expert Group on Artificial Intelligence*, con acronimo IA HLEG, che è un gruppo di esperti del mondo delle imprese, della ricerca, dell'Università e della pubblica amministrazione, nominato dalla Commissione europea nel 2008, allo scopo di fornire consulenza sulla sua strategia in tema di intelligenza artificiale proprio al fine di supportare l'implementazione di tale strategia. E v., al riguardo, il *considerando* n. 27), dove sono declinati i principi sull'IA.

⁽⁴⁶⁾ Alla cui tutela appaiono declinati sempre i principi in tema di IA per cui v. nota precedente.

⁽⁴⁷⁾ Così come affermato dal *considerando* n. 6).

francese, dove si usano i termini “*Specific purpose*” (art. 5)/“*intended purpose*” (art. 3) e “*à cette fin spécifique*” (art. 5)/“*destination*” (art. 3) e ciò è da stigmatizzare, perché le norme cui si riferiscono devono utilizzare la definizione normativa, atteso che le definizioni hanno significato precettivo e quindi non vanno modificate.

La pratica di IA vietata, *sub* lettera d), sempre di cui all’art. 5, par. 1, comma 1, dell’AIA consiste nella immissione sul mercato, la messa in servizio per tale finalità specifica, o l’uso di un sistema di IA per effettuare valutazioni del rischio relative a persone fisiche che commettano un reato, unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della sua personalità; tale divieto non si applica ai sistemi di IA utilizzati a sostegno della valutazione umana del coinvolgimento di una persona in un’attività criminosa, relativa alla commissione di un reato, che si basa già su fatti oggettivi e verificabili nonché direttamente connessi, come detto, ad attività criminosa.

La *ratio* del divieto è di impedire un controllo sociale pubblico o d’ordine pubblico basato unicamente su valutazioni o previsioni come quelle qui in oggetto che siano fondate esclusivamente su sistemi di IA. E del resto anche il trattamento di dati che la lettera qui interessata implica proprio perché di per sé invasivo della persona fisica (o dei soggetti interessati) ricordiamo che deve risultare necessariamente circoscritto per mezzo del sistema di IA al raggiungimento di “finalità determinate” (*ex* art. 8, co. 2 della Carta ed *ex* art. 5, par. 1, lett. b), GDPR). Nel caso di specie, il trattamento di dati avviene attraverso il sistema di IA interdetto per valutare o prevedere il rischio che una persona fisica commetta un reato sulla base esclusiva della sua profilazione, sui tratti della sua personalità o delle sue caratteristiche. Stabilisce in proposito il *considerando* n. 42): «In linea con la presunzione di innocenza, le persone fisiche nell’Unione dovrebbero sempre essere giudicate in base al loro comportamento effettivo. Le persone fisiche non dovrebbero mai essere giudicate sulla base di un comportamento previsto dall’IA basato unicamente sulla profilazione, sui tratti della personalità, o su caratteristiche quali la cittadinanza, il luogo di nascita, il luogo di residenza, il numero dei figli, il livello di indebitamento, o il tipo di automobile, senza che vi sia un ragionevole sospetto che la persona sia coinvolta in un’attività criminosa sulla base di fatti oggettivi verificabili e senza una valutazione umana al riguardo. Pertanto, dovrebbero (*rectius*: devono)⁽⁴⁸⁾ essere vietate le valutazioni del

(48) Parentesi mia.

rischio effettuate in relazione a persone fisiche intese a determinare la probabilità che queste ultime commettano un reato, o volte a prevedere il verificarsi di un reato effettivo o potenziale unicamente sulla base della sola loro profilazione o della valutazione dei loro tratti della personalità o delle loro caratteristiche (...)».

In altri termini, è bene precisarlo nuovamente, la disciplina normativa vieta un sistema di IA diretto alla valutazione del rischio di commissione o di previsione di un reato sulla unica ed esclusiva base di profilazione, o della valutazione dei tratti, o di specifici tratti, come ad esempio, in passato, nei casi di lombrosiana memoria, o/e delle caratteristiche della personalità (sono a dire il vero i “pregiudizi” che, in quanto tali, si possono basare esclusivamente sulle caratteristiche delle persone fisiche, quali ad esempio, il pregiudizio, che persone per la loro cittadinanza perché provenienti dall’Afghanistan, o perché nate in Sicilia possano commettere un reato di terrorismo, o un reato di mafia). Il divieto si giustifica, perché altrimenti si darebbe luogo ad un’invasione non giustificata della sfera inerente la persona fisica, e inoltre del tutto avulsa dal principio di proporzionalità, per il caso di specie, al di fuori di qualsiasi fatto concreto rilevante al riguardo che possa giustificare un’intrusione nella sfera della persona stessa, con in parallelo, un controllo sociale così spaventoso al fine di valutare o prevenire un reato, come abilmente scritto anche nel famoso “1984” di Orwell. Qui la ragione giustificatrice del divieto è chiara, si vuole impedire un controllo sociale che risulterebbe, se permesso, in violazione di qualsiasi principio proprio di uno Stato di diritto ed al contempo tale da risultare fortemente ed ingiustificatamente lesivo della dignità umana, anche costituzionalmente protetta, degli interessati coinvolti. Nelle Linee guida della Commissione europea è riportato, fra gli altri, l’esempio di un’autorità preposta all’applicazione della legge che utilizzi un sistema di intelligenza artificiale per prevedere il comportamento criminale per reati come il terrorismo, basandosi esclusivamente sull’età, sulla nazionalità, sull’indirizzo, sul tipo di auto e sullo stato civile delle persone. Con questo sistema, si potrebbe ritenere che gli individui abbiano maggiori probabilità di commettere in futuro reati che non hanno ancora commesso, solo in ragione alle loro caratteristiche personali. Un simile sistema sarebbe chiaramente vietato, sulla base dell’articolo 5, paragrafo 1, comma 1, lettera d) ⁽⁴⁹⁾.

⁽⁴⁹⁾ V. EU Commission, *op. cit.*, p. 69.

Parallelamente, la parte finale della stessa disposizione consente invece la valutazione o la previsione del suddetto rischio di commissione di un reato da parte di una persona fisica, correlato a sistemi di IA, quando il corrispondente sistema di IA è utilizzato (non unicamente come nell'ipotesi di prima vietata dalla legge), ma invece "a sostegno" di una valutazione necessariamente umana, che sia relativa al coinvolgimento di soggetti in attività criminosa, e tutto ciò, sulla base di fatti che sono, innanzi tutto fatti, ossia comportamenti commissivi (od eventualmente omissivi, se rilevanti) qualificati dall'essere «oggettivi e verificabili», nonché «direttamente connessi ad un'attività criminosa». Le Linee guida riportano quale esempio di pratica in oggetto consentita quella di un sistema di polizia predittiva basato sull'IA che fornisca un punteggio della probabilità di criminalità in diverse aree di una città, sulla base dei tassi di criminalità precedenti per zona e di altre informazioni di supporto, come le mappe stradali, per evidenziare l'elevato rischio di specifici tipi di reati, quali furti con scasso, crimini con armi da taglio, ecc., e aiutare le autorità di polizia a determinare i luoghi in cui dispiegare più o meno pattuglie, al fine di svolgere un'azione di polizia efficace, volta ad interrompere e/o scoraggiare le attività criminali. Si tratterebbe di un sistema di IA che svolgerebbe previsioni di criminalità basate sulla geolocalizzazione e altri dati oggettivi e che, pertanto, sarebbe escluso dall'ambito di applicazione del divieto⁽⁵⁰⁾.

L'esclusione dal divieto nei termini precisati si fonda su due profili. Innanzitutto, la profilazione da parte del sistema di IA per la valutazione del rischio che una persona fisica commetta un reato o per la previsione di un reato deve risultare tale da risolversi in uno strumento d'ausilio di una valutazione che non può essere sottratta agli umani e che deve necessariamente rimanere una valutazione umana. Sotto altro aspetto, la valutazione che possa essere commesso un reato o la previsione di un reato non può essere estranea da una valutazione condotta e fondata sull'esistenza effettiva di elementi probatori, o di mezzi di prova da cui, in uno stato di diritto, oggi, non è possibile prescindere. Tutto ciò, in definitiva, si rivela funzionale e strumentale per facilitare un controllo pubblico che è e deve rimanere sostanzialmente umano e che, di nuovo, proprio perché delimitato nel senso precisato dalla legge, risponda ai principi cui è improntato ogni stato democratico che fa parte dell'Unione europea.

Solo un'osservazione critica marginale: forse in proposito una qualche perplessità rimane nel caso in cui il divieto si applica (o la corrispondente

⁽⁵⁰⁾ *Ibidem*, p. 72.

esclusione non si applica) ai sistemi di IA interessati, quando si tratti, eventualmente, di attività criminosa rivolta alla perpetrazione di reati per così dire bagatellari⁽⁵¹⁾, perché la formula legislativa nulla ci dice in realtà al riguardo ed altrimenti potrebbe riscontrarsi un eventuale possibile, contrasto con il principio di proporzionalità.

Rimane non inclusa nel divieto, perché non ricompresa nella lettera stessa del medesimo e perché non riconducibile sotto il profilo degli interessi tutelata alla ragione giustificatrice del divieto stesso, l'analisi del rischio non fondata sulla profilazione delle persone fisiche, o su tratti della personalità e sulle caratteristiche delle persone. La parte finale del *considerando* 42) stabilisce al riguardo: «È il caso, ad esempio, dei sistemi di IA che impieghino l'analisi dei rischi per valutare il rischio di frode finanziaria da parte di imprese sulla base di transazioni sospette o di strumenti di analisi del rischio per prevedere la probabilità di localizzazione di stupefacenti o merci illecite da parte delle autorità doganali, ad esempio sulla base di rotte di traffico conosciute».

La pratica di IA vietata *sub* lettera e), par. 1, comma 1, art. 5 dell'AIA consiste nell'immissione sul mercato, messa in servizio per tale finalità specifica o uso di sistemi di IA che creano o ampliano le banche dati di riconoscimento facciale mediante *scraping* non mirato di immagini facciali da internet o da filmati di telecamere a circuito chiuso. Con chiarezza si stabilisce nel *considerando* n. 43), a giustificazione del divieto di questa pratica di IA, con espressione che vale la pena riprendere per intero e senza più aggiungere nulla in proposito: «tali sistemi di IA dovrebbero (*rectius*: devono)⁽⁵²⁾ essere vietati, in quanto la suddetta pratica accresce il senso di sorveglianza di massa e può portare a gravi violazioni dei diritti fondamentali, compreso il diritto alla vita privata».

Può essere opportuno precisare che il fenomeno dello *scraping* è di per sé è molto utile perché permette di prelevare con estrazione, ad esempio dai siti *web*, un massivo quantitativo di data, classificandoli, suddividendoli ed archiviandoli all'interno di un *database*. In questo caso, tuttavia, l'illiceità del procedimento e quindi il suo divieto colpisce sistemi di IA che creano o allargano, se già esistenti, *database* di riconoscimento facciale, attraverso uno *scraping* massivo di immagini facciali che non è “mirato”, ossia avviene senza un obiettivo specifico o senza filtrare i contenuti in modo preciso e, quindi, non preordinato a finalità

⁽⁵¹⁾ Perché la formula legislativa nulla ci dice in realtà al riguardo.

⁽⁵²⁾ Parentesi mia.

determinate come anche la disciplina del GDPR invece prescrive⁽⁵³⁾, con immagini facciali prese da Internet o da telecamere a circuito chiuso, rendendo possibili le gravi violazioni ricordate (una grave invasione della sfera privata, incluso il furto di identità facciale, già finora a dire il vero perpetrato, in alcuni casi esemplificativi nazionali, di recente impatto, oggetto di discussione).

Sembra utile riportare alcuni casi chiarificatori tra i numerosi segnalati *sub* lettera *e*) delle Linee guida della Commissione prima citate: una società di *software* di riconoscimento facciale raccoglie immagini di volti. Le fotografie detenute dalla società sono raccolte dai *social media* (ad esempio Facebook, YouTube, Twitter, Venmo) con uno “*scraper* automatico di immagini” che cerca su Internet e rileva immagini contenenti volti umani. Raccoglie queste immagini con tutte le informazioni associate (come la fonte dell’immagine (URL), la geolocalizzazione e talvolta i nomi delle persone). I tratti del volto vengono quindi estratti dalle immagini e trasformati in rappresentazioni algoritmiche, che vengono sottoposte a *hash* per l’indicizzazione e il confronto futuro. Quando un utente carica l’immagine di un individuo sul sistema di intelligenza artificiale, quest’ultimo determina se l’immagine corrisponde a un volto presente nel *database*. L’immagine caricata subirà la stessa trasformazione algoritmica delle immagini “raschiate” in modo automatico⁽⁵⁴⁾.

Nella prospettiva di tutela della *privacy* delle persone fisiche, può essere opportuno ricordare che il Garante della *privacy* è già intervenuto per sanzionare fenomeni di *scraping* illecito. A titolo indicativo, si può ricordare il noto caso della società statunitense Clearview AI⁽⁵⁵⁾, che ha sviluppato un motore di ricerca per il riconoscimento facciale che, tra-

⁽⁵³⁾ V. art. 5, par. 1, lett. *b*) e *c*), GDPR, relative, rispettivamente al principio di finalità e di minimizzazione dei dati. Che il GDPR si applichi anche alle attività di *web scraping* è chiaro e lo si evince, tra l’altro, dall’art. 3, par. 2, lett. *b*), GDPR che dispone: «Il presente regolamento si applica al trattamento dei dati personali di interessati che si trovano nell’Unione, effettuato da un titolare del trattamento o da un responsabile del trattamento che non è stabilito nell’Unione, quando le attività di trattamento riguardano: (...) il monitoraggio del loro comportamento nella misura in cui tale comportamento ha luogo all’interno dell’Unione».

⁽⁵⁴⁾ V. EU Commission, *op. cit.*, p. 79.

⁽⁵⁵⁾ V. GPDP, 10 febbraio 2022, doc. *web* n. 9751362. Cfr. anche GPDP, 17 maggio 2023, doc. *web* n. 9903067 sull’attività di *web scraping* effettuata dalla piattaforma Trovanumeri. Nel 2023, il Garante *privacy* ha poi avviato un’indagine conoscitiva sul fenomeno (se ne dà conto nel Prov. del GPDP, 21 dicembre 2023, doc. *web* n. 9972593) che si è conclusa con la pubblicazione della Nota informativa «in materia di *web scraping* per finalità di addestramento di intelligenza artificiale generativa e di possibili azioni di contrasto a

mite *web scraping*, raccoglie immagini da fonti pubbliche *online* e le elabora con algoritmi biometrici, trasformandole in rappresentazioni numeriche indicizzate per facilitarne la ricerca. La società non si limita, dunque, a raccogliere immagini, ma le trasforma in dati biometrici altamente strutturati, offrendo così un servizio di riconoscimento facciale avanzato. Il Garante italiano ha, in proposito, rilevato plurime violazioni della disciplina del GDPR in Italia e ha irrogato, a carico della società, la sanzione amministrativa massima di 20 milioni di euro. Fra le tante violazioni, come già convenuto, spiccano quelle di più immediato riconoscimento. Infatti, come è dato leggere nella decisione del Garante, «la raccolta di dati personali liberamente disponibili in Internet mediante tecniche di *web scraping* costituisce un trattamento di dati personali, che deve trovare legittimazione in una delle basi giuridiche previste dall'art. 6 del Regolamento. (...) Si ritiene che l'interesse legittimo della Società alla libera iniziativa economica non possa che flettere rispetto ai diritti e alle libertà degli interessati, in particolare alla grave messa in pericolo del diritto alla riservatezza, al divieto di essere sottoposti a trattamenti automatizzati e al principio di non discriminazione insiti in un trattamento di dati personali come quello effettuato dalla Società. In conclusione, si ritiene che Clearview non possa vantare alcuna valida base giuridica su cui fondare la liceità del trattamento di dati personali posto in essere». E «con riferimento al trattamento di dati posto in essere da Clearview, non solo deve ritenersi insussistente alcuna valida base giuridica di cui all'art. 6 del Regolamento, ma risulta violato anche il divieto generale di trattamento di categorie particolari di dati (con riferimento ai dati biometrici)», *ex art.* 9.

La pratica di IA vietata dalla lett. f), par. 1, comma 1, di cui all'art. 5, consiste nell'immissione sul mercato, messa in servizio per tale finalità specifica o uso di sistemi di IA per inferire le emozioni di una persona fisica nell'ambito di luogo di lavoro e degli istituti di istruzione. Stabilisce in proposito il *considerando* n. 44): «Sussistono serie preoccupazioni in merito alla base scientifica dei sistemi di IA diretti ad identificare o inferire emozioni, in particolare perché l'espressione delle emozioni varia notevolmente in base alle culture e alle situazioni e persino in relazione a una stessa persona. Tra le principali carenze di tali sistemi figurano la limitata affidabilità, la mancanza di specificità e la generalizzabilità che rimane

tutela dei dati personali» (si tratta del Provv. del GPDP, 20 maggio 2024, doc. *web* n. 10020316).

davvero limitata. Pertanto, i sistemi di IA che identificano o inferiscono emozioni o intenzioni di persone fisiche sulla base dei loro dati biometrici possono portare a risultati discriminatori e possono essere invasivi dei diritti e delle libertà delle persone interessate. Considerando lo squilibrio di potere nel contesto del lavoro o dell'istruzione, combinato con la natura invasiva di tali sistemi, questi ultimi potrebbero (*rectius*: possono)⁽⁵⁶⁾ determinare un trattamento pregiudizievole o sfavorevole di talune persone fisiche o di interi gruppi di persone fisiche. È pertanto opportuno vietare l'immissione sul mercato, la messa in servizio o l'uso di sistemi di IA destinati a essere utilizzati per rilevare lo stato emotivo delle persone in situazioni relative al luogo di lavoro e all'istruzione (...)». La parte finale della disposizione interessata esclude tuttavia dal divieto i sistemi di IA immessi sul mercato per la finalità qui in considerazione, per ragioni esclusivamente terapeutiche o di sicurezza.

Nelle Linee guida della Commissione europea sono riportati alcuni esempi di pratiche vietate e di pratiche ammesse nel luogo di lavoro e presso gli istituti di istruzione. Tra gli altri, si segnala, l'uso di sistemi di intelligenza artificiale per il riconoscimento delle emozioni durante i colloqui finalizzati all'assunzione che è senz'altro vietato, così come è proibito l'uso di telecamere da parte di un supermercato, al fine di monitorare le emozioni dei dipendenti. Al contrario, è consentito l'impiego di telecamere da parte di una banca per individuare clienti sospetti, ad esempio al fine di stabilire che qualcuno è sul punto di commettere una rapina. Ciò, purché si garantisca che nessun dipendente venga monitorato e vi siano sufficienti garanzie. Nel contesto degli istituti di istruzione, si riporta l'ipotesi dell'utilizzo di un sistema di IA per il riconoscimento delle emozioni, allo scopo di dedurre l'interesse e l'attenzione degli studenti. Tale utilizzo è vietato, così come è vietato l'utilizzo di sistemi di questo tipo durante i *test* di ammissibilità per i nuovi studenti. Per contro, un istituto di istruzione potrebbe utilizzare un *software* di tracciamento oculare basato sull'intelligenza artificiale durante un esame *online* degli studenti, per tracciare il punto di fissazione e il movimento degli occhi. Tale fattispecie non è vietata, perché il sistema di IA qui non identifica, né deduce le emozioni⁽⁵⁷⁾. Possiamo anche pensare, con riferimento ad un istituto scolastico, ove magari un anno prima uno studente ha compiuto un massacro, che

⁽⁵⁶⁾ Parentesi mia.

⁽⁵⁷⁾ V. EU Commission, *op. cit.*, pp. 86-87.

nell'anniversario, a fini di sicurezza, si scrutinino le emozioni di studenti sospetti, per evitare reati di emulazione.

La pratica di IA vietata *sub* lett. g), par. 1, comma 1, art. 5 concerne immissione sul mercato, messa in servizio per tale finalità specifica, o uso di sistemi di categorizzazione biometrica volti alla classificazione individuale di persone fisiche, sulla base di dati biometrici, come, ad esempio, impronte digitali, scansioni dell'iride o della retina al fine di trarne deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose, o filosofiche, vita o orientamento sessuale; il divieto non concerne tuttavia etichettatura o filtraggio di *set* di dati biometrici acquisiti legalmente⁽⁵⁸⁾, come le immagini sulla base di dati biometrici o della categorizzazione di dati biometrici nel settore delle attività di contrasto. E la chiarezza significativa del precetto risulta facilitata da alcune delle definizioni di cui all'art. 3, come ad esempio quella di "sistemi di categorizzazione biometrica", di cui al n. 40)⁽⁵⁹⁾.

Può essere ricordato nuovamente che i dati biometrici sono stati originariamente definiti nel GDPR, come la stessa definizione è ripresa anche in questo Regolamento, *ex* art. 3, n. 34)⁽⁶⁰⁾, come dati personali che, ottenuti attraverso trattamento tecnico specifico, ricomprendono dati personali che riguardano non solo caratteristiche fisiche o fisiologiche come negli esempi di cui sopra, e come forse è più immediato pensare, ma anche caratteristiche comportamentali come le particolari andature del passo di un soggetto, o di sua battitura al computer, che sono finalizzati all'identificazione della persona fisica interessata ed è opportuno rammentare che il

(58) Il termine "*set*" indica un insieme strutturato di dati, ossia una collezione di elementi (in questo caso dati biometrici) organizzati secondo specifici criteri. Si veda, in proposito, la definizione di «*data set*» riportata nel Dizionario Treccani *online*: «In informatica, insieme di dati organizzati in forma relazionale. Ha una struttura tabellare, dove di solito ogni colonna rappresenta una variabile e ogni riga corrisponde a una osservazione».

(59) Si tratta di «un sistema di IA che utilizza i dati biometrici di persone fisiche al fine di assegnarle a categorie specifiche, a meno che non sia accessorio a un altro servizio commerciale e strettamente necessario per ragioni tecniche oggettive».

(60) V. art. 4, par. 1, n. 14), GDPR: «"dati biometrici": 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». Diverge solo marginalmente la definizione di cui all'art. 3, par. 1, n. 34), AIA: «"dati biometrici": i dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica, quali le immagini facciali o i dati dattiloscopici»; a cui si aggiunge quella di «identificazione biometrica», riportata nel n. 35: «il riconoscimento automatizzato delle caratteristiche umane fisiche, fisiologiche, comportamentali o psicologiche allo scopo di determinare l'identità di una persona fisica confrontando i suoi dati biometrici con quelli di individui memorizzati in una banca dati».

trattamento di dati biometrici in quanto dati sensibili è soggetto alla disciplina particolare di cui all'art. 9 del GDPR⁽⁶¹⁾. Del resto, anche una delle definizioni di questo Regolamento, *ex art. 3, n. 37*), richiama "categorie particolari di dati personali", ossia quelle oggetto dell'art. 9, par. 1, del GDPR. In altri termini, l'art. 9 del GDPR stabilisce un generale divieto di trattamento di dati biometrici, e prevedendo al riguardo specifiche eccezioni. I dati biometrici sono strumentali e funzionali all'identificazione delle persone fisiche e, in relazione alla tutela della *privacy* nell'ottica del GDPR, risultano emanati in tema sia provvedimenti da parte dell'Autorità Garante della *privacy*, sia decisioni della Cassazione, con esito da parte di ciascuno dei due originariamente non sempre perfettamente coincidente, ma dove il Garante si è poi successivamente allineato all'orientamento della Cassazione.

È accaduto, ad esempio, con riferimento a casi di trattamenti di dati biometrici in ambito lavorativo, laddove il Garante *privacy* ha adottato un approccio particolarmente restrittivo, vietando in modo assoluto l'uso di sistemi di riconoscimento facciale per il controllo delle presenze⁽⁶²⁾; mentre la Corte di Cassazione, in una sua significativa pronuncia, ha dichiarato illegittimo un sistema di rilevazione biometrica di presenza, tramite impronta digitale, atteso che il consenso espresso dai lavoratori non poteva, nella specie, reputarsi specifico (e quindi ammettendolo, implicitamente, nel caso contrario di consenso specifico, espresso dai lavoratori)⁽⁶³⁾. Si segnala, altresì, il recente caso relativo all'impiego di un sistema di supervisione da remoto, nell'ambito dello svolgimento delle prove scritte d'esame di studenti universitari, al fine di identificare questi ultimi o di verificarne il corretto comportamento durante lo svolgimento della prova d'esame in video conferenza. Qui le posizioni di Garante e Cassazione sono apparse allineate⁽⁶⁴⁾, offrendo un'interpretazione estensiva della definizio-

⁽⁶¹⁾ L'articolo 9 del GDPR stabilisce, infatti, il divieto di trattare categorie particolari di dati personali, ossia quelli che rivelino l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, l'appartenenza sindacale, i dati genetici, nonché i dati biometrici intesi a identificare in modo univoco una persona fisica, quelli sulla salute, sulla vita sessuale o sull'orientamento sessuale della persona. In dottrina, v., per tutti, *Dati personali. Protezione, libera circolazione e governance*, vol. 1. *Principi*, a cura di Bravo, Pisa, 2023, *passim*.

⁽⁶²⁾ V. GDPD, 22 febbraio 2024, doc. *web* nn. 9995680, 9995701, 9995741, 9995762, 9995785.

⁽⁶³⁾ V. Cass., 19 maggio 2023, n. 13873, in *Diritto & Giustizia* 2023, 22 maggio.

⁽⁶⁴⁾ V., rispettivamente, GDPD, 16 settembre 2021, doc. *web* n. 9703988 e Cass., 13 maggio 2024, n. 12967, in *Diritto & Giustizia* 2024, 14 maggio.

ne di trattamento di dati biometrici, divergente da quella sostenuta dal Tribunale di Milano in appello al provvedimento del Garante⁽⁶⁵⁾.

Il divieto nei termini indicati dalla lettera g) di cui alla disposizione interessata può dar luogo a gravi discriminazioni o può di per sé risultare discriminatorio, in palese contrasto con il principio di uguaglianza. La parte finale lascia la possibilità di trattare attraverso etichettatura o filtraggio di *set* di dati biometrici acquisiti legalmente, come le immagini sulla base di dati biometrici o della categorizzazione di dati biometrici nel settore delle attività di contrasto⁽⁶⁶⁾.

Nelle Linee guida, tra i sistemi di IA vietati *sub* art. 5, lettera g), è riportato il caso di un sistema di categorizzazione biometrica che promette di poter dedurre la razza di un individuo dalla sua voce, ovvero quello di un sistema di categorizzazione biometrica che permetta di dedurre l'orientamento religioso di un individuo dai suoi tatuaggi o dal suo volto⁽⁶⁷⁾. Al contempo, sono riportati alcuni esempi chiarificatori di etichettatura o filtraggio consentiti, quali l'etichettatura dei dati biometrici per evitare che un membro di un gruppo etnico abbia minori possibilità di essere invitato a un colloquio di lavoro, perché l'algoritmo è stato addestrato sulla base di dati in cui quel particolare gruppo ha ottenuto risultati peggiori, ossia ha ottenuto risultati peggiori rispetto ad altri gruppi; nonché la categorizzazione dei pazienti in base al colore della pelle o degli occhi che può essere importante per la diagnosi medica, ad esempio per la diagnosi del cancro⁽⁶⁸⁾.

Il «settore delle attività di contrasto» riguarda, come abbiamo già detto, la sfera penale e la sfera dell'ordine pubblico correlata. Le Linee guida riportano, in proposito, l'esempio dell'utilizzo da parte di un'autorità preposta all'applicazione della legge di un sistema di IA che consenta di etichettare e filtrare una serie di dati sospettati di contenere materiale pedopornografico. In una prima fase, le forze dell'ordine si avvarrebbero del supporto dei sistemi di intelligenza artificiale per individuare e cancellare i dati sensibili dalle immagini. Inoltre, il filtraggio delle etichettature in base al sesso, all'età, a dati biometrici, come il colore degli occhi e dei capelli, alle cicatrici e altri segni potrebbe aiutare ad identificare le vittime

⁽⁶⁵⁾ V. Trib. Milano, 20 ottobre 2022, in *DeJure*.

⁽⁶⁶⁾ E in aiuto, per chiarire il significato che ha per oggetto l'esclusione dal divieto, sovengono alcune delle definizioni contenute nell'art. 3, come quella di «sistema di categorizzazione di dati biometrici», di cui n. 40), già ricordata, e quella di «attività di contrasto», *sub* n. 46).

⁽⁶⁷⁾ V. EU Commission, *op. cit.*, pp. 93-94.

⁽⁶⁸⁾ *Ibidem*, p. 94.

o a creare collegamenti con altri casi. Allo stesso modo, è consentito filtrare ed etichettare le mani degli abusatori in base a caratteristiche specifiche, come la lunghezza delle dita o qualsiasi altro segno distintivo o tatuaggio per aiutare ad identificare possibili persone sospette⁽⁶⁹⁾.

4. – Il divieto di pratiche di sistemi di IA di cui alla lettera *b*), par. 1, comma 1, *ex art* 5, si riferisce all'uso e quest'ultimo è di per sé sufficiente (non occorrendo una immissione sul mercato, né una messa in servizio come invece nelle lettere precedenti), di sistemi di identificazione biometrica remota in tempo reale, in spazi accessibili al pubblico a fini di attività di contrasto a meno che e nella misura in cui tale uso sia strettamente necessario per uno degli obiettivi seguenti⁽⁷⁰⁾:

i) la ricerca mirata di specifiche vittime di sottrazione, di tratta di esseri umani, o di sfruttamento sessuale di esseri umani, nonché la ricerca di persone scomparse;

ii) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di una minaccia reale e attuale, o reale e prevedibile di un attacco terroristico;

iii) la localizzazione o l'identificazione di una persona sospettata di aver commesso un reato ai fini dello svolgimento di un'indagine penale o dell'esercizio di un'azione penale o l'esecuzione di una sanzione penale per i reati di cui all'Allegato II⁽⁷¹⁾, punibile nello stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno 4 anni.

La norma pone, dunque, all'interno dell'ampia categoria dei sistemi di IA volti all'identificazione biometrica, un divieto per le pratiche di sistemi di IA di identificazione biometrica remota in tempo reale, per fini di attività di contrasto. Il divieto, tuttavia, non è un divieto assoluto, perché

⁽⁶⁹⁾ *Ibidem*, p. 95.

⁽⁷⁰⁾ E per chiarire con più certezza il significato della norma, sovengono alcune definizioni di cui all'art. 3, quali, ad esempio, quella *sub* n. 42) di «sistema di identificazione biometrica remota in tempo reale»; quella *sub* n. 44) di «spazio accessibile al pubblico»; quella *sub* n. 45) di «autorità di contrasto»; e quella *sub* n. 46) di «attività di contrasto».

⁽⁷¹⁾ Si tratta, in particolare, dei seguenti reati: terrorismo, tratta di esseri umani, sfruttamento sessuale di minori e pornografia minorile, traffico illecito di stupefacenti o sostanze psicotrope, traffico illecito di armi, munizioni ed esplosivi, omicidio volontario, lesioni gravi, traffico illecito di organi e tessuti umani, traffico illecito di materie nucleari e radioattive, sequestro, detenzione illegale e presa di ostaggi, reati che rientrano nella competenza giurisdizionale della Corte penale internazionale, illecita cattura di aeromobile o nave, violenza sessuale, reato ambientale, rapina organizzata o a mano armata, sabotaggio, partecipazione a un'organizzazione criminale coinvolta in uno o più dei reati elencati sopra.

per il raggiungimento dello scopo indicato dalla legge, rispettivamente *sub* punti *i*)-*iii*), se strettamente necessario, l'uso di sistemi di IA di identificazione biometrica remota in tempo reale è permesso.

La stessa disposizione sancisce espressamente che la lettera *b*) «lascia impregiudicato» l'art. 9 del GDPR, in tema, rispettivamente, di divieto di trattamento e di eccezioni al divieto stesso per i dati biometrici diretto a finalità diverse dall'attività di contrasto.

Stabilisce il *considerando* 32), nell'illustrare la ragione giustificatrice del divieto: «l'uso di sistemi di IA di identificazione biometrica remota in tempo reale delle persone fisiche in spazi accessibili al pubblico a fini di attività di contrasto è particolarmente invasivo dei diritti e delle libertà delle persone interessate, nella misura in cui potrebbe avere ripercussioni sulla vita privata di un'ampia fetta della popolazione, farla sentire costantemente sotto sorveglianza e scoraggiare in maniera indiretta l'esercizio della libertà di riunione e di altri diritti fondamentali (...)». Le attività di contrasto a cui si riferisce il divieto interessano lo svolgimento di quelle attività cui è di regola preposta la pubblica Autorità per combattere violazioni delle norme di carattere penale. E si continua nel *considerando* 32): «Le inesattezze di carattere tecnico dei sistemi di IA destinati all'identificazione biometrica remota delle persone fisiche possono determinare risultati distorti e comportare effetti discriminatori. Tali possibili risultati distorti ed effetti discriminatori sono particolarmente importanti per quanto riguarda l'età, l'etnia, la razza, il sesso o le disabilità. L'immediatezza dell'impatto e le limitate opportunità di eseguire ulteriori controlli o apportare correzioni in relazione all'uso di tali sistemi che operano "in tempo reale" comportano inoltre un aumento dei rischi per quanto concerne i diritti e le libertà delle persone interessate nell'ambito delle attività di contrasto o che sono da queste condizionate». Per circoscrivere e comprendere meglio la portata del divieto qui in considerazione è opportuno tenere presente alcune delle definizioni di cui all'art. 3 dell'AI Act, qui significative. Più in particolare, si tratta delle definizioni che chiariscono, in riferimento ai sistemi di IA, alcuni concetti relativi all'uso (o uso proprio ed improprio), come ad esempio *sub* 14) ⁽⁷²⁾, *sub* n. 34) «dati biometrici»: «i dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali, di una persona fisica quali le immagini facciali o i dati dattiloscopici» ⁽⁷³⁾; va però ricordato, che

⁽⁷²⁾ V. *supra*, nella Premessa.

⁽⁷³⁾ I dati dattiloscopici sono le informazioni biometriche relative alle impronte digitali di una persona che vengono raccolti, analizzati e utilizzati principalmente per l'identifica-

nella definizione sempre più rilevante di «identificazione biometrica» di cui *sub* n. 35) vengono ricompresi anche i dati psicologici; ma qui ne importano soprattutto due: la definizione di «identificazione biometrica remota in tempo reale» di cui al n. 42) (dell'art. 3), secondo cui si tratta di «un sistema di identificazione biometrica remota in cui il rilevamento dei dati biometrici, il confronto e l'identificazione avvengono senza ritardi significativi, il quale comprende non solo le identificazioni istantanee, ma anche quelli che avvengono con brevi ritardi al fine di evitare l'elusione»⁽⁷⁴⁾; la definizione non precisa il tempo in termini di minuti od ore in cui l'identificazione avviene rispetto al rilevamento, sancendo solo che essa avvenga senza ritardi significativi e ciò preclude che si tratti di settimane perché queste rientrano di sicuro nel sistema di identificazione biometrica remota a posteriori che anche per finalità di contrasto è sempre permessa; e la definizione di «spazio accessibile al pubblico», di cui *sub* n. 44) vale a dire «qualsiasi luogo fisico di proprietà pubblica o privata accessibile ad un numero indeterminato di persone fisiche, indipendentemente dal fatto che possano applicarsi determinate condizioni di accesso e indipendentemente dalle potenziali restrizioni di capacità»; e, ancora, la definizione di «attività di contrasto», di cui *sub* n. 46), vale a dire «attività svolte dalle autorità di contrasto o per loro conto a fini di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse»; si tratta dunque di definizioni rilevanti perché entrambe integrano requisiti della norma di divieto in oggetto e quindi ne delimitano l'ambito di applicazione.

La parte finale della disposizione interessata consente l'uso di sistemi di IA volti all'identificazione biometrica remota in tempo reale, nei tre casi indicati nella lettera *b*) da *i*) a *iii*), che siano consentiti attraverso autorizzazione dell'Autorità nazionale, osservando le condizioni di tutela stabilite dai paragrafi da 2) a 7) dell'art. 5.

Per quanto riguarda in sé il divieto di cui alla presente lettera *b*), complessivamente si richiedono cumulativamente 5 requisiti: 1) un sistema di IA che sia un sistema di identificazione biometrica remota; 2) il divieto

zione e la verifica dell'identità individuale in ambiti forensi, di sicurezza e amministrativi. Si veda, in proposito, la definizione di «dattiloscopia» riportata nel Dizionario Treccani *online*: è la disciplina che si occupa del rilevamento e dello studio delle creste cutanee papillari presenti sui polpastrelli delle dita, comunemente note come impronte digitali. Queste creste formano disegni unici per ogni individuo e rimangono invariati per tutta la vita, rendendo le impronte digitali uno strumento affidabile per l'identificazione personale.

⁽⁷⁴⁾ In antitesi, vale la definizione di cui all'art. 3, par. 1, n. 43).

come già detto concerne l'uso, qualsiasi utilizzazione o impiego ma qui, a differenza che nelle precedenti lettere, non si richiede né immissione sul mercato, né messa in servizio; 3) il sistema di IA effettua l'identificazione biometrica remota in tempo reale; 4) in spazi accessibili al pubblico; 5) per gli scopi previsti dalla legge.

Una chiarificazione in punto. Il divieto riguarda la finalità dell'attività di contrasto, secondo la formula legislativa. Sono le pratiche di identificazione biometrica remota in tempo reale di cui ai punti *i-iii*) che avvengono ciascuna per lo scopo specifico indicato dalla legge, dove lo scopo previsto dalla legge e dal Regolamento serve non già a giustificare il divieto ma a giustificare l'eccezione al divieto, cioè la possibilità di utilizzare il sistema di identificazione biometrica remota in tempo reale corrispondente.

In riferimento ai sistemi di identificazione biometrica da remoto è innanzitutto e a titolo di premessa utile ricordare quanto contemplato dal *considerando* 15), in riferimento alla nozione di «identificazione biometrica». Qui, dopo aver specificato che essa consiste nel riconoscimento automatico di caratteristiche fisiche, fisiologiche o comportamentali quali il volto, il movimento degli occhi, la voce, l'odore, la pressione esercitata sui tasti, ecc., «allo scopo di determinare l'identità della persona interessata attraverso il confronto dei suoi dati con quelli di altri individui memorizzati in una banca dati di riferimento (...) indipendentemente dal fatto che la persona abbia fornito il suo consenso», aggiunge che sono esclusi dal concetto di identificazione biometrica «i sistemi di IA destinati ad essere impiegati per la verifica biometrica, che include l'autenticazione, la cui *unica* finalità è confermare che una data persona fisica è la persona che dice di essere (come, per esempio, il confronto fra il volto di un viaggiatore scansionato ad un *e-gate* e l'immagine del volto contenuta nel passaporto) ⁽⁷⁵⁾ e/o confermare l'identità di una persona fisica al solo scopo di accedere ad un servizio, sbloccare un dispositivo o disporre dell'accesso di sicurezza a locali» (come, per esempio, l'accesso ai locali di una centrale elettrica attraverso la tecnologia di scansione facciale) ⁽⁷⁶⁾. Nel requisito n. 1 che il sistema di IA sia un sistema di identificazione biometrica remota sovviene la corrispondente definizione di cui *sub* art. 3, n. 41), secondo cui è tale «un sistema di IA finalizzato all'identificazione di persone fisiche, senza il loro coinvolgimento attivo tipicamente a distanza mediante il confronto di dati biometrici di una persona con i dati biome-

⁽⁷⁵⁾ Parentesi mia.

⁽⁷⁶⁾ Parentesi mia.

trici contenuti in una banca dati di riferimento». L'ambito in considerazione implica dunque l'assenza di un "attivo coinvolgimento della persona interessata" e l'effetto di "catturare" le caratteristiche strumentali all'identificazione "a distanza" e dove, per dar luogo all'identificazione, occorre anche la comparazione dei dati biometrici ottenuti a distanza con i dati contenuti in una banca dati, come ad esempio un *criminal database* contenente le immagini facciali, o profili o modelli di sospetti. Con riferimento all'assenza di un attivo coinvolgimento della persona interessata, viene specificato nelle Linee guida che non è necessario che la persona sia presente e sia informata, ad esempio, della presenza di telecamere, ma occorre che essa attivamente e consapevolmente si posizioni di fronte alle telecamere con una partecipazione, quindi, attiva.

Ricadono, invece, nella definizione di identificazione biometrica, i sistemi di IA usati per "seguire" una persona fisica, ad esempio per vedere che direzione ha preso un sospetto che scappi; ciò non è contraddetto dalle citate Linee guida che riportano l'esempio di un caso nel concreto in realtà sottratto al divieto di cui *sub iii*) in cui la polizia debba localizzare un sospettato di un grave attacco terroristico: qui, proprio a fronte dell'avvenuto attacco terroristico, l'uso delle tecnologie di riconoscimento facciale in tempo reale può essere consentito, sempre purché conforme all'art. 5, par. 1, comma 1, lettera *h*).

Secondo il requisito indicato *sub 3*) di cui al divieto *ex* lettera *h*) che si riferisce ai sistemi di IA (per l'identificazione biometrica) "a distanza" non rientrano nemmeno i casi di cui abbiamo già parlato, sottratti alla nozione di identificazione biometrica *sub 1*), diretti al solo scopo di verifica o autenticazione biometrica, perché qui è vero ad esempio che, per accedere al locale della cabina per l'energia elettrica, il confronto avviene con l'immagine del volto scansionata all'ingresso che riconosce il soggetto che presenta la corrispondente tessera di riconoscimento, ma la persona è anche presente e può essere coinvolta attivamente eventualmente per confermare di essere proprio il soggetto raffigurato nella tessera, al momento di accedere al locale interessato per lo svolgimento del corrispondente servizio. E ciò finisce per sminuire l'impatto dell'uso del sistema di IA, come sistema a distanza. Si stabilisce in proposito al *considerando 17*): «i sistemi di identificazione biometrica remota a distanza sono generalmente utilizzati per percepire più persone o il loro comportamento simultaneamente al fine di facilitare in modo significativo l'identificazione di persone fisiche senza il loro coinvolgimento attivo. Sono esclusi i sistemi di IA destinati ad essere impiegati per la verifica biometrica per confermare l'identità di una persona al solo scopo di accedere ad un servizio (sbloccare

un dispositivo o disporre dell'accesso di sicurezza a locali). Una tale esclusione è giustificata dal fatto che detti sistemi hanno un impatto minore sui diritti fondamentali delle persone coinvolte rispetto ai sistemi di identificazione biometrica remota, che possono invece essere utilizzati per il trattamento dei dati biometrici di un numero elevato di persone senza il loro coinvolgimento attivo (...). Secondo gli esempi in proposito indicati nelle Linee guida, rientrano nel divieto di cui alla lettera *b*) in quanto a distanza, i sistemi di IBR che vengono utilizzati nelle telecamere installate sulle pareti o sul soffitto delle stazioni della metropolitana a scopo di sorveglianza. Viceversa, sono esclusi proprio perché non soddisfano il requisito attinente a costituire un sistema da remoto (cioè a distanza), i sistemi di IA per dare accesso alla stazione della metropolitana come i biglietti biometrici, perché qui le persone sono attivamente coinvolte e si avvicinano (consapevolmente) al sensore biometrico per ottenere l'accesso. In realtà, questi stessi esempi possono valere, rispettivamente, secondo me, anche per giustificare l'inclusione nel divieto di cui alla lettera *b*) (o l'esclusione dal medesimo) con riferimento alla circostanza che il sistema di identificazione biometrica remota sia usato in tempo "reale" o meno, sempre secondo il numero 3) qui (già) in oggetto. È infatti proprio a quest'ultimo riguardo che il *considerando* n. 17), nella parte finale a seguire di quella poc'anzi già riportata nel testo, stabilisce: «I sistemi di IA "in tempo reale" comportano l'uso di materiale "dal vivo" o "quasi dal vivo" come ad esempio filmati, generati da una telecamera, o da altro dispositivo con funzioni analoghe. Nel caso di identificazione biometrica "a posteriori" invece (di cui alla corrispondente definizione *ex art.* 3, n. 43)⁽⁷⁷⁾, i dati biometrici sono già stati rilevati e il confronto o l'identificazione avvengono solo con un ritardo significativo, ad esempio uno o due giorni. Si tratta, in quest'ultimo caso, di materiale, come immagini o filmati generati da telecamere a circuito chiuso o da dispositivi privati, che è stato generato prima che il sistema fosse usato in relazione alle persone fisiche interessate».

Per la condizione indicata *sub* 4, rispetto al divieto di cui alla lett. *b*) ossia che i sistemi di IA siano usati per l'identificazione biometrica remota in tempo reale, in "spazi accessibili al pubblico" sovviene, come già accennato, la corrispondente definizione di cui all'art. 3, n. 44), già ricordata, ove per «spazio accessibile al pubblico» si intende: «qualsiasi luogo fisico di proprietà pubblica o privata accessibile ad un numero indeterminato di

⁽⁷⁷⁾ Parentesi mia.

persone, fisiche, indipendentemente dal fatto che possano applicarsi determinate condizioni di accesso e indipendentemente dalle potenziali restrizioni di capacità». Anche qui indicazioni esemplificative possono essere tratte dagli esempi già fatti in riferimento ad altri requisiti di cui alla lettera *b*) (come le telecamere a circuito chiuso installate per l'uso di identificazione biometrica nella stazione ferroviaria principale di una città, o sistemi di identificazione biometrica remota che vengono impiegati nelle telecamere installate sulle pareti della metropolitana a scopo di sorveglianza, o ancora, per fare altri esempi, sistemi di IBR usati in una piazza, o in un parco ove si è programmato un certo evento, come un festival, un concerto, ecc., per finalità di contrasto).

Secondo le Linee guida⁽⁷⁸⁾, non sono comprese nel divieto quelle pratiche poste in essere da sistemi di IA che non soddisfino tutti e ciascuno dei requisiti indicati. In altri termini, è sufficiente che un sistema di IA non soddisfi uno dei requisiti costitutivi del divieto, perché, come risulta dagli esempi fatti nel testo, il sistema di IA non rientri nel divieto di cui alla lettera *b*) dell'art. 5.

Uno dei requisiti più significativi richiesti per il divieto di sistema di IA indicati *sub* lett. *b*) attiene allo scopo perseguito dai sistemi di IA di IBR qui in considerazione, ossia il suo essere preordinato a finalità di contrasto, o, come invece detto nelle Linee guida, per lo scopo previsto dalla legge. Ribadiamo quanto detto sopra. Lo scopo per cui è espresso il divieto è l'attività di contrasto, ossia, come detto, l'insieme di attività svolte dalle autorità di contrasto o per loro conto a fini di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse. Lo scopo previsto dalla legge di cui parlano le Linee guida è lo scopo specifico rispetto al quale sono consentite le pratiche di identificazione biometrica remota in tempo reale, come indicato *sub* *i*)-*iii*).

Per comprendere meglio tutte le accezioni di questa condizione va precisato che le stesse si devono innanzi tutto ricavare, *a contrario*, per esclusione, dall'uso consentito dalla legge di sistemi di IBR come quelli qui in oggetto, di cui alla lett. *b*), consentito nei tre casi non ricompresi nel divieto stesso *sub* lett. *b*) di cui, rispettivamente, ai punti *i*), *ii*), *iii*).

Alcune delucidazioni derivano da specifici *considerando* di cui al Regolamento sull'intelligenza artificiale. Per quanto riguarda la ragione giustificatrice del divieto espressa nel *considerando* n. 32) si è già parlato dove,

(78) V. EU Commission, *op. cit.*, p. 97.

alla fine del *considerando* stesso, veniva chiarito che il rischio di risultati distorti e di effetti discriminatori soprattutto per quanto riguarda l'età, l'etnia, la razza, il sesso, o le disabilità delle persone fisiche potevano acuirsi proprio per il fatto che si tratta di identificazione biometrica remota "in tempo reale", vale a dire in cui, da un lato, l'immediatezza dell'impatto, e d'altro lato, le limitate possibilità di eseguire ulteriori opportuni controlli possono comportare un sensibile aumento dei rischi di lesione dei diritti e delle libertà fondamentali delle persone interessate da questa specifica tipologia di attività di contrasto correlata ai sistemi di IA. Qui pare più opportuno soffermarsi sul *considerando* 33). Quest'ultimo prende in considerazione i tre casi definiti dal legislatore nei punti *i)* *ii)* e *iii)* della lettera *b)* e sottratti al corrispondente divieto in quanto l'uso del sistema di IA per l'identificazione biometrica remota in tempo reale sia "strettamente necessario" al risultato perseguito in ognuno dei tre punti indicati ossia per tutelare "(...) un interesse pubblico rilevante"; in quest'evenienza viene infatti sottolineata l'importanza dell'interesse che giustifica il mancato operare del divieto, che prevale sui rischi, secondo il bilanciamento di interessi in conflitto operato dal medesimo legislatore "anche statale".

E si prosegue: «Tali situazioni comprendono la ricerca di determinate vittime di reato, comprese le persone scomparse, determinate minacce per la vita o l'incolumità fisica delle persone o un attacco terroristico nonché la localizzazione o l'identificazione degli autori o dei sospettati di reati elencati in allegato del presente Regolamento, qualora tali reati siano punibili nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà personale della durata massima di almeno quattro anni e sono definiti conformemente al diritto di tale Stato membro (...)». In altre parole, la soglia indicata per la pena o la misura di sicurezza privativa della libertà personale, secondo quanto stabilito dal diritto nazionale, «(...) contribuisce a garantire che il reato sia sufficientemente grave da giustificare potenzialmente l'uso di sistemi di identificazione biometrica remota in tempo reale». E si aggiunge: «Inoltre l'elenco dei reati di cui all'allegato considera 32 reati che sono stati elencati nella decisione quadro 2002/584/GAI del Consiglio⁽⁷⁹⁾, tenendo conto che alcuni di essi risultano più pertinenti di altri, poiché il grado di necessità e proporzionalità del ricorso all'identificazione biometrica remota in tem-

(79) V. Consiglio UE, *Decisione quadro del 13 giugno 2002, relativa al mandato d'arresto europeo e alle procedure di consegna tra Stati membri – Dichiarazioni di alcuni Stati membri sull'adozione della decisione quadro*, 2002/584/GAI, Bruxelles, G.U.C.E., n. L 190 del 18 luglio 2002.

po reale potrebbe risultare prevedibilmente molto variabile per quanto concerne il perseguimento in concreto della localizzazione o dell'identificazione di un autore o sospettato dei vari reati elencati e con riguardo alle possibili differenze in termini di gravità, probabilità e portata del danno o delle eventuali conseguenze negative. Si afferma anche che una grave minaccia per la vita e l'incolumità delle persone potrebbe derivare da un grave danneggiamento della infrastruttura critica, quale definita dall'art. 2, punto 4, della direttiva UE 2022/2557⁽⁸⁰⁾ e dove un tale danneggiamento o distruzione della infrastruttura critica può comportare una minaccia "imminente" per la vita o l'integrità fisica di una persona anche in ragione di un grave danno alla fornitura di forniture di base alla popolazione o alla funzione essenziale dello Stato».

Si vedano ora più specificamente i casi *ex* punti *i*), *ii*) *iii*), della lettera *b*), ove, se strettamente necessario e con rispetto del principio di proporzionalità, l'uso di sistemi di IA diretti all'identificazione biometrica remota in tempo reale è consentito. Preliminarmente va puntualizzato che l'uso di sistemi di identificazione biometrica remota in tempo reale deve risultare strettamente necessario e proporzionato (ed al riguardo, si devono eventualmente indicare le misure alternative al sistema di IA volto all'IBR in tempo reale in spazi accessibili al pubblico, per finalità di contrasto, meno lesive per il loro impatto sui diritti fondamentali in discussione).

Il caso *i*), come già indicato, riguarda la ricerca mirata di specifiche vittime di sottrazione, tratta di esseri umani o sfruttamento sessuale di esseri umani, nonché la ricerca di persone scomparse. È naturalmente chiaro che la lista di reati che costituiscono l'eccezione al divieto di pratica di sistema di IA di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico deve intendersi solo come tassativa, non potendo essere previste altre figure criminose al riguardo. Nelle Linee guida è riportato l'esempio di un bambino che viene rapito e ci sono indicazioni concrete che il rapitore intende portarlo da un luogo all'altro in auto. Qui la polizia potrà utilizzare un sistema di rilevazione biometrica in tempo reale per la ricerca mirata del bambino, ma, a tal fine, dovrà definire

⁽⁸⁰⁾ V. Direttiva (UE) 2022/2557 del Parlamento europeo e del Consiglio del 14 dicembre 2022 relativa alla resilienza dei soggetti critici e che abroga la direttiva 2008/114/CE del Consiglio. Anche il presente Regolamento contiene una definizione di «infrastruttura critica», *ex* art. 3, n. 61), in cui viene operato espresso rinvio alla definizione indicata nella direttiva del '22. Le infrastrutture critiche sono, in sostanza, sistemi e servizi essenziali per la società (come energia, trasporti e sanità), in cui il malfunzionamento dei sistemi di IA potrebbe avere gravi conseguenze per la sicurezza e il benessere pubblico.

esattamente il perimetro di impiego e la durata di utilizzo di tale sistema di IA ⁽⁸¹⁾.

Il caso *ii*), come già detto, riguarda la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di una minaccia reale e attuale o reale e prevedibile di un attacco terroristico. Nelle Linee guida viene riportata l'ipotesi della polizia che viene informata che una persona si aggira in un parco alla ricerca di persone da aggredire con un coltello, mentre urla *slogan* violenti ed estremisti solitamente legati ad attacchi terroristici e a gruppi terroristici. In questa ipotesi, se lo Stato membro ha autorizzato l'uso di un sistema di rilevazione biometrica in tempo reale, conformemente al disposto dell'art. 5, lettera *b*), *ii*), AIA, le autorità di contrasto potranno utilizzare quel sistema di IA per identificare e localizzare la persona all'interno e nei dintorni del parco e, in questo modo, prevenire l'attacco terroristico, sempre purché siano soddisfatte le altre condizioni, di cui all'articolo 5, par. da 2 a 7 ⁽⁸²⁾.

Il caso *iii*) si riferisce, infine, lo si è detto, alla localizzazione o all'identificazione di una persona sospettata di aver commesso un reato, ai fini dello svolgimento di un'indagine penale, o dell'esercizio di un'azione penale o dell'esecuzione di una sanzione penale per i reati di cui all'allegato II, punibile nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno quattro anni. Le Linee guida quale esempio di uso consentito di sistemi di identificazione biometrica in tempo reale, *sub* art. 5, lettera *b*), *iii*), riportano il caso di un intervento di polizia successivo ad un grave attacco terroristico in un mercatino di Natale con 12 morti. Qui la polizia potrebbe utilizzare tecnologie di riconoscimento facciale in tempo reale per identificare l'aggressore e monitorare sue possibili vie di fuga, vagliando, tramite il sistema di IA, la vicina stazione ferroviaria e le stazioni di destinazione dei treni in partenza da lì poco dopo l'attacco ⁽⁸³⁾, proprio perché sia il divieto, sia in questo caso l'eccezione al divieto consente anche la localizzazione quale oggetto di identificazione biometrica remota.

Completano, come accennato, le condizioni d'uso dei sistemi di identificazione biometrica remota in tempo reale le condizioni di cui ai paragrafi da 2 a 7 dell'art. 5.

⁽⁸¹⁾ V. EU Commission, *op. cit.*, p. 106.

⁽⁸²⁾ *Ibidem*, p. 109.

⁽⁸³⁾ *Ibidem*, p. 111.

Il paragrafo 2 è suddiviso in due parti. La prima parte stabilisce che l'uso dei sistemi di IA in discussione è consentito solo per confermare l'identità della persona "specificamente" interessata tenendo conto: *a*) della natura della situazione che dà luogo al possibile uso, in particolare la gravità, la probabilità e l'entità del danno che sarebbe causato in caso di mancato uso del sistema; *b*) delle conseguenze dell'uso del sistema per i diritti e le libertà di tutte le persone interessate, in particolare la gravità, la probabilità, e l'entità di tali conseguenze. La disposizione di cui alla appena ricordata prima parte si riferisce alla circostanza che, qualora il sistema di IRB in tempo reale per attività di contrasto sia sottratto al divieto di cui alla lettera *b*), l'uso consentito, come già detto, diviene ammissibile a condizione che sia strettamente necessario e sia contestualmente conforme al principio di proporzionalità e venga operata a tal fine una valutazione di elementi come quelli appena indicati. Orbene, le condizioni a riguardo espresse nel caso sono le due appena riportate: la prima concernente la natura della situazione che rende ammissibile il possibile uso, ossia più in particolare relativa al danno conseguente al mancato utilizzo del sistema stesso, per quanto attiene a gravità (ad es. la perdita della vita o gravi lesioni all'incolumità fisica delle persone presenti nel parco dei mercatini di natale oggetto del sistema di IA), la probabilità del danno stesso (ossia la percentuale di possibilità correlate al verificarsi del danno) e la sua entità (ossia la misura di grandezza quantitativa dell'entità del danno); la seconda condizione d'uso del sistema di identificazione biometrica remota in tempo reale, di cui a questa prima parte del par. 2 dell'art. 5, valutata come adeguata rispetto alla stretta necessità e proporzionalità dell'ipotesi in discussione, considera le conseguenze d'uso del sistema nel loro possibile impatto per i diritti e le libertà di tutti coloro che sono stati non direttamente oggetto dell'IBR, ma comunque coinvolti nell'utilizzazione del sistema stesso. Chiarisce al riguardo il *considerando* n. 34): «Al fine di garantire che tali sistemi siano impiegati in modo responsabile e proporzionato, è altresì importante stabilire che in ciascuna delle situazioni elencate in modo esaustivo e definite rigorosamente, (*sub* punti *i*), *ii*) e *iii*)⁽⁸⁴⁾ è opportuno tener conto di taluni elementi, in particolare per quanto riguarda la natura della situazione all'origine della richiesta e le conseguenze dell'uso per i diritti e le libertà di tutte le persone interessate, nonché le tutele e le condizioni previste per l'uso».

(84) Parentesi mia.

La seconda parte della disposizione contenuta nel par. 2 dell'art. 5, stabilisce che l'uso di un sistema di IBR in tempo reale in spazi accessibili al pubblico per finalità di contrasto per uno qualsiasi degli obiettivi indicati al par. 1, comma 1, lett. *b*), *sub* punti *i*) *ii*) *iii*) deve rispettare la tutela e le condizioni d'uso necessarie e proporzionate in relazione all'uso secondo il diritto nazionale che lo autorizza, in particolare per quanto riguarda le limitazioni temporali geografiche e personali. A tal fine si stabilisce inoltre che l'uso è autorizzato solo se l'autorità di contrasto⁽⁸⁵⁾ ha completato una valutazione d'impatto sui diritti fondamentali come previsto all'art. 27 e ha registrato il sistema nella banca dati UE in conformità all'art. 49. Tuttavia, si aggiunge, che «in situazioni d'urgenza debitamente giustificate, è possibile iniziare a usare tali sistemi senza registrazione nella banca dati UE, a condizione che siffatta registrazione sia completata senza indebito ritardo».

E il *considerando* n. 34) illumina in proposito che l'uso dei sistemi di IBR in tempo reale in spazio accessibile al pubblico per attività di contrasto è consentito proprio perché preordinato a quanto strettamente necessario deve risultare limitato e circoscritto entro un limitato periodo di tempo e in relazione all'ambito geografico e personale «con particolare riguardo ad indicazioni od elementi probatori relativi a minacce, vittime o autori di reati». E il *considerando* si chiude con lo stabilire che l'uso può essere autorizzato solo se l'autorità di contrasto abbia «completato una valutazione d'impatto sui diritti e le libertà fondamentali e, salvo disposizione contraria del presente Regolamento, abbia altresì registrato il sistema nella banca dati di cui al presente Regolamento». Ancora, non si dimentica di sottolineare come l'indicata banca dati debba risultare «adeguata per ogni caso d'uso in ciascuna delle situazioni di cui sopra».

Il par. 3 concerne l'autorizzazione preventiva da parte dell'Autorità nazionale competente ivi indicata all'uso dei sistemi di IBR in tempo reale in spazi accessibili al pubblico per finalità di contrasto ove consentiti dalla sorveglianza umana correlativa. La disposizione è divisa in due commi. Nell'ultima parte del secondo comma, si stabilisce il precetto più signifi-

⁽⁸⁵⁾ E in punto sovviene la corrispondente definizione, di cui all'art. 3, n. 45), AIA: «“autorità di contrasto”: *a*) qualsiasi autorità pubblica competente in materia di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse; oppure *b*) qualsiasi altro organismo o entità incaricati dal diritto dello Stato membro di esercitare l'autorità pubblica e i poteri pubblici a fini di prevenzione, indagine, accertamento o perseguimento di reati o esecuzione di sanzioni penali, incluse la salvaguardia contro le minacce alla sicurezza pubblica e la prevenzione delle stesse».

cativo in proposito ossia che «nessuna decisione che produca effetti giuridici negativi su una persona può essere presa unicamente sulla base dell'*output* del sistema di identificazione biometrica remota "in tempo reale"». Ovvero, come anche in analoghe corrispondenti situazioni di uso di un sistema di IA ad alto rischio non più inaccettabile, il sistema di IA è essenziale strumento che può aiutare una valutazione che in proposito non può essere sottratta agli umani e che deve sempre risultare una valutazione umana. Il primo comma, come accennato, concerne l'autorizzazione preventiva cui è subordinato l'uso consentito dall'art. 5, par. 1, comma 1, lett. *b*) del sistema di IBR in tempo reale per fini di attività di contrasto. L'autorizzazione è rilasciata da un'Autorità giudiziaria o un'autorità amministrativa indipendente le cui decisioni sono vincolanti nello Stato membro in cui deve avvenire l'uso, rilasciata su richiesta motivata e in conformità alle regole dettagliate del diritto nazionale di cui al successivo paragrafo 5. Viene tuttavia precisato che «in una situazione di urgenza debitamente giustificata, è possibile iniziare ad usare il sistema senza autorizzazione» se la stessa sia richiesta senza «indebito ritardo» entro le 24 ore. In questo caso, se l'autorizzazione è respinta «l'uso è interrotto con effetto immediato e tutti i dati, nonché i risultati, gli *output* di un siffatto uso sono immediatamente eliminati e cancellati».

Il secondo comma stabilisce che l'Autorità giudiziaria o amministrativa indipendente competente rilascia l'Autorizzazione in oggetto, solo se ha accertato «sulla base di prove oggettive o indicazioni chiare che le sono state presentate» che l'uso del sistema di IBR in tempo reale per attività di contrasto è necessario e proporzionato e «in particolare rimane limitato a quanto strettamente necessario per quanto riguarda il periodo di tempo, l'ambito geografico e personale». Nel prendere la decisione sul rilascio dell'autorizzazione l'Autorità «tiene conto degli elementi di cui al paragrafo 2».

In proposito all'autorizzazione preventiva di cui al par. 3 chiarisce il *considerando* n. 35): «tale autorizzazione in linea di principio dovrebbe essere ottenuta prima dell'uso del sistema di IBR in tempo reale in gioco. Eccezioni a tale regola dovrebbero essere ammesse in situazioni di urgenza debitamente giustificate, vale a dire quelle situazioni in cui la necessità di utilizzare i sistemi interessati è tale da far sì che sia effettivamente ed oggettivamente impossibile ottenere un'autorizzazione prima di iniziare ad utilizzare il sistema». E si continua specificando che in siffatte situazioni di urgenza, l'uso del sistema dovrebbe risultare limitato al minimo indispensabile e subordinarlo a tutele e condizioni adeguate secondo il diritto nazionale, nel contesto specificato di ogni singolo caso di uso urgente da

parte dell'Autorità di contrasto competente. «In tali situazioni, inoltre, l'Autorità di contrasto nel richiedere l'autorizzazione dovrebbe contestualmente indicare i motivi per cui non ha potuto richiederla prima, e comunque con richiesta senza indebito ritardo, entro le 24 ore». Nel caso in cui la richiesta sia respinta l'uso del sistema va interrotto immediatamente e tutti i dati ottenuti da tale uso devono essere eliminati e cancellati. Tali dati, si specifica, comprendono sia i dati di *input* acquisiti da un sistema di IA nel corso dell'uso, nonché i risultati (c.d. intermedi) o gli *output* dell'uso correlati all'autorizzazione non accolta⁽⁸⁶⁾.

Il par. 4 dell'art. 5 richiede che, «fatto salvo il paragrafo 3», ogni uso consentito di sistema di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico per finalità di attività di contrasto debba essere notificato alla pertinente Autorità di vigilanza del mercato e all'Autorità nazionale per la protezione dei dati⁽⁸⁷⁾ in conformità alle regole nazionali di cui al par. 6. La notifica contiene «almeno» le informazioni di cui al par. 6, e non include dati operativi sensibili. La notificazione qui in oggetto deve avvenire successivamente ad ogni uso consentito al fine di rendere possibile la comunicazione di *report* sul numero di autorizzazioni e sul loro risultato. La notificazione non richiede di includere i dati operativi sensibili rispetto ai quali sovviene la corrispondente definizione di cui all'art. 3, n. 38) del presente AI Act, art. 3, n. 38) di «dati operativi sensibili»: «dati operativi relativi ad attività di prevenzione, accertamento, indagine o perseguimento di reati, la cui divulgazione potrebbe compromettere l'integrità dei procedimenti penali».

Secondo il par. 5, lo Stato nazionale può decidere di prevedere la possibilità di autorizzare, *in toto* ma anche solo in parte, l'uso di sistemi di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico a fini di attività di contrasto o entro i limiti ad alle condizioni di cui al par. 1, comma primo, lettera *b*) e ai paragrafi 2 e 3. Nel caso, gli Stati membri emanano nel proprio diritto nazionale le necessarie regole dettagliate per la richiesta, il rilascio, l'esercizio delle autorizzazioni di cui al par. 3, nonché per le attività di controllo e di comunicazione ad esse relative. Siffatte regole devono specificare per quale degli obiettivi elencati al par. 1, comma primo, lettera *b*), inclusi i reati di cui alla lettera *b*), punto *iii*), le

⁽⁸⁶⁾ Per comprendere meglio in punto si rinvia alla definizione di sistema di IA di cui al mio scritto in corso di pubblicazione su *Contratto e Impresa/Europa*.

⁽⁸⁷⁾ Anche perché il rischio di impatto sulle persone in oggetto all'identificazione di cui si tratta, c.d. FRIA di cui al par. 2, non sostituisce l'analogo procedimento in materia di dati, di cui ad esempio all'art. 35 GDPR.

autorità competenti possono essere autorizzate ad impiegare tali sistemi di IA a fini di attività di contrasto. Successivamente alla adozione di queste regole, al più tardi 30 giorni dopo, gli Stati membri le notificano alla Commissione (la notifica di per sé non vale quale presunzione di conformità della disciplina nazionale all'AI Act. A sua volta il corrispondente AI Office invierà una conferma di ricevimento). Gli Stati membri possono introdurre disposizioni più restrittive sull'uso dei sistemi di identificazione biometrica remota.

In altri termini, l'uso nazionale di sistemi di IBR per finalità di contrasto qui in discussione deve rispettare il diritto europeo ma "può" (non deve) essere consentito anche nel territorio dello Stato, secondo la legge nazionale. E le norme nazionali sull'uso devono risultare dettagliate entro i limiti di ammissibilità ed alle condizioni stabiliti dal diritto dell'AI Act, così, ad esempio, non può essere estesa la lista dei reati di cui al punto *i*), ma la loro specificità o il loro essere dettagliate devono essere tali da consentire un controllo sull'uso che rimanga sempre comunque strettamente necessario e proporzionato (e non lo è se sono possibili misure che presentano, agli stessi fini di contrasto, un impatto di rischio minore sui diritti fondamentali rispetto al sistema di IBR in tempo reale). Ma lo Stato può, rispetto ad un tale uso, introdurre eventualmente regole ancora più restrittive rispetto a quelle di cui all'art. 5, lettera *b*) e di quelle di cui ai paragrafi da 2 a 7 qui in discussione.

A riguardo del par. 5 qui considerato è opportuno chiarire che il corrispondente *considerando* n. 37) ricorda che gli Stati membri restano liberi, a norma del presente regolamento, di non autorizzare affatto con regole nazionali dettagliate una tale possibilità di uso entro lo Stato membro o di prevederla soltanto rispetto ad alcuni (e non rispetto a tutti) degli obiettivi idonei ad autorizzare l'uso, secondo quanto stabilito dall'art 5, par.1, lettera *b*).

Secondo il paragrafo 6, le Autorità nazionali di vigilanza sul mercato e le Autorità nazionali per la protezione dei dati degli Stati membri cui è stato notificato l'uso di sistemi di IBR in tempo reale, in spazi accessibili al pubblico ai fini di attività di contrasto, secondo quanto visto nel paragrafo 4, presentano alla Commissione relazioni annuali su tale uso. A tal fine la Commissione fornisce agli Stati membri e alle Autorità nazionali di vigilanza del mercato e di protezione dei dati un modello comprendente informazioni sul numero di decisioni che le autorità giudiziarie competenti o un'autorità amministrativa indipendente la cui decisione è vincolante hanno adottato in risposta alle richieste di autorizzazione a norma del paragrafo 3 e del loro esito. Solo la relazione dell'Autorità nazionale per

la protezione dei dati coprirà il periodo di tempo fra il 2 febbraio 2025 e il 2 agosto 2025, in quanto prima di tale data non occorre, secondo i diversi termini di tempo fissati dall'AI Act per l'inizio della sua efficacia, che sia già nominata e operativa l'Autorità nazionale per la vigilanza sul mercato.

Infine, il par. 7 impone che la Commissione pubblichi relazioni annuali sull'uso di sistemi di identificazione biometrica remota in tempo reale in spazi accessibili al pubblico a fini di attività di contrasto, fondate su dati aggregati negli Stati membri sulla base delle relazioni annuali di cui al paragrafo 6. Anche in questo caso le relazioni non includono dati operativi sensibili delle relative attività di contrasto.

In definitiva, tutte le condizioni di cui ai paragrafi 2-7 dell'art. 5 dell'AIA valgono quando a norma dell'art. 5, lettera *b*), *sub* punti *i*) *ii*) *iii*) consentano l'uso di sistemi biometrici remoti in tempo reale. È chiaro tuttavia che mentre nel caso del divieto si tratta di sistemi di IA di alto rischio inaccettabile, nel caso poc'anzi indicato, il rischio non è più inaccettabile ma è sempre e rimane comunque un alto rischio, e di conseguenza l'uso consentito, in quanto strettamente necessario e proporzionato, deve essere sottoposto a tutte le particolari misure di tutela proprie di questa categoria di sistemi di IA, per limitare la possibilità di un danno grave ai diritti ed alle libertà fondamentali in gioco, con una tutela che è caratteristica particolare del modello di protezione europeo in tema di intelligenza artificiale, e permettendo così di presentare l'AI Act come un modello che vuole risultare *leader* mondiale nella protezione corrispondente⁽⁸⁸⁾.

Il divieto impedisce immissione sul mercato, messa in servizio o uso di un sistema di IA che dia luogo ad una delle pratiche vietate. Ogni violazione al divieto di pratiche di IA di cui all'art. 5 è soggetta ad una sanzione amministrativa pecuniaria molto elevata fino a 35 000.000 EUR, o se l'autore dell'illecito è un'impresa, una sanzione fino al 7% del fatturato mondiale totale annuo dell'esercizio precedente qualora superiore ai 35000000. La cifra è sufficientemente alta da scoraggiare per davvero di porre in essere una delle pratiche di IA vietate ed in questo senso potrebbe risultare più che sufficiente allo scopo per la tutela del sistema valoriale all'apice delle fonti correlato ai sistemi di IA.

⁽⁸⁸⁾ Mi permetto di rinviare ancora al mio scritto in corso di pubblicazione su *Contratto e impresa/Europa*.

5. – Appare importante ribadire nuovamente che il pregio maggiore che le otto tipologie di pratiche vietate nel paragrafo 1, comma 1, dell'art. 5 dalla lettera *a*) alla lettera *b*) evidenziano tutte una caratteristica significativa. Esse, per quanto stabilite in norme precise e tassative anche nella loro formulazione (consentendo spesso di operare rinvio a definizioni normative che contribuiscono a definirne più precisamente il relativo ambito di applicazione sia per quanto attiene al divieto, ma parallelamente, anche per quanto attiene alla sottrazione al divieto stesso), non stabiliscono divieti assoluti. Ognuna delle pratiche viste rende anche possibile individuare eccezioni al divieto rendendo la corrispondente pratica una pratica anche ad alto rischio accettabile, e quindi ammissibile nella sua immissione sul mercato messa in servizio o uso con le opportune cautele e misure di protezione che l'AI Act stabilisce per fronteggiare per l'appunto l'alto rischio non inaccettabile. Questa fluidità propria delle pratiche di cui all'art. 5 è un punto di forza, perché permette di pensare ad una possibile uscita dall'art. 5 quando anche per il progresso della tecnica la corrispondente pratica non presenti più quel rischio grave di lesione a diritti fondamentali delle persone come invece oggi essa presenta e che rende giustificato il suo divieto. Così, ad esempio, è vero che i sistemi di identificazione biometrica remota in tempo reale, in spazi accessibili al pubblico per finalità di attività di contrasto, attualmente acquiscono i rischi di lesione grave dei diritti fondamentali di un numero significativo di persone fisiche, vale a dire non solo l'autore o sospettato del reato interessato e la vittima del medesimo, ma anche tutte le persone ipoteticamente coinvolte. Ma non dobbiamo dimenticare che l'intelligenza artificiale è ancora agli albori. È quindi possibile aspettarsi che un domani anche non così lontano lo sviluppo tecnologico dell'IA consenta di immaginare corrispondenti sistemi che risultino immuni da quegli "errori" che oggi sono invece presenti e tali da far inserire quella pratica di IA nel gruppo delle pratiche vietate. In tal modo, in definitiva, è immaginabile una lista di pratiche oggi vietate che nel corso del tempo possa risultare oggetto di modifiche ma che non porti quel numero di pratiche vietate ad ingigantirsi nel corso del tempo, ma rimanendo sempre un numero circoscritto solo oggetto a contenuto diverso.

Un ultimo punto: la Commissione ha pubblicato, come ripetutamente indicato nel testo, un numero considerevole di Linee guida destinate prioritariamente agli operatori economici di sistemi di IA per chiarire i precetti di cui all'art. 5 con un'elencazione esemplificativa sorprendente sia di casi rientranti nel divieto sia di casi non rientranti nel medesimo e a cui talvolta anche in questo scritto si è operato proficuamente riferimento. Tuttavia,

forse proprio per la loro valenza quantitativa, le Linee guida non risultano prive di sviste. Così ad esempio in riferimento ad alla lettera *a*) ricorre un caso di pratica di sistema di IA “volutamente manipolativa”, consistente nella manipolazione sensoriale, in cui un sistema di IA distribuisce audio o immagini di sottofondo che portano ad alterazioni d’umore, aumentando magari l’ansia e il disagio mentale che influenzano il comportamento degli utenti fino a creare danni significativi⁽⁸⁹⁾, quando la disciplina normativa rende chiaro sia per la formulazione delle disposizioni in oggetto, sia per espressa ammissione di un *considerando*⁽⁹⁰⁾ al riguardo, che non occorre la presenza di un elemento soggettivo intenzionale, vale a dire un elemento soggettivo quale il dolo, essendo sufficiente che per oggetto o per effetto la pratica in sé risulti manipolativa. E quindi vietata.

⁽⁸⁹⁾ V. EU Commission, *op. cit.*, p. 22.

⁽⁹⁰⁾ Si tratta del citato *considerando* n. 29).