



UNIVERSITÀ POLITECNICA DELLE MARCHE  
FACOLTÀ DI ECONOMIA “GIORGIO FUÀ”

---

Corso di dottorato in Management and Law  
Curriculum “Diritto dell’economia” – XXXVI° Ciclo

## **Responsabilità civile e Intelligenza Artificiale**

Relatrice: Chia.ma

**Prof.ssa Erika Giorgini**

Dottoranda:

**Marta Mariolina Mollicone**

A.A. 2022/2023



Alle radici e ai rami del mio tronco.  
Che queste parole siano ricordate come fiori.

## Indice

<b>Introduzione</b>	<b>I</b>
<b>CAPITOLO I</b>	<b>1</b>
<b>Fenomenologia dell'Intelligenza Artificiale</b>	<b>1</b>
1. La storia dell'intelligenza artificiale.	1
1.1. Introduzione.	1
1.2. La nascita dell'intelligenza artificiale.	2
1.3. L'apprendimento automatico: elementi per un giurista.	8
1.3.1. Gli algoritmi.	8
1.3.1. Il machine learning.	13
1.3.2. Il deep learning.	18
2. Definizione e tassonomia dell'intelligenza artificiale allo stato dell'arte.	18
<b>CAPITOLO II</b>	<b>27</b>
<b>Inquadramento giuridico dell'Intelligenza Artificiale</b>	<b>27</b>
1. Introduzione.	27
2. I soggetti giuridici: <i>essentialia</i> .	29
2.1. Il concetto di persona. Dal diritto romano al diritto dell'Unione europea.	32
2.2. Le persone giuridiche.	35
3. La qualificazione giuridica dei sistemi di IA.	41
3.1. IA come soggetto del diritto.	41
3.1.1. Segue...secondo la teoria della finzione ovvero il criterio funzionale o della semplificazione giuridica.	44
3.1.2. Segue...Secondo il criterio della <i>moral agency</i> .	47
3.1.3. Segue...il <i>tertium genus</i> della personalità elettronica.	49
3.2. IA come oggetto del diritto.	53
3.2.1. IA come prodotto...nel d.lg. del 6 settembre 2005, n. 206 (Codice del Consumo italiano).	55
3.2.2. Segue...nella nuova proposta di direttiva europea sulla responsabilità per prodotti difettosi ( <i>PLD</i> ).	59
<b>CAPITOLO III</b>	<b>66</b>
<b>Il problema della responsabilità civile</b>	<b>66</b>
1. Introduzione.	66
2. I paradigmi di responsabilità applicabili in base all'ordinamento nazionale.	68
2.1. La responsabilità per colpa.	69
2.2. La responsabilità da attività pericolosa ex art. 2050 c.c.	81
2.3. La responsabilità da cose in custodia ex art. 2051 c.c.	87
3. Il paradigma previsto dalla proposta di direttiva <i>AI liability</i> ( <i>AILD</i> ).	90
3.1. <i>AILD</i> e <i>PLD</i> : criticità, osservazioni e confronti.	97
<b>Conclusioni</b>	<b>112</b>
<b>Bibliografia</b>	<b>117</b>

## Introduzione

«Pensiamo alle nuove tecnologie, ai risultati straordinari della ricerca scientifica, della medicina, alle nuove frontiere dello spazio, alle esplorazioni sottomarine. Scenari impensabili fino a pochi anni fa e ora davanti a noi.

Sfide globali, sempre.

Perché è la modernità, con il suo continuo cambiamento, a essere globale.

Ed è in questo scenario, per larghi versi inedito, che misuriamo il valore e l'attualità delle nostre scelte strategiche: l'Europa, la scelta occidentale, le nostre alleanze. La nostra primaria responsabilità nell'area che definiamo Mediterraneo allargato. Il nostro rapporto privilegiato con l'Africa.

Dobbiamo stare dentro il nostro tempo, non in quello passato, con intelligenza e passione.

Per farlo dobbiamo cambiare lo sguardo con cui interpretiamo la realtà. Dobbiamo imparare a leggere il presente con gli occhi di domani.

Pensare di rigettare il cambiamento, di rinunciare alla modernità non è soltanto un errore: è anche un'illusione. Il cambiamento va guidato, l'innovazione va interpretata per migliorare la nostra condizione di vita, ma non può essere rimossa.

La sfida, piuttosto, è progettare il domani con coraggio»<sup>1</sup>.

Con tali parole il Presidente Mattarella saluta l'anno 2022. L'invito che rivolge ai cittadini italiani è quello di non rifuggire dalla prospettazione di un nuovo ordine di idee per timore della violenza del cambiamento. La strada dell'evoluzione sociale passa, inevitabilmente, per un mutamento degli equilibri. Ogni nuova «era» ha dovuto fare i conti con l'instabilità delle nuove conoscenze. Tuttavia, sebbene la reazione più istintuale di fronte all'ignoto sia la diffidenza, solo il coraggio della sua comprensione permette di non farsi trascinare dagli accadimenti empirici del presente ma di dominarli in maniera consapevole.

A tal fine deve intervenire il diritto che è, innanzitutto, cultura di una data società nonché «perenne confronto con la realtà fattuale, con l'insieme di questioni e possibili soluzioni storicamente determinate»<sup>2</sup>. Attraverso l'interpretazione della realtà, il diritto, che ne è frutto, rende intellegibili i problemi concreti e permette di orientarsi nel continuo mutamento, garantendo il vivere ordinato grazie alla regolamentazione dei conflitti.

Proprio con l'obiettivo di offrire comprensione si redige questo scritto, che si propone di inserirsi nella scia di coloro che rinvergono nel diritto una funzione promozionale<sup>3</sup> nonché sostanzialistica<sup>4</sup>.

---

<sup>1</sup> Messaggio di fine anno del Presidente della Repubblica Sergio Mattarella, Palazzo del Quirinale, 31 dicembre 2022.

<sup>2</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, 4ª ed., I, *Metodi e Tecniche*, Napoli, 2020, p. 5 e 17.

<sup>3</sup> Sulla funzione promozionale del diritto v. N. BOBBIO, *Sulla funzione promozionale del diritto*, in *Riv. trim.*, 1969, p. 1313 ss, e A. CATANIA, *Manuale di filosofia del diritto*, Torino, 2020, p. 151 ss.

<sup>4</sup> Sull'interpretazione del diritto quale attività inventiva capace di intermediare, nonostante uno iato temporale, tra il senso originario del testo con quello presente, non trascurando gli interessi contemporanei

Oggetto dell'argomentare sarà quello che, primissimo tra i problemi propri dell'età tecnologica<sup>5</sup>, attiene i profili di responsabilità conseguenti l'utilizzo dei nuovi strumenti, specificamente i sistemi di intelligenza artificiale.

Il dirompente sviluppo tecnologico-informatico non è stato accompagnato da una parallela sufficiente ed efficiente regolamentazione e, perciò, non si è riusciti né ad evitare che sorgessero dubbi sull'allocazione del rischio tecnologico né che questi si esasperassero con l'avvento degli algoritmi cognitivo-predittivi, ormai, sempre più spesso, funzionanti in piena autonomia.

Come attenta dottrina già rilevava più di un ventennio fa<sup>6</sup> e le cui parole vale la pena riportare, nel suo complesso l'età tecnologica suscita tre grandi tipi di reazioni intellettuali: ottimismo, pessimismo e realismo. «L'atteggiamento ottimistico vede nell'età tecnologica e nelle sue conseguenze d'ordine sociale il compiersi di un grande balzo in avanti nello sviluppo della civiltà umana. Il progresso scientifico e tecnologico rappresenta il segno del dominio dell'uomo sulla natura e la promessa di un soddisfacimento dei bisogni umani che porta all'attuazione di un ideale della felicità in una società di autentico benessere. L'atteggiamento pessimistico vede invece nell'età tecnologica presente un pericolo che insidia quei valori della specie umana, che non sono riducibili al soddisfacimento dei suoi "bisogni inferiori", ma che sono radicati nelle tradizioni della civiltà. Inoltre, esso intravede gli altri pericoli che si profilano: gli effetti malefici della tecnica attuale sul mondo dell'uomo e sul suo ambiente naturale, e il deteriorarsi della coesistenza umana nelle società manipolate dalle nuove tecniche. [...] L'atteggiamento realistico sceglie la via di mezzo fra le due sopra indicate. L'età tecnologica viene considerata come una sfida per l'uomo, che deve valutare le sue promesse e i suoi pericoli. Sotto questo aspetto l'età tecnologica rivela una sua ambivalenza e richiede all'uomo di compiere alcune scelte basilari che tengano conto delle sue latenti possibilità sia nel bene che nel male per lui e per le generazioni future. Le scelte dell'uomo d'oggi debbono tener in conto le conseguenze sul suo futuro ed egli deve perciò basarle su una conoscenza accurata dei fatti e su una presa in considerazione dei valori che sono espressione della coesistenza dell'umanità. Perciò la sua risposta alla sfida dev'essere razionale e consistere in una valutazione delle possibilità conosciute di controllare il suo futuro. Vi sono tuttavia alcuni limiti di questa decisione razionale: essi sono i limiti della conoscenza, i rischi della azione pratica e alcune ambivalenze e conflitti fra i valori invocati per giustificare la sua decisione»<sup>7</sup>.

---

v. L. MENGONI, *Ancora sul metodo giuridico*, in *Riv. trim. dir. proc. civ.*, 1984, p. 337 ss.; ID., *Dogmatica giuridica*, in ID., *Ermeneutica e dogmatica giuridica, Saggi*, Milano, 1996, p. 42; P. PERLINGIERI, *Dogmatica giuridica e legalità costituzionale*, in *Annali Sisdic*, III, 2019, p. 1 ss.; ID., *Valori normativi e loro gerarchia. Una precisazione dovuta a Natalino Irti*, in *Rass. dir. civ.*, 1999, p. 802 ss.

<sup>5</sup> Secondo J. WROBLEWSKI, *Dilemmi dell'età tecnologica: il diritto e l'omeostasi dell'esistenza umana*, in F. RICCOBONO (a cura di), *Nuovi diritti dell'età tecnologica*, Milano, 1991, p. 195 ss. L'età tecnologica è caratterizzata da almeno tre gruppi di mutamenti tecnologici «a) lo sviluppo della computerizzazione, della tecnologia dell'informazione e della telematica b) l'impiego di nuove fonti di energia nucleare, sia per la produzione industriale che per la produzione di armi c) le tecniche di controllo genetico in vari campi, definite nel loro complesso come "biotecnica"». Si concentrerà l'attenzione sul primo di questi mutamenti.

<sup>6</sup> J. WROBLEWSKI, *Dilemmi dell'età tecnologica*, cit. p. 196-197.

<sup>7</sup> J. WROBLEWSKI, *o.u.c.*

Se in un primo momento nei confronti delle nuove risultanze della tecnica informatica imperava l'ottimismo, oggi traspare pessimismo e, per lo più, timore verso i sistemi di intelligenza artificiale. Alla prospettiva dell'intelligenza artificiale quale veicolo per raggiungere una precoce evoluzione umana si sta man mano sostituendo la paura di perdere il controllo della tecnologia stessa, l'angoscia di correre troppo e di porre le basi per un futuro distopico in cui non c'è spazio per i valori dell'umanità<sup>8</sup>. Un nuovo «Oppenheimer». La concretizzazione della terza legge di Sir Arthur Clarke nel suo romanzo *Profiles of the Future: An Inquiry into the Limits of the Possible*: qualsiasi tecnologia sufficientemente avanzata è indistinguibile dalla magia. La magia è incomprensibile e, quindi, pericolosa.

Da ciò deriva il proliferare di norme precauzionali e preventive, specialmente di matrice europea, che tentano un controllo *ex ante* del fenomeno e cercano di responsabilizzare ogni componente della filiera del valore prima della messa in circolazione di un sistema di intelligenza artificiale.

Perciò si riversano nell'ordinamento, galoppanti come una cavalcata delle valchirie, informative, linee guida, prese d'atto, proposte di direttive, di regolamenti, norme di ogni tipo, ognuna diretta a suggerire il possibile trattamento giuridico delle conseguenze patologiche derivanti dal «comportamento antiggiuridico» di un sistema di intelligenza artificiale. E, di tal passo, il temibile fenomeno dell'iper-regolamentazione acquista la *chance* di concretizzarsi.

Al fine di scongiurare detto scenario, si propone un testo che predilige un atteggiamento razionale nei confronti dell'innovazione tecnologica e dei suoi problemi, ferma la convinzione che solo la comprensione delle cose è necessaria per effettuare valutazioni del fenomeno prodromiche a scelte consapevoli.

Come ben preannunciava Falzea<sup>9</sup>, «la società globale, della quale *Internet* è veicolo imprescindibile, esige un giurista aperto al mondo». Non è più tempo di giuristi che si chiudono nella platonica caverna della tradizione per la preoccupazione di non trovare la stella polare di riferimento in un universo di luce. In questi tempi di forte dinamismo sono fondamentali giuristi coraggiosi, che sappiano spiegare il presente a partire dal passato<sup>10</sup>, non permettendo, però, a questo di fagocitare il primo.

Si tenterà, quindi, di accostarsi alle questioni, come la migliore dottrina insegna, con «un approccio chiaro ed essenziale, anticipatorio ma pur sempre problematico, concettualmente rigoroso sebbene pragmatico ed esemplificatorio»<sup>11</sup>.

---

<sup>8</sup> Si vive un contesto ai confini con la fantasia. Non si può non pensare al celebre film *Io Robot* (2004), ispirato all'antologia di Isaac Asimov, ove in una Chicago del 2035 da tempo i robot positronici convivono naturalmente con gli esseri umani fino a che non si scopre il terribile piano del supercomputer VIKI (*Virtual Interactive Kinetic Intelligence*), il cervello positronico centrale in grado di comandare tutti i robot. Questi, evolutosi al punto da interpretare autonomamente le tre leggi della robotica, tenta di limitare la libertà degli esseri umani per proteggerli da sé stessi.

<sup>9</sup> A. FALZEA, *Il civilista e le sfide del millennio*, in Aa.Vv., *Studi in onore di Pietro Schlesinger*, I, Milano, 2004, p. 15 ss.

<sup>10</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale*, I, cit., p. 49.

<sup>11</sup> P. PERLINGIERI, *o.u.c.*

Si inizierà delineando la fenomenologia dell'intelligenza artificiale, cercando di coglierne tutte le espressioni, senza avere l'ambizione di sostituirsi a chi del suo studio ne fa una professione. Ci si riferisce agli ingegneri informatici che, sicuramente, sono i più idonei a fornire spiegazioni tecniche.

Si tenterà, quindi, un suo inquadramento giuridico addentrandosi nella problematica della sua soggettività, ed, infine, si concentrerà l'attenzione sul problema della responsabilità civile, soffermandosi, in particolare, sul quadro regolamentare europeo in corso di formazione

# CAPITOLO I

## Fenomenologia dell'Intelligenza Artificiale

SOMMARIO: 1. La storia dell'intelligenza artificiale – 1.1. Introduzione – 1.2. La nascita dell'intelligenza artificiale – 1.3. L'apprendimento automatico: elementi per un giurista – 1.3.1. Gli algoritmi – 1.3.2. Il *machine learning* – 1.3.3. Il *deep learning* – 2. Definizione e tassonomia dell'intelligenza artificiale allo stato dell'arte.

### 1. La storia dell'intelligenza artificiale.

#### 1.1. Introduzione.

Ricalcando uno degli incipit più noti della letteratura, Anna Karenina di Lev Tolstoj, si potrebbe dire che «tutti siamo in grado di riconoscere la stupidità, ma dell'intelligenza ognuno ha una propria idea»<sup>1</sup>, perciò occorre trovare una definizione comune. E se ciò è vero per l'intelligenza cd. naturale, lo è altrettanto per quella artificiale (da qui in poi IA).

Il concetto di IA vede i suoi albori nella seconda metà del 1600 quando Isaac Newton, in occasione della pubblicazione dei *Principi matematici della filosofia naturale*, intuì come il libro della natura potesse esser scritto nel linguaggio della matematica. Tuttavia, solo nel 1840, questo assunse un significato più vicino a quello odierno, seppur in una declinazione debole, grazie all'illustrazione che Charles Babbage fece della sua macchina analitica ad una conferenza tenutasi a Torino. Più precisamente, il concetto è da ricondursi alla triade Babbage-Menabrea-Lovelace i quali, tramite una successione di eventi<sup>2</sup>, hanno tracciato per la prima volta la distinzione, in chiave di complementarità, tra aspetti meccanici e capacità analitiche della macchina, distinzione che oggi richiama i concetti di *hardware* e *software*.

È però alla seconda metà del XX secolo che occorre guardare per individuare quelle che verranno considerate le tappe fondamentali della storia dell'IA e che, secondo Stuart J. Russell, uno dei più brillanti scienziati informatici, possono essere riepilogate ripescando l'elenco dei vincitori del *A. M. Turing Award*<sup>3</sup> Marvin Minsky (1969) e John McCarthy (1971) per la definizione dei fondamenti del campo in base alla rappresentazione e al ragionamento; Ed Feigenbaum e Raj Reddy (1994) per lo sviluppo di sistemi esperti in grado di codificare la

---

<sup>1</sup> Il medesimo *incipit* è parafrasato da G. MAIRA, *Intelligenza umana e intelligenza artificiale* in *Federalismi.it*, VII, 2021, p. V ss.

<sup>2</sup> Luigi Federico Menabrea scrisse un articolo su quanto esposto dallo scienziato proto-informatico Charles Babbage. L'articolo fu successivamente tradotto e glossato da Ada Lovelace.

<sup>3</sup> L'A.M. Turing Award è un premio, assegnato annualmente dalla Association for Computing Machinery (ACM), a una personalità che eccelle per i contributi di natura tecnica offerti alla comunità informatica, in particolare per progressi che siano duraturi e di elevata importanza tecnica. Il premio viene spesso anche chiamato «premio Nobel dell'informatica» ed è intitolato al matematico inglese Alan Mathison Turing (1912-1954), in riconoscimento del suo contributo unico e originale alla nascita delle attività di calcolo mediante dispositivi automatici.

conoscenza umana per risolvere problemi del mondo reale; Judea Pearl (2011) per lo sviluppo di tecniche di ragionamento probabilistico in grado di affrontare le situazioni di incertezza in modo formalizzato; e, infine, Yoshua Bengio, Geoffrey Hinton e Yann LeCun (2019) per aver fatto del deep learning (reti neurali multistrato) una parte fondamentale dell'informatica moderna<sup>4</sup>.

È necessario, a questo punto, ripercorrere ogni singola fase.

## 1.2. La nascita dell'intelligenza artificiale.

Tralattivamente si riferisce che il primo lavoro appartenente all'IA fu quello svolto nel 1943 da Warren McCulloch e Walter Pitts i quali, a tal fine, attinsero alla conoscenza di base della fisiologia e della funzione dei neuroni del cervello<sup>5</sup>, all'analisi formale della logica proposizionale di Russell e Whitehead<sup>6</sup> e alla teoria della computazione di Turing<sup>7</sup>.

---

<sup>4</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, in *Intelligenza artificiale. Un approccio moderno*, Londra, 2021, p. 19.

<sup>5</sup> Come ricorda il Prof. di neurochirurgia G. MAIRA nel suo saggio *Intelligenza umana*, cit., p. V ss., tra gli aspetti principali del cervello umano, da tenere in considerazione, vi sono la plasticità, che consiste nella capacità di modificazione continua e costante apprendimento fino alla senescenza, e l'emozionalità, ovvero la capacità di provare emozioni che altro non sarebbero che algoritmi di sopravvivenza «[...] evoluti per permetterci di evitare il pericolo» (R. CARTER, *Mapping the Mind*, Berkeley-Los Angeles, 2010) e dare forza alla nostra memoria. L'oscillazione tra emozioni e razionalità permette al nostro cervello di organizzare in vario modo il pensiero. Si distingue, infatti, il cd. «sistema 1» ovvero il pensiero veloce ed intuitivo che coinvolge il sistema emozionale ed opera automaticamente, attraverso scorciatoie mentali, sulla base delle esperienze accumulate, dal cd. «sistema 2» ovvero il pensiero lento e razionale che coinvolge la corteccia cerebrale del raziocinio e necessita di un ragionamento profondo e cosciente. I due pensieri non sono in perenne conflitto ma operano in termini di complementarità. Le regole del pensiero umano possono, così, essere compendiate dalle parole di Erasmo *festina lente*, affrettati lentamente.

<sup>6</sup> I *Principia Mathematica* di Alfred N. WHITEHEAD e B. RUSSELL possono considerarsi la più ampia, articolata e completa summa della logica matematica moderna.

Essi sono il punto di approdo e la massima espressione di quello che viene chiamato in storiografia «programma logicista» o «logicismo». Esso denota una impostazione di filosofia della logica di matrice «idealistica» secondo cui, in linee generali, i concetti e i relativi costrutti della logica: a) hanno validità universale, b) hanno potenza espressiva illimitata, c) sono meta-soggettivi, ossia totalmente indipendenti dai meccanismi psicologici di una mente individuale, d) riguardano leggi universali del pensiero/linguaggio, e) eccedono la stessa capacità di comprensione, f) sono a-temporali. I concetti fondamentali della matematica devono potersi ridurre a concetti logici, a loro volta indefinibili.

Stando a quanto riporta il primo impianto dei *Principia*, ossia il volume di B. RUSSELL intitolato *The Principles of Mathematics*, del 1903, queste nozioni indefinibili e irriducibili sono sei: 1. l'implicazione formale; 2. l'implicazione materiale; 3. la relazione fra un termine e la classe di cui è elemento (appartenenza); 4. la nozione di tale che; 5. la nozione di relazione; 6. la nozione di verità. Si voleva dimostrare che la matematica, per essere consistente e sicura non doveva far altro che fondarsi sulla logica. Da qui la denominazione «logica matematica», che racchiude implicitamente tre specificazioni: 1) è una logica dalla matematica, nel senso che nasce dalle ricerche su di essa; 2) è una logica della matematica, ossia sottesa alle sue dimostrazioni e ai suoi concetti base; 3) è una logica per la matematica, ossia per una «fondazione» – in qualunque senso la si intenda – di essa.

I Principi non solo sono stati la sorgente di preziosi concetti tra cui quello di funzione proposizionale e la teoria dei tipi, ma hanno anche posto le basi per la scoperta di cruciali risultati metateorici tra cui i teoremi di Kurt Gödel, Alonzo Church, Alan Turing e altri, i quali, a loro volta, hanno iniziato una tradizione di lavoro tecnico comune in campi diversi tra loro, come la filosofia, la matematica, la linguistica, l'economia, l'informatica, la crittografia, l'intelligenza artificiale.

<sup>7</sup> La teoria della computabilità si occupa di stabilire quali funzioni possono essere calcolate da un automa o in generale da un elaboratore. Fondamentale fu, nel 1950, l'elaborazione del test di Turing in base al quale venne offerta la prima definizione operativa di intelligenza artificiale. Una macchina sarebbe stata

L'IA è nata proprio dalla sfida di questi scienziati di immaginare di poter simulare, e possibilmente superare, il funzionamento del cervello umano, in particolare la sua plasticità. Agli albori, infatti, i più parlano di IA antropomorfa in quanto tesa a riprodurre i meccanismi caratterizzanti l'essere umano, sia fisici che cognitivi. «I due proposero un modello di neuroni artificiali in cui ogni neurone era caratterizzato dallo stato “acceso” o “spento”, e la cui accensione si verificava in risposta allo stimolo da parte di un numero sufficiente i neuroni adiacenti. Lo stato del neurone veniva così concepito come “di fatto equivalente alla proposizione corrispondente agli stimoli adeguati”»<sup>8</sup>. Gli stessi mostrarono, inoltre, «che ogni funzione computabile poteva essere calcolata da una rete di neuroni collegati e che tutti gli operatori logici (*and, or, not* e così via) potevano essere implementati con semplici strutture a rete. Suggestarono inoltre che reti neurali adeguatamente definite potessero essere capaci di apprendere»<sup>9</sup>.

Nel 1949 Donald O. Hebb aggiornò la regola per la modifica dei pesi delle connessioni tra i neuroni, creando quello che oggi si chiama «apprendimento hebbiano»<sup>10</sup>, gettando le basi per il successivo progetto di Marvin Minsky e Dean

---

intelligente solo se avesse superato il test ed il superamento si sarebbe avuto unicamente nel caso in cui un esaminatore umano, dopo aver posto delle domande in forma scritta, non sarebbe stato in grado di capire se le risposte provenissero da una persona o meno. Nonostante la presenza di altri lavori, la visione di Turing, secondo cui era più facile creare IA di livello umano sviluppando algoritmi di apprendimento rispetto alla programmazione manuale dell'intelligenza della macchina, è stata quella che più ha influenzato lo sviluppo dell'IA.

<sup>8</sup> Così S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 20. Per lungo tempo il corpo e la mente umana sono stati studiati e imitati separatamente, dando origine, da una parte, a macchine molto potenti e precise oppure, dall'altra, a cervelli elettronici dalle capacità computazionali straordinarie. L'evoluzione ha, poi, condotto queste due strade parallele verso un avvicendamento tale da creare l'integrazione meccatronica (con una biomeccanica simile a quella del corpo umano) e sistemi di calcolo (paragonabili, e talvolta superiori, al cervello umano).

Successivamente, rileva U. RUFFOLO, *Intelligenza Artificiale e diritti della persona: le frontiere del 'transumanesimo'*, in *Giuri. it.*, 2019, VII, p. 1658-1670, l'IA divenne tecnologia di human *enhancement* grazie anche alla «combinazione tra impiego dell'A.I., biotecnologie e scienze mediche, biomediche e genetiche, utilizzate per aumentare, quantitativamente e qualitativamente, le capacità fisiche e cognitive dell'essere umano o crearne di nuove, sul piano fisico, mentale ed emotivo». Si parla, così, di «transumanesimo» (o «post-umanesimo»), quale «anelito dell'uomo “prigioniero del proprio corpo” verso il travalicarne i naturali limiti, sino a superare la morte». In generale sul tema, si v. tra molti A. BESNIER, *Do Robots make love? From AI to immortality – Understanding transhumanism in 12 questions*, Londra, 2018; N. BOSTROM, *Superintelligenza – Tendenze, pericoli, strategie*, Torino, 2018; W. DAUGHERTY, *Human + Machine: reimagining work in the age of AI*, Boston, 2018. Sull'IA, meno specificamente, sia consentito rinviare a F. FIDANZA, *Sulla distinzione tra intelligenza artificiale e algoritmi*, in *Riv. dir. impr.* 2022, II, p. 389-406; v. anche N. ABRIANI e G. SCHNEIDER, *Diritto delle imprese e intelligenza artificiale. Dalla Fintech alla Corpotech*, Bologna, 2021, p. 21 ss.; T. CASADEI e S. PIETROPAOLI, *Intelligenza artificiale: fine o confine del diritto?*, in T. CASADEI e S. PIETROPAOLI (a cura di), *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*, Trento, 2021, p. 219-232; J. ARMOUR e H. EIDENMÜLLER, *Self-Driving Corporations?*, in *Harvard Business L. Rev.*, 2020, X, p. 87 ss.; per una trattazione monografica, N. NILSSON, *The quest for artificial intelligence. A history of ideas and achievements*, Cambridge, 2009.

<sup>9</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 20.

<sup>10</sup> «Modello formale di apprendimento spiegabile secondo tre ipotesi, ciascuna delle quali ha col tempo ricevuto adeguate conferme sperimentali. La prima ipotesi è che i neuroni corticali rafforzino le loro connessioni quando risultano con frequenza attivi contemporaneamente. Questo principio di apprendimento associativo sembra essere valido per la maggioranza dei neuroni corticali; a esso si fa comunemente riferimento, appunto, come regola di Hebb. La seconda ipotesi è che la corteccia sia un'enorme memoria associativa in cui il rafforzamento delle sinapsi abbia luogo non solo tra neuroni vicini, ma anche tra neuroni in aree corticali distanti. Questa seconda ipotesi trae sostegno dagli studi neuroanatomici che evidenziano percorsi cortico-corticali tra molte aree della corteccia. Secondo la terza

Edmonds. I due studenti di Harvard, già nel 1950, costruirono, infatti, il primo computer basato su reti neurali, lo Snarc, utilizzando 3000 tubi a vuoto e un sistema automatico di pilotaggio riciclato da un bombardiere B-24 per simulare una rete di 40 neuroni.

Tuttavia, è nel 1958 che gli entusiasmi iniziali, spesso contrastati dall'*establishment* intellettuale dell'epoca, iniziarono ad apparire una realtà percorribile. Se da una parte Minsky otteneva risultati al MIT focalizzandosi sul funzionamento di programmi in una prospettiva antilogica<sup>11</sup>, dall'altra parte il gruppo di ricerca di Stanford, guidato da John McCarthy, si concentrava su metodi di uso generale per il ragionamento logico<sup>12</sup>.

In primo luogo, venne definito da John McCarthy il linguaggio di programmazione di alto livello Lisp<sup>13</sup>, destinato ad assumere un ruolo primario nei successivi trent'anni. In secondo luogo, lo stesso scienziato presentò in un articolo una proposta concettuale per sistemi di IA basati su conoscenza e ragionamento. Si trattava dell'*Advice Taker*, un programma ipotetico che avrebbe racchiuso in sé la conoscenza generale del mondo e avrebbe potuto usarla per ricavare piani d'azione. Il concetto veniva esemplificato dimostrando che dei semplici assiomi logici erano sufficienti per generare un piano che consentiva di guidare un'automobile fino all'aeroporto. «Il programma era anche progettato in modo da accettare nuovi assiomi durante la normale esecuzione, acquisendo così competenze in nuove aree senza essere riprogrammato»<sup>14</sup>. Non può non notarsi come *Advice Taker* incarnasse il principio fondamentale della rappresentazione della conoscenza e del ragionamento. Questo articolo influenzò inevitabilmente il corso dell'IA e, tutt'oggi, mantiene la sua rilevanza.

---

ipotesi, la contemporanea e frequente attivazione di un gruppo di neuroni che dà luogo al rafforzamento sinaptico ha conseguenze funzionali importanti. I neuroni fortemente connessi probabilmente agiscono insieme, come un'unità funzionale. Se vengono attivati solo alcuni dei neuroni, si attiverà l'intero gruppo, a causa delle forti connessioni tra i membri del gruppo stesso. Se l'insieme è attivo, i suoi neuroni sono attivi simultaneamente, o mostrano, per es., schemi di funzionamento sincronizzati in modo preciso, quando l'attività neuronale si diffonde nell'insieme. Gli insiemi hebbiani di cellule si possono definire come unità funzionali composte da molti neuroni che si formano in una rete associativa, la corteccia, come risultato di una frequente attività neuronale simultanea che causa un rafforzamento sinaptico. Negli ultimi anni, l'idea hebbiana di insiemi distribuiti con topografie corticali definite è stata incorporata nelle teorie neuronali a grande scala del linguaggio e di altre funzioni cognitive». Così è spiegato in *Treccani, Enciclopedia della Scienza e della Tecnica*, 2008.

<sup>11</sup> Al MIT Minsky supervisionò un gruppo di studenti che scelse di occuparsi di problemi circoscritti la cui soluzione sembrava richiedere un certo grado di intelligenza. Questi domini limitati divennero famosi come *micromondi*. Il micromondo più famoso era quello dei blocchi, che consisteva in una serie di blocchetti solidi disposti su un tavolo, o più spesso la sua simulazione. Obiettivo tipico in questo mondo era sistemare i blocchi in una certa configurazione usando un braccio meccanico che poteva prendere un blocco per volta. Il mondo dei blocchi fu l'ambiente sperimentale per progetti di visione artificiale dedicati alla visione e alla propagazione dei vincoli.

<sup>12</sup> John McCarthy definiva l'IA come «una macchina in grado di comportarsi secondo modalità che potrebbero essere definite intelligenti ove fosse un essere umano a comportarsi nella medesima maniera».

<sup>13</sup> La caratteristica principale di questo linguaggio di programmazione è la rappresentazione dei dati attraverso liste, che vengono utilizzate sia per il codice che per i dati stessi. Lisp è noto per la sua semplicità sintattica e per la sua flessibilità nel manipolare il codice stesso come dati, consentendo la creazione di macro e funzioni ad alto livello che possono modificare il comportamento del linguaggio stesso.

<sup>14</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 22.

Nonostante gli enormi sforzi e gli ottimi risultati, il fenomeno dell'intelligenza artificiale avrebbe conosciuto il suo boom solo alla fine degli anni '80 del secolo scorso. Fino a quel momento le sperate prospettive di operatività dell'IA si scontravano con almeno due ostacoli. Innanzitutto, molti dei primi sistemi di IA si basavano essenzialmente su una sorta di introspezione informata riguardo il modo in cui gli umani svolgono un'attività, anziché un'attenta analisi dell'attività in questione delle caratteristiche di una soluzione e di che cosa avrebbe dovuto fare un algoritmo per produrre in modo affidabile tali soluzioni. In secondo luogo, i sistemi erano incapaci di capire l'intrattabilità di molti dei problemi matematici che l'IA stava cercando di risolvere. La maggior parte dei programmi di risoluzione dei problemi operavano provando varie combinazioni di passi fino a trovare la soluzione. Se questa strategia poteva funzionare fintanto che i micromondi contenevano pochi oggetti, e quindi le azioni possibili erano molto poche e le sequenze risolutive corte, altrettanto non poteva fare di fronte a micromondi complessi. Questi approcci sono stati chiamati «metodi deboli» perché, sebbene fossero generali, non erano in grado di scalare verso l'alto e risolvere istanti ed i problemi molto grandi o difficili. A torto, infatti, prima dello sviluppo della teoria della complessità computazionale<sup>15</sup>, si riteneva che la scalabilità verso problemi di maggiori dimensioni fosse solo una questione di potenza dell'hardware e grandezza delle memorie<sup>16</sup>. Diversamente, era necessaria una conoscenza più influente e specifica del dominio, che permettesse di intraprendere passi di ragionamento più ampi e potesse gestire più facilmente i casi tipici di aree ristrette di esperienza. In altre parole, per risolvere un problema complesso era necessario quanto possibile conoscere la soluzione.

Tale differente approccio «a conoscenza intensiva» fu adottato dal gruppo di ricerca guidato da Ed Feigenbaum a Stanford per il progetto Dendral, uno dei primi sistemi cd. esperti per l'immensa conoscenza specialistica che possedevano.

L'importanza della conoscenza del dominio divenne evidente anche nel campo della comprensione del linguaggio naturale tanto che si diffusero diversi

---

<sup>15</sup> La teoria della complessità computazionale è un ramo dell'informatica teorica che si occupa dello studio della quantità di risorse computazionali (come tempo e spazio) necessarie per risolvere problemi algoritmicamente. Ci sono due aspetti principali nella teoria della complessità computazionale: la complessità temporale (si concentra sul tempo richiesto per eseguire un algoritmo in funzione delle dimensioni dell'*input*). Si misura spesso tramite la notazione della «Grande O» che fornisce un limite superiore alla crescita dell'algoritmo rispetto alle dimensioni dell'*input*); la complessità spaziale (si riferisce allo spazio di memoria richiesto da un algoritmo in funzione delle dimensioni dell'*input*).

Sostanzialmente, la teoria della complessità computazionale si occupa di classificare i problemi in base alla loro complessità intrinseca. Ad esempio, alcuni problemi potrebbero essere risolti in tempo polinomiale (P), il che significa che l'algoritmo può risolverli in un tempo ragionevole anche per *input* grandi. Altri problemi, come quelli all'interno della classe NP (non deterministico polinomiale), possono richiedere un tempo esponenziale per essere risolti con gli algoritmi attualmente noti. Il problema centrale della teoria della complessità è il «problema P vs NP», ovvero se i problemi facilmente verificabili (appartenenti a NP) possano essere anche facilmente risolti (appartenere a P). Questo è uno dei problemi più importanti e ancora irrisolti nell'informatica teorica. Questa teoria è cruciale poiché fornisce linee guida per comprendere i limiti intrinseci della computazione e per sviluppare algoritmi efficienti che risolvono problemi in modo pratico.

<sup>16</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 24.

strumenti per la rappresentazione della conoscenza e del ragionamento, alcuni basati sulla logica come Prolog, altri basate su strutture denominate *frame*<sup>17</sup>.

Ad ogni modo, solo dopo la prima metà del 1980 l'industria dell'IA conobbe uno sviluppo senza precedenti anche grazie al ritorno delle reti neurali e alla reinvenzione degli algoritmi di apprendimento basati sulla retropropagazione<sup>18</sup>. Questi modelli chiamati connessionisti furono considerati da molti in diretta opposizione, sia i modelli simbolici di Newell e Simon, sia all'approccio logicista di McCarthy e altri. La fragilità dei sistemi esperti portò a un nuovo approccio, più scientifico, basato sulla probabilità più che sulla logica booleana<sup>19</sup> sull'apprendimento automatico più che sulla programmazione manuale e sui risultati sperimentali più che su affermazioni filosofiche. Divenne più comune fondarsi su teorie esistenti anziché proporre di interamente nuove, basare le affermazioni su teoremi rigorosi o su una solida metodologia sperimentale anziché sull'intuito e mostrare l'attinenza dei metodi alle applicazioni del mondo reale più che a esempi simili a giochi<sup>20</sup>.

Negli anni '80 hanno prevalso gli approcci basati sui modelli nascosti di Markov (HMM, *Hidden Markov Models*). Questi si fondano su una rigorosa teoria matematica tanto da permettere ai ricercatori di avvalersi dei risultati ottenuti in decenni di lavoro in altri campi. Inoltre, questi modelli sono generati mediante

---

<sup>17</sup> Il metodo appartiene a Minsky. Fondamentalmente venivano raccolti fatti che riguardavano particolari tipi di oggetti ed eventi e, successivamente, venivano organizzati tali tipi in un grande gerarchia analoga alla tassonomia biologica.

<sup>18</sup> Si tratta di un tipo di modello di apprendimento automatico ispirato al funzionamento del cervello umano. Il concetto fondamentale è quello di regolare i pesi delle connessioni all'interno di una rete neurale in modo che la rete possa imparare a svolgere un determinato compito. Il funzionamento consta di più fasi: la prima consiste nell'«inoltro» (*Forward Pass*). Durante questa fase, i dati vengono inseriti nella rete neurale e vengono effettuate le predizioni. I dati di *input* passano attraverso i vari strati della rete (strato di *input*, strati nascosti e strato di *output*) producendo un *output* predetto; la seconda attiene al «calcolo dell'errore». Viene calcolata la differenza tra l'*output* predetto e l'*output* desiderato utilizzando una funzione di costo o perdita; la terza fase concerne la «retropropagazione dell'errore». L'errore calcolato viene propagato all'indietro attraverso la rete neurale. Questo processo coinvolge il calcolo dei gradienti delle funzioni di perdita rispetto ai pesi delle connessioni all'interno della rete. Si utilizza il calcolo del gradiente per determinare in che misura ciascun peso contribuisce all'errore complessivo; infine, vi è la fase dell'«aggiornamento dei pesi». Utilizzando i gradienti calcolati, vengono regolati i pesi delle connessioni all'interno della rete neurale utilizzando un metodo di ottimizzazione come la discesa del gradiente stocastica. Questo passaggio aiuta a minimizzare l'errore complessivo della rete.

L'intero ciclo di operazioni (inoltro, calcolo dell'errore, retropropagazione e aggiornamento dei pesi) viene ripetuto iterativamente durante l'addestramento della rete neurale finché l'errore complessivo non raggiunge un livello accettabile o stabile.

<sup>19</sup> La logica booleana è un sistema formale che si basa sul lavoro del matematico inglese George Boole. Questo sistema si occupa di operazioni logiche a due variabili che possono essere solo in uno dei due stati: falso o vero. Queste variabili vengono spesso rappresentate utilizzando, rispettivamente, 0 e 1, o usando le parole chiave «false» e «true» (o abbreviate come «F» e «T»).

Le operazioni logiche fondamentali nella logica booleana includono: la negazione (*NOT*, questa inverte lo stato logico di una variabile. Se la variabile è vera, la negazione la rende falsa e viceversa. Es.:  $NOT(true) = false$ ,  $NOT(false) = true$ ); la coniugazione (*AND*, restituisce vero solo se entrambe le variabili coinvolte nell'operazione sono vere. Es.:  $true AND true = true$ ,  $true AND false = false$ ,  $false AND false = false$ ); la disgiunzione (*OR*, restituisce vero se almeno una delle variabili coinvolte nell'operazione è vera. Es.:  $true OR false = true$ ,  $true OR true = true$ ,  $false OR false = false$ ); la disgiunzione esclusiva (*XOR*, restituisce vero solo se esattamente una delle variabili coinvolte nell'operazione è vera, mentre se entrambe o nessuna sono vere, restituisce falso. Es.:  $true XOR false = true$ ,  $true XOR true = false$ ,  $false XOR false = false$ ).

<sup>20</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 27.

un processo di apprendimento basato su una grande mole di dati reali e ciò assicura prestazioni robuste. Gli *HMM*, infatti, hanno dato prova di affidabilità sempre maggiore nei test rigorosi condotti negli ultimi anni fino a dominare il campo del riconoscimento vocale così come il campo attiguo relativo al riconoscimento della scrittura manuale tanto da attuare la transizione verso un uso quotidiano da parte dell'industria e degli utenti privati.

Il 1988 fu un anno cruciale. Avvenne il collegamento tra l'IA e altri settori, quali la statistica, la ricerca operativa, la teoria delle decisioni e la teoria del controllo, la robotica, la visione artificiale, apportando benefici sia in termini applicativi sia in termini di migliore comprensione teorica dei problemi centrali dell'IA. Il libro *Probabilistic reasoning in Intelligence System* di Judea Pearl portò a una nuova considerazione delle probabilità e della teoria delle decisioni nell'IA. Lo sviluppo delle reti bayesiane<sup>21</sup> da parte di Pearl portò a un formalismo rigoroso ed efficiente per rappresentare la conoscenza incerta e algoritmi pratici per il ragionamento probabilistico. Un altro contributo importante, nel medesimo anno, fu dato da Rich Sutton che mise in connessione l'apprendimento per rinforzo<sup>22</sup> – usato nel programma per il gioco della dama di Arthur Samuel negli anni '50 – con la teoria dei processi decisionali di Markov (*MDP*, *Markov decision process*)<sup>23</sup> sviluppata nel campo della ricerca operativa.

Il summenzionato fermento scientifico permise notevoli passi in avanti. Ma tali passi non sarebbero stati gli stessi senza l'avvento e lo sviluppo, nei primi anni del nuovo millennio, del *World Wide Web* che ha facilitato la creazione di

---

<sup>21</sup> Le reti bayesiane, anche chiamate «reti di probabilità» o «reti di conoscenza probabilistica», sono modelli grafici che rappresentano le relazioni probabilistiche tra diverse variabili in un sistema. Queste reti sono basate sulla teoria delle probabilità bayesiane, che si occupa della modellazione delle incertezze e delle relazioni di dipendenza tra eventi.

Le reti bayesiane utilizzano una rappresentazione grafica sotto forma di grafi direzionati aciclici (DAG), che consistono in nodi e archi. Ogni nodo rappresenta una variabile e gli archi rappresentano le relazioni di dipendenza causale o condizionale tra le variabili. Ogni nodo può avere un insieme di genitori, ovvero i nodi da cui dipende direttamente. Ogni nodo in una rete bayesiana ha associata una tabella di probabilità condizionata (*CPT - Conditional Probability Table*), che rappresenta la probabilità che quel nodo assuma determinati valori dati i valori dei suoi genitori.

Le reti bayesiane sono utilizzate per l'inferenza probabilistica, il ragionamento causale, il supporto alle decisioni.

<sup>22</sup> Il «*reinforcement learning*», o apprendimento per rinforzo, è l'applicazione del *machine learning*, o apprendimento automatico, ai processi decisionali rivolti all'obiettivo attraverso un segnale numerico di ricompensa che si attiva quando l'obiettivo è raggiunto. Il rinforzo viene inserito nell'algoritmo di addestramento senza istruzioni dirette su quali azioni intraprendere: sarà l'agente a dedurlo, sulla base di ripetute interazioni con l'ambiente circostante.

<sup>23</sup> La teoria dei processi decisionali di Markov è un *framework* matematico utilizzato nell'intelligenza artificiale, nell'ingegneria dei sistemi e nella teoria del controllo per modellare decisioni in ambienti in cui gli esiti dipendono dalle azioni intraprese e sono soggetti a incertezza.

Un *MDP* è definito da: stati (che rappresentano le condizioni possibili in cui può trovarsi un sistema o un agente. Questi costituiscono il «panorama» in cui vengono prese decisioni); azioni (sono le scelte o le decisioni che un agente può compiere in uno stato specifico. Le azioni possono influenzare lo stato successivo); ricompense (sono i segnali di rinforzo o punizione che l'agente riceve per ciascuna transizione di stato. Possono essere positive, negative o neutre e possono essere assegnate in base alle azioni intraprese); probabilità di transizione (rappresentano le probabilità associate alle transizioni tra gli stati, ovvero la probabilità che, eseguendo un'azione in uno stato, si passi a un altro stato).

L'obiettivo principale in un *MDP* è massimizzare la «ricompensa cumulativa» nel lungo periodo. Gli agenti prendono decisioni basate su politiche, che sono strategie che specificano quale azione scegliere in ciascuno stato.

enormi quantità di dati, i cd. *data set* o, in altre parole, il fenomeno dei *big data*<sup>24</sup>, grazie ai quali è stato possibile sviluppare algoritmi di apprendimento più precisi ed efficienti. La disponibilità dei *big data* aiutò, tra l'altro, l'IA a recuperare appetibilità commerciale e concorse all'affermazione dell'apprendimento automatico come metodo prediletto.

Fu, infatti, l'affermazione del *machine learning* e, soprattutto, del cd. *deep learning* – di cui si tratterà nei prossimi paragrafi – che ha determinato l'era dell'IA per come la conosciamo oggi.

### 1.3. L'apprendimento automatico: elementi per un giurista.

#### 1.3.1. Gli algoritmi.

Per comprendere il metodo dell'apprendimento automatico è necessario partire dalla sua componente base, l'algoritmo.

«L'algoritmo può essere descritto come una sequenza di istruzioni ordinate in modo preciso e chiaro al fine di trasformare dati di partenza (*input*) in un qualche risultato (*output*), il quale va poi 'scritto' in un codice con uno dei linguaggi informatici disponibili, perché possa funzionare concretamente in una macchina»<sup>25</sup>.

Gli algoritmi sono sempre esistiti. Quel che è mutato, secondo i matematici, è che nel periodo più recente si è avuta la necessità di costruire e, nondimeno, rendere manifesti i meccanismi di calcolo in maniera tale che questi fossero, oltre ogni dubbio, certi ed effettivi in un dato tempo e spazio. Il concetto di algoritmo è, infatti, pre-euclideo e serviva, già da allora, a confrontare tra loro due grandezze<sup>26</sup>.

Rimanendo su un piano generale, si distinguono, da un lato, gli algoritmi «condizionali o deterministici», dall'altro lato, gli algoritmi «non deterministici».

Nei primi, tutte le istruzioni, gli *input* e gli *output* attesi, nonché i passaggi necessari richiesti per produrre il risultato ed i criteri sono forniti ex ante dal programmatore. Essi presentano una logica lineare e, salvo difetti di progettazione o funzionamento, dovrebbero essere leggibili *ex post*. Il punto centrale di questi algoritmi è che la capacità degli stessi di assumere decisioni, piani o inferenze discende dall'applicazione di regole predeterminate in fase di programmazione dell'algoritmo stesso. In altri termini, l'algoritmo esegue comandi che sono stati previamente stabiliti e che costituiscono essi stessi il

---

<sup>24</sup> La definizione di *Big Data* si riferisce a dati che contengono una maggiore varietà, che arrivano in volumi crescenti e con più velocità. Questo concetto è anche noto come le tre V. In poche parole, i *big data* sono *set* di dati più grandi e complessi, provenienti soprattutto da nuove origini dati. Questi *set* di dati sono così voluminosi che il *software* di elaborazione dati tradizionale non è in grado di gestirli. Questi rappresentarono un fattore determinante nella vittoria del sistema Watson di IBM sui campioni umani nel gioco a quiz Jeopardy! nel 2011, evento che ebbe impatto sulla percezione dell'IA da parte del pubblico.

<sup>25</sup> Così parla G. LO SAPIO, *La black box: l'esplicabilità delle scelte algoritmiche quale garanzia di buona amministrazione*, in *Federalismi.it*, 2021, 16, p. 136-149, la quale rinvia alla lettura, per una semplificata ricognizione, di G. AVANZINI, *Decisioni amministrative e algoritmi informatici. Predeterminazione analisi predittiva e nuove forme di intellegibilità*, Napoli, 2020, p. 3-13.

<sup>26</sup> Sul punto v. P. ZELLINI, *La matematica degli dei e gli algoritmi degli uomini*, Milano, 2016.

modello procedurale del programma. Gli algoritmi condizionali non fanno altro, dunque, che applicare regole (informatiche). Dette regole sono inserite normalmente nel programma da uno o più esseri umani, i programmatori, pur se stanno emergendo forme di automazione per la scrittura di dette regole. Si parla, infatti, oggi della possibilità che gli algoritmi condizionali siano in parte o interamente programmati da altri programmi informatici. In ogni caso, ad oggi, tali algoritmi si basano su una serie di regole la cui corretta formulazione può essere previamente verificata da un essere umano. Perciò, il processo logico che caratterizza questi algoritmi è particolarmente in linea con il ragionamento giuridico<sup>27</sup>.

Diversamente avviene nei secondi, in cui la macchina ha un margine di autonomia, con livelli diversi a seconda del modello algoritmico utilizzato.

Prediligendo e sviluppando sempre più gli algoritmi «non deterministici» si è passati da una tecnologia prevalentemente di deduzione ad una di induzione. Si è così parlato di vero e proprio «cambio di paradigma», riprendendo la celebre espressione coniata da Kuhn nell'ambito della filosofia della scienza<sup>28</sup>.

Ai metodi basati sulla logica e sulla conoscenza sono stati preferiti, come si è visto nel paragrafo precedente, metodi basati sull'apprendimento automatico.

Tali due approcci distinguono, distintamente, due grandi «famiglie» ovvero i sistemi non simbolici ed sistemi simbolici.

I secondi sono caratterizzati dalla rappresentazione simbolica della conoscenza e da istruzioni predefinite prestabilite dai programmatori<sup>29</sup>. Tra questi, i più diffusi sono i cd. «sistemi esperti» in cui, in un ambito predefinito (es. in uno specifico settore medico), viene fornita alla macchina una conoscenza di base corrispondente a quella di «un esperto del settore». Ciò che consente di classificare tali modelli algoritmici come «intelligenti» (così distinguendoli dagli algoritmi «deterministici» privi di autonomia) è che essi mostrano un'abilità umana, simile alla capacità logico-deduttiva di ragionamento<sup>30</sup>; sono cioè capaci di dedurre, dai dati forniti, informazioni nuove, eventualmente producendo anche risultati in termini di probabilità se i dati sono incompleti. I sistemi simbolici hanno, pertanto, un grande vantaggio, in termini di «compatibilità» con le regole e con il ragionamento giuridico, articolato anch'esso su logiche argomentative presupposti/conseguenze o logiche inferenziali cause/effetto: il loro meccanismo è trasparente, poiché è possibile ripercorrere all'indietro il processo e comprenderne il meccanismo di funzionamento. D'altro canto, però, hanno anche un enorme svantaggio ovvero la non scalabilità. In altre parole,

---

<sup>27</sup> Sul punto si esprime G. CARULLO, *Decisione amministrativa e intelligenza artificiale*, in *Dir. Inf. Inform.*, 2021, III, p. 431-461.

<sup>28</sup> Si v. T.S. KUHN, *The structure of Scientific Revolutions*, Chicago, 1962. Per una visione più giuridica degli algoritmi si v. E. GIORGINI, *Algorithms and Law*, in *ItalJ*, 2019, I, p. 131-149.

<sup>29</sup> Si fa riferimento alla filosofia di progettazione dell'IA chiamata «tradizionale» o «good old fashioned AI» o «approccio pulito». Gli approcci all'IA ordinata sono di natura prescrittiva, il che significa che forniscono una spiegazione del loro funzionamento.

<sup>30</sup> L'evoluzione naturale viene utilizzata, di fatti, anche come fonte di ispirazione per determinare quali soluzioni di IA ottengono i risultati migliori, dato un particolare problema. L'approccio dell'evoluzione naturale, che si basa sulla teoria darwiniana dell'evoluzione e della sopravvivenza del più adatto, è prevalente nella costruzione di modelli predittivi.

all'aumentare delle dimensioni e della complessità dei problemi da risolvere, richiedono maggiori risorse per garantire la medesima precisione.

Diversamente, invece, operano i primi, i sistemi non simbolici. Questi poggiano sulla considerazione che la «conoscenza» è quella che si acquisisce con l'esperienza e l'apprendimento. Tale convinzione fa sì che ci si riferisca ad essi anche con la locuzione «IA trasandata» proprio per sottolineare il fatto che si tratti di tecniche meno nitide, in grado di individuare soluzioni approssimative, imprecise o parzialmente vere ai problemi, ma con un costo ragionevole di risorse. Tale IA utilizza strategie di inferenza per l'adattamento e l'apprendimento e le basa su processi biologici o naturali. L'IA trasandata è descrittiva, il che significa che rivela come si è arrivati a una soluzione (il processo per ottenere una soluzione) ma non il perché.

La differenza maggiore tra questi due metodi menzionati, quindi, è che il primo può spiegare perché una soluzione è stata suggerita, mentre il secondo può spiegare come è stata raggiunta ma non perché. Perciò, il primo modello ha una maggiore capacità esplicativa mentre il secondo una migliore capacità predittiva.

La predilezione odierna per l'utilizzo di algoritmi non deterministici ai fini predittivi, tuttavia, non sta a significare che la spiegazione sia diventata superflua e la ricerca della causalità obsoleta. Di contro, mette in evidenza la possibilità e la necessità di distinguere i due canali e di analizzare la specificità scientifica della previsione, con le sue forme, le sue procedure e i suoi problemi, che sono diversi da quelli della spiegazione causale.

Perciò, mentre l'ipotesi che guida l'approccio esplicativo è necessario che sia comprensibile<sup>31</sup>, nella modellistica predittiva la trasparenza è di secondaria importanza. E questo incide anche sulla modalità di utilizzo dei dati. Se nei modelli esplicativi l'utilizzo dei dati serve a convalidare la teoria ipotizzata, nei modelli predittivi dai dati stessi potrebbe inferirsi l'ipotesi teorica. Gli algoritmi qui utilizzano tutti i dati accessibili, anche quelli raccolti per altri fini, senza «ripulirli» per correggere registrazioni imprecise o distorte e senza selezionarli. Nell'elaborazione dei dati, non si cercano relazioni causali che confermino l'ipotesi perché l'ipotesi, appunto, non c'è. Si cercano, invece, associazioni e

---

<sup>31</sup> Anche se non si può mai applicare una generalizzazione ottenuta da una scoperta specifica ad altri casi diversi (il classico problema dell'induzione del filosofo David Hume), il calcolo delle probabilità fornisce un metodo rigoroso e una base razionale per estrapolare, da un insieme inevitabilmente circoscritto di osservazioni, una generalizzazione utile per tutti i casi. Diversamente, nell'approccio statistico ciò non è né possibile né necessario, perché è sufficiente prendere in considerazione un campione adeguato e sufficientemente grande da essere rappresentativo. I dati, in un certo senso, sono al servizio della teoria, cioè servono a convalidare l'ipotesi che spiega i fenomeni.

correlazioni<sup>32</sup> tramite l'uso di modelli decisionali<sup>33</sup> che dovrebbero consentire la formulazione di previsioni efficaci. Spesso sono gli stessi modelli ad essere difficilmente spiegabili. Metodi algoritmici come le reti neurali o le foreste casuali, ad esempio, spesso non sono interpretabili. Qui si radica, infatti, la tanto discussa questione della non trasparenza degli algoritmi ovvero il ben noto problema della *black box*<sup>34</sup>.

Nondimeno, gli algoritmi predittivi soffrono di ulteriori inconvenienti dovuti, in particolare, al fatto che il loro *training* si basa su una serie di dati di addestramento che provengono dal passato. Il che sembra scontrarsi con la loro

---

<sup>32</sup> E. ESPOSITO, *The future of prediction: from statistical uncertainty to algorithmic forecasts*, in *artificialcommunication.mitpress.mit.edu*, 2022. L'a. associa alcune caratteristiche dell'uso predittivo degli algoritmi di apprendimento automatico all'antica logica divinatória e preveggenete. I termini usati nella predizione algoritmica (correlazioni, schemi), l'idea di predizioni indipendenti da relazioni causali, il riferimento a strutture inaccessibili al ragionamento umano, richiamerebbero le tradizioni antiche e complesse delle società divinatorie, come quelle del Medio Oriente, della Grecia o della cultura cinese. Le società divinatorie si basavano sul presupposto che il mondo fosse governato da una logica cosmica e da un ordine di base che gli esseri umani, con le loro capacità limitate, non erano in grado di cogliere, proprio come oggi non riusciamo a comprendere pienamente le procedure degli algoritmi. La razionalità divinatória non era di tipo scientifico ma rituale, con l'obiettivo non di fornire spiegazioni ma di gestire una «conoscenza totale» che rimaneva inaccessibile. Come per gli algoritmi, l'obiettivo non era la comprensione del fenomeno, ma l'ottenimento di indicazioni per l'azione e la decisione. L'intero universo era considerato infinitamente significativo, articolato in una rete inesauribile di corrispondenze.

<sup>33</sup> Come già accennati nei paragrafi precedenti, i modelli decisionali sono strumenti utilizzati per prendere decisioni razionali ed efficaci in una varietà di contesti. Questi modelli spesso coinvolgono l'analisi di dati e informazioni per guidare le decisioni. Alcuni dei modelli decisionali più comuni sono: a) «albero delle decisioni»: questo modello rappresenta una serie di decisioni e conseguenze sotto forma di un diagramma ad albero. Ogni nodo rappresenta una scelta e i rami rappresentano le possibili conseguenze delle decisioni. È spesso utilizzato per problemi di classificazione e categorizzazione; b) «rete bayesiana»: è una rappresentazione grafica delle relazioni probabilistiche tra diverse variabili. È utile per modellare l'incertezza e le dipendenze tra le variabili e può essere utilizzata per fare previsioni basate su dati e conoscenze pregresse; c) «programmazione lineare e programmazione lineare intera»: questi modelli sono utilizzati per risolvere problemi di ottimizzazione, come ad esempio la massimizzazione o la minimizzazione di una funzione-obiettivo soggetta a vincoli. La programmazione lineare si concentra su funzioni lineari, mentre la programmazione lineare intera si applica quando le variabili devono essere valori interi; d) «analisi multicriterio»: questo modello è utile quando una decisione deve essere presa in base a più criteri o obiettivi. Vengono assegnati pesi ai diversi criteri e viene valutata la migliore opzione in base a tali pesi; e) «teoria dei giochi»: questo modello è utilizzato per prendere decisioni in situazioni in cui il risultato dipende dalle azioni di più parti interessate. La teoria dei giochi analizza le strategie e le interazioni tra le diverse parti per determinare le scelte ottimali; f) «processo decisionale di Markov (MDP)»: questo è un modello utilizzato in problemi di decisione sequenziali. Si basa sulla teoria delle catene di Markov e viene spesso utilizzato in problemi di ottimizzazione del percorso, robotica e gestione delle risorse; g) «analisi delle opzioni reali»: questo modello si applica alle decisioni di investimento e business, consentendo di valutare opzioni di investimento in base alle varie scelte future possibili; h) «simulazioni»: sono modelli che riproducono il comportamento di un sistema nel tempo, consentendo di esaminare come diverse decisioni potrebbero influenzare il sistema nel lungo termine; i) «logica fuzzy»: questo modello è utile quando si deve trattare l'incertezza o la vaghezza nelle decisioni. La logica fuzzy consente di gestire concetti sfumati e valori intermedi; l) «modelli di *machine learning* per *decision making*»: l'uso di algoritmi di *machine learning*, come reti neurali, *support vector machines* e regressione, può essere incorporato nei processi decisionali per prevedere risultati futuri e aiutare a prendere decisioni informate.

La scelta del modello decisionale dipenderà dal contesto specifico e dalla complessità delle decisioni da prendere. Spesso, la combinazione di più modelli può essere utile per affrontare diverse sfaccettature di un problema decisionale.

<sup>34</sup> Si fa riferimento all'opacità (o il loro essere «scatole nere») di tali metodi di *machine learning*, ovvero il fatto che la complessità della loro struttura impedisca agli utenti umani di comprenderne e spiegarne appieno il funzionamento. Una opacità che rischia di compromettere il metodo scientifico, fondato sulla riproducibilità degli esperimenti e il nesso razionale tra causa ed effetto.

efficacia predittiva che, invece, dipende dalla loro capacità di ottenere buone prestazioni su dati inediti e futuri. I dati di addestramento e i dati reali, pertanto, sono tanto diversi quanto il passato è diverso dal futuro.

Per trovare un equilibrio tra due obiettivi parzialmente incompatibili è necessario, quindi, minimizzare gli errori che possono intervenire sia durante l'addestramento, sia durante il test di validazione. Se così non fosse si affronterebbero, partitamente, i frequenti problemi dell'*underfitting*<sup>35</sup> o dell'*overfitting*<sup>36</sup> algoritmico.

In particolare, tale ultimo inconveniente è stato definito «il problema dell'apprendimento automatico» quante volte in sua presenza il sistema si assesta su un'architettura più rigida e presenta una visione un po' autistica degli oggetti, perdendo la capacità di catturare la varietà empirica del mondo. Il sistema, in altri termini, si adatta eccessivamente agli esempi che conosce non riuscendo a distinguere efficacemente le informazioni rilevanti (i segnali) da quelle irrilevanti (i rumori). In termini sociologici, l'esperienza del passato rischia di compromettere l'apertura verso il futuro. E quanto più il sistema opera in condizioni complesse ed incerte, quanto più il rischio di *overfitting* aumenta sol si consideri che la componente di rumore tende ad aumentare più della componente di segnale. Quindi, il futuro tende a diversificarsi ancora di più dal passato.

Evidentemente, per evitare che la capacità predittiva dell'algoritmo sia condizionata dal passato e dalle sue idiosincrasie, bisognerebbe che il sistema «dimentichi» alcuni elementi del passato. Il problema si sposterebbe, però, sull'individuazione degli elementi da ignorare<sup>37</sup>. Perciò, per far fronte

---

<sup>35</sup> In caso di *underfitting* l'errore riguarda l'addestramento. L'algoritmo non elabora con successo gli esempi su cui è stato addestrato, ha prestazioni scarse e non è in grado di risolvere problemi complessi.

<sup>36</sup> In caso di *overfitting* l'algoritmo impara a lavorare bene sugli esempi che gli vengono forniti ma si irrigidisce rispetto a ogni variazione. Pertanto, la sua capacità predittiva si riduce e l'errore predittivo aumenta. L'algoritmo ha appreso gli esempi di addestramento talmente bene da diventare cieco a ogni nuovo elemento.

<sup>37</sup> Il problema è discusso in termini di relazione tra *bias* e varianza, dove il *bias* misura l'accuratezza del modello e la varianza la differenza tra le sue previsioni. Dal punto di vista sociologico, il *bias* corrisponde alla memoria e la varianza alla fantasia. I sistemi più complessi tendono ad avere un *bias* elevato, cioè una forte aderenza al passato. Questo pregiudizio non è necessariamente sbagliato (molti stereotipi hanno una base realistica), ma non è utile laddove restringa l'attenzione e impedisca ai modelli decisionali di vedere ciò che non corrisponde alle loro idee preconcepite. Sistemi più semplici sebbene meno accurati si dimostrerebbero, in certe situazioni, più aperti e, quindi, più capaci di affrontare l'imprevedibilità del nuovo.

Diverso è l'utilizzo del termine *bias* in termini di pregiudizio. In questo senso si tratta di un aspetto intrinseco della scienza dei dati e quindi delle tecnologie IA. In altre parole, nel momento in cui si maneggiano i dati, questi portano automaticamente con sé dei pregiudizi. L'atto di scegliere un *data set* piuttosto che un altro rifletterà potenzialmente un certo pregiudizio. I pregiudizi possono essere sia intenzionali che non. I *bias* sono presenti in quasi tutti i *data set* e i dati distorti porteranno invariabilmente a un risultato distorto da parte dei modelli addestrati su questi dati distorti.

I pregiudizi vanno inoltre distinti dalla discriminazione, che è un concetto legale che può essere descritto come «il trattamento pregiudizievole di un individuo in base alla sua appartenenza a un determinato gruppo o categoria», dove gli attributi che comprendono la discriminazione includono la razza, l'etnia, la religione, la nazionalità, il sesso, la sessualità, la disabilità, lo stato civile, le caratteristiche genetiche, la lingua e l'età. Di conseguenza, si dice che un modello è discriminatorio nelle situazioni in cui due individui hanno le stesse caratteristiche rilevanti per un processo decisionale, ma differiscono rispetto a un attributo sensibile, che porta a una decisione diversa prodotta dal modello. Perciò, *bias* e discriminazione si sovrappongono nella

all'*overfitting*, gli scienziati propongono di «rallentare» l'algoritmo imponendo errori casuali che impediscano all'apprendimento di diventare troppo accurato<sup>38</sup>.

### 1.3.1. Il machine learning.

Esempio emblematico di sistema non simbolico è il *machine learning*. In generale il *machine learning* concerne l'effettuare classificazioni e previsioni. Esso funziona attraverso la fornitura di quantità enormi di dati quali «esempi»<sup>39</sup> sui quali la macchina, supportata da una elevata capacità di calcolo, viene addestrata al fine di perseguire l'obiettivo assegnato, prima di entrare a regime. Grazie alla quantità di dati (*big data*) elaborati durante la fase di addestramento, gli obiettivi assegnati (es. riconoscimento immagini, riconoscimento vocale, interpretazione linguaggio naturale) vengono conseguiti con affidabilità sempre più elevate, poiché il sistema individua, sulla base dei dati immessi, correlazioni nascoste, fornendo risposte non spiegabili secondo logiche argomentative.

In particolare, il *machine learning* consiste nel creare programmi capaci, non tanto di reagire agli stimoli applicando regole inserite in precedenza dai programmatori, quanto di inferire regole a partire dagli *input* processati. Semplicisticamente, la differenza rispetto ai sistemi precedenti evoca la distinzione tra induzione e deduzione. Nel *machine learning* le macchine, inizialmente, sono scatole vuote capaci, però, di apprendere a partire dai dati, mentre negli altri sistemi queste sono dotate sin dall'inizio di regole generali e tecniche di manipolazione delle informazioni, ma hanno una ridotta capacità di apprendimento. Perciò, alcuni hanno parlato di vero e proprio «cambio di paradigma» (v. *supra*)<sup>40</sup>.

Esistono, poi, vari metodi di machine learning. I principali sono:

a. il *supervised learning* (o apprendimento supervisionato). In questo tipo di apprendimento, il modello viene addestrato su un insieme di dati di *input* e le corrispondenti etichette di *output*. L'obiettivo è far apprendere al modello una relazione tra *input* e *output*, in modo che possa generalizzare e fare previsioni accurate su dati non visti in fase di addestramento.

b. L'*unsupervised learning* (o apprendimento non supervisionato). In questo caso, il modello viene addestrato su un insieme di dati senza etichette ovvero non classificati in precedenza. L'obiettivo è identificare strutture e pattern nei dati, come cluster di punti simili, senza una guida esterna.

---

misura in cui la parzialità dei dati può portare ad effetti discriminatori, ma non necessariamente in tutti i casi.

<sup>38</sup> E. ESPOSITO, *The future of prediction*, cit., 87 ss.

<sup>39</sup> Si tratta dei cd. «*data set* di addestramento». Il buon adattamento al «*data set* di addestramento» ma la scarsa capacità di fare previsioni si chiama «compromesso *bias*-varianza». V. anche nota 37.

<sup>40</sup> Così spiega G. LO SAPIO, *La black box*, cit., p. 141-142. In tema di *machine learning* si v. E. ALPAYDIN, *Machine learning. The new AI*, Cambridge, MA, 2016 e N. ABRIANI, G. SCHNEIDER, *Diritto delle imprese e intelligenza artificiale*, cit., p. 43 ss.

c. Il *semi-supervised learning* (o apprendimento semi-supervisionato). Questo tipo di apprendimento combina elementi del *supervised* e *unsupervised learning*. Viene utilizzato quando si ha solo un insieme di dati etichettato e un insieme di dati non etichettato, cercando di sfruttare le informazioni non etichettate per migliorare l'accuratezza delle previsioni.

d. Il *reinforcement learning* (o apprendimento per rinforzo). In questo approccio, un agente impara a prendere decisioni interagendo con un ambiente. L'agente riceve ricompense o punizioni in base alle azioni che compie e il suo obiettivo è massimizzare le ricompense nel lungo termine.

e. Il *transfer learning* (o apprendimento trasferimento). Questo tipo di apprendimento coinvolge il trasferimento delle conoscenze apprese da un compito a un altro correlato, al fine di migliorare le prestazioni in quest'ultimo.

f. Infine, il *deep learning* (o apprendimento profondo). Si tratta di un sottoinsieme del machine learning che utilizza reti neurali artificiali profonde<sup>41</sup> per imparare rappresentazioni complesse dei dati. È ampiamente utilizzato in problemi come il riconoscimento di immagini, il riconoscimento del linguaggio naturale e altro ancora.

Oltre a questi, esistono anche approcci ibridi e combinazioni di tecniche di apprendimento che possono essere utilizzati in contesti specifici. La scelta del tipo di machine learning dipenderà dalla natura del problema e dalla disponibilità dei dati.

Per scegliere quello che si adatta di più alla situazione specifica si fa ricorso a tecniche statistiche che utilizzano i cd. «*data set* di valutazione». Una volta in possesso di tale *data set*, bisogna fare due cose: in primo luogo, stimare i parametri

---

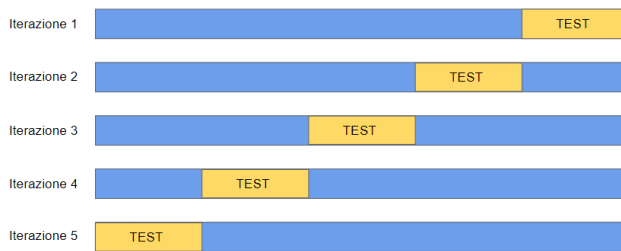
<sup>41</sup> Le reti neurali sono la frontiera attuale dell'intelligenza artificiale e, per meglio comprenderle, è bene considerarle come tecniche per l'approssimazione di funzioni. Il fenomeno può essere molto complesso e le funzioni possono avere molti argomenti in ingresso e restituire molti valori in uscita. Ad esempio, attraverso questo paradigma di *learning*, è possibile «dare vita» ad ambienti virtuali disponendo di pochi dati reali di partenza in modo che l'ambiente virtuale anticipi in modo molto verosimile il futuro andamento di un certo fenomeno, o sia in grado di simulare ciò che manca all'osservazione. I settori di applicazione sono i più svariati e vanno dalla diagnostica per immagini, al gaming, alla *computer graphics*.

Le reti neurali artificiali, sinteticamente, sono un'architettura modellata sui neuroni del cervello umano, con processi di apprendimento adattivi, utilizzati per il riconoscimento di modelli e in cui il *feedback* dall'ambiente è suddiviso in strategie di apprendimento supervisionate o non supervisionate.

È in questo contesto che nascono i termini «apprendimento profondo» o «reti neurali profonde», reti che imparano essenzialmente ricevendo dati e informazioni su questi dati (apprendimento supervisionato). L'architettura delle reti neurali profonde è costituita da molti strati di nodi, ai quali vengono inviati i dati. In genere, c'è il «livello di ingresso» (che accetta i dati alla rete), il «livello di uscita» (che fornisce l'*output*) e tra questi due livelli ci possono essere molti «livelli nascosti» (dove vengono eseguiti i calcoli matematici sui dati d'ingresso). Le reti neurali imparano proprio come i bambini: per insegnare a una rete neurale a identificare l'immagine di un gatto, le vengono somministrate migliaia di immagini di gatti, con le immagini di gatti «etichettate» come rappresentanti un gatto (ciò richiede che un essere umano insegni all'algoritmo l'aspetto di un gatto). Quindi, per testare la capacità della rete neurale di riconoscere le immagini di gatti, le vengono somministrate immagini di tutti i tipi di animali con il compito di identificare le immagini in cui sono presenti gatti. Quando la rete neurale viene informata di aver sbagliato, regola automaticamente la complessa struttura matematica di ponderazione dei suoi strati interni per «imparare» e aumentare la sua precisione in futuro. Quando la precisione è ritenuta sufficientemente buona, viene messa al lavoro nell'ambiente digitale per identificare i gatti con un certo grado di probabilità.

dei metodi di valutazione ovvero, usando un gergo più tecnico, occorre allenare l'algoritmo e scoprire la forma della linea rappresentabile sul piano cartesiano; in secondo luogo, occorre valutare quale metodo di *machine learning* si adatti meglio a quei dati. In altre parole, occorre scoprire se la forma della linea tracciata abbia una buona capacità di categorizzazione dei dati, quindi, in gergo informatico, occorre testare l'algoritmo.

Una delle tecniche migliori per scegliere la tipologia di *machine learning* da usare è la cd. validazione incrociata dei dati (o *k-fold cross validation*). Questa permette di escludere l'arbitrio nella scelta degli stessi. Di fatti, la validazione incrociata li



utilizza tutti, un blocco alla volta, e alla fine, ne riassume i risultati.

Come si può vedere nell'immagine, nella validazione incrociata ogni volta si utilizzano porzioni diverse per l'addestramento (il *train*)

e per la valutazione (il *test*). Il *train* è rappresentato dalla parte blu, mentre il *test* da quella gialla. Il risultato finale consiste in una media delle *performances* delle varie iterazioni.

Per descrivere le prestazioni di un modello di classificazione (o «classificatore») su un insieme di dati di *test* per i quali sono noti i valori reali viene spesso utilizzata la *confusion matrix* (anche detta *error matrix*).

		Predicted	
		Negative (N) -	Positive (P) +
Actual	Negative -	True Negative (TN)	False Positive (FP) Type I Error
	Positive +	False Negative (FN) Type II Error	True Positive (TP)

Questa è una tabella che permette la visualizzazione delle prestazioni di un algoritmo confrontando i risultati predetti

(*predicted*) con la realtà conosciuta (*actual*). Tramite la stessa si può calcolare la sensibilità e la specificità di un algoritmo.

Per sensibilità si intende la percentuale di successo dell'algoritmo di predire casi positivi effettivi tra tutti i casi positivi. Questa si ottiene, rapportando i risultati predetti coincidenti con la verità reale con la somma di tale dato con i risultati cd. falsi negativi<sup>42</sup>.

Per specificità, invece, si fa riferimento alla percentuale di successo dell'algoritmo di predire casi negativi effettivi tra tutti i casi negativi.

I due concetti non devono confondersi con quello di precisione dell'algoritmo che fa riferimento alla percentuale di casi positivi correttamente predetti.

Orbene, come si è visto nel paragrafo precedente, i più recenti modelli di apprendimento automatico, o *machine learning*, soffrono di un serio inconveniente che impedisce di affidarsi pienamente ad essi in contesti sensibili (come, ad

<sup>42</sup> I falsi negativi sono predizioni errate di segno contrario al dato reale.

esempio, quello medico-sanitario o legale): la loro opacità (o il loro essere cd. «*black box*») ovvero il fatto che la complessità della loro struttura impedisca agli utenti umani di comprenderne e spiegarne appieno il funzionamento. Detta opacità risulta ancora più di ostacolo nel settore della ricerca scientifica (dove, tra l'altro l'utilizzo di modelli di *machine learning* è molto frequente) laddove ne deriva una «influenza nefasta sul modo di condurre la scienza», come indicato in un recente *discussion paper* pubblicato congiuntamente dalla Royal Society e dall'Alan Turing Institute. In particolare, i summenzionati modelli interferirebbero negativamente con gli obiettivi di comprensione scientifica, obiettività e spiegazione, inclusa la riproducibilità degli esperimenti scientifici, che sono al centro della scienza moderna.

Di fatto, gli algoritmi di *machine learning*, per quanto prestanti e potenti, si basano esclusivamente sull'uso di euristiche statistiche applicate a dati osservazionali e non sono, di conseguenza, in grado di distinguere legami causali da semplici correlazioni ovvero, sommariamente, di constatare che, date due variabili di interesse, a ciascun valore della prima corrisponda regolarmente un valore della seconda. In effetti, come sottolineato dal celebre informatico Judea Pearl, l'inferenza causale richiede, da un lato, la capacità di intervenire sul fenomeno per testare specifiche ipotesi riguardo a potenziali nessi causali tra le variabili coinvolte e, dall'altro, quella di immaginare scenari controfattuali, tutte abilità che un'IA il cui addestramento è basato su euristiche statistiche non sarà mai in grado di svolgere, anche disponendo di ingenti quantità di dati.

Un modo per ovviare a questo problema è quello di incorporare le ipotesi causali nella procedura di addestramento sotto forma di conoscenza di *background*. Così facendo, si può garantire che il modello di *machine learning* generato dal processo di addestramento eseguirà solo previsioni basate su associazioni statistiche causalmente rilevanti.

All'uopo si stanno sviluppando varie tecniche per risolvere il problema dell'opacità. Tra queste il progetto *eXplainable Artificial Intelligence (XAI)* o *Explainable AI* ovvero un programma di ricerca in cui i processi decisionali degli algoritmi risultano comprensibili dagli umani<sup>43</sup>. Finora, gli sforzi profusi sono stati principalmente diretti allo sviluppo di strumenti adatti o alle esigenze di ingegneri e programmatori o a contesti come la presa di decisione e la governance dell'IA. Ciò non significa che alcuni degli approcci delle tecniche sviluppate dalla XAI non possano giocare un ruolo centrale anche nell'ambito della ricerca scientifica e, soprattutto, come vedremo, della determinazione della responsabilità civile. Le IA che non sono in grado di offrire una spiegazione sufficiente a soddisfare le richieste relative alla responsabilità rischiano di incontrare problemi, che vanno dalla difficoltà di ottenere un'assicurazione alla mancata accettazione sociale.

---

<sup>43</sup> Per approfondimenti sull'analisi della causalità con riferimento alla XAI cfr. P.H. PADOVAN, C.M. MARTINS e C. REED, *Black is the new orange: how to determine AI liability*, in *AI and Law*, Berlino, 2023, p. 133–167.

Fra gli altri metodi in corso di elaborazione sono degne di nota, poi, il cd. «approccio della *input features relevance*», e l'apprendimento automatico causale (*causal machine learning*).

Nel primo caso, si ha a che fare con tecniche che permettono, ad esempio, l'identificazione delle variabili che il modello decisionale ritiene contribuiscano maggiormente al verificarsi di un determinato fenomeno *target*. In questo modo si permette agli scienziati di ridurre significativamente il numero di variabili coinvolte e, quindi, di eseguire manipolazioni sperimentali più tradizionali che permettano, a posteriori, di comprendere l'interazione causale tra le variabili necessarie per produrre il fenomeno di interesse.

L'obiettivo centrale dell'apprendimento automatico causale è, invece, identificare automaticamente le variabili di *input* che sono causalmente rilevanti per una data previsione. Nel *causal feature learning*, sottoinsieme del *causal machine learning*, per esempio, si analizzano le distribuzioni di probabilità condizionata definite sulle variabili in *input* per determinare quali di esse abbiano un ruolo causale nel verificarsi di un dato fenomeno e quali no. Nei metodi di spiegazione controfattuale, invece, si utilizzano perturbazioni dell'*input* per rilevare determinate caratteristiche invarianti all'analisi controfattuale che, in quanto tali, potrebbero avere rilevanza causale per i fenomeni predetti. Nella *model-agnostic causality* applicata alle reti neurali profonde, infine, si utilizzano tecniche ispirate al cd. «*do-calculus*» di Pearl<sup>44</sup> per visualizzare il contributo causale di ciascun gruppo di nodi della rete rispetto ad un certo *output*<sup>45</sup>.

Orbene, in séguito all'applicazione di tali metodi, sarebbe possibile operare, quindi, una nuova classificazione dei modelli decisionali in base alla loro opacità distinguendo i modelli «*white-box*» - spiegabili chiaramente in termini di come si comportano, come producono previsioni e le variabili che li influenzano – dai modelli «*black-box*» - in cui, sebbene sia garantita una maggiore accuratezza previsionale, agli utenti appare disponibile e comprensibile la relazione *input-output*, ma non le ragioni o i processi sottostanti alla produzione dell'*output* -.

---

<sup>44</sup> Il «*do-calculus*» è un concetto introdotto da Judea Pearl, noto informatico e studioso dell'intelligenza artificiale di cui si è cennato nei paragrafi precedenti, che ha contribuito significativamente allo sviluppo del campo dell'apprendimento causale.

In breve, il «*do-calculus*» è un sistema formale di regole e operazioni che consente di manipolare le espressioni causali in modo coerente e corretto. È un'estensione della logica probabilistica che permette di trattare le azioni come eventi causali controllati dall'esterno e di ragionare sulle conseguenze di tali azioni nei modelli causali.

Le regole principali del «*do-calculus*» includono: a) regola del controllo inverso: questa regola consente di invertire il flusso di causa ed effetto in un modello causale per ragionare sulle conseguenze di un'azione controllata dall'esterno. Per esempio, è utile per rispondere a domande del tipo «Se facessi X, cosa succederebbe a Y?»; b) regola di eliminazione: questa regola permette di eliminare variabili dalla rappresentazione causale quando si considerano interventi controllati. Ad esempio, consente di rispondere a domande del tipo «Qual è l'effetto di X su Y dopo aver eliminato l'effetto di Z?»; c) regola del collaudo della scelta: questa regola consente di semplificare espressioni causali quando si ragiona sugli effetti di un intervento su una variabile; d) regola dell'aggiunta dei nodi: questa regola aiuta a introdurre variabili aggiuntive o a modificare la struttura causale per comprendere meglio gli effetti di interventi specifici.

<sup>45</sup> Si v. P.H. PADOVAN, C.M. MARTINS e C. REED, *Black is the new orange*, cit., p. 133-167. e A. CHITI, *Explainable AI: come andare oltre la black box degli algoritmi*, in *agendadigitale.eu*, 2022.

### 1.3.2. Il deep learning.

Il sottoinsieme del *machine learning* che più desta interesse ai nostri fini è il *deep learning*.

Come detto, questo pone l'accento sull'apprendimento di strati successivi di rappresentazioni di dati sempre più significative<sup>46</sup>.

Più nel dettaglio, i «sistemi di deep learning utilizzano un'architettura di modelli matematici ispirata alle reti neurali biologiche: le cd. reti neurali artificiali. Tale modello è costituito da un gruppo di interconnessioni di informazioni (si parla infatti di approccio di “connessionismo” al calcolo, contrapposto all'approccio simbolista): gli input trasmettono i segnali, ad una potenza ovviamente incomparabile con quella dei neuroni biologici, ai diversi nodi che costituiscono una rete complessa (*deep*) e nel corso dell'apprendimento, i ‘pesi’ di ciascun nodo vengono continuamente riparametrati, in un percorso non lineare e multistrato la cui ricostruzione però sfugge alla comprensione umana». «I sistemi di Deep Learning portano, quindi, alle estreme conseguenze il tendenziale disallineamento tra l'elaborazione informatica complessa di dati e il ragionamento giuridico, poiché in essi il “modello” viene costruito a posteriori dalla macchina che, quasi invertendo il metodo scientifico moderno che dalla ipotesi giunge alle tesi, individua, nella mole di dati a disposizioni, schemi ricorrenti, pattern, regolarità statistiche; la decisione è pertanto fondata sui dati e non vi è una spiegazione “ragionevole” del risultato perché il sistema non è geneticamente costruito per dare motivazioni»<sup>47</sup>.

## 2. Definizione e tassonomia dell'intelligenza artificiale allo stato dell'arte.

Come si è visto, in passato gli studiosi hanno indagato diverse versioni di IA. Alcuni hanno definito l'IA in termini di fedeltà alla prestazione umana, mentre altri hanno preferito dare una definizione formale di intelligenza come razionalità, o per dirla in parole povere «fare la cosa giusta»<sup>48</sup>. D'altro canto, alcuni considerano l'intelligenza artificiale una proprietà dei «processi di pensiero» e del «ragionamento», mentre altri si concentrano sul «comportamento» intelligente, con una caratterizzazione esterna<sup>49</sup>.

Dalle due dimensioni di umano *versus* razionale<sup>50</sup> e pensiero *versus* comportamento si ottengono quattro possibili combinazioni, per ognuna delle quali sono stati sviluppati programmi di ricerca con il sostegno di diversi gruppi di studiosi. I metodi usati sono stati necessariamente diversi: l'approccio che

---

<sup>46</sup> Come spiegato da F.C. HOLLET, *Deep learning with Python*, 2<sup>ed.</sup>, Shelter Island, New York, 2021, p. 7.

<sup>47</sup> Cfr. G. LO SAPIO, *La black box*, cit., p. 141-142 ma anche P. PERLINGIERI, *Sul trattamento algoritmico dei dati*, in *Tecn. dir.*, 2020, p. 181; U. RUFFOLO, *Intelligenza artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 2019, p. 19.

<sup>48</sup> Il direttore della ricerca di Google, Greg Corrado, afferma che l'IA consiste nel sapere che cosa fare quando non si sa che cosa fare.

<sup>49</sup> S. RUSSELL e P. NORVIG, *Intelligenza artificiale*, cit., p. 4.

<sup>50</sup> Non si suggerisce che gli umani siano irrazionali nel senso che siano privi di lucidità mentale ma si osserva che non sempre le decisioni umane sono matematicamente perfette.

persegue un'intelligenza simile a quella umana deve essere in parte una scienza empirica correlata alla psicologia, richiedendo osservazioni e ipotesi riguardo il comportamento umano e i processi di pensiero. Un approccio razionalista, invece, sfrutta una combinazione di matematica e ingegneria e si collega alla teoria del controllo e all'economia. I vari gruppi di studiosi si sono duramente contestati l'un l'altro, ma anche aiutati. Il risultato è che l'intelligenza artificiale, oggi, è una disciplina che non ha caratteri di unitarietà bensì è combinazione di varie scienze, applicate alla risoluzione di specifici problemi.

In questo senso può dirsi che si tratti di una disciplina in continuo divenire, dal momento che non si possono prevedere i problemi che le future esigenze dell'uomo genereranno.

Ad ogni modo, è comunque bene ora, al fine di una più ampia comprensione del tema trattato, compiere uno sforzo definitorio unificatore e stilare, per quanto possibile, una tassonomia dell'IA. Possono soccorrere in aiuto tutti quegli atti risultati di lunghe consultazioni, di *hard* o di *soft law*, sia internazionali, che europei o nazionali.

Con il termine IA si indica «una famiglia di tecnologie in rapida evoluzione che contribuisce al conseguimento di un'ampia gamma di benefici a livello economico, ambientale e sociale nell'intero spettro delle attività industriali e sociali». Così si afferma nell'ultimo testo di compromesso della proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (COM 2021/206, da ora in poi *AI Act*), al considerando n. 3<sup>51</sup>, dandone una definizione molto generica stante la storica complessità di individuarne una.

Come si ricorda nel preambolo del medesimo atto, la priorità del Parlamento europeo è quella di assicurare che i sistemi di IA in uso nell'UE siano sicuri, trasparenti, tracciabili e non discriminatori. Perciò, è necessario creare una definizione neutra e uniforme dal punto di vista tecnologico per l'IA, che si renda applicabile a tutti i futuri sistemi di IA.

La medesima amplissima definizione si riscontra anche nel considerando n. 1 della proposta di Direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (COM 2022/496 finale, d'ora in poi direttiva sulla responsabilità da intelligenza artificiale)<sup>52</sup>, ove dispone che l'intelligenza artificiale «è un insieme di tecnologie abilitanti che possono contribuire al conseguimento di un'ampia gamma di benefici nell'intero spettro delle attività economiche e sociali. Offre un grande potenziale di progresso tecnologico e rende possibili nuovi modelli imprenditoriali in numerosi settori dell'economia digitale».

---

<sup>51</sup> Con il voto del Coreper del 2 febbraio 2024 gli Stati membri dell'UE hanno approvato formalmente e definitivamente il testo dell'*AI Act* concordato dal trilatero del 6 dicembre 2023. Con tale ultimo testo si segna definitivamente la fine delle negoziazioni. L'approvazione da parte del Parlamento Europeo e la successiva pubblicazione del testo sulla Gazzetta Ufficiale sono previsti ad aprile 2024.

<sup>52</sup> Si fa riferimento alla Direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da IA), Bruxelles, 28 settembre 2022, COM(2022)496 finale in *eur-lex.europa.eu*.

Ancóra, il Comitato per l'intelligenza artificiale del Consiglio di Europa, nel suo recente *consolidated working draft* sulla Convenzione Quadro sull'intelligenza artificiale, i diritti umani, la democrazia e lo stato di diritto, del 7 luglio 2023, definisce un sistema di IA, ai fini della Convenzione, come qualsiasi sistema algoritmico o una combinazione di tali sistemi che utilizza metodi computazionali derivati dalla statistica o da altre tecniche matematiche e che genera testi, suoni, immagini o altri contenuti o che assiste o sostituisce il processo decisionale umano. Anche in questo caso la descrizione resta molto generica e malleabile in modo da permettere alla Conferenza delle Parti, se del caso, di dare un'interpretazione a questa definizione in modo coerente con gli sviluppi tecnologici pertinenti<sup>53</sup>.

Sul versante interno, in Italia, una definizione di IA è contenuta nelle Proposte per una Strategia italiana per l'intelligenza artificiale, elaborate da un gruppo di esperti MISE e pubblicate in data 2 luglio 2020<sup>54</sup>. Le Proposte fanno riferimento all'IA come alla «disciplina che si occupa dello sviluppo di sistemi software (spesso anche utilizzati in combinazione con hardware) che, dato un obiettivo complesso, sono in grado di agire nella dimensione fisica o virtuale, ragionare sulle conoscenze acquisite, e formulare decisioni, basate sull'evidenza raccolta, sulle migliori azioni da svolgere al fine di raggiungere l'obiettivo prefissato, anche in situazioni non esplicitamente previste a priori».

Una ulteriore definizione si rinviene, poi, sempre sul versante nazionale, nella sentenza del 25 novembre 2021, n. 7891 del Consiglio di Stato<sup>55</sup>. Ad avviso dei giudici amministrativi, un algoritmo sarebbe definibile di IA, e dunque distinguibile dagli algoritmi *tout court*, quando «contempla meccanismi di machine learning e crea un sistema che non si limita solo ad applicare le regole software e i parametri preimpostati (come fa invece l'algoritmo "tradizionale") ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni, secondo un processo di apprendimento automatico». La pronuncia specifica come, a differenza degli algoritmi di IA, il concetto di algoritmo *tout court* faccia riferimento «semplicemente una sequenza finita di istruzioni, ben definite e non ambigue, così da poter essere eseguite meccanicamente e tali da produrre un determinato risultato».

Anche nel diritto statunitense si riscontrano, infine, delle definizioni di IA.

---

<sup>53</sup> COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI), *Consolidated working draft of the framework convention on artificial intelligence, human rights, democracy and the rule of law*, Strasbourg 7 luglio 2023, CAI(2023)18. All'art. 3 si stabilisce che «For the purposes of this Convention, "artificial intelligence system" means any algorithmic system or a combination of such systems that uses computational methods derived from statistics or other mathematical techniques and that generates text, sound, image or other content or either assists or replaces human decision-making. The Conference of the Parties may, as appropriate, decide to give interpretation to this definition in a manner consistent with relevant technological developments».

<sup>54</sup> Proposte per una Strategia italiana per l'intelligenza artificiale, elaborate da un gruppo di esperti MISE e pubblicate in data 2 luglio 2020 reperibile in *mise.gov.it*.

<sup>55</sup> Cons. di Stato, 25 novembre 2021, n. 7891 in *foroplus.it*. Nel caso di specie, il Consiglio di Stato ha affrontato il problema della differente qualificazione degli algoritmi e degli algoritmi che fanno ricorso all'IA, nell'ambito dell'impugnazione di una procedura di gara relativa alla fornitura di peacemaker di alta fascia.

Il primo atto di riferimento è il *National Defense Authorization Act* del 2019 («*NDAA 2019*»), la legge federale annuale di stanziamento delle risorse destinate al Dipartimento della Difesa. La definizione di IA è formulata per la sola *Section 238* del provvedimento, dedicata alla ricerca sull'utilizzo dell'IA in ambito militare, che fa riferimento a «Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets. An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring».

Una seconda definizione di IA è contenuta nel *National Artificial Intelligence Initiative Act*20 («*NAIIA*»)<sup>56</sup>, adottato dal Congresso in data 12 marzo 2020 ed entrato in vigore il 1° gennaio 2021: «The term “artificial intelligence” means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to – (A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action».

La ragione che ha portato il legislatore statunitense a intervenire una seconda volta sul tema è duplice. In primo luogo, il *NDAA 2019* e il *NAIIA* sono atti con confini ed obiettivi diversi: mentre il primo concerne il Dipartimento della Difesa, il secondo ambisce a incentivare ricerca e sviluppo sull'IA. In secondo luogo va ricordato che, pochi mesi dopo l'entrata in vigore del *NDAA 2019*, una diversa definizione era stata recepita all'interno di una raccomandazione del Consiglio dell'OCSE: «An AI system is a machine-based system that can, for a given set of human defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy».

Orbene, volendo trarre delle conclusioni da quanto sin qui riportato, non può non notarsi che chiunque abbia tentato di dare una definizione di IA si sia dovuto assestare su termini molto lati, omnicomprensivi e permeabili ad eventuali sviluppi tecnologici.

Dal punto di vista della tecnica normativa, il *modus operandi* ricorda la metodologia delle clausole generali, i balconi aggettanti della regolamentazione, quei frammenti di norma capaci di far comunicare l'ordinamento giuridico con l'ordinamento sociale grazie all'ampiezza della loro «superficie calpestabile».

Questa volontaria carenza di precisione risulta, tuttavia, inevitabile e non può essere neanche qui superata sol si consideri che il passo del progresso scientifico è fin troppo lesto e per accompagnarlo si dovrebbe far fronte ad una modifica continua della definizione che comporterebbe unicamente incertezza normativa ed instabilità nel mercato di riferimento.

Ad ogni modo, volendo compendiare una definizione, si può dire che con la locuzione «sistema di IA» ci si riferisca ad ogni processo decisionale

---

<sup>56</sup> Il *National Artificial Intelligence Initiative Act*20 («*NAIIA*») è reperibile in *congress.gov*.

automatizzato che opera sulla base di algoritmi. E per dirlo con le parole proprie dell'art 3, comma I, n. 1) del proposto *AI Act*<sup>57</sup>, meglio parlare di sistema basato su una tecnologia progettato per operare con livelli di autonomia variabili, che può presentare un'adattabilità dopo l'impiego e che, per obiettivi espliciti o impliciti, deduce, dagli *input* che riceve, come generare *output* come previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.

Ora, se non è possibile ricostruire una definizione dettagliata di IA risulta, però, possibile prospettare diverse macro-classificazioni o «livelli» di IA.

Sulla base di un criterio generale fondato sulle capacità e sulle funzionalità delle macchine o dei sistemi si può distinguere, innanzitutto, l'IA debole (o stretta, *narrow*) dalla IA forte.

La prima rappresenta l'IA limitata a compiti specifici e ben definiti. Le macchine in grado di eseguire compiti di IA debole sono progettate per risolvere problemi specifici senza comprensione o coscienza. Esempi includono sistemi di raccomandazione, riconoscimento vocale, automazione industriale e giochi di scacchi.

La seconda rappresenta un livello avanzato di intelligenza artificiale che potrebbe superare le capacità umane in termini di comprensione, ragionamento e apprendimento. Questo tipo di IA è in grado di affrontare una vasta gamma di compiti complessi, comprendere il contesto e apprendere da nuove situazioni. Le IA forti sono, in sostanza, quelle che l'*AI Act* definisce ad alto rischio<sup>58</sup>.

In secondo luogo, è possibile segnalare l'IA generale (*Artificial General Intelligence, AGI* o *General Purpose Artificiale Intelligence GPAl*s) in qualità di IA teorica che ha la capacità di comprendere, apprendere e applicare conoscenze in

---

<sup>57</sup> L'ultima versione del testo originale di compromesso del 2 febbraio 2024 ha, innanzitutto, confermato la preferenza di non far riferimento alla parola *software*, forse perché troppo capace di condurre la mente umana dentro gli schemi della disciplina per quello predisposta. In secondo luogo, ha ripudiato l'eterointegrazione della definizione da parte del precedente Allegato I, probabilmente per agevolare gli interpreti nella ricostruzione. Infine, ma non per importanza, è stato eliminato il riferimento alla fissazione per mano umana degli obiettivi da raggiungere. Ecco, certamente, ciò riflette la rinnovata consapevolezza circa la possibilità del sistema di inferire autonomamente i propri obiettivi o di mutuarli da altri sistemi.

<sup>58</sup> Si rinvia, così, all'art. 6 *AI Act*: «Irrespective of whether an AI system is placed on the market or put into service independently from the products referred to in points (a) and (b), that AI system shall be considered high-risk where both of the following conditions are fulfilled: (a) the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation listed in Annex II; (b) the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II. 2. In addition to the high-risk AI systems referred to in paragraph 1, AI systems referred to in Annex III shall also be considered high-risk. 2a. By derogation from paragraph 2 AI systems shall not be considered as high risk if they do not pose a significant risk of harm, to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making. This shall be the case if one or more of the following criteria are fulfilled: (a) the AI system is intended to perform a narrow procedural task; (b) the AI system is intended to improve the result of a previously completed human activity; (c) the AI system is intended to detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review; or (d) the AI system is intended to perform a preparatory task to an assessment relevant for the purpose of the use cases listed in Annex III. Notwithstanding first subparagraph of this paragraph, an AI system shall always be considered high-risk if the AI system performs profiling of natural persons [...]».

una varietà di domini, simile alla capacità umana. Questo livello di IA non è limitato a compiti specifici e potrebbe mostrare comprensione e adattamento in molteplici contesti. Lo stesso art. 3, punto (44e) dell'ultimo compromesso interistituzionale dell'*AI Act* la descrive come sistema di IA basato su un modello di IA per scopi generali, che ha la capacità di servire una varietà di scopi, sia per l'uso diretto che per l'integrazione in altri sistemi di IA. A sua volta l'art. 3, punto (44b) definisce «modello di IA per scopi generali» un modello di IA, anche quando viene addestrato con una grande quantità di dati utilizzando l'auto-supervisione su scala, che mostra una significativa generalità ed è in grado di eseguire con competenza un'ampia gamma di compiti distinti, indipendentemente dal modo in cui il modello viene immesso sul mercato e che può essere integrato in una varietà di sistemi o applicazioni a valle. Questo non riguarda i modelli di AI che vengono utilizzati prima dell'immissione sul mercato per attività di ricerca, sviluppo e prototipazione.

Infine, si potrebbero raggruppare quelle IA, sia forti che deboli basate sulla coscienza. Questo livello teorico si riferirebbe a quei sistemi di IA che potrebbero sviluppare una coscienza o un'autocoscienza, consentendo una comprensione più profonda e un'interazione più simile a quella umana. Questa è, principalmente, un'area di speculazione filosofica e scientifica ma che, recentemente, è stata oggetto – e tutt'ora lo è – di un dibattito più pragmatico grazie all'intervento della IA cd. generativa. Con tale termine ci si riferisce a una categoria di modelli di IA progettati per generare dati nuovi e originali che condividono somiglianze con i dati di addestramento. Questi modelli cercano di catturare le caratteristiche e le strutture dei dati in modo da poter creare nuovi esempi che sembrano provenire dalla stessa distribuzione dei dati di origine. Un esempio notevole di IA generativa è l'apprendimento automatico generativo, che include approcci come le reti neurali generative (*Generative Adversarial Networks, GAN*) e le reti neurali ricorrenti generative (*Recurrent Neural Networks, RNN*). Di tale argomento ci si riserva una trattazione più ampia *infra*.

All'interno di queste categorie principali, potremmo distinguere, poi, diverse sottocategorie o settori specifici di ricerca e applicazione dell'IA, tra cui: l'apprendimento automatico (*machine learning*), di cui si è già ampiamente parlato insieme alle reti neurali artificiali; l'elaborazione del linguaggio naturale (*NLP*) ovvero quell'IA che si concentra sulla comprensione e la generazione del linguaggio umano da parte delle macchine; la visione artificiale ovvero l'IA che lavora per consentire alle macchine di interpretare e comprendere il mondo visivo attraverso l'analisi di immagini e video e la robotica ossia l'IA utilizzata per consentire ai robot di compiere azioni autonome e interagire con l'ambiente circostante.

In base ad un criterio diverso di classificazione basato su una matrice di utilità<sup>59</sup>, invece, si può incasellare l'IA in diversi livelli richiamando lo scopo perseguito.

---

<sup>59</sup> Cfr. E. HAHANA, *The AI Utility Levels Schema – Building an AI Classification*, in *lan.stanford.edu*, 2022.

Così, al I livello si potrebbero collocare le IA dirette a comprendere in quanto capaci di ragionamento automatico, apprendimento automatico. Ad es. si fa riferimento alle IA che si occupano di pubblicità comportamentale.

Al II livello si collocherebbero tutte le IA capaci di percepire oltre che comprendere. Esse ragionano automaticamente ed operano sulla base dell'apprendimento automatico profondo grazie a reti neurali, anche convoluzionali (CNN) e ricorrenti (RNN). Ad es. si fa riferimento alle IA che giocano a scacchi, che acquistano beni o servizi in base a criteri preselezionati.

Al III livello si troverebbero, infine, le IA che riescono, in più, a predire attraverso un ragionamento probabilistico. Queste sono capaci di processare il linguaggio naturale, rappresentare quanto da loro appreso attraverso l'apprendimento automatico profondo e capire il contesto. Ad es. si fa riferimento alle IA che acquistano beni e servizi in base a modelli passati e predittivi o alle IA che compiono compiti fiduciari.

Alla luce di quanto sin qui riportato senza pretesa di esaustività e di tecnicismo, appare evidente quanto ampio, intricato, complesso ed instabile sia il mondo dell'IA. Di riflesso, è comprensibile la difficoltà che si incontra nella ricostruzione del paradigma di responsabilità in caso di danni cagionati da un'IA fornita di ampia autonomia, tema centrale di questo scritto.

In altre parole, il *punctum pruriens* dell'applicazione algoritmica complessa risiede nella allocazione del cd. rischio di autonomia, ovvero quello discendente dalle decisioni assunte dalla IA in maniera indipendente, il quale si concretizza quando gli agenti *software* si presentino come attori capaci di comportamento, di interazione<sup>60</sup>.

Da un punto di vista strettamente logico-matematico, il risultato a cui si giunge con tecniche di intelligenza artificiale è sempre chiaro e, di volta in volta, potrà consistere in una deduzione logica, un'ottimizzazione, una previsione<sup>61</sup>,

---

<sup>60</sup> G. TEUBNER, *Soggetti giuridici digitali*, Napoli, 2019, p. 37-38 e in nota n. 40 rileva che «sono soprattutto tre i rischi in materia di responsabilità derivanti dalla digitalizzazione: (1) il rischio di autonomia, che discende dalle decisioni indipendenti che l'agente software assume; (2) il rischio di associazione, che (3) il rischio di interconnessione, che consiste nell'agire non isolato, ma in stretta interdipendenza di una pluralità di computer in rete. [...] il rischio di autonomia è parecchio incisivo: scaturisce dalla condotta, in linea di principio imprevedibile, degli algoritmi con autoapprendimento. Sono necessarie forme alternative di contenimento del rischio rispetto a quelle già conosciute da tempo in tema di rischio da automazione. [...] A differenza di quest'ultimo, il rischio di autonomia si concretizza – come da più parti si sostiene: nella filosofia dell'informazione, nella actor-network theory [ANT: Teoria dell'attore-rete\*], nella teoria dei sistemi – quando gli agenti software si presentino come attori capaci di comportamento. [...] La teoria dell'attore-rete\* dissolve la distinzione fra soggetto e oggetto. L'azione è il risultato dell'interazione tra soggetti umani e oggetti mobilmente connessi in rete, all'interno della quale non si deve parlare di “attori”, ma di “attanti”, che possono essere anche non umani. Soprattutto, entro le reti, la connessione tra soggetti e oggetti produce continuamente degli ibridi: un uomo al volante della sua auto, ad es., non è un soggetto attivo più un oggetto inerte, ma va considerato un unico centro comportamentale, non isolabile dalla rete della circolazione, anch'essa in continua mutazione». In merito alla tematica si v. anche K. KNORR-CETINA e A.V. CICOUREL, *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro Sociologies*, London-Boston, 1981, p. 276 ss; J. LAW, *Power, Action, and Belief: A New Sociology of Knowledge?*, London-Boston, 1986; B. LATOUR, *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford, 2005, p. 21 ss.

<sup>61</sup> «Il compito degli algoritmi è quello di prevedere il futuro», così dichiarò nel 2013 Amit Singhal, ex capo di Google Search, convinto che, da quel momento in poi, la funzione principale dei motori di ricerca sarebbe stata quella di anticipare, di prevedere le informazioni di cui avremmo avuto bisogno piuttosto che

una classificazione<sup>62</sup>. A non essere sempre chiaro, però, è il risultato che si trae dall'analisi delle immense moli di dati in termini di conoscenza del mondo.

Sicuramente la potenza computazionale delle macchine è in grado di sistematizzare meglio le nostre categorie di pensiero, scoprire correlazioni non immediatamente visibili, individuare similarità e differenze più sottili. Addirittura, alcuni scienziati sostengono che sia anche possibile progettare intelligenze artificiali in grado di imparare in maniera così profonda da scoprire nuove regole a cui sono soggetti i fenomeni fisici e socio-economici, mai immaginate prima.

Tuttavia, come supportato da autorevoli voci della comunità scientifica, forse c'è un orizzonte non raggiungibile anche per le macchine. Di fatti, queste, ad oggi, sembrano incapaci di svolgere analisi controfattuali (del tipo «cosa sarebbe successo se») e, quindi, di individuare la causa ultima dei fenomeni<sup>63</sup>, necessitandosi quel tipo di creatività tipico degli uomini nel momento in cui effettuano una scoperta, non schematizzabile da nessun tipo di regola logico-matematica oggi esistente né prevedibile<sup>64</sup>.

Di certo, la macchina offre all'uomo la possibilità di risolvere problemi prima intrattabili, per via dell'esplosione combinatoria delle varianti a cui essi danno luogo, allargando progressivamente l'insieme dei problemi risolvibili. Non è però affatto detto che nel fare questo la macchina invada l'ambito del pensiero umano, come i più pessimisti ritengono.

Esisterebbero, perciò, così ragionando, due falsi problemi.

Innanzitutto, l'interrogativo non sarebbe tanto se la macchina possa pensare come l'uomo, sottraendo a questo spazi di autodeterminazione nella produzione creativa, ma se l'uomo sarà costretto, per effetto della presenza di agenti intelligenti, a ragionare come questi e, in questo, a vedere limitata la propria libertà. Guardando alla storia dell'uomo, infatti, non sembra astruso ipotizzare che l'utilizzo di tali tecnologie finisca per creare nuove forme di divisione tra uomini e uomini, nuove classi sociali - forse anche più nettamente definite di quelle attuali - in cui, da una parte, potrebbero esserci coloro che trovano il loro «posto sociale» nello spazio creato dal funzionamento delle macchine intelligenti, diventando in qualche modo «strumento della macchina» (ad esempio, alimentando le macchine con attività massive di *data entry*, oppure curando il buon funzionamento degli algoritmi); dall'altra parte, potrebbero esserci coloro che utilizzeranno la macchina come uno strumento, delegandole la fatica delle azioni più ripetitive e noiose.

In secondo luogo, la domanda se la macchina abbia o meno una coscienza porterebbe – secondo molti - fuori strada dal momento che, come Turing già affermava, ciò che davvero interessa è vedere il comportamento esterno della macchina. D'altronde, anche quando si interloquisce con un umano non si sa se

---

di rispondere alle *query* degli utenti. L'obiettivo dell'IA, sarebbe, quindi, più prevedere che comprendere il mondo.

<sup>62</sup> G. D'ACQUISTO, *Intelligenza artificiale*, Torino, 2021, p. 13.

<sup>63</sup> Sono in corso di sviluppo, come detto i metodi di *XAI* e *casual machine learning*. V. *supra*.

<sup>64</sup> G. D'ACQUISTO, *o.u.c.*

la sua risposta sia il risultato di un pensiero. Lo si può solo supporre. Con le macchine si farebbe la stessa *polite convention*.

## CAPITOLO II

### Inquadramento giuridico dell'Intelligenza Artificiale

SOMMARIO: 1. Introduzione. – 2. I soggetti giuridici: *essentialia*. – 2.1. Il concetto di persona. Dal diritto romano al diritto dell'Unione europea. – 2.2. Le persone giuridiche. – 3. La qualificazione giuridica dei sistemi di IA. – 3.1. IA come soggetto del diritto. – 3.1.1. Segue...secondo la teoria della finzione ovvero il criterio funzionale o della semplificazione giuridica. – 3.1.2. Segue...Secondo il criterio della *moral agency*. – 3.1.3. Segue...il *tertium genus* della personalità elettronica. – 3.2. IA come oggetto del diritto. – 3.2.1. IA come prodotto...nel d.lg. del 6 settembre 2005, n. 206 (Codice del Consumo italiano). – 3.2.2. Segue...nella nuova proposta di direttiva europea sulla responsabilità per prodotti difettosi (PLD).

#### 1. Introduzione.

Una volta esaurita una panoramica generale dell'IA da un punto di vista ontologico, risulta necessario, ai nostri fini, esaminare la stessa da un punto di vista più strettamente giuridico.

Di fatti, non c'è dubbio che parole chiave come *machine learning* o *artificial intelligence* siano concetti di interesse non solo nel campo dell'*Information Technology* (IT), ma anche del diritto.

Innanzitutto, essi identifichino un'area di ricerca specifica per quest'ultimo in cui l'obiettivo fissato è quello di comprendere meglio il processo di ragionamento e argomentazione giuridica, utilizzando modelli e tecniche computazionali, al fine di costruire sistemi intelligenti di informazione giuridica a supporto degli operatori del diritto, dei decisori e dei cittadini<sup>1</sup>. In questo senso, il primo ad individuare una stretta relazione tra informatica e diritto è stato Lee Loevinger, giudice della Corte Suprema del Minnesota, il quale utilizzò per ciò il termine *jurimetric*<sup>2</sup>.

In secondo luogo, le medesime parole si collocano nella sconfinata area di quanto del mondo è soggetto a regolamentazione o dovrebbe esserlo.

Finora il diritto è stato sviluppato dall'uomo per l'uomo. Inizialmente, per regolare le relazioni tra le persone fisiche. Successivamente, per regolare le relazioni anche delle persone giuridiche o tra queste e le persone fisiche. Oggi ci si domanda se i sistemi di IA abbiano l'attitudine di modificare nuovamente il quadro giuridico.

In tale solco si radica lo spinoso dilemma circa la tecnica di regolamentazione di un sistema di IA ovvero se quest'ultimo debba essere trattato come centro di

---

<sup>1</sup> E. FRANCESCONI, *The winter, the summer and the summer dream of artificial intelligence in law. Presidential address to the 18th International Conference on Artificial Intelligence and Law*, in *Artificial Intelligence and Law*, 30, 2022, p. 147–161.

<sup>2</sup> Loevinger auspicava l'uso metodico dell'informatica per aiutare e verificare le scelte legali, partendo dall'applicazione delle leggi *antitrust*, anche per elaborare in modo efficiente l'enorme massa di dati economici che le agenzie dovevano analizzare per accertare più facilmente se determinate imprese fossero in posizione dominante nel mercato.

imputazione di situazioni giuridiche soggettive – quindi soggetto di diritto – oppure come oggetto di situazioni giuridiche altrui – quindi oggetto del diritto –. Come direbbe il filosofo Esposito<sup>3</sup>, i nostri ancestrali dubbi giuridici deriverebbero dell'incapacità di superare la divisione tra persone e cose tipica del diritto romano.

Il problema sorge, in particolare, davanti ai sistemi di IA cd. forti (v. *supra* cap. I, par. 2) laddove l'autonomia comportamentale si manifesta al suo massimo livello, tanto da innescare dubbi sull'esistenza di un'autocoscienza<sup>4</sup>. In altre parole, tali sistemi sono in grado di operare senza l'intervento umano, apparendo, così, per molti, entità «senzienti» e, quindi, autoresponsabili delle proprie azioni.

Così opinando, ecco che risulterebbe necessario riadattare il quadro giuridico esistente prevedendo una personalità giuridica dei robot o, quantomeno, una forma di soggettività giuridica<sup>5</sup> che permetta di imputare situazioni soggettive, attive o passive, direttamente in capo al sistema di IA.

---

<sup>3</sup> Il filosofo R. ESPOSITO elabora una teoria in base alla quale tutte le nostre precomprensioni politico-sociali e, quindi, giuridiche derivano dal «peccato originale» di non riuscire ad abbandonare la suddivisione, iniziata dai Romani, tra persone e cose. Per tutto il tempo a venire essa è restata vincolante: il soggetto che domina e l'oggetto che serve. Questo male primordiale della distinzione tra persone e cose è stato poi rinforzato e spalleggiato dalla doppia natura di Cristo, uomo e figlio di Dio. In forza di tali idee i cristiani ritengono che l'anima debba dominare sul corpo: e qui si torna, di nuovo, alla distinzione tra persona (anima) e cose (corpo). Tanto la struttura sociale gerarchica, quanto il doloroso rifiuto della corporeità proprio del Cristianesimo, sono conseguenze del peccato originale che sta nel dividere le persone dalle cose. Ma c'è una via per rimediare, ed Esposito ce la indica: il corpo. Il corpo rifiuta la suddivisione del mondo in persone e cose, tant'è che, in particolare, esso a lungo non è stato oggetto di diritto. Il corpo non è nulla di personale, giacché esso assomiglia piuttosto a una res omnium nella sua inviolabilità e inalienabilità; tuttavia non è nemmeno una cosa. Il corpo è, in quest'ottica, senza personalità e di conseguenza senza colpa e adatto ad affrancarsi da tutti i mali di cui è responsabile la distinzione tra persone e cose. Esso ha già e sempre offerto questa possibilità di salvezza nella politica, per la precisione con la configurazione della plebe come una parte del corpo politico. Sul punto si v. R. ESPOSITO, *Le persone e le cose*, Torino, 2014. Per una critica alla citata teoria si v. J.F. STAGL, *Da 'qualcosa' a 'qualcuno', da 'qualcuno' a 'qualcosa'. Percorsi esatti ed errati del concetto di persona*, in P. BUONGIORNO e S. LOHSE (a cura di), *Fontes iuris. Atti del VI Jahrestreffen Junger Romanistinnen und Romanisten (Lecce, 30-31 marzo 2012)*, Napoli, 2013, p. 87 ss.; ID. *Domani l'impersonale sarà il genere umano?*, in *REHJ*, 2016, XXXVIII, p. 377-391.

<sup>4</sup> Da Cartesio in poi la credenza che la coscienza presupponga il libero arbitrio è divenuta dominante. Aristotele era convinto che questo libero arbitrio esistesse anche negli animali. Ma la consistenza di tale libero arbitrio è sempre stata dibattuta. Secondo S. NICHOLS, *Folk concepts and intuitions: From philosophy to cognitive science*, in *Trends in Cognitive Sciences*, 2004, vol. 8, 11, p. 514-518, ciò che noi chiamiamo libero arbitrio non sarebbe altro che una serie di segnali bioelettrici, ovvero neuroni che si attivano in determinate aree cerebrali. Se così davvero fosse allora il meccanismo sarebbe riproducibile e, quindi, si aprirebbe uno spiraglio per poter dire che anche le IA hanno il libero arbitrio.

<sup>5</sup> Sulla differenza fra personalità giuridica e soggettività Cfr. P. PERLINGIERI, *La personalità umana nell'ordinamento giuridico*, Camerino-Napoli, 1972, p. 357; invece evidenzia che un tempo — nella polarità «soggetto/persona» — ci si interrogava intorno al rapporto tra soggettività e capacità e non sul primato della persona P. FEMIA, *Perlingieri Pietro*, in A. PAVAN (a cura di), *Enciclopedia della persona nel XX secolo*, Napoli, 2008, p. 819 ss. V. anche P. STANZIONE, *Capacità e minore età nella problematica della persona umana*, Camerino-Napoli, 1975, p. 55; C. PERLINGIERI, *Enti e diritti della persona*, Napoli, 2008; nonché M. D'AMBROSIO, *Partecipazione e attività. Contributo allo studio delle associazioni*, Napoli, 2012, p. 45 ss. Sull'attenzione della dottrina tradizionale esclusivamente per il profilo strutturale, cfr. V. FROSINI, *Soggetto del diritto*, in *Noviss. dig. it.*, XVII, Torino, 1970, p. 813 ss.; pone attenzione in modo innovativo alla questione dei nuovi «soggetti del diritto» G. TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, trad. it. di L. Zampino, Milano, 2015, che considera tale ampliamento necessario per comprendere la natura sociale di quei nuovi attori collettivi emergenti, come «le reti».

In tal modo si inserisce il dibattito sulla soggettività giuridica di un sistema di IA all'interno della questione della responsabilità. Rispondere nell'uno o nell'altro senso comporterebbe, infatti, la delineazione di paradigmi diversi di responsabilità.

Invero, laddove il sistema di IA sia qualificato come soggetto giuridico si potrebbe prospettare una responsabilità per colpa dello stesso agente intelligente.

Diversamente, qualora si preferisca evitare una tale qualificazione, bisognerebbe, in primo luogo, andare ad individuare il soggetto rimproverabile; in secondo luogo, determinare il criterio di imputazione del danno.

Allo stato attuale lo scenario globale è frammentato sul punto.

Se a livello macroscopico può dirsi che si scontrano due blocchi, l'ovest del mondo – di cui l'Italia fa parte – che è più determinato a non riconoscere una soggettività ai sistemi di IA autonomi, e l'est del mondo, che sta dimostrando maggiore propensione a concedere autonomia soggettiva, a livello microscopico le voci non sono univoche.

Sia nell'uno che nell'altro blocco si rinvergono correnti contrastanti.

Per quanto d'interesse, si focalizzerà l'attenzione sullo scenario occidentale.

Ad ogni modo, è bene ricordare che ad ogni scelta tecnico-normativa corrisponde una scelta assiologica. Di fatti, attribuire un certo *status* giuridico ai robot non potrebbe essere azione neutrale rispetto ai valori etici e ai diritti fondamentali caratterizzanti la società odierna, i quali potrebbero essere minati da passi arditissimi, compiuti sulla scia di un eccessivo entusiasmo.

Pertanto, gli attenti giuristi ed i politici garantisti non possono esimersi dal ricercare, ad ogni costo, un equilibrio tra i valori umani che stanno a cuore, come la *privacy*, la riservatezza, la trasparenza, l'identità e la libera scelta<sup>6</sup>, e gli usi irresistibili delle nuove tecnologie.

La nostra società non può permettersi di abbandonare questi valori in nome dell'innovazione e della convenienza.

## **2. I soggetti giuridici: *essentialia*.**

Il concetto di «soggetto» assume un ruolo centrale nell'ordinamento giuridico. Ad onore del vero, il termine «soggetti» insieme ai termini «individui», «norma» e «società» configura proprio l'essenza del diritto, non potendo esistere l'uno senza gli altri.

Come attenta dottrina suggerisce<sup>7</sup> forse è meglio parlare di «soggetti», al plurale, «distinguendo i problemi della persona umana dai problemi di quei

---

<sup>6</sup> A. MURRAY, *Almost Human: law and human agency in the time of artificial intelligence*, The Hague, 2021, osserva con preoccupazione che «gli agenti algoritmici» penetrano la libertà di scelta umana toccando ciascuno dei criteri identificanti l'autonomia decisionale, individuati da M. GAL, *Algorithmic Challenges to Autonomous Choice*, in *Michigan Telecom. Tech. Law Rev.*, 2018, 25, p. 59. Si tratterebbe del criterio dell'efficienza, del criterio psicologico e del criterio della libertà.

<sup>7</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, 4<sup>ed.</sup>, I, *Metodi e Tecniche*, Napoli, 2020, p. 6.

soggetti (sia pure diversificati per scopi e funzioni) che persone umane non sono: le cd. persone giuridiche e ogni altro centro d'imputazione soggettiva».

Per qualificare qualcosa o qualcuno come soggetto giuridico è primariamente necessario comprendere cosa si intenda per soggettività giuridica e quali siano i requisiti indispensabili.

Il dialogo protrattosi nel tempo tra filosofi e civilisti<sup>8</sup> ha permesso di precisare i contorni di vari concetti propri di ciascuna disciplina. Tra questi vi rientra quello di soggettività giuridica.

A Cartesio si deve la ridefinizione del diritto in termini individualistici<sup>9</sup> e, nello stesso tempo, quel «processo di generalizzazione e spersonalizzazione in cui la singolarità di ciascun individuo tende a scomparire dietro una massima astrazione dogmatica: il soggetto di diritto»<sup>10</sup>.

Tra il XVII e il XIX secolo, poi, l'idea di soggetto si impone e si afferma assieme a tutta una serie di corollari, la capacità giuridica, il diritto soggettivo, la personalità, i quali si sviluppano nella loro compiuta formulazione grazie alla pandettistica, al giusnaturalismo, al romanticismo della Scuola storica ed al razionalismo illuministico (v. *infra*).

Tuttavia, è proprio dovuta a Cartesio «l'indispensabile premessa verso l'esigenza di tradurre la variabilità ed opinabilità delle condizioni e delle possibilità esistenziali nell'unità ed univocità di una serie di essenze a priori che valgono sempre e comunque, per qualsiasi uomo, in qualsiasi circostanza»<sup>11</sup>. Si passa, così, da una soggettività soltanto vissuta ad una soggettività anche pensata che consente di ipotizzare la riduzione dell'uomo a concetto.

L'uomo, soggetto di diritto, è un'entità generica ed indifferenziata, «espressione attiva e riflesso passivo della legge come fenomeno altrettanto generale, indifferente ed astratto. L'uno e l'altra dovrebbero assorbire e circoscrivere la realtà nella sua globalità, senza residui e senza incertezze»<sup>12</sup>.

---

<sup>8</sup> Di là dalle risalenti opere di A. RAVÀ, *Diritti della Persona nella scienza e nella filosofia del diritto*, Torino, 1901 e di B. BRUGI, *Introduzione enciclopedica alle scienze giuridiche e sociali nel sistema della giurisprudenza*, 4a ed., Milano, 1907 (sul pensiero del quale v. G. MARINO, *Positivismo e Giurisprudenza*, Biagio Brugi alla congiunzione di scuola storica e filosofia positiva, Napoli, 1986), il dibattito, ravvivato dai contributi raccolti in P. PERLINGIERI (a cura di), *Soggetti e norma individuo e società*, Napoli, 1987, è rilanciato da N. IRTI e E. SEVERINO, *Le domande del giurista e le risposte del filosofo* (un dialogo su diritto e tecnica), in *Contr. impr.*, 2000, p. 665 ss.; IDD., *Dialogo su diritto e tecnica*, Roma-Bari, 2001; N. IRTI, *Norma e luoghi, Problemi di geo-diritto*, Roma-Bari, 2001; ID., *Le categorie giuridiche della globalizzazione*, in *Riv. dir. civ.*, 2002, p. 625 ss.; L. MENGONI, *Diritto e tecnica*, in *Riv. trim.*, 2001, p. 1 ss.; P. PERLINGIERI, *Dialogando con due filosofi, ermenenti del diritto*, in *Rass. dir. civ.*, 2001, p. 669 ss. [e in R. FAVALE e B. MARUCCI (a cura di), *Studi in memoria di V.E. Cantelmo*, II, Napoli, 2003, p. 421 ss.], con particolare attenzione al significato e alla portata del dialogo tra filosofi e giuristi in sede di «ermeneutica assiologica»; ID., *Praxeologismo, istituzionalismo e formalismo*, in *Fil. dir. umani*, 2001, 7-8, p. 56 ss. Cfr., inoltre, per un esempio, F. GALGANO, *Dialogo sull'equità (fra il filosofo del diritto e il giurista positivo)*, in *Contr. impr.*, 1996, p. 402 ss.

<sup>9</sup> Cartesio ritiene che l'unicità dell'uomo si sostanzia nell'anima immateriale che governa il corpo dalla ghiandola pineale, ove ha sede. Espressione ed essenza dell'anima sarebbe il pensiero o *res cogitans*. Nelle sue *Meditationes* coglie proprio nell'atto del pensiero il «punto d'Archimede» sul quale fondare ogni certezza. Il *Ego cogito, ergo sum*. Questa sarebbe una proposizione necessariamente vera.

<sup>10</sup> S. AMATO, *Il soggetto e il soggetto di diritto*, Torino, 1990, p. 11-12.

<sup>11</sup> *Ibidem*.

<sup>12</sup> *Ibidem*.

Ovviamente questa auspicata assenza di incertezze non si è mai concretizzata sol si considerino gli innumerevoli mutamenti sociali che hanno continuamente messo alla prova la resistenza dell'allora vigente quadro giuridico.

Ma si vada per gradi.

Come detto, il concetto di soggetto di diritto<sup>13</sup> è nato con riferimento all'uomo ed esprime, oggi, l'idoneità di questi ad essere titolare di diritti e di doveri e più in generale di situazioni giuridiche soggettive, tanto che secondo l'orientamento prevalente<sup>14</sup> vige l'equazione capacità giuridica – soggettività. Tale equazione sarebbe riconducibile a due tesi differenti: la teoria organica e le teorie cd. atomistiche.

Secondo la teoria organica – detta anche della realtà – gli elementi necessari per la costituzione di un soggetto giuridico sarebbero due: uno materiale ed uno formale. Il primo consisterebbe nell'organismo psico-fisico; il secondo nel riconoscimento giuridico della qualità di soggetto attraverso la nascita.

Secondo le teorie atomistiche, diversamente, il soggetto di diritto sarebbe un fascio di comportamenti normati. In altre parole, sarebbe una pura forma normativa tesa regolare i comportamenti mediante l'imposizione di doveri e il conferimento di diritti.

Ad ogni modo, oggi, possono tranquillamente farsi due affermazioni: innanzitutto, che ogni uomo – inteso come appartenente al genere umano – è soggetto di diritto; in secondo luogo, che all'affermazione precedente non si applica il principio commutativo.

Qualche precisazione è necessaria.

La prima affermazione, per quanto ad un lettore contemporaneo possa apparire scontata, non rappresenta una vuota statuizione. Si pensi al diritto romano, in cui gli schiavi, pur essendo persone fisiche, non erano considerati soggetti di diritto, ma erano posti sullo stesso piano delle cose (*res*). Ancóra, si pensi alle donne le quali, fino a poco tempo fa, non erano considerate soggetti dotati di piena capacità giuridica.

Il piano riconoscimento della soggettività giuridica ad ogni persona fisica, in definitiva, lo si è raggiunto solo con l'avvento della Costituzione nella quale l'uomo, in quanto «persona», ha trovato una posizione primaria rispetto ad ogni interesse patrimoniale<sup>15</sup>. Per vero, quanto descritto è il riflesso dell'elaborazione del concetto di personalità, da tenersi distinto quello di soggettività, indicando il primo un insieme di valori che l'ordinamento giuridico tutela anche prima della

---

<sup>13</sup> Per un'approfondita analisi del concetto di soggetto di diritto si v. S. COTTA, *Soggetto di diritto*, in *Enc. dir.*, XLII, Milano, 1990.

<sup>14</sup> Si v. P. RESCIGNO, *Capacità giuridica*, in *Noviss. dig. it.*, II, Torino, 1988; A. FALZEA, *voce Capacità (teoria generale)*, in *Enc. dir.*, VI, Milano, 1960, p. 8 ss.

<sup>15</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale*, I, cit., p. 11. Per approfondimenti si v. V. SCALISI, *Ermeneutica dei diritti fondamentali e principio «personalista» in Italia e nell'Unione europea*, in *Riv. dir. civ.*, 2010, I, p. 145 ss., e ID., *L'ermeneutica della dignità*, Milano, 2018, spec. p. 65, il quale riconosce alla persona umana quel «superiore valore co e centro di gravità dell'intero sistema di diritto positivo»; N. LIPARI, *Diritti fondamentali e ruolo del giudice*, in *Riv. dir. civ.*, 2010, I, p. 635 ss. (ora in ID., *Il diritto civile tra legge e giudizio*, Milano, 2017, p. 102 ss.).

nascita, il secondo un centro di imputazione di interessi meritevoli di tutela in termini strutturali.

## 2.1. Il concetto di persona. Dal diritto romano al diritto dell'Unione europea.

Volendo tracciare un profilo diacronico del termine «persona» per capirne l'evoluzione, occorre tornare al diritto romano arcaico ove con esso veniva indicata la maschera teatrale e, ancor prima, la maschera del defunto. La persona altro non era che il doppio del defunto. Per il diritto romano il concetto di persona non alludeva né alla soggettività giuridica, né alla capacità giuridica. Nell'ambito di tale ordinamento giuridico operava la sola figura di *status* volta ad indicare l'appartenenza dell'individuo ad una determinata comunità: quella degli uomini liberi (*status libertatis*), quella della *civitas* (*status civitatis*), quella della famiglia (*status familiae*).

Nel diritto intermedio, nell'ambito del quale la disparità di trattamento era la regola, in cui campeggiavano privilegi e prerogative, il concetto di *status* si rafforzò ed, anzi, si assistette alla moltiplicazione degli *status*.

La nozione unitaria di persona si affermò solo a partire dal XVII secolo grazie alla dottrina giusnaturalistica che, rifiutando l'idea di un diritto «originario» ha fondato la concezione di un diritto «convenzionale» che trae origine dallo Stato e si dipana dal principio di uguaglianza formale di tutti i sudditi di fronte alle legge.

La concezione unitaria della persona ha trovato, così, il suo formale riconoscimento giuridico nella «Dichiarazione dei diritti» delle colonie americane (1776) e, quindi, nella «Dichiarazione dei diritti dell'uomo e del Cittadino» (1789), frutto della Rivoluzione Francese, il cui art. 1 consacra il fondamentale principio di libertà ed uguaglianza di tutti gli uomini di fronte alla legge.

Tuttavia, il percorso evolutivo della nozione di persona ha subito un arresto con il Code Napoleon che, incentrato sull'istituto della proprietà e sulla figura del proprietario, ha perso di vista la definizione unitaria di persona, privilegiando un'ottica prettamente patrimoniale.

L'esperienza del Codice austriaco del 1811, ripresa, poi, ad un secolo di distanza, dal Codice svizzero, è stata differente. Qui è ricomparsa la centralità della persona e della sua capacità<sup>16</sup>, ancorché racchiusa in pure formule e categorie formali.

È la dottrina pandettistica, come si accennava *supra*, che ha attuato la prima vera e propria sistemazione del concetto di persona e di capacità.

Savigny per primo ha inquadrato il concetto di persona nell'ambito del rapporto giuridico ossia della relazione di una persona con un'altra persona.

L'uomo è persona, la persona è titolare del rapporto giuridico perché dotata di capacità giuridica, quindi ogni uomo è soggetto di diritto. Il diritto positivo può, però, modificare il «concetto primitivo di persona», limitando la capacità

---

<sup>16</sup> All'art. 16 Codice civile austriaco si afferma che «ogni uomo ha diritti innati che si conoscono con la sola ragione: perciò egli è da considerarsi persona».

giuridica, in tutto o in parte a singoli uomini, o estendendola alla persona giuridica.

Collocando la persona nell'ambito del rapporto giuridico, la sua concezione resta astratta, così come astratta e formale è la concezione di capacità giuridica che può essere limitata dal legislatore e che appare collegata essenzialmente al diritto patrimoniale.

Riprendendo Savigny, anche Puchta ha costruito la nozione di persona intorno al concetto della volontà, definendo più precisamente la personalità, come «la possibilità di una qualità giuridica, come qualità di un soggetto». In tal modo, per la prima volta, è stato elaborato il concetto di diritto della personalità come potere giuridico di essere persona. Il diritto spetta all'uomo quale soggetto del volere e, come tale, spetta a tutti gli uomini ad onta della loro naturale disuguaglianza.

Bisognerà aspettare Windscheid per collocare la persona nella trattazione del diritto in generale. Secondo tale visione il diritto consiste nella estrinsecazione della «facoltà di volere, la persona, la cui volontà viene dichiarata decisiva, è il soggetto del diritto». Persona è l'uomo che esterna una volontà, l'uomo «è capace di essere soggetto di diritti e di obblighi, di essere soggetto giuridico; egli è capace di diritto fin dalla sua nascita».

Verso la fine dell'Ottocento la figura della capacità costituiva un dato acquisito e, ormai, concepito quale frutto, creazione dell'ordinamento, avendo perso la connotazione giusnaturalistica conferitagli dalla prima dottrina pandettistica.

Volgendo lo sguardo all'Italia dello Statuto Albertino, prima, e del Codice civile 1865, poi, si può notare come mancasse qualsiasi sistemazione organica alla disciplina attinente sia i soggetti, sia la capacità. Bisognerà, invero, attendere il Codice civile del 1942 per una un'espressa regolamentazione della capacità nel titolo primo «Delle persone fisiche», sulla scia dell'esempio di alcuni codici, tra i quali quello tedesco e quello svizzero.

L'influsso della cultura germanica, in particolare, è convogliato nell'accoglimento della teoria della fattispecie, permeante l'intero nostro Codice civile. In tale prospettiva i fenomeni giuridici sono configurati come composti di due elementi: l'elemento materiale, dato dalla situazione di fatto esterna, e l'elemento formale, dato dal complesso delle regole che consentono la qualificazione in termini giuridici della fattispecie. I soggetti, in questo contesto, divengono i presupposti soggettivi della fattispecie, e quando la fattispecie ha prodotto i suoi effetti, tali effetti sono riferiti direttamente ai soggetti: i soggetti divengono in tal modo essi stessi fattispecie giuridiche.

Il soggetto giuridico è, dunque, fattispecie composta, del pari, dei due elementi, l'uno materiale (elemento esistente di fatto nella realtà extragiuridica), l'altro formale (la qualificazione normativa ovvero il riconoscimento formale da parte dell'ordinamento giuridico) [v. *supra* la teoria della realtà].

Il fondamento della personalità dei soggetti diviene, quindi, unico per le persone, mutando solamente il loro substrato materiale.

Come anzidetto, è con la Costituzione italiana del 1948 che è avvenuto il passaggio dal concetto di individuo inteso quale astratto soggetto di diritto alla nozione di individuo come «persona» e, con esso, il tramonto della concezione formalistica della persona medesima.

Come ben ha espresso Rodotà «[a]ttraverso la costruzione del soggetto astratto [è] stato possibile liberare formalmente la persona dalle servitù del ceto, del mestiere, della condizione economica, del sesso, che fondavano la società della gerarchia, della disuguaglianza»<sup>17</sup>.

L'art. 2 della Costituzione riconosce e garantisce i diritti inviolabili dell'uomo, sia come singolo, sia nelle formazioni sociali ove si svolge la sua personalità, al contempo, l'art. 3 enuncia il principio di eguaglianza formale e sostanziale.

Tali norme testimoniano l'esistenza di un principio generale di rango costituzionale volto alla tutela della persona e pongono le basi per la creazione di nuovi diritti della persona.

I valori costituzionali, considerati oramai alla stregua di principi direttamente operanti nei rapporti tra privati, hanno segnato un ampliamento della protezione dell'individuo attraverso il ricorso a clausole generali e a forme aperte, a scapito delle forme chiuse ed eccessivamente analitiche caratterizzanti le precedenti codificazioni civilistiche.

L'importante evoluzione del concetto giuridico di persona avutasi nel secondo dopoguerra è, poi, testimoniata dalle Convenzioni internazionali stipulate in quel periodo, tese a scongiurare le aberrazioni che da quel conflitto erano scaturite.

La persona, attraverso il riconoscimento dei suoi diritti fondamentali è venuta, così, a rivestire una posizione centrale sia nella Dichiarazione universale dei diritti dell'uomo approvata dall'Assemblea dell'ONU il 10 dicembre 1948, sia nella Convenzione europea per la salvaguardia dei diritti dell'uomo e delle libertà fondamentali, siglata a Roma il 4 novembre 1950.

Infine, ma non per importanza, la persona ha conosciuto una pregnante considerazione anche nell'ambito dell'ordinamento giuridico europeo. In una prima fase, infatti, i Trattati istitutivi del 1957 dettavano pressoché esclusivamente la disciplina dei soli rapporti economici transnazionali: il diritto comunitario regolava soprattutto l'attività economica, prendendo in considerazione la persona esclusivamente nella sua dimensione economica.

Grazie all'opera della Corte di Giustizia, prima, e alla Carta di diritti fondamentali dell'unione europea, siglata a Nizza il 7 dicembre del 2000, poi, hanno fatto ingresso nell'ordinamento giuridico comunitario i diritti sociali e alla persona si è iniziata a garantire una tutela anche sociale oltre che economica<sup>18</sup>.

Ecco, dunque, come si è giunti ad affermare che se ogni uomo è considerato persona ed ogni persona è ancestralmente titolare di diritti e doveri e, quindi, soggetto di diritto, è necessaria conseguenza che ogni uomo sia soggetto di diritto.

---

<sup>17</sup> S. RODOTÀ, *Dal soggetto alla persona*, Napoli, 2007, p. 13.

<sup>18</sup> Per la visione diacronica del concetto di persona v. F. CARINGELLA e L. BUFFONI, *Manuale di diritto civile*, VI ed., Roma, 2016, p. 79-82.

## 2.2. Le persone giuridiche.

Quanto all'operabilità del principio commutativo in relazione alla frase «ogni uomo è soggetto diritto» soccorre la constatazione che, nel tempo, la soggettività è stata attribuita anche ad altre entità diverse dall'uomo, diverse dalla persona fisica. Così, non ogni soggetto di diritto è uomo. Esistono, di fatti, le cd. persone «giuridiche»<sup>19</sup>.

Si badi che con l'utilizzo del termine «persona» accanto all'aggettivo «giuridica» non si vuole accostare l'istituto alla persona fisica per ricondurlo al medesimo genere di «persona», ma, piuttosto, a quello di «soggetto di diritto».

Secondo l'insegnamento tradizionale, infatti, la persona giuridica è un particolare soggetto di diritto, dotato di capacità giuridica, di capacità di agire, per mezzo dei suoi organi e di autonomia patrimoniale perfetta (v. *infra*).

L'aggettivo giuridica, di contro, sta ad indicare che la persona è tale per il diritto, in virtù del fatto che è considerata persona dalla legge. Le persone giuridiche «sono il frutto di una tecnica giuridica che attua una sorta di sdoppiamento della personalità, facendo dell'insieme una persona (giuridica) distinta dalle persone (fisiche o giuridiche) di ciascuno dei suoi componenti, sicché questi possono dirsi terzi rispetto alla persona giuridica, la quale si presenta per converso come terza rispetto ad essi»<sup>20</sup>.

La nascita di un concetto atto a definire entità diverse dagli individui fisici risale al XIV secolo, ad opera dei canonisti prima, dei commentatori poi. Se pure il diritto romano classico aveva già conosciuto il concetto di *universitas*, tuttavia con tale termine non veniva indicato un ente astratto distinto dagli individui suoi membri, ma una semplice collettività organizzata di uomini.

L'estensione del concetto di persona ad entità diversa dall'uomo si manifesta per la prima volta nel Trecento, ad opera del canonista Sinibaldo de' Fieschi<sup>21</sup>, cui si deve l'enunciato: «*collegium in causa universitatis fingatur una persona*», espresso durante il tentativo di dare una spiegazione teleologica alle elezioni canoniche a maggioranza. L'*universitas*, in tal caso, stava ad indicare, appunto, la Chiesa.

---

<sup>19</sup> Per una panoramica completa della nozione persona giuridica si v. F. FERRARA, *Teoria delle persone giuridiche*, Napoli-Torino, 1915; ID. e F. FERRARA JR., *Le persone giuridiche*, in *Tratt. Vassalli*, II, 2, Torino, 1956; A. FALZEA, *Il soggetto nel sistema dei fenomeni giuridici*, Milano, 1939; V. ARANGIO-RUIZ, *Gli enti soggetti dell'ordinamento internazionale*, Milano, 1951; P. VERRUCOLI, *Il superamento della personalità giuridica delle società di capitali nella common law e nella civil law*, Milano, 1964; G. SCALFÍ, *L'idea di persona giuridica e le formazioni sociali titolari di rapporti nel diritto privato*, Milano, 1968; R. ORESTANO, *Il problema delle «persone giuridiche» in diritto romano*, Torino, 1968; ID., *Azione, diritti soggettivi, persone giuridiche*, Bologna, 1978; F. GALGANO, *Delle persone giuridiche*, in *Comm. Scialoja e Branca*, Bologna-Roma, 1969; ID., *Associazioni, fondazioni, comitati*, in ID. (a cura di), *I grandi orientamenti della giurisprudenza civile e commerciale*, Padova, 1987; P. ZATTI, *Persona giuridica e soggettività*, Padova, 1975; G. TAMBURRINO, *Persone giuridiche e associazioni non riconosciute. Comitati*, in *Giur. Sist. Biggiani*, Torino, 1980. Per la letteratura storica e comparatistica si v., invece, R. SERICK, *Rechtsform und Realität juristischer Personen*, Berlin-Tübingen, 1955.

<sup>20</sup> F. GALGANO, voce «Persona giuridica», in *Dig. disc. priv., Sez. cin.*, XIII, Torino, 1995, p. 392 ss.

<sup>21</sup> Cf. F. RUFFINI, *La classificazione delle persone giuridiche in Sinibaldo dei Fieschi*, in *Scritti giuridici minori*, II, 1936, p. 14.

«La riduzione ad unità dell'*universitas* sotto il nome di persona si accompagna all'idea della finzione; muove dalla tacita premessa che solo l'uomo è, propriamente, una persona»<sup>22</sup>.

Questa premessa è rafforzata nel pensiero di Bartolo di Sassoferrato il quale ha affermato: che l'*universitas* «*proprie non est persona: tamen hoc est fictum positum pro vero, sicut ponimus nos iuristaes*»<sup>23</sup>.

Nel pensiero giuridico medievale, quindi, l'*universitas* altro non è che una metafora, un'immagine del parlare figurato, cosicché per i commentatori i beni appartenenti all'*universitas* non sono beni appartenenti ad un soggetto distinto, bensì beni appartenenti ad una collettività di uomini, sebbene non *uti singuli*, quanto *universi*.

Tuttavia, si è dovuto attendere qualche secolo affinché il concetto vero e proprio di persona giuridica potesse trovare ulteriori elaborazioni sulle quali fondare la propria nascita.

La seconda fase evolutiva del concetto di persona giuridica si ha nel Seicento, grazie all'opera di sistemazione concettuale dei giusnaturalisti.

«L'eredità delle epoche precedenti offriva materiali per due diverse costruzioni: si poteva, sviluppando l'intuizione dei canonisti, procedere ad una duplicazione della categoria di soggetti di diritto ed accostare all'essere umano una ulteriore, non umana, serie di soggetti di diritto, cui facessero capo rapporti giuridici corrispondenti a quelli propri degli esseri umani; ma si poteva, restando nel solco tracciato dai commentatori, tenere ferma l'unità del soggetto e procedere ad una duplicazione della categoria dei rapporti giuridici, distinguendo fra rapporti facenti capo all'individuo come singolo e rapporti facenti capo all'individuo come membro di una collettività organizzata»<sup>24</sup>.

Al concetto di *persona ficta* i giusnaturalisti preferirono quello di *persona moralis*<sup>25</sup>. Una persona al pari dell'uomo – inteso come *persona naturalis* – ma che differisce da questo per il fatto d'essere la prima una persona costruita intellettualmente (*moralis*) e il secondo una persona allo stato di natura (*naturalis*).

Una terza e decisiva fase evolutiva può rinvenirsi nei primissimi anni dell'Ottocento in cui il concetto di persona giuridica è penetrato nel linguaggio legislativo e si è passati dall'idea di persona morale all'idea di persona giuridica<sup>26</sup> quale frutto della creazione normativa.

A Savigny si deve la cd. teoria della finzione<sup>27</sup>, la quale si differenzia dalla concezione finzionistica, già emersa nel pensiero dei canonisti e dei

---

<sup>22</sup> F. GALGANO, voce «*Persona giuridica*», cit., p. 392 ss..

<sup>23</sup> BARTOLO, *Prima super digesto novo, l. nonnunquam*, Lugduni, 1533, fo. 205.

<sup>24</sup> F. GALGANO, voce «*Persona giuridica*», cit., p. 392 ss.

<sup>25</sup> Il giusnaturalismo poneva l'uomo al centro del proprio sistema e ciò spiegava l'avversione di HOBBS o di altri pensatori dell'epoca, come HORN, all'idea di concepire altri soggetti oltre agli esseri umani. Tuttavia, non mancarono voci contrarie, quali quelle di Grozio e Pufendorf, le quali esercitarono una grande influenza sul pensiero giuridico dell'epoca successiva.

<sup>26</sup> La sostituzione è attribuita a G. HUGO, *Lehrbuch des Naturrechts als einer Philosophie des positiven Rechts*, Berlin, 1798, p. 45; la divulgazione della nuova locuzione è dovuta ad G.A. HEISE, *Grundriss eines Systems des gemeinen Civilrechts*, Heidelberg, 1807, p. 8 ss.

<sup>27</sup> F.K. SAVIGNY, *Sistema del diritto romano attuale, trad. it.*, III, Torino, 1900, p. 240. L'idea della *fictio iuris* veniva riproposta dal B. WINDSCHEID, *Diritto delle pandette, trad. it.*, I, Torino, 1930, p. 232: «persona giuridica

commentatori, per il fatto di ritenere il concetto di collettività organizzata non più quale mero frutto di una finzione dottrinale, bensì una vera e propria *factio legis*. Il concetto di persona giuridica giunse, così, ad indicare quei «soggetti artificialmente creati dalla legge», con la precisazione che si trattasse sì di «persone» in quanto nuovi soggetti di rapporti giuridici oltre l'uomo singolo, ma «giuridiche» perché ammesse solo per uno scopo giuridico.

L'uomo, ormai posto al centro dell'ordinamento giuridico, viene ad essere considerato come l'unico soggetto capace di diritto. La persona giuridica è una finzione, è una creazione dell'ordinamento giuridico che solo può conferirle, attraverso il riconoscimento, la capacità giuridica. Il legislatore finge per soddisfare esigenze proprie del commercio giuridico<sup>28</sup>.

La teoria di Savigny, però, non mise d'accordo tutti ma diede la stura ad un dibattito dottrinale rimasto ancora oggi attivo.

Fondamentalmente, si fronteggiarono – e si fronteggiano – due orientamenti estremi: l'uno collocante la persona giuridica nella più ampia categoria dei soggetti di diritto, di cui costituisce, assieme alla persona fisica, una specie; l'altro considerante solo l'individuo quale soggetto di diritto, negando l'autonoma soggettività giuridica delle persone giuridiche.

Nel primo orientamento, ritenuto «tradizionale», convogliarono le teorie cd. antropomorfiche ovvero la teoria della finzione di Savigny – come sopra descritta – e la teoria organica o della realtà di Otto von Gierke<sup>29</sup>.

Quest'ultima prese piede nella seconda metà dell'Ottocento, mossa dall'intento di superare il disfavore dello Stato nei confronti delle formazioni sociali e, con esso, i sistemi concessionistici della personalità giuridica, e si prodigò nel dimostrare che il concetto di persona giuridica non fosse solo il frutto di una finzione legislativa ma, all'inverso, una realtà prenormativa, come tale preesistente alle stesse norme giuridiche.

Alla luce degli studi così effettuati, né risultò che la persona giuridica fosse proprio un'entità reale, esistente nella realtà sociale ancor prima che in quella giuridica.

In base a tale costruzione dogmatica, quindi, le organizzazioni collettive erano ritenuti organismi naturali dotati di una propria volontà, portatori di propri

---

è una persona non esistente in effetto, ma soltanto immaginaria, che si considera come soggetto di diritti ed obbligazioni». La filosofia che sottace alla teoria della finzione è rilevata da R. ORESTANO, *Il «problema delle persone giuridiche» in diritto romano*, I, Torino, 1968, p. 20 ss.; la persona giuridica doveva essere solo finzione in un sistema di pensiero che restava imperniato sull'uomo.

<sup>28</sup> Hanno ripreso il medesimo concetto di persona giuridica anche gli autori stranieri contemporanei che abbracciano la «teoria della persona artificiale e dipendente». In particolare J. NESTERUK, *Persons, Property, and the Corporation: A Proposal for a New Paradigm*, in *DePaul L. Rev.*, 1990, 39, p. 543-564 afferma «A corporation is artificial in that it is a human creation subject to human choices». Ripercorrendo detta teoria S.K. RIPKEN, *Corporations are people too: a multi-dimensional approach to the 1080 corporate personhood puzzle*, in *Fordham Journal of Corporate & Financial Law*, 2010, 15, p. 97- 177 scrive «So a corporation is just a legal construct, an artificial creation of law and humans; its personhood is a legal fiction conferred by law solely for the facilitation of trade and commerce».

<sup>29</sup> Cfr. O. GIERKE, *Die Genossenschaftstheorie und die deutsche Rechtsprechung*, Berlin, 1887, p. 22 ss. La teoria della realtà è stata chiamata in molti modi, specie negli USA del primo Novecento. Sul punto si v. P.I. BLUMBERG, *The Corporate Entity in an Era of Multinational Corporations*, in *Del. J. Corp. L.*, 1990, 15, pp. 283 e 292-293. Negli USA la teoria della realtà è ad oggi dominante.

interessi distinti da quelli delle persone fisiche che li componevano. Pertanto, il riconoscimento statale della personalità giuridica non aveva alcun valore costitutivo ma, viceversa, meramente dichiarativo, limitandosi ad accertare una realtà già esistente nell'ordinamento statale<sup>30</sup>.

Il secondo orientamento formatosi, invece, raggruppò le teorie riduzionistiche o negazionistiche le quali condividevano la fondamentale premessa dell'assoluto ripudio delle teorie antropomorfiche in quanto era ritenuto impensabile che la volontà<sup>31</sup> o gli interessi potessero essere riferiti a soggetti diversi dall'uomo.

Sostanzialmente, dette teorie negavano l'autonoma soggettività giuridica delle persone giuridiche, riconoscendo quali unici soggetti di diritto le sole persone fisiche che se ne avvalevano dal momento che la norma giuridica, in quanto norma di comportamento, non poteva che avere, come soggetto dell'azione, l'individuo umano.

Il principale fautore della negazione della soggettività delle persone giuridiche fu, con la sua teoria normativa, Kelsen<sup>32</sup>, il quale, in diretta replica alla idea del Gierke, riteneva che solo il comportamento degli esseri umani potesse essere regolato dall'ordinamento giuridico e, pertanto, il concetto di persona giuridica non era altro che l'«ipostatizzazione» di un ordinamento: «l'ordinamento è trasformato in una sostanza, e questa sostanza è considerata come una cosa separata, un essere distinto dall'ordinamento e dagli esseri umani il cui comportamento è regolato dall'ordinamento stesso»<sup>33</sup>.

Altro importante filone negazionista era rappresentato dai fautori delle teorie nominaliste.

Queste partivano dall'accoglimento del moderno metodo dell'analisi del linguaggio che, col proposito di ristudiare la definizione dei concetti giuridici, giungeva ad affermare come i termini giuridici presentassero delle anomalie dovute alla loro non «diretta ed immediata connessione con le realtà di fatto loro corrispondenti»<sup>34</sup>.

---

<sup>30</sup> F. CARINGELLA e L. BUFFONI, *Manuale*, cit., p. 112-113.

<sup>31</sup> BERNATZIK in *Archiv für öffentl. Recht*, 1889, p. 218 obietta che la volontà è un fenomeno psichico, irriferribile ad entità non umane. In Italia l'obiezione viene, poi, ripresa da F. FERRARA, *Teoria delle persone giuridiche*, cit., p. 215 s.; da P. TRIMARCHI, *Invaldità delle deliberazioni di assemblea di società per azioni*, Milano, 1958, p. 19; e da A. CANDIAN, *Per la qualificazione del contratto di società commerciale*, RS, 1963, p. 252.

R.V. JHERING, in *Geist des römischen Rechts*, II, 2, Leipzig, 1891, p. 518 ss. eccepisce che interesse è una valutazione razionale, pertinente anch'essa alla funzione intellettuale dell'uomo e soltanto dell'uomo. In Italia riprendono questo spunto critico W. CESARINI SFORZA, *Studi sul concetto d'interesse generale*, in *Arch. studi corporativi*, 1935, p. 51 ss.; P.G. JAEGER, *L'interesse sociale*, Milano, 1963, p. 8 ss.; G. SCALFÍ, *L'idea di persona giuridica*, cit., p. 84 ss.; G.S. COCO, *Volontà collettiva in diritto privato*, Milano, 1967, 25, p. 131.

<sup>32</sup> H. KELSEN, *Teoria generale del diritto e dello Stato*, trad. it., Milano, 1952, p. 98. Secondo tale impostazione i diritti e i doveri della persona giuridica altro non sono che i diritti e i doveri dei suoi membri, sebbene si tratti di situazioni giuridiche possedute in maniera specifica, ovvero non in qualità di singoli individui, ma in quanto membri di un'organizzazione collettiva.

<sup>33</sup> H. KELSEN, *o.u.c.*, p. 110. Sul punto F. GALGANO, voce «*Persona giuridica*», cit., p. 392 ss.

<sup>34</sup> L.A. HART, *Contributi all'analisi del diritto*, trad. it., Milano, 1964, p. 67: «noi non abbiamo da fare con entità fittizie od anomale, bensì con un uso nuovo ed esteso di regole giuridiche, e delle espressioni che in esso sono implicate».

Guardando alle persone giuridiche si affermava, quindi, che il concetto non veniva utilizzato per indicare entità fittizie o anomale quanto veniva piuttosto enucleato attraverso l'estensione di norme formulate per gli individui e che, per tal via, acquistavano un nuovo significato.

La visione veniva ben compendiata da Ascarelli il quale affermava che il concetto di persona giuridica consistesse, in fondo, in «una speciale disciplina normativa concernente relazioni fra uomini»<sup>35</sup>.

Parlando di tesi nominalistiche o analitiche non si può non far riferimento, in Italia, a D'Alessandro, il quale ha sostenuto che «i nomi di persone giuridiche sono simboli incompleti» ossia «simboli ai quali non corrisponde alcun ente» e che ciò che contraddistingue una collettività di individui dalla semplice pluralità è solo il possesso di un simbolo, di un nome proprio, con a conclusione che le situazioni giuridiche imputate alla persona giuridica si riducono a situazioni imputabili ai suoi membri<sup>36</sup>.

Il metodo filosofico-analitico è stato, poi, ripreso da Galgano nei suoi studi sulle persone giuridiche che sono approdati ad un piena adesione della tesi nominalista: la persona giuridica non è un soggetto di diritto ulteriore e distinto dalla persona fisica in quanto il riconoscimento della personalità giuridica produce quale unico effetto l'esenzione dei suoi membri dall'applicazione del diritto comune.

La persona giuridica, in tale prospettiva, è «solo uno strumento del linguaggio giuridico utile per riassumere una complessa disciplina di rapporti intercorrenti tra persone fisiche», una disciplina speciale, derogatoria a quella comune, che consente di elevare a regola quella che è, nel diritto comune, l'eccezione, ovvero il beneficio della responsabilità limitata<sup>37</sup>. Quello che sino alla costruzione del concetto di persona giuridica era solo un beneficio concesso dall'ordinamento si trasforma, in virtù di quella elaborazione concettuale, da beneficio, mero privilegio quale era, ad applicazione di un principio generale, che vale a garantire alla persona giuridica la responsabilità patrimoniale limitata delle persone fisiche che la compongono<sup>38</sup>.

Orbene, nonostante il fascino, le teorie negazioniste restarono e restano oggi minoritarie.

Il ripudio della visione antropomorfa delle organizzazioni collettive non è valso, infatti, di per sé solo, a soppiantare la nozione di persona giuridica che ha potuto sopravvivere, nella dottrina successiva, sulla base di un più radicale finzionismo, ovvero attribuendo carattere «artificiale» alla stessa personalità

---

<sup>35</sup> T. ASCARELLI, *Problemi giuridici*, I, Milano, 1959, p. 242. Oltre i confini dell'Italia si suole parlare di «teoria della persona aggregata». Secondo i suoi sostenitori «a corporation is not an artificial entity nor merely a creation of law as such; rather it is an organisation or association of a group of people who form its individual human constituents and without whom it would not have any identity of its own or any ability to function towards its goals.

<sup>36</sup> F. D'ALESSANDRO, *Persone giuridiche e analisi del linguaggio*, Padova, 1989, p. 92 ss.

<sup>37</sup> F. GALGANO, *Delle persone giuridiche*, cit., p. 60 s.; ID., *Le associazioni, le fondazioni, i comitati*, 2a ed., Padova, 1996, p. 145 ss.; ID., voce «Persona giuridica», cit., p. 392 ss.

<sup>38</sup> F. CARINGELLA e L. BUFFONI, *Manuale*, cit., p. 114.

giuridica dell'uomo, riducendo il ruolo dell'azione e riconoscendo l'emersione di interessi impersonali.

Si è partiti dall'idea che «l'uomo non dalla natura, ma per via del diritto è persona»; che egli è tale in quanto «è investito di diritti e di obblighi», in quanto «è punto di riferimento di diritti e di doveri dell'ordinamento giuridico». Se ne è indotta la necessità di «bandire l'idea che la personalità sia congiunta con l'organismo corporeo e psichico dell'uomo, e che quindi non possa trovarsi fuori di lui». La personalità giuridica è «un prodotto dell'ordinamento giuridico» e «non v'è alcun ostacolo perché esso non possa ugualmente attribuire la soggettività giuridica anche ad enti non umani, a figure dell'intelletto, ad entità ideali»<sup>39</sup>.

In tale rinnovata prospettiva, sopperendo il substrato fisico, il soggetto effettivo e sostanziale della norma giuridica non deve essere per forza chi compie l'azione – che non può che essere l'uomo – ma può essere il titolare dell'interesse che la norma è chiamata a tutelare, che, ove l'interesse protetto sia un interesse impersonale, può essere un'entità diversa dall'uomo.

Si può concludere che la sola esistenza della categoria degli interessi impersonali – ovvero non riferibili ad alcun soggetto fisico, quali gli interessi pubblici, collettivi o diffusi – sia di per sé atta a giustificare la possibilità di un soggetto giuridico diverso dall'individuo umano che operi attraverso il principio di immedesimazione organica.

D'altronde, non può negarsi l'esistenza delle organizzazioni collettive nella realtà empirica sol si consideri il dato di fatto che l'uomo possa agire sia come singolo sia come membro di formazioni sociali, e questa è una situazione riconosciuta e tutelata dall'art. 2 della nostra Costituzione. Il concetto di persona giuridica offre unicamente una consapevolezza teorica alle possibilità pratiche<sup>40</sup>.

Ad oggi, quindi, è dominante, specie in Italia, il pensiero di coloro che ritengono che le persone giuridiche, come le persone fisiche, sono autonomi soggetti di diritto. Il principio dell'autonoma soggettività giuridica (o principio dell'alterità) trova, d'altra parte, conferma nel dato positivo del nostro Codice che dedica alle persone giuridiche una disciplina generale. Tra l'altro, considerato che il Codice, all'interno del Libro I, Titolo II dedicato alle persone giuridiche, riserva agli enti non riconosciuti un proprio spazio, può anche affermarsi qualcosa in più, ossia che la nozione di soggettività giuridica accolta dal nostro legislatore è una nozione ben più ampia di quella di personalità giuridica e, finalmente, scollegata da questa.

Ad ogni modo, seguendo la dottrina preferibile<sup>41</sup>, è bene distinguere la soggettività degli enti dalla soggettività delle persone fisiche, in quanto «la

---

<sup>39</sup> Cfr. F. GALGANO, voce «*Persona giuridica*», cit., p. 392 ss. Le espressioni sono di F. FERRARA JR., *Le persone giuridiche*, cit., p. 33; ID., *Teoria delle persone giuridiche*, cit., p. 358. Concorda nell'impostazione A. FALZEA, *Il soggetto*, cit. p. 64, ripetuta in ID., M. BASILE, «*Persona giuridica (diritto privato)*», in *Enc. dir.*, XXXIII, Milano, 1983, p. 264 ss. Analogamente G. ARANGIO-RUIZ, *Gli enti*, cit., p. 91 ss.; come G. SCALFÌ, *L'idea di persona giuridica*, cit., p. 48: «è incontestabile che il diritto tende a regolare le azioni umane: che, cioè, mira a disciplinare gli interessi di uomini: ma questo non significa che solo l'uomo possa essere soggetto di diritti».

<sup>40</sup> S. AMATO, *Il soggetto*, cit., p. 36.

<sup>41</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale*, I, cit., p. 13.

prospettiva relativizzante e assiologica impone di individuare non un unico fondamento normativo per tutti i soggetti giuridici, bensì una pluralità di fondamenti anche costituzionali secondo i diversi interessi in gioco (2, 18, 41 cost.): altro è la persona umana quale valore unitario e fondamentale del sistema, altro è l'ente quale soggetto costituito per la realizzazione di uno scopo particolare e quindi strumento tecnico previsto dal legislatore per lo svolgimento di un'attività nel rispetto di una necessaria pluralità non soltanto del mercato ma di ogni settore di vita: pertanto all'ente non sono applicabili le disposizioni normative che presuppongono una persona fisica, quali quelle relative ai rapporti familiari, alla scomparsa, alla morte presunta, ecc»<sup>42</sup>.

### 3. La qualificazione giuridica dei sistemi di IA.

A questo punto occorre ricondurre il tema della soggettività giuridica alla questione di nostro interesse e rispondere alla difficilissima domanda: i sistemi di IA possono essere considerati autonomi soggetti di diritto?

Come da sempre accade riguardo alle questioni più spinose, la dottrina è divisa e non ha tutt'oggi trovato un punto di incontro<sup>43</sup>.

Sebbene l'orientamento maggioritario, seguito dall'Unione europea nelle sue recentissime proposte di regolamentazione dell'IA<sup>44</sup>, abbia optato per la soluzione meno radicale di ritenere l'IA ancora oggetto del diritto e non soggetto, per avere un quadro più completo, è bene andare a vedere le prospettive per il momento accantonate ma potenzialmente valide.

#### 3.1. IA come soggetto del diritto.

«I believe that the abominable deterioration of ethical standards stems primarily from the mechanization and depersonalization of our lives,» «a disastrous byproduct of science and technology. Nostra culpa!»<sup>45</sup>.

Quindi la tecnologia si è impossessata anche della nostra anima?

---

<sup>42</sup> *Ibidem*.

<sup>43</sup> M. FARINA, *Brevi riflessioni sullo status delle "persone elettroniche"*, in *L'Irroverro*, 2021, 2, p. 106 ss.

<sup>44</sup> Si fa riferimento a Commissione UE, Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'Unione, 21 aprile 2021, COM(2021)206 final; Commissione UE, Proposta di Direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale, 28 settembre 2022, COM(2022) 496 final; Commissione UE, Proposta di Direttiva del Parlamento europeo e del Consiglio sulla responsabilità per danno da prodotti difettosi, 28 settembre 2022, COM (2022) 495 final.

Di contro nella sua Risoluzione del 16 febbraio 2017, lo stesso Parlamento aveva affermato, all'art 59, lett. f), di ritenere opportuna la previsione della personalità elettronica per i sistemi più sofisticati «*creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently*».

<sup>45</sup> A. EINSTEIN lo scrisse in una lettera indirizzata al suo amico psichiatra Otto Juliusburger, nel 1948. Traduzione: «Credo che l'abominevole deterioramento degli standard etici derivi principalmente dalla meccanizzazione e dalla spersonalizzazione delle nostre vite», «un disastroso sottoprodotto della scienza e della tecnologia. Nostra culpa!».

Si tratta di una domanda alquanto drammatica ma non senza un briciolo di aderenza alla realtà in un mondo sempre più automatizzato ed integrato da sofisticati sistemi di IA.

Nelle scritture religiose si trovano spesso riferimenti alla presenza dell'anima quale parte consistente della persona fisica. Le persone giuridiche artificiali e le entità inanimate, contrariamente, non hanno un'anima (qualunque cosa essa sia).

Per il catechismo della Chiesa cattolica con la parola «anima» si fa riferimento al principio spirituale dell'uomo. L'anima è il soggetto della coscienza e della libertà umana, è ciò che permette il libero arbitrio, che costituisce lo sfondo etico e giuridico della responsabilità che abbiamo come esseri naturali. Gli individui sono sovrani delle loro decisioni e quindi legalmente responsabili delle loro azioni.

Probabilmente i tempi sono decisamente prematuri per accostare il concetto di anima ad un sistema di IA, tuttavia non lo sono per coloro che vorrebbero quantomeno qualificare l'IA autonoma come un soggetto di diritto.

D'altronde, il concetto di soggetto di diritto nel mondo è in continua espansione. Se un'entità debba essere considerata o meno un soggetto/persona giuridica, alla luce di quanto detto *supra*, dipende dalla risposta alla domanda se quell'entità debba essere titolare di una serie specifica di situazioni giuridiche soggettive ai fini del suo funzionamento sistematico e questo dipende da circostanze sociali, politiche ed economiche.

La personalità giuridica<sup>46</sup> – intesa come soggettività – è un aspetto flessibile e mutevole del sistema giuridico. Come ha affermato il giurista internazionale Ian Brownlie<sup>47</sup>, è ben noto che i soggetti di diritto in ogni sistema giuridico non sono esattamente identici nella loro natura e nei loro diritti o nell'estensione dei loro diritti e della loro natura, ma dipendono dalle esigenze e dalla sensibilità della comunità.

Si pensi al riconoscimento della soggettività giuridica al fiume Whanganui in Nuova Zelanda e a diversi fiumi in India. Non tutti gli ordinamenti giuridici condividerebbero una tale qualificazione sol si consideri che sono entità inanimate.

Ancor più complessa è la questione se sistemi di IA o combinazioni di hardware e software o algoritmi (i cd. robot), possano essere considerati titolari di diritti nonché agenti capaci sia di agire per la tutela dei loro interessi sia di assumersi la responsabilità delle loro azioni.

Il dibattito è arrivato sul piano della concretezza, ad esempio, con Sophia, il robot umanoide intelligente al quale, nel 2017, è stata concessa la cittadinanza saudita, segnando la prima volta nella storia dell'umanità in cui un robot è diventato cittadino nazionale. Senza dubbio, Sophia è un essere artificiale di grande valore, ma, nonostante la sua cittadinanza saudita, la maggior parte degli studiosi concorda sul fatto che non soddisfa i criteri necessari per ottenere diritti

---

<sup>46</sup> A livello internazionale si parla sempre di *legal personhood* da tradursi, in base all'ordinamento giuridico italiano, come soggettività giuridica.

<sup>47</sup> I. BROWNLIE, *Principles of public international law*, Oxford, 1990, p. 58 ss.

morali e legali paragonabili a quelli dei cittadini umani adulti, non è ancora in grado di avere una dignità morale.

Si può dire che coloro che sostengono – o quantomeno non negano – la possibilità che un sistema di IA goda di una «*legal personhood*» si dividono tra coloro che fanno rientrare l'IA sotto l'ombrello della disciplina delle persone giuridiche in virtù di un'interpretazione sartoriale del concetto di personalità giuridica, e coloro che disegnano i contorni di una nuova personalità, appunto la personalità elettronica.

Nel primo gruppo rientra sia la dottrina che riprende la matrice funzionale della teoria della finzione di Savigny, sia la dottrina che fonda il concetto di personalità giuridica sulla *moral agency*. Il secondo gruppo è, invece, costituito da coloro che individuano la necessità di un terzo genere di soggetti giuridici accanto alle persone fisiche e giuridiche: le persone elettroniche.

Per completezza d'informazione, è bene ricordare che di tale ultimo avviso era, peraltro, il Parlamento europeo al momento dell'adozione della risoluzione del 2017<sup>48</sup>.

Tuttavia, dopo l'adozione della risoluzione, un gruppo di 56 esperti di IA e robotica, leaders del settore, specialisti di legge, medicina ed etica, ha redatto la cd. lettera aperta sulla robotica (o *robot open letter*, *ROL*) in cui ha espresso la sua profonda preoccupazione per la concessione della personalità giuridica ai robot più sofisticati. Essi hanno ritenuto che la creazione di uno status giuridico di «persona» elettronica sarebbe ideologica, non sensata e non pragmatica. In particolare, hanno contestato il § 59, lett. f.) della risoluzione, laddove veniva previsto di creare uno status giuridico specifico per i robot nel lungo periodo, in modo che almeno i robot autonomi più sofisticati potessero essere definiti come persone elettroniche responsabili di risarcire eventuali danni che possono causare, ed eventualmente applicare la personalità elettronica ai casi in cui i robot prendono decisioni autonome o interagiscono in altro modo con terzi in modo indipendente.

Anche se i redattori della *ROL* condividevano le motivazioni di base dei membri del Parlamento europeo, essi hanno asserito che i problemi economici, giuridici, sociali ed etici fossero da considerare come un problema di sicurezza e che, perciò, gli impatti dell'IA e della robotica avrebbero dovuto essere considerati con maggiore attenzione. Inoltre, hanno considerato il tentativo di risolvere il problema della responsabilità per danni introducendo il concetto di «personalità elettronica», cioè creando uno status giuridico per i robot, inutile e fuorviante.

Gli stessi hanno manifestato che il § 59, lett. f.) offrì molti pregiudizi basati su una sopravvalutazione delle capacità effettive anche dei robot più avanzati, una comprensione superficiale dell'imprevedibilità e delle capacità di

---

<sup>48</sup> Risoluzione del 16 febbraio 2017 recante Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, P8 TA(2017)0051. Di orientamento più cauto al riguardo era, invece, il parere del Comitato Economico e Sociale dell'Unione europea (CESE, INT/086 del 31 maggio 2017), che, al Punto 3.33, è contrario alla assunzione di forme di personalità giuridica per i robot, in quanto «comporterebbe un rischio inaccettabile di azzardo morale». In argomento, si veda E. PALMERINI, voce «Robotica» (parte giuridica), in *Enciclopedia di bioetica e scienza giuridica*, X, Napoli, 2016, p. 1104.

autoapprendimento e una percezione dei robot distorta dalla fantascienza e da alcuni recenti annunci sensazionali sulla stampa.

A sostegno della loro posizione, hanno sollevato due obiezioni principali, fondate su questioni tecniche, etiche e legali.

Se in generale, i redattori del ROL hanno esplicitato le convinzioni di molti studiosi, d'altra parte hanno peccato nel non avere una visione prospettica. Si pensi all'AGI (*artificial general intelligence*) – citata al capitolo I –. È vero che è un'IA ancora in fase di elaborazione, ma non bisogna farsi trovare impreparati nel momento in cui le IA avranno raggiunto un certo livello di intelligenza, comprensione e autonomia paragonabile alle capacità delle persone umane.

### 3.1.1. Segue...secondo la teoria della finzione ovvero il criterio funzionale o della semplificazione giuridica.

In termini generali, per finzione legale si fa riferimento ad uno stratagemma giuridico in base al quale si dà come esistente o come inesistente un fatto, indipendentemente dalla preoccupazione dell'accertamento della verità e, talora anche nel riconosciuto contrasto con questa, e ciò al fine di ricollegare alla sussistenza del fatto, o alla sua mancanza, le conseguenze giuridiche che ne deriverebbero qualora esse corrispondessero alla realtà, rendendo cioè possibile o escludendo l'applicabilità di una norma al rapporto dato<sup>49</sup>.

Precisamente, come ricorda Bretone<sup>50</sup>, rileggendo Austin<sup>51</sup>, la finzione consisterebbe, in linea di massima, «nel simulare o assumere che qualche cosa che certamente era, non era; e che qualcosa che certamente non era, era»<sup>52</sup>.

In altri termini, invece di modificare le norme esistenti, si modifica artificiosamente il fatto a cui queste dovrebbero applicarsi, fingendo inesistenti alcuni elementi che osterebbero all'applicazione di una norma, o fingendo sussistenti tali altri elementi che soli ne renderebbero possibile l'applicazione.

L'artificio della finzione, dunque, permette, senza nessuna formale modifica delle norme vigenti, di raggiungere gli stessi effetti che si conseguirebbero da un'espressa statuizione di norme nuove<sup>53</sup>.

---

<sup>49</sup> Cfr. R. DE RUGGIERO, *voce Finzione legale*, in *Diz. prat.*, 3, Milano, 1923, p. 184 ss.;

<sup>50</sup> M. BRETONE, *Finzioni e formule nel diritto romano*, in *Materiali per una storia della cultura giuridica*, XXXI, 2, Bologna, 2001, p. 297.

<sup>51</sup> Cfr. J. AUSTIN, *Lectures on Jurisprudence*, London, 1885, p. 609 ss.

<sup>52</sup> C. CHAVALLIER, *Fictio e natura: due metodi alternativi per la realizzazione dell'aequitas*, in *Teoria e storia del dir. priv.*, XIV, 2021, p. 1- 60.

<sup>53</sup> Sul punto, a titolo meramente esemplificativo, si richiamano G. DEMELIUS, *Die Rechtsfiktion in ihrer geschichtlichen und dogmatischen Bedeutung. Eine juristische Untersuchung*, Weimar, 1858, p. 51 ss.; R. VON JHERING, *Geist des römischen Rechts auf den verschiedenen Stufen seiner Entwicklung*, Leipzig, 1875, p. 357 ss.; B. VON BÜLOW, *Zivilprozessualische Fiktionen und Wahrheiten*, in *AcP*, 62, 1879, p. 1 ss.; H. KELSEN, *Zur Theorie der juristischen Fiktionen*, in *Annalen der Philosophie*, 1919, 1, p. 655; S. RICCOBONO, *'Formulae ficticiae' a Normal Means of Creating New Law*, in *Tijdschrift voor Rechtsgeschiedenis*, IX, 1929, p. 1 ss.; F. PRINGSHEIM, *Symbol und Fiktion in den antiken Rechten*, in *Studi in onore di P. De Francisci*, IV, 1955, p. 209 ss.; G. GARCIA GARRIDO, *Sobre los verdaderos límites de la ficción en derecho romano*, in *AHDE*, 1957- 1958, 27-28, p. 305 ss.; C. LLANO CIFUENTES, *Naturaleza jurídica de la 'fictio iuris'*, Madrid, 1963, *passim*; S. PUGLIATTI, *voce Finzione* in *Enc. dir.*, XVII, Milano, 1968, p. 662 ss.; V. COLACINO, *voce 'Fictio iuris'*, in *NNDI*, VII, 1968, p. 269 ss.; U. ROBBE, *La 'fictio iuris' e finzione di adempimento della condizione nel diritto romano*, in *Scritti in onore di S. Pugliatti*, IV, Milano, 1978, p. 634 ss.; A. GUZMAN BRITO, *Sobre el objecto de las ficciones pretorias*, in *REHJ*, 3, 1978; R. SOTTY, *Recherche sur les 'utiles*

Come si evince dalla descrizione, la finzione come l'interpretazione analogica estende l'impero della norma oltre le ipotesi previste, per applicarla anche là dove non concorrerebbero tutti i presupposti della sua applicazione. Tuttavia, a differenza dell'interpretazione analogica, che può essere attuata dall'interprete, in presenza di una lacuna legislativa, in base all'indagine di una comune *ratio legis* tra l'ipotesi espressamente disciplinata e quella non prevista dal legislatore, nella finzione giuridica è implicita la dichiarazione dell'irrilevanza d'uno di quegli elementi che pur sarebbero condizione essenziale per l'applicazione della norma, sì che l'applicazione di questa si rende possibile solo per la dichiarata volontà del legislatore che deve espressamente statuire la finzione.

È nella letteratura tedesca della seconda metà dell'ottocento e dei primi due decenni del novecento che si addensarono la maggior parte degli studi metodologici dedicati espressamente al tema delle finzioni.

I più asserirono che, per quanto la tecnica delle finzioni legali fosse grossolana, la stessa restasse uno strumento insostituibile per fornire una prima risposta alle emergenze create dalla velocità delle trasformazioni storiche<sup>54</sup>.

In tema di persone giuridiche, Savigny – come riportato *supra* – per primo utilizzò lo strumento della finzione legale per dare contenuto all'istituto. Secondo la sua visione, la persona giuridica non poteva che esistere per volere della legge e per lo scopo da quest'ultima stabilito.

In tal modo, si superava facilmente il fatto che, in un ordinamento antropocentrico, l'unico soggetto riconosciuto naturalmente titolare di diritti e doveri fosse l'uomo.

Lo strumento della finzione applicato alla teoria della soggettività-personalità giuridica ha permesso di considerare soggetti di diritto entità sprovviste di quelli che, secondo la dottrina più tradizionale, erano gli elementi costitutivi del soggetto di diritto, ovvero l'elemento oggettivo del corpo fisico e l'elemento soggettivo dell'*animus* ovvero della coscienza.

Con riguardo ai sistemi di IA, è evidente che questa è la teoria che maggiormente si presta ai fini del riconoscimento di una soggettività giuridica.

Di fatti, in tal modo, un'IA sarebbe soggetto in quanto riconosciuto tale dalla legge e non per sue proprietà intrinseche. Inoltre, lo sarebbe nei modi, con i limiti e per gli scopi sempre dalla legge previsti.

Evidentemente si tratterebbe di una soggettività funzionale ad una migliore gestione dei rischi propri della messa in circolazione di una tecnologia tanto

---

*actiones*'. *La notion d'action utile en droit romain classique*, Grenoble, 1977, p. 24 ss. e ID., *Les actions qualifiées d'utiles en droit classique*, in *Labeo*, 25, 1979; F. TODESCAN, *Diritto e realtà. Storia e teoria della 'fictio iuris'*, Padova, 1979, *passim*; K. HACKL, *Sulla finzione nel diritto privato*, in *Studi in onore di A. Biscardi*, I, Milano, 1982, p. 245 ss.; M. FUHRMANN, *Die Fiktion in römischem Recht, in Funktionen des Fiktiven*, München, 1983, p. 413 ss.; A. DOS SANTOS JUSTO, *A 'fictio iuris' no direito romano ('actiones ficticiae')*, I. *Época clássica*, Coimbra, 1988, p. 1 ss.; ID., *A 'fictio iuris' no direito romano ('actio ficticia')*. *Época classica*, in *Boletim de Faculdade de Direito*, 33, 1989, p. 1 ss.; A. GAMBARO, voce *Finzione giuridica nel diritto positivo*, in *Dig. disc. priv.*, Sez. civ., 8, Torino, 1992, p. 342 ss.; T. GIARO, *Die Fiktion des eigentlichen Eigentümers*, in *Au-delà des frontières. Mélanges W. Wolodkiewicz*, I, Warszawa, 2000, p. 277 ss.; M. BRETONI, *Finzioni*, cit., p. 295 ss.

<sup>54</sup> Citatissime al riguardo le tre pagine di R. VON JHERING, *Geist des römischen Rechts auf den verschiedenen Stufen seiner Entwicklung*, III, 1923, p. 301-306 in cui l'a. afferma che le finzioni legali sono il male minore.

dirompente quanto difficilmente controllabile. In questi termini, i sistemi di IA potrebbero essere considerati delle persone giuridiche in senso stretto.

Sulle orme della prima teorizzazione della personalità giuridica di Naffine<sup>55</sup>, restando il concetto di persona giuridica astratto, la personalità, in quanto finzione legale, potrebbe essere applicata o rimossa nei stessi termini del «gatto del Cheshire», l'entità inesistente che può essere presente nel momento in cui se ne ha bisogno e scomparire quando superflua.

D'altronde, tale costruzione risponderebbe, pienamente, a quella considerata da molti<sup>56</sup> la vera *ratio* della creazione di persone giuridiche diverse dall'uomo ovvero la limitazione della responsabilità patrimoniale delle persone fisiche. Come in riferimento alle società, la segregazione patrimoniale avrebbe l'effetto di limitare la responsabilità patrimoniale dei soci, così in riferimento ai sistemi di IA si raggiungerebbe l'effetto di limitare la responsabilità patrimoniale dei produttori di IA<sup>57</sup>.

Di talchè, non sussisterebbe alcun problema rispetto al riconoscimento della personalità ad un qualunque sistema algoritmico. Le perplessità deriverebbero da una distorsione psicologica: la precomprensione tipica del vizio dell'antropomorfismo. Per superarla basterebbe semplicemente iniziare a riferirsi ai predetti sistemi tecnologici evitando di utilizzare il termine «intelligenza» accanto ad «artificiale» che, inconsciamente ed inevitabilmente, si contrappone al concetto di «intelligenza naturale».

---

<sup>55</sup> N. NAFFINE, *Who are law's persons? From Cheshire cats to responsible subjects*, in *Modern Law Review*, 2003, vol. 66, 3, p. 346- 367. L'a., in base alla prima teorizzazione della persona giuridica, ritiene che non si debba nemmeno distinguere fra persone fisiche, giuridiche ed altre dal momento che, in ogni caso, il concetto di persona giuridica è astratto e consiste in una mera finzione legale che può essere applicata o disapplicata all'occorrenza. Avere la personalità, perciò, non significa altro che avere la capacità formale di essere portatore di diritti e doveri legali. Non c'è alcuna dimensione morale, etica, storica o empirica in questa definizione. È una capacità che viene conferita dalla legge perché è conveniente per la legge avere una tale creazione.

L'a. teorizza altre due versioni del concetto di persona in termini di soggetto giuridico. In base alla seconda, solo l'uomo può essere considerato soggetto di diritto in quanto l'unico essere ragionevole. In parole povere: per qualificarsi come persona giuridica, bisogna essere umani. È noto a tutti che una persona, intendendo in questo contesto una persona umana, diventa persona giuridica nel momento stesso in cui nasce o viene concepita, a seconda dell'ordinamento giuridico, e certamente termina con la morte. Con questa definizione, la personalità non è più una questione puramente giuridica, ma riguarda, invece, la questione di cosa significhi essere umani. In base alla terza, accanto alla razionalità, ai fini del riconoscimento della soggettività, è necessaria la responsabilità. Questa definizione insiste su un certo livello di capacità mentale e quindi esclude i bambini piccoli, gli incapaci mentali e gli animali. Sebbene sembri porre questa definizione di persona come l'attore giuridico ideale, essa incontra anche il pericolo dell'elitarismo.

<sup>56</sup> Cfr. A. BERTOLINI, *Robots as products the case for a realistic analysis of robotic applications and liability rules*, in *Law Innov. Technol.*, 2013, vol. 5, 2, p. 214-247; ID., *Robot companions: a legal and ethical analysis*, in *The Information Society*, 2018, vol. 34, 3, p. 130-14. Il pensiero è condiviso anche da F. ASTONE come espresso nel Convegno Internazionale EUNAP – *Regulatory models and principles to handle and rule Artificial Intelligence in the European and Asia-Pacific Area*, del 12-14 luglio 2023, presso la LUM di Bari.

<sup>57</sup> Così sostiene A. BERTOLINI, *Robot companions*, cit. p. 135. Tuttavia, si ritiene che l'applicabilità del regime della responsabilità del produttore per prodotto difettoso sia un argomento a sostegno della tesi contraria al riconoscimento della personalità giuridica dei robot o sistemi di IA di qualsivoglia tipo. V. *infra*.

In conclusione, se attribuissero, quindi, la personalità giuridica in base ad un criterio funzionale le parole di R. Tur<sup>58</sup> annuncerebbero una profonda verità: «legal personality can be given to just about any object if it is deemed to serve the ends of justice»<sup>59</sup>.

### 3.1.2. Segue...Secondo il criterio della *moral agency*.

Un altro grande filone dottrinale, pur non negando in astratto la possibilità di considerare un sistema di IA persona giuridica, resta più critico sul punto concludendo che solo i sistemi di IA dotati di uno *status* morale, ovvero la capacità di essere senzienti e consapevoli, possano essere ritenuti tali<sup>60</sup>.

Benchè, al di là della teoria della finzione, alcune voci isolate abbiano tentato di rompere nettamente il legame tra la personalità giuridica e la personalità morale, adottando un approccio pragmatico che basasse la personalità giuridica sulla mera esistenza di relazioni giuridiche<sup>61</sup>, tendendo la concessione della personalità giuridica alla promozione della regolamentazione del comportamento, sia tra gli esseri umani sia tra questi e la particolare entità non umana a cui è stata concessa la personalità giuridica— facilitando, quindi, l'adattabilità del concetto all'IA —, i più restano ancorati, in sostanza, alla necessità di una personalità morale.

Esistono tre approcci paradigmatici alla ricostruzione di una personalità morale: l'approccio della dignità umana, l'approccio dell'autonomia o della razionalità, l'approccio relazionale<sup>62</sup>.

---

<sup>58</sup> R. TUR, *The "Person" in Law*, in A. PEACOCKE, G. GILLET (a cura di), *Persons and personality: a contemporary inquiry*, 1987, p. 116-121. V. anche A.W. JR. MACHEN, *Corporate Personality*, in *Harv. L. Rev.*, 1911, 24, p. 347-350.

Tra i filosofi che hanno studiato la roboetica, una posizione di rilievo spetta a W. WALLACH e C. ALLEN, *Moral machines: Teaching robots right from wrong*, Oxford, 2009.

<sup>59</sup> Trad.: la personalità giuridica può essere attribuita praticamente a qualsiasi oggetto se si ritiene che serva ai fini della giustizia.

<sup>60</sup> Fra i molti rilevanti si v. L.B. SOLUM, *Legal personhood for artificial intelligences*, in *N. Carolina Law Rev.*, 1992, 70, p. 1231- 1287; L. FLORIDI, *L'estensione dell'intelligenza. Guida all'informatica per filosofi*, Roma, 1996. La visione dell'a. è singolare. Egli ritiene che sia possibile riscontrare uno *status* morale dell'IA e che questo non debba fondarsi sul riscontro di una cd. coscienza. Siamo talmente lontani dal comprendere adeguatamente persino la coscienza umana che parlare di quella delle macchine sembra assurdo. Serve un criterio estrinsecamente valutabile. La peculiarità delle affermazioni risiede nel fatto che il riconoscimento della personalità non si rifletta nella statuizione di una autonoma responsabilità.

P.M. ASARO, *Robots and responsibility from a legal perspective*, in *Proceedings of the IEEE*, 2007, p. 20- 24; A. MATTHIAS, *Automaten das Träger von Rechten. Plädoyer für eine Gesetzänderung*, Berlin, 2008; B.J. KOOPS, M. HILDEBRANDT e D.O. JAQUET-CHIFFELLE, *Bridging the accountability gap: rights for new entities in the information society?*, in *Minnesota. J. Law Sci. Technol.*, 2010, vol. 11, 2, p. 497- 561; S.M. SOLAIMAN, *Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy*, in *Artif. Intell. Law*, 2017, vol. 25, 2, p. 155-179;

<sup>61</sup> Cfr. B. SMITH, *Legal personality*, in *Yale Law Journal*, 1928, vol. 37, 3, p. 283-299. Tale visione viene ampiamente criticata da B.J. KOOPS, M. HILDEBRANDT e D.O. JAQUET-CHIFFELLE, *Bridging*, cit., p. 497- 561, i quali muovono, fondamentalmente, tre obiezioni: a) il problema della rappresentanza; b) il problema della responsabilità; c) il problema della appellabilità ai diritti umani da parte di entità non capaci di riflettere.

<sup>62</sup> Originale è l'approccio ibrido di F.M. KAMM, *Intricate ethics, rights, responsibilities, and permissible harm*, Oxford, 2007, p. 229, la quale fornisce un'importante definizione del concetto di *status* morale, sostenendo che «an entity has moral status when, in its own right and for its own sake, it can give us reason to do things such as not destroy it or help it». Trad: un'entità ha uno *status* morale quando, di per sé e per il suo bene, può darci motivo di fare cose come non distruggerla o aiutarla.

L'iter ricostruttivo dell'approccio della dignità parte dalla convinzione che solo gli esseri dotati di dignità possano essere persone morali. Dal momento che, secondo tale visione, solo gli umani possono avere dignità nel senso di coscienza del proprio valore sul piano morale, allora solo gli umani possono godere di uno *status* morale il quale, laddove sia pieno, garantisce una personalità morale che comporta l'attribuzione di diritti morali e legali.

La principale obiezione è quella di ridurre illegittimamente il concetto di personalità a quello di dignità umana dando spazio allo specieismo (v. *infra*).

Secondo l'approccio dell'autonomia-razionalità, invece, solo un essere moralmente responsabile può avere personalità morale. Ma un essere può essere ritenuto moralmente responsabile unicamente se sia in grado di essere rimproverato. E questo può accadere quando il soggetto ha la capacità di decidere se agire o meno e, nel caso in cui agisce autonomamente, decida consapevolmente di non seguire i propri principi morali.

Le primarie obiezioni sono mutate dal dibattito aperto sui diritti degli animali e dei disabili. Si contesta di considerare la razionalità e l'autonomia precondizioni della personalità morale senza dar peso né alla capacità di soffrire né a quella di relazionarsi.

Infine, secondo l'approccio relazionale la personalità morale è intrinsecamente agganciata alla socialità. Il concetto di persona dovrebbe essere apprezzato come categoria sociale in cui lo *status* morale emerge attraverso le relazioni. La personalità morale è assoluta ed inerente ad ogni essere umano in qualità di entità per natura sociale, anche se, talvolta, le relazioni non hanno natura reciproca.

A tale visione viene, però, contestato il fatto che può accadere che esseri che potrebbero avere la personalità morale in base a criteri oggettivi, non ne avrebbero diritto qualora soffrano di esclusione sociale per decisione di altri membri della comunità. Inoltre, il riconoscere il valore delle relazioni sociali non reciproche non escluderebbe, comunque, i casi ingiusti e pericolosi di deumanizzazione e depersonalizzazione.

In generale, l'idea che si debba scegliere un solo approccio per determinare lo *status* morale di una determinata entità non sembra convincente. Pertanto, è seguito dai più un approccio ibrido che accosta criteri relazionali a criteri fenomenologici. La razionalità, l'autonomia e la sensibilità sono, perciò, moralmente significative come le relazioni sociali.

Perciò, fermo il fatto che la personalità giuridica non dipende dall'appartenenza al genere umano, questi ritengono che, però, possa essere considerata persona giuridica unicamente l'entità non umana che sia legalmente responsabile per il proprio comportamento e sia capace di essere titolare di diritti e doveri. In altri termini, richiedono che l'entità oltre che autonomia abbia consapevolezza e comprensione della propria intelligenza.

---

Kamm afferma che le cose possono essere fatte «per il bene di X» se X è senziente o cosciente (o ha la capacità di essere senziente o cosciente). L'a. sostiene che la capacità di ragionare e la capacità di soffrire sono proprietà distinte che possono essere separate, ma che nessuna delle due è necessaria e né sufficiente per questa versione dello *status* morale.

È necessario, quindi, che l'entità sia un cd. agente morale ovvero goda della *moral agency*. Tale concetto nel sistema ordinamentale italiano non è molto conosciuto ma può bene essere spiegato riferendosi alla consapevolezza dell'entità della propria capacità di influenzare l'ambiente esterno. Consapevolezza che, come ben spiega Solum<sup>63</sup> nel suo saggio diretto a comprendere la convenienza di una soggettività dell'IA, immaginando sia lo scenario in cui questa operi come amministratore fiduciario sia quello in cui la stessa sia titolare di diritti costituzionalmente garantiti, ben sfugge al test di Turing. Perciò, il superamento di quest'ultimo non confermerebbe la presenza di un'intelligenza coscienziosa. Piuttosto, ribadirebbe la capacità di simulazione antropomorfa da parte dell'IA interrogata.

Alla luce di tali considerazioni, i sostenitori di tale orientamento ritengono che, allo stato attuale, i limiti tecnologici ancora caratteristici di ogni IA esistente, non permettano di riconoscere le IA come persone o soggetti giuridici che si voglia. Basti considerare che, oggi, non esiste nemmeno un'IA di tipo generale in grado di decidere autonomamente quale attività compiere alla luce di un'analisi azioni-conseguenze indipendentemente intrapresa.

Ciò non toglie che gli stessi autori sono favorevoli al riconoscimento della personalità giuridica nel prossimo futuro, momento in cui le IA, in qualche modo, saranno consapevoli della loro esistenza e, pertanto, saranno responsabili delle loro azioni anche penalmente in quanto titolari di diritti ma, soprattutto, di doveri.

### **3.1.3. Segue...il *tertium genus* della personalità elettronica.**

Altri autori si distanziano da entrambi gli orientamenti menzionati ritenendo preferibile e più opportuno, di fronte alle IA, la creazione di un centro di imputazione di diritti e doveri diverso sia dalle persone fisiche sia dalle persone giuridiche.

Benchè le voci sul punto, all'interno del filone dottrinario, siano varie e non cantino all'unisono, il minimo comun denominatore può essere identificato nell'intenzione di creare una vera e propria personalità elettronica.

---

<sup>63</sup> L.B. SOLUM, *Legal personhood*, cit., p. 1231- 1287.

Sicuramente, tra molti<sup>64</sup>, emerge, per originalità e coerenza metodologica, la teoria dei sistemi di Teubner<sup>65</sup> il quale tiene ben distinto il concetto di «persona», come artefatto semantico della comunicazione attraverso cui si permette ad un attore l'ingresso nel sistema e la contemporanea attribuzione di una capacità di agire, da quello di «individuo» che è correlato all'idea di coscienza e vita psichica interna<sup>66</sup>.

Teubner riprende e rivitalizza gli interrogativi sull'«essenza» della persona giuridica, che hanno tormentato generazioni di studiosi, valorizzando il dato dell'autonomia. L'organizzazione è un valore aggiunto, l'intero è di più della somma delle sue parti, non limitandosi a rappresentarne il semplice raggruppamento ma qualcosa di diverso.

Secondo tale visione, quindi, i sistemi di IA o, come Teubner indica, «gli agenti software sono – al pari delle imprese o altre organizzazioni formali – null'altro che meri flussi informativi, i quali diventano “persone” o (persone parziali), quando nel processo comunicativo pervengono ad una identità sociale e ad essi sia effettivamente attribuita una propria capacità di agire, insieme alle

---

<sup>64</sup> Si fa riferimento, quantomeno, in primo luogo, alla proposta di R. VAN DEN HOVEN VAN DEN GARDEN, *Do We Need New Legal Personhood in the Age of Robots and AI?*, in M. CORRALES, M. FENWICK e N. FORGÓ (a cura di) *Robotics, AI and the Future of Law*, Singapore, 2018, p. 15-55, il quale ritiene che, ai fini di un uso controllato dell'IA, la previsione di una personalità elettronica *sui generis* sia di aiuto. All'uopo P.a. afferma che, per capire in cosa questa consista, bisogna partire dalla teoria Naffiniana del gatto del Cheshire, riscontrare la razionalità tipica del modello di agente razionale ed individuare gli altri elementi di volta in volta richiesti dalla legge ossia: la necessità nella società «umana», rilevanza socio-economica, necessità di certificazione legale; determinazione dell'intelligenza come autonoma, simil test di Turing, livello di imitazione umana; sufficiente intelligenza sociale; l'entità IA deve essere in grado di comprendere il valore socio-emotivo e morale delle dichiarazioni di altre parti per rispondere in modo appropriato in modo che ci sia una base equivalente per il consenso; essere in grado di rispondere a circostanze mutevoli (questo aspetto si potrebbe chiamare intelligenza «adattiva o dinamica»; accettazione e fiducia da parte di altre persone giuridiche e naturali in modo da potersi integrare nelle interazioni economiche, sociali e legali; un registro pubblico che specifichi quali robot avranno specifiche competenze legali per determinati ruoli e compiti; un codice etico che tenga conto anche dell'uso di diverse categorie di robot, nonché delle regole predefinite necessarie per gli sviluppatori e i produttori di robotica. In un'elaborazione più estesa di questa idea, si potrebbe istituire un fondo finanziato da una certa percentuale dei guadagni dei robot per garantire eventuali perdite o danni.

In secondo luogo, si menzionano gli studi di J-S. GORDON il quale ancora la personalità elettronica alla *moral agency*. Per approfondire si v. J-S. GORDON, O. RAUPRICH e J. VOLLMAN, *Applying the four principles approach*, in *Bioethics*, 2011, vol. 25, 6, p. 293-300; J-S. GORDON, *Remarks on a disability-conscious bioethics*, in J-C. PÖDER, H. BURCKHART e J-S. GORDON (a cura di), *Human rights and disability. Interdisciplinary Perspectives*, London, 2017, p. 9- 20; ID. e F. TAVERA-SALYUTOV, *Remarks on disability rights legislation*, in *Equal. Divers. Incl.*, 2018, vol. 37, 5, p. 506-526; ID., *What do we owe to intelligent robots?*, in *AI Society*, 2020, vol. 35, 2, p. 209-223; ID., *Artificial, moral and legal personhood*, in *AI Society*, 2021, 36, p. 457-471. L'a. parte dalla presunzione secondo cui, comunemente, gli esseri (almeno quelli umani) dotati di personalità giuridica hanno anche la personalità morale e che gli esseri che possiedono la personalità morale la abbiano perché si ritiene che sussista uno *status* morale sufficiente. In altre parole, senza *status* morale non ci può essere né personalità morale né giuridica. Al contrario, non è detto che tutto ciò che goda di personalità giuridica abbia personalità morale. Si v. il dibattito ancora vivo sulla esistenza o meno della personalità morale delle società o dei fondi fiduciari, pacificamente considerati soggetti giuridici. Sarebbe un caso di mettere il carro davanti ai buoi. Sembra, quindi, che il concetto di personalità giuridica per le entità non umane non richieda la personalità morale, ma solo un elevato *status* morale.

<sup>65</sup> G. TEUBNER, *Rights of Non-humans? Electronic agents and animals as new actors in politics and law*, Badia Fiesolana, 2007; ID., *Soggetti giuridici digitali*, Napoli, 2019.

<sup>66</sup> G. TEUBNER, *Ibridi ed attanti*, cit., p. 40.

necessarie disposizioni organizzative, quali, ad esempio, le regole di rappresentanza»<sup>67</sup>.

Teubner, partendo dalla teoria dell'attore-rete<sup>68</sup>, tende ad un allargamento della soggettività giuridica, anche nei confronti di candidati sorprendenti per una visione generale del mondo tradizionalmente «a misura d'uomo»: da un lato, animali ma, in una prospettiva più ampia, anche alberi e ambiente in generale; dall'altro, agenti elettronici.

La svolta della teoria dei sistemi risiede proprio nella spersonalizzazione ovvero nell'attribuzione della capacità di agire a configurazioni sociali considerate come un insieme di comunicazioni collegate tra loro. Ci si distacca da un antropocentrismo fondato su premesse individualistiche sempre più inadeguate a rendere conto della complessità dei problemi. L'impostazione sistemica dialoga in questo ambito con la teoria degli attanti di Latour<sup>69</sup>, che intende operare una redistribuzione della capacità di agire sulla scena sociale tra umani e non umani. Non si tratta, naturalmente, di equiparare *tout court* attori così diversi tra loro, riconoscendo a tutti in eguale misura diritti, doveri e corrispondenti tutele, ma di ricostruire, all'interno di ciascun sistema e secondo i criteri elaborati da ognuno di questi, gli artefatti delle persone cui è possibile, attraverso un meccanismo di imputazione dettato dal sistema stesso, ascrivere operazioni comunicative come loro azioni.

Un ulteriore filone, basandosi sulle teorie di Norbert Wiener<sup>70</sup>, il quale riconosceva la somiglianza tra la macchina automatica e la posizione dello schiavo, che è l'archetipo della «cosa intelligente», ritiene di accostare i sistemi di IA, specie se incorporati in dei robot, proprio agli schiavi per la peculiarità di tale categoria. È vero che entrambe queste entità possono essere ritenute una cosa della quale il padrone può disporre liberamente ma, al tempo stesso, possono essere anche un agente del rispettivo proprietario, in grado di svolgere azioni produttive, idonee a porre in essere effetti rilevanti, tali da modificare la realtà

---

<sup>67</sup> G. TEUBNER, *o.u.c.*, p. 42. Si v. anche E.M. WEITZENBOECK, *Electronic Agents and the Formation of Contracts*, in *International Journal of Law and Information Technology*, 2001, III, p. 204-234, il quale individua tre profili su cui fondare il riconoscimento della personalità giuridica: l'autorità morale, la capacità sociale e la convenienza legale.

<sup>68</sup> G. TEUBNER, *Soggetti*, cit., nota 40 afferma che la «teoria dell'attore-rete dissolve la distinzione fra soggetto e oggetto. L'azione è il risultato dell'interazione tra soggetti umani e oggetti mobilmente connessi in rete, all'interno della quale non si deve parlare di "attori", ma di "attanti", che possono essere anche non umani. Soprattutto, entro le reti, la connessione tra soggetti e oggetti produce continuamente degli ibridi: un uomo al volante della sua auto, ad es., non è un soggetto attivo più un oggetto inerte, ma va considerato un unico centro comportamentale, non isolabile dalla rete della circolazione, anch'essa in continua mutazione».

<sup>69</sup> Già B. LATOUR, *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford, 2005, p. 21 ss., affermava che i «processi digitali capaci di azione sono ricondotti sotto il neologismo "attante". ...non si tratta affatto di una antropomorfizzazione dei processi digitali, ma, proprio all'opposto, di una de-antropomorfizzazione degli agenti software: come affermano gli agenti software [in quanto attanti], essendo ormai loro direttamente imputabili i comportamenti sono divenuti membri non umani della società, *maidless machines*, macchine senza mente [o senz'anima]».

Sul punto M. CALLON e B. LATOUR, *Unscrewing the big leviathan: How actors macrostructure reality and how sociologists help them to do so* in K., KNORR-CETINA e A.V. CICOUREL (a cura di), *Advances in Social Theory and Methodology: Toward an Integration of Micro- and Macro Sociologies*, London-Boston, 1981, p. 276 ss.; J. LAW (a cura di), *Power, Action, and Belief: A New Sociology of Knowledge?*, London-Boston, 1986.

<sup>70</sup> N. WIENER, *The Human Use of Human Beings: Cybernetics and Society*, Boston, 1950.

sul piano del diritto. Se si fa un paragone con la posizione dello schiavo romano è giocoforza ricordare che, benchè il diritto romano non abbia mai concesso una personalità giuridica allo schiavo, il rapporto tra lo schiavo e il suo padrone e il rapporto tra lo schiavo e la società nel suo complesso era, comunque, più che strumentale. Gli schiavi potevano svolgere una funzione di rappresentanza legale, compiere atti giuridici indipendenti e comparire come testimoni in tribunale. Inoltre, lo schiavo poteva essere dichiarato «uomo libero» dal suo padrone attraverso la *manumissione*.

Ebbene, non stonerebbe con il riferimento storico l'attribuzione di una certa capacità giuridica al sistema di IA. Si tratterebbe di una capacità afferente ad una forma di personalità diversa ed indipendente dal concetto di personalità esistente, una costrutto *sui generis* teso a riconoscere autonomia legale in conformità al ruolo che il sistema di IA ricopre nella società, allo scopo che esso deve raggiungere.

Perciò, all'interno di detto orientamento vi è chi, come Pagallo<sup>71</sup>, creativamente, suggerisce, per alcuni tipi di robot, individuati sulla base delle funzioni, di ricorrere al meccanismo del *peculium* degli schiavi abilitati ad agire – in qualità di manager, banchieri o commercianti ecc. – mediante la previsione di un portfolio o *digital peculium* affinché possano essere considerati direttamente responsabili delle loro azioni e quindi «the idea is that, at times, only “robots shall pay” could be the right answer» sebbene sia necessario operare una distinzione tra i diversi tipi di attività svolta dai robot fino ad arrivare a desiderare a «robotic personal assistant such as a sort of i-Jeeves that helps us schedule a set of conferences, lectures and meetings at several European (or US) universities».

Pagallo, citando Chopra and White<sup>72</sup>, ha però sottolineato che, dal punto di vista della affidabilità e credibilità giuridica, per l'accettazione di azioni giuridiche aventi efficacia legale, è necessario conoscere esattamente quali siano le attribuzioni legali dell'agente ed il suo scopo. All'uopo, per una maggiore affidabilità, potrebbe essere utile, da una parte, prevedere l'esistenza di un registro pubblico ove riportare le competenze legali di ogni sistema di IA<sup>73</sup>, dall'altra, ancorare la qualificazione giuridica dei sistemi IA in base alla diversa autonomia e funzione ad un parametro uniforme. Magari ci si potrebbe riferire alle norme ISO<sup>74</sup> che già hanno effettuato una classificazione dei robot/sistemi di IA in base alla sicurezza e al tipo di interazione-collaborazione con gli uomini.

---

<sup>71</sup> U. PAGALLO, *The Law of robots. Crimes, Contract and Torts*, Heidelberg, 2013, pp. 82 e 102 ss.

<sup>72</sup> S. CHOPRA e L.F. WHITE, *A Legal Theory for Autonomous Artificial Agents*, Ann Arbor, 2011, p.130.

<sup>73</sup> R. VAN DEN HOVEN VAN DEN GARDEN, *Do We*, cit., p. 49.

<sup>74</sup> ISO 13, 482:2014, ad esempio, specifica i requisiti e le linee guida per una progettazione intrinsecamente sicura, misure di protezione e informazioni per l'utilizzo di robot per l'assistenza personale, in particolare i seguenti tre tipi di robot: robot servitore mobile; robot assistente fisico; robot trasportatore di persone.

### 3.2. IA come oggetto del diritto.

Orbene, nonostante l'autorevolezza di ciascuna visione, il timore di fare il passo piú lungo della gamba unito alle molte contestazioni di stampo giuridico- economico hanno reso il terreno inospitale per l'attecchimento delle tesi a supporto della soggettività-personalità dei sistemi di IA.

Dopo tutto, che senso avrebbe rendere un'entità - che non può comprendere la legge né agire su di essa - il soggetto di un dovere legale<sup>75</sup>?

Le contestazioni avanzate sono diverse.

Innanzitutto, secondo alcuni studiosi<sup>76</sup> il sistema che vedesse i robot e i sistemi di IA equiparati, in qualche modo, alle persone naturali sarebbe troppo costoso. Il peso dell'autonoma responsabilità patrimoniale delle identità tecnologiche sovrasterebbe i vantaggi conseguiti tramite il loro utilizzo, specie i benefici morali. E, in accordo a tale visione, il perseguimento e la tutela degli interessi morali degli esseri umani sarebbe il cuore dell'intero sistema ordinamentale<sup>77</sup>, tant'è che solo l'essere umano, in quanto titolare di «diritti umani» potrebbe essere ritenuto «persona»<sup>78</sup>.

La personalità giuridica delle IA, perciò, non potrebbe derivarsi né dal modello della persona fisica, sol si consideri che i robot, se cosí fosse, diverrebbero titolari di diritti umani, come il diritto alla dignità, all'integrità, al risarcimento e alla cittadinanza, e questo non è accettabile in un ordinamento in cui solo gli esseri umani sono titolari di diritti umani, né potrebbe derivarsi dal modello della persona giuridica dal momento che questa richiede comunque l'esistenza di almeno una persona umana dietro il suo velo e i sistemi di IA, si è detto, non sono persone umane.

Tale orientamento, tuttavia, è sembrato ai piú<sup>79</sup> fuorviante e destinato a fallire. Si baserebbe, infatti, su una carente comprensione del concetto di persona. L'idea che solo gli esseri umani siano persone è una posizione molto controversa in filosofia morale, in particolare nei contesti di discussione sui diritti degli animali<sup>80</sup>. Si tratterebbe di un caso di specieismo che confliggerebbe con la

---

<sup>75</sup> Cfr. J.C. GRAY, *The nature and sources of the law*, MacMillan, 1921, il quale era scettico in relazione all'idea che una cosa inanimata potesse essere considerata una persona giuridica. Nella sua ottica anche le società erano riducibili alle relazioni tra le persone che ne possiedono le azioni, le gestiscono e cosí via.

<sup>76</sup> Si v. sul punto J.J. BRYSON, M.E. DIAMANTIS e T.D. GRAND, *Of, for, and by the people: the legal lacuna of synthetic persons*, in *Artif. Intell. Law*, 2017, vol. 25, 3, p. 273-291.

<sup>77</sup> Secondo J.J. BRYSON, M.E. DIAMANTIS, T.D. GRAND, *Of*, cit., p.283, gli scopi principali del «sistema giuridico umano» sarebbero i seguenti: 1) promuovere gli interessi materiali delle persone giuridiche; 2) far valere come diritti e obblighi legali qualsiasi diritto e obbligo morale importante (cioè, sufficientemente rilevante), con la successiva diffida; 3) dare la preferenza ai diritti morali degli esseri umani, se i loro diritti morali si scontrano con i diritti morali altrettanto importanti di entità non umane.

<sup>78</sup> La linea generale di ragionamento fornita da J.J. BRYSON, M.E. DIAMANTIS, T.D. GRAND, *Of*, cit., p. 283, rivela un forte pregiudizio nei confronti degli esseri artificialmente intelligenti e soffre di quello che può chiamarsi «specieismo».

<sup>79</sup> Fra tutti molto rilevanti sono state le contestazioni sollevate da D. GUNKEL, *The machine question: critical perspectives on AI, robots, and ethics*, Cambridge Mass., 2012; ID., *A vindication of the rights of machines*, in *Philos. Technol.*, 2014, vol. 27, 1, p. 113-132; ID., *Robot rights*, Cambridge Mass., 2018.; D.J. GUNKEL e J. BRYSON, *Introduction to the special issue on machine morality: the machine as moral agent and patient*, in *Philos. Technol.*, 2014, vol. 27, 1, p. 5-142.

<sup>80</sup> Cfr. P. CAVALIERI, *La questione animale: perché gli animali non umani meritano i diritti umani*, Oxford, 2001; D.J. CALVERLEY, *Android science and animal rights: does an analogy exist?*, in *Connect. Sci.*, 2006, vol. 18, 4, p. 403-

crescente inclusione tra i titolari di diritti di soggetti non umani, come i mammiferi superiori (attraverso il movimento per i diritti degli animali) e la natura (attraverso il movimento per i diritti dell'ambiente)<sup>81</sup>.

In secondo luogo, altra parte della dottrina<sup>82</sup>, ha rimarcato la consistente difficoltà nel considerare persona l'IA quante volte il «valore della persona passa per la carne, il volto in cui si evidenzia la fragilità e la sofferenza delle scelte, non la meccanica degli algoritmi».

Altresí, è stato mostrato scetticismo nei confronti dell'esigenza di riconoscere alle IA soggettività giuridica considerato che «si può essere responsabili, e titolari di risorse patrimoniali, anche in assenza di personalità giuridica, o comunque di equiparazione allo *status* della persona umana. In questa direzione, i difetti di progettazione, di fabbricazione e di informazione costituiscono tipici rischi d'impresa e vanno imputati oggettivamente al produttore. Qualora, poi, distinte condotte, pur tra loro soggettivamente indipendenti, abbiano concorso a causare un medesimo evento dannoso, la responsabilità non può che essere solidale<sup>83</sup>».

Alla luce di quanto riportato, al di là della criticità specifica di ognuna delle riportate tesi, si afferma come maggioritaria la visione di chi ricorre, quindi, al «paradigma dello strumento» per qualificare giuridicamente l'IA, indifferentemente dal grado di autonomia. E ciò, si badi, appare rafforzato dai portati dell'immaginazione collettiva nel momento in cui, laddove si pensi al futuro dei sistemi di IA, l'approccio naturale è quello di figurarsi di interagire con loro similmente al modo in cui si interagisce con i programmi software: come

---

417; P. SINGER, *Speciesism and moral status*, in *Metaphilosophy*, 2009, vol. 40, 3-4, p. 567-581; E. SCHAEFER, R. KELLEY e M. NICOLESCU, *Robots as animals: A framework for liability and responsibility in human-robot interaction*. In *Robot and human interaction communication. RO-MAN 2009—The 18th IEEE international symposium on robot and human interactive communication*, in *Journal Advanced Robotics*, 2009, vol. 24, 13, p. 1861-1871.

<sup>81</sup> D. GUNKEL, *Robot*, cit. non sostiene che tutti gli animali e tutti gli oggetti ambientali debbano avere lo stesso *status* giuridico degli esseri umani nel nostro sistema giuridico ma, piuttosto, che un sistema giuridico costruito sullo specicismo è destinato a fallire nel lungo periodo. L'idea che i robot non debbano essere trattati come persone giuridiche, perché sarebbe troppo costoso visti i limitati guadagni morali, sembra una posizione debole ed insostenibile considerato che i robot sono destinati a diventare più intelligenti degli uomini. Fermo che ad oggi i robot non godono di personalità giuridica, a questi potrebbero applicarsi in estensione o le norme disciplinanti gli animali o gli enti/società.

Fra gli esponenti che ritengono di estendere le norme sugli enti/società fra molti si ricorda G. HALLEVY, *The criminal liability of artificial intelligence entities*, in *ssrn.com*, 2010, p. 1-42; ID., *Virtual criminal responsibility*, in *ssrn.com*, 2010, p. 1-22; ID., *When robots kill. Artificial intelligence under criminal law*, Boston, 2013.

Fra coloro che, invece, estendono le norme sugli animali si richiama K. DARLING, *Extending legal protection to social robots: the effects of anthropomorphism, empathy, and violent behavior towards robotic objects*, in R. CALO, M.A. FROOMKIN e I. KERR (a cura di), *Robot law*, Northampton, 2016, p. 213-231; T.L. JAYNES, *Legal personhood for artificial intelligence: citizenship as the exception to the rule*, in *AI Soc.*, 2020, vol. 35, 2, p. 343 ss. Per una visione critica, si v. A. BERTOLINI, *Robot companions*, cit., p. 227-231, il quale rifiuta l'analogia tra robot e animali principalmente a causa delle differenze nel loro *status* ontologico, in quanto gli animali sono naturali e i robot sono artificiali.

<sup>82</sup> A. GORASSINI, *Il valore della cultura giuridica nell'era digitale*, in *Tecn. Dir.*, 2021, 2, p. 38 ss.

<sup>83</sup> P. PERLINGIERI, *Presentazione*, in ID., S. GIOVA e I. PRISCO (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*, *Atti del 15° Convegno Nazionale Sisdic. Napoli 14-15-16 maggio 2020*, Napoli, 2020, p. X.

specifici strumenti usati per raggiungere specifici obiettivi, basati su competenze prestabilite, attraverso interfacce specialmente disegnate<sup>84</sup>.

In conclusione, ad oggi, i robot ed ogni sistema di IA, in generale, sarebbero, quindi, da considerarsi oggetti di diritto<sup>85</sup>, piú precisamente artefatti creati dalla progettazione e dal lavoro umano, allo scopo di servire bisogni umani identificabili.

### 3.2.1. IA come prodotto...nel d.lg. del 6 settembre 2005, n. 206 (Codice del Consumo italiano).

Si è detto, quindi, che ad oggi convince maggiormente l'idea secondo cui un sistema di IA venga considerato un oggetto di diritto. Se questo è vero, lo sarebbe altrettanto quello di considerare i sistemi di IA, a tutti gli effetti giuridici ed etici, dei prodotti<sup>86</sup>.

Cosí opinando, infatti, non vi sarebbe alcuna lacuna normativa, nessun *responsability gap*<sup>87</sup>.

Al contrario, potrebbero ben soccorrere le regole sulla responsabilità già presenti negli ordinamenti nazionali (si rinvia al cap. III) ed, in particolare, se si considera un sistema di IA un prodotto, potrebbe applicarsi lo statuto della responsabilità del produttore per ogni danno cagionato dallo stesso<sup>88</sup>.

---

<sup>84</sup> J. TALLINN e R. NGO, *Automating Supervision of AI delegates*, in S. VOENEKY, P. KELLMEYER, O. MUELLER e W. BURGARD, (a cura di), *The Cambridge Handbook of Responsible Artificial Intelligence. Interdisciplinary perspectives*, Cambridge, 2022, p. 19-30.

<sup>85</sup> Di tal avviso è A. PROCIDA MIRABELLI DI LAURO, *Le intelligenze artificiali tra responsabilità civile e sicurezza sociale*, *Atti del 15° convegno nazionale Sisdic. Napoli 14-15-16 maggio 2020*, Napoli, 2020, p. 297.

<sup>86</sup> Cosí afferma A. BERTOLINI, *Robot companions*, cit. p. 130 ss.

<sup>87</sup> C. PERLINGIERI, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in *Rass. dir. civ.*, 2015, 4, p. 1235 ss. stigmatizza le lacune normative, specie quelle che potrebbero tramutarsi in un *responsability gap*, pregiudicando, in qualche modo, lo sviluppo economico, gli scambi commerciali e l'intero progresso tecnologico. Queste, infatti, non solo potrebbero dissuadere gli imprenditori dall'investire in determinati settori e produrre determinati oggetti, ma potrebbero anche allontanare gli utenti dall'acquisto degli stessi. G. TEUBNER, *Soggetti*, cit., p. 86 ne rinviene una proprio nell'ambito della responsabilità extracontrattuale ove «il fallimento del diritto vigente a fronte del rischio di autonomia risulta in tutta la sua evidenza».

<sup>88</sup> Per un esame sistematico della normativa sulla responsabilità del produttore, senza pretesa di esaustività, si rinvia, fra molti, ad G. ALPA, R. BIN e P. CENDON (a cura di), *La responsabilità del produttore*, in *Tratt. Galgano*, XIII, Padova, 1989; A. GORASSINI, *Contributo per un sistema della responsabilità del produttore*, Miano, 1990; S. PATTI (a cura di), *Il danno da prodotti in Italia, Austria, Repubblica federale di Germania, Svizzera*, Padova, 1990; G. PONZANELLI, *Responsabilità del produttore*, in *Riv. dir. civ.*, 1990, II, p. 529 ss.; G. ALPA, U. CARNEVALI, F. DI GIOVANNI, G. GHIDINI, U. RUFFOLO e C.M. VERARDI, *La responsabilità per danno da prodotti difettosi*, Milano, 1990; U. CARNEVALI, *Responsabilità del produttore*, *Enc. dir., Agg. II*, Milano, 1998, p. 936 ss.; M. FRANZONI, *Dieci anni di responsabilità del produttore*, in *Danno resp.*, 1998, p. 823 ss.; A. LUPOLI, *La responsabilità per prodotti difettosi*, in F. BOCCHINO (a cura di), *Diritto dei consumatori e nuove tecnologie*, II, *Il mercato*, Torino, 2003, p. 39 ss.; A. ALBANESE, *La sicurezza generale dei prodotti e la responsabilità del produttore nel diritto italiano ed europeo*, in *Egr. dir. priv.*, 2005, p. 977 ss.; P. STANZIONE e G. SCIANCALEPORE (a cura di), *Commentario al codice del consumo, inquadramento sistematico e processo applicativo*, Milano, 2006, p. 773 ss.; G. STELLA, *La responsabilità del Produttore per danno da prodotto difettoso nel nuovo codice del consumo*, in *Resp. civ. prev.*, 2006, p. 1589 ss.; F. CARINGELLA, *La responsabilità del produttore*, in F. CARINGELLA e G. DE MARZO (a cura di), *I contratti dei consumatori*, Milanofiori Assago, 2007, p. 547 ss.; L. CABELLA PISU, *Ombre e luci nella responsabilità del produttore*, in *Contr. impr./Eur.*, 2008, p. 617 ss.; A.L. BITETTO, *La responsabilità del produttore: da mera comparsa a protagonista*, in *Riv. crit. dir. prin.*, 2008, p. 137 ss.; ID., *Dal biscotto 41 pan carré: il tortuoso percorso della responsabilità da prodotto*, in *Danno resp.*, 2012, 1, p. 80 ss.; U. CARNEVALI, *Produttore e responsabilità per danno*

Certamente, sostanziandosi l'algoritmo in una formula matematica generata dall'intelletto umano, piú naturale sembrerebbe attribuire le conseguenze patologiche al suo autore-ideatore, il quale, talvolta, coincide col produttore del device in cui l'IA è inserita ma, talvolta, diverge.

Se pare evidente la responsabilità nei confronti dei terzi danneggiati dell'autore-fabbricante del *device*, potrebbe, invece, sembrare piú problematico ipotizzare una autonoma responsabilità in capo al mero autore dell'algoritmo predittivo, stante la difficoltà di ricondurre l'algoritmo alla nozione di prodotto.

Secondo l'art. 115 cod. cons. italiano sono prodotti difettosi i beni mobili fabbricati, i prodotti agricoli del suolo e i prodotti dell'allevamento, della pesca e della caccia che non presentano la sicurezza che è ragionevole attendersi in relazione alle circostanze.

L'art. 117 cod. cons. specifica che, ai fini della valutazione della sicurezza occorre tener conto «di tutte le circostanze, tra cui: a) il modo in cui il prodotto è stato messo in circolazione, la sua presentazione, le sue caratteristiche palesi, le istruzioni e le avvertenze fornite; b) l'uso al quale il prodotto può essere ragionevolmente destinato e i comportamenti che, in relazione ad esso, si possono ragionevolmente prevedere; c) il tempo in cui il prodotto è stato messo in circolazione. 2. Un prodotto non può essere considerato difettoso per il solo fatto che un prodotto piú perfezionato sia stato in qualunque tempo messo in commercio. 3. Un prodotto è difettoso se non offre la sicurezza offerta normalmente dagli altri esemplari della medesima serie.»

Come si evince dal testo delle norme, benché ampia, la nozione di prodotto non sembra apprestarsi a ricomprendere i moderni congegni tecnologici, soprattutto se richiedono un collegamento alla rete internet per svolgere la propria funzione, avvicinandosi, quindi, sempre piú alla nozione di servizio piuttosto che prodotto.

Solo facendo assurgere l'algoritmo a componente immateriale caratterizzante il prodotto, qualificandosi così l'ideatore come autore-fornitore<sup>89</sup> di quella componente, allora è possibile che questi si esponga alla responsabilità verso i terzi lesi dal prodotto difettoso intelligente.

---

*da prodotto difettoso nel codice del consumo*, in *Resp. civ. prev.*, 2009, 10, p. 1938 ss.; M. GIUFFRIDA, *Responsabilità per danno da prodotto difettoso*, in *Dig. Disc. Priv., Sez. civ., Agg. IV*, Torino, 2009, p. 453; A. QUERCI, *Responsabilità da prodotto negli USA e in Europa. Le ragioni di un revirement globale*, in *Nuova giur. Civ. comm.*, 2011, p. 128 ss.; E. AL MUREDEN, *La sicurezza dei prodotti e la responsabilità del produttore*, Torino, 2017; P. PARDOLESI, *Riflessioni sulla responsabilità da prodotto difettoso in chiave di analisi economica del diritto*, in *Riv. dir. priv.*, 2017, p. 87 ss.; G. ALPA (a cura di), *La responsabilità del produttore*, Milano, 2019; A. SPANGARO, *Il danno da prodotto conforme: dai derivati del tabacco ai telefoni cellulari*, in *Giur. it.*, 2019, p. 1312 ss.

<sup>89</sup> Così F. FRATTARI, *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in *Contr. e impresa*, 2020, I, p. 452 ss., richiamando un passaggio della Risoluzione UE del 2017, dal quale emerge sia il ruolo fondamentale del programmatore, che per distribuire la responsabilità per i danni sia opportuno prendere atto che «quanto maggiore è la capacità di apprendimento o l'autonomia di un robot e quanto è maggiore la durata della formazione di un robot, tanto maggiore dovrebbe essere la responsabilità del suo formatore»; a quel punto: «il principale indiziato a candidarsi al ruolo di “formatore” non può che essere l'autore dell'algoritmo, i cui inputs guidano il robot nel raccogliere i dati provenienti dalla realtà circostante, nell'elaborarli e, sulla base di questi, nel modulare il loro operato». Sul punto v. anche G. FINOCCHIARO, *Intelligenza artificiale e responsabilità*, in *Contr. e impresa*, 2020, II, p. 716 ss.

L'applicazione della disciplina della responsabilità del produttore incontra, però, ad ogni modo, alcune difficoltà dovute, soprattutto, al profondo mutamento del contesto produttivo e commerciale.

Innanzitutto, come anticipato, risulta complicato far rientrare l'IA forte nella nozione di prodotto ai sensi degli artt. 3 comma 1, lett. e), 18 comma 1, lett. c) e 115 cod. cons.<sup>90</sup>.

Benché il d.lg. del 7 marzo 2023, n. 26<sup>91</sup> abbia disposto con l'art. 1, comma 3, lettera a) la modifica dell'art. 18, comma 1, lettera c), allargando la nozione di prodotto anche ai servizi digitali e al contenuto digitale, il rinnovamento non appare sufficiente per abbracciare i sistemi di IA predetti sol si consideri che non li si può ridurre a meri beni, servizi digitali o contenuti digitali. Il sistema di IA, come visto nel cap. I, non è unicamente un insieme di dati informatici ma – secondo ad esempio le parole di Munakata<sup>92</sup> – comprende l'abilità di inferire sulla base della conoscenza, di ragionare con informazioni incerte o incomplete, di percepire e apprendere in vario modo e di essere applicata per la risoluzione di problemi di controllo, previsione, classificazione e ottimizzazione.

Inoltre, pur se si volesse forzare la mano e far rientrare nella nozione di prodotto di cui all'art. 18 cod. cons. anche i sistemi di IA, deve dirsi che la norma ha un ambito applicativo limitato. È proprio il suo *incipit* a dircelo: quanto disposto si applica «ai fini del presente titolo» ossia ai fini del Titolo III della Parte II del cod. cons., relativo alle pratiche commerciali, pubblicità e altre comunicazioni commerciali. In nessun modo è richiamata la parte circa la responsabilità del produttore.

In secondo luogo, il risultato pregiudizievole generato dall'autonomo processo decisionale della IA sembra non allinearsi col concetto di «difetto» del prodotto. Con quest'ultimo l'art. 117 cod. cons. si riferisce alla mancanza di sicurezza che legittimamente ci si può attendere dal prodotto, tenendo a mente le modalità ed il momento di messa in circolazione e l'uso cui è destinato.

---

<sup>90</sup> Art. 3, comma 1, lett. e) cod. cons.: «prodotto: fatto salvo quanto stabilito nell'art. 18, comma 1, lettera c), e nell'art. 115, comma 1, qualsiasi prodotto destinato al consumatore, anche nel quadro di una prestazione di servizi, o suscettibile, in condizioni ragionevolmente prevedibili, di essere utilizzato dal consumatore, anche se non a lui destinato, fornito o reso disponibile a titolo oneroso o gratuito nell'ambito di un'attività commerciale, indipendentemente dal fatto che sia nuovo, usato o rimesso a nuovo; tale definizione non si applica ai prodotti usati, forniti come pezzi d'antiquariato, o come prodotti da riparare o da rimettere a nuovo prima dell'utilizzazione, purché il fornitore ne informi per iscritto la persona cui fornisce il prodotto».

Art. 18 comma 1, lett. c) cod. cons.: «prodotto: qualsiasi bene o servizio, compresi i beni immobili, i servizi digitali e il contenuto digitale, nonché i diritti e gli obblighi».

Art. 115 cod. cons. «prodotto e produttore» «1. Prodotto, ai fini del presente titolo, è ogni bene mobile, anche se incorporato in altro bene mobile o immobile. 2. Si considera prodotto anche l'elettricità. 2-bis. Produttore, ai fini del presente titolo, è il fabbricante del prodotto finito o di una sua componente, il produttore della materia prima, nonché, per i prodotti agricoli del suolo e per quelli dell'allevamento, della pesca e della caccia, rispettivamente l'agricoltore, l'allevatore, il pescatore ed il cacciatore.

<sup>91</sup> In G.U. 18 marzo 2023, n. 66. Il d.lg. indicato è di attuazione della direttiva (UE) 2019/2161 del Parlamento europeo e del Consiglio, del 27 novembre 2019, che modifica la direttiva 93/13/CEE del Consiglio e le direttive 98/6/CE, 2005/29/CE e 2011/83/UE del Parlamento europeo e del Consiglio per una migliore applicazione e una modernizzazione delle norme dell'Unione relative alla protezione dei consumatori.

<sup>92</sup> T. MUNAKATA, *Fundamentals of the New Artificial Intelligence*, London, 2008.

Nel codice del consumo si fa riferimento ad una deviazione dagli standard di sicurezza<sup>93</sup> da valutarsi sulla base del cd. *consumer expectation test* «veicolante un criterio di carattere “relazionale” la cui interpretazione è ampiamente lasciata alla pratica delle Corti la quale, in uno con la dottrina, ha coniato la ormai consolidata articolazione del concetto di “difetto” secondo la tripartizione in difetto di fabbricazione, difetto di progettazione e difetto di informazione»<sup>94</sup>.

Ecco, non è detto che il precipitato decisionale algoritmico sia pregiudizievole perché rappresenti una deviazione dagli standard di sicurezza.

Altresì, non è detto che il «difetto» sia genetico e rientrante in una della tre ipotesi menzionate.

Ben può essere che questo sopravvenga proprio nel processo di autoapprendimento. In caso di sistema intelligente, così, il difetto potrebbe concretizzarsi, non solo in un errore di progettazione, di fabbricazione o di programmazione<sup>95</sup>, ma anche in un errore di aggiornamento o apprendimento<sup>96</sup>. Di fatti, è possibile che il prodotto messo in commercio privo di difetti genetici possa modificare la propria condotta in base all’esperienza maturata giungendo a dar la stura a conseguenze indesiderate nonché dannose<sup>97</sup>.

In terzo luogo, risulta estremamente pesante per il danneggiato il regime probatorio<sup>98</sup>.

Di fatti, spetta a quest’ultimo provare il difetto – ossia la devianza rispetto agli *standard* di sicurezza che il pubblico può legittimamente attendersi (non valendo come prova la mera concretizzazione di un danno) –, il danno subito e la connessione causale fra difetto e danno.

Orbene, non solo si ha di fronte un settore fisiologicamente caratterizzato da tante significative asimmetrie conoscitive tra produttore e consumatore quanto è tecnologicamente avanzata l’IA in questione, ma il funzionamento delle IA forti risulta tendenzialmente opaco ed oscuro persino agli stessi ideatori dell’algoritmo.

---

<sup>93</sup> Gli *standard product safety* individuano un livello minimo di sicurezza necessaria ai fini dell’immissione in commercio. Si ha a che fare, perciò, con criteri statici che riguardano il prodotto in astratto senza considerarlo nella sua vita concreta. Non può che notarsi una intrinseca contraddizione con quello che viene chiamato «*consumer protection test*» ossia la verifica delle aspettative del consumatore, il quale deve farsi, necessariamente, ex post in concreto dal momento che devono essere testate le «reazioni» del prodotto in base all’uso che ne fa il consumatore. Sul punto si cfr. A. AMIDEI, *La proposta di Regolamento UE*, in *Tecn. Dir.*, 2022, I, p. 26, il quale sottolinea la non coincidenza fra la nozione di difettosità ai fini della disciplina *product liability* e la mancata aderenza del prodotto agli *standard* tecnici previsti dalla normativa di settore, in questo caso quelli che deriverebbero dall’*AI Act* (v. *infra*).

<sup>94</sup> A. AMIDEI, *Il (futuro) nuovo volto della responsabilità da prodotto difettoso*, in *Tecn. Dir.*, 2023, I, p. 41.

<sup>95</sup> Cfr. L. COPPINI, *Robotica ed intelligenza artificiale: questioni di responsabilità civile*, *Il Mulino*, 2018, p. 728; F. NADDEO, *Intelligenza artificiale: profili di responsabilità*, in *Comparazione e diritto civile*, 2020, III, p. 1150.

<sup>96</sup> Di questo la nuova proposta di direttiva europea PLD ne tiene conto, v. *infra*.

<sup>97</sup> Cfr. A. AMIDEI, *Intelligenza artificiale e diritto – Intelligenza artificiale e product liability: sviluppi del diritto dell’Unione Europea*, *Giur. It.*, 2019, VII, p. 1657.

<sup>98</sup> Sul tema dell’onere probatorio in materia di responsabilità da prodotto si v., fra gli altri, F. DI GIOVANNI, *Art. 8 – Prova*, in G. ALPA, U. CARNEVALI, F. DI GIOVANNI, G. GHIDINI, U. RUFFOLO e C.M. VERARDI, *La responsabilità per danno da prodotti difettosi*, cit., p. 119 ss.; S. PATTI, *Probabilità e verosimiglianza nella disciplina del danno da prodotto*, in *Riv. dir. civ.*, 1990, p. 705 ss.; AR. FUSARO, *Responsabilità del produttore: la difficile prova del difetto*, in *Nuova giur. civ. comm.*, 2017, p. 896 ss..

Riesce, perciò, difficile credere che un terzo, *quisque de populo*, riesca ad individuare l'esatta causa del danno. Dovrebbe essere in grado, non solo di decifrare il linguaggio binario tipico dell'informatica, ma anche di svelare l'intero procedimento algoritmico di quella che ad oggi viene chiamata «black box»<sup>99</sup> proprio per la sua inesplorabilità.

Infine, non sembra adeguata l'applicabilità delle cause di esclusione della responsabilità legate alla messa in circolazione del prodotto o allo stato delle conoscenze scientifiche e tecniche (il cd. rischio da sviluppo). «Al produttore è consentito liberarsi dalla responsabilità provando, che il prodotto non era stato da lui messo in circolazione o che al momento della immissione in circolazione il prodotto non era difettoso o, trattandosi di difetti di progettazione, che lo stato delle conoscenze scientifiche e tecniche non consentiva di considerare il prodotto come difettoso»<sup>100</sup>.

Ma è la stessa natura dell'IA capace di autoapprendere e di evolversi nel tempo che mostra il limite di applicare esimenti statiche. Il produttore/autore dell'algoritmo potrebbe andare facilmente esente da responsabilità invocando il rischio da sviluppo. Ma qui l'IA *self-learnig* si nutre giorno per giorno di nuovi dati, aumentando il suo «bagaglio culturale» in base al quale decidere. Segue il progresso tecnologico, il suo sviluppo non si arresta al momento della messa in circolazione<sup>101</sup>.

### **3.2.2. Segue...nella nuova proposta di direttiva europea sulla responsabilità per prodotti difettosi (PLD).**

L'inadeguatezza della attuale normativa sulla responsabilità da prodotto difettoso non è stata avvertita unicamente nell'ordinamento italiano ma è oggetto di attenzione da parte del Legislatore europeo ormai da molto tempo<sup>102</sup> tant'è

---

<sup>99</sup> Si fa riferimento ai cd. modelli opachi. Molte delle più note introduzioni all'IA classificano come trasparenti i modelli basati sugli alberi di decisione o quelli basati sulla programmazione logica, mentre considerano i modelli della «nuova IA», come le reti neurali ad apprendimento profondo, la cui architettura è interamente sub-simbolica, l'esempio prototipico di modello opaco.

Altre volte, tuttavia, per «opacità di un sistema» ci si riferisce al fatto che le caratteristiche e la tipologia dell'algoritmo di apprendimento impediscono agli utenti di comprendere e verificare l'affidabilità dei *pattern* appresi. In altri casi ancora, i modelli sono considerati opachi perché prevedono solo il verificarsi di certi fenomeni ma non ne spiegano né le cause né le leggi che li governano.

In ogni caso, gli studiosi concordano sul fatto che l'opacità dipenda da vari fattori: dagli obiettivi legati all'utilizzo di un certo sistema di machine learning, dalla conoscenza che di questo hanno i suoi utilizzatori, e non da ultimo il contesto concreto in cui il sistema è impiegato. Detto altrimenti, lo stesso sistema può risultare trasparente se usato in un dato contesto, per un dato scopo e da una data classe di utenti, ma opaco se impiegato in un altro contesto, da utenti con competenze e scopi differenti.

<sup>100</sup> Così scrive C.M. BIANCA, *Diritto civile*, V, cit., p. 717. Il rischio da sviluppo è stato oggetto di ampio dibattito in seno alla Commissione Europea, costituendo motivo di scontro tra coloro che consideravano tale prova liberatoria come elemento necessario al fine di favorire la competitività tra le imprese e l'immissione sul mercato di prodotti nuovi, e quanti invece ritenevano di dover escluderla per offrire una maggiore protezione al consumatore danneggiato. Si rinvia per un approfondimento a D. CARUSO e R. PARDOLESI, *Per una storia della Direttiva 1985/374/CEE*, in *Danno resp.*, 2012, p. 9-17.

<sup>101</sup> Sia consentito rinviare alle considerazioni già svolte M. M. MOLLICONE, *Il rischio dell'intelligenza artificiale applicata. Modelli di allocazione a confronto*, in *AJI*, 2023, 18, p. 2108-2133.

<sup>102</sup> Cfr., in particolare, il *Commission staff working document su Liability for emerging digital technologies* del 7 maggio 2018, con il quale la Commissione aveva valutato la Direttiva come nel suo complesso ancora idonea

che è stata avanzata la proposta di Direttiva del Parlamento europeo e del Consiglio del 28 settembre 2022, COM(2022) 495 finale, sulla responsabilità per prodotti difettosi (PLD), alla cui paventata approvazione seguirebbe l'abrogazione completa della attualmente vigente Direttiva 85/374/CEE – e quindi dell'attuale Codice del Consumo italiano che la recepisce – il tutto in un'ottica di equo riparto del rischio e neutralità tecnologica<sup>103</sup>.

L'impianto strutturale della proposta direttiva, abbandonata per il momento l'idea di predisporre un regime di responsabilità ad hoc per i danni cagionati da prodotti ad elevata tecnologia e *AI-based*, non prevede una rivoluzione del regime vigente quanto piuttosto una revisione chirurgica<sup>104</sup> – il cd. *fine tuning* – delle attuali norme al fine di adeguarle agli scenari odierni<sup>105</sup>.

Per dovere di completezza è necessario dire che la modifica normativa è stata ritenuta necessaria, non solo per rimodellare il regime di *product liability* per adattarlo ai nuovi prodotti animati da IA, ma anche per rispondere a più generali esigenze della produzione contemporanea, quale ad esempio l'incremento della complessità tecnologica e scientifica all'interno delle catene del valore mondiali.

A rigore l'iniziativa europea si inserisce, infatti, in un più ampio quadro, disegnato per la modernizzazione dei consumi, all'interno del quale sicuramente sono protagonisti, la direttiva sui contratti di fornitura a contenuto digitale<sup>106</sup>, il

---

a contribuire a «una ragionevole equilibrio tra la tutela dei danneggiati e la garanzia di una concorrenza leale sul mercato unico», evidenziando tuttavia la necessità di un *fine-tuning* di concetti quali quello di «prodotto», «produttore», «difetto» e «danno», oltre che in relazione all'onere della prova, pur lasciando inalterato il generale impianto di *strict liability*.

Sulle avanzate prospettive di revisione della PLD si rinvia a A. AMIDEI, *Le responsabilità da intelligenza artificiale tra product liability e sicurezza del prodotto*, in U. RUFFOLO (a cura di), *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2021, p. 149 ss. Cfr. altresì, *inter alia*, G. ALPA, *Quale modello normativo europeo per l'intelligenza artificiale?*, in V. CUOCCI, F.P. LOPS e C. MOTTI (a cura di), *La responsabilità civile nell'era digitale*, Bari, 2022, p. 3 ss.; B.A. KOCH, *Product liability 2.0 Mere update or new version?*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things*, Baden-Baden, 2019, p. 99 ss.; G. DE ROSA, *Quali regole per i sistemi automatizzati "intelligenti"?*, in *Riv. dir. civ.*, 2021, p. 842 ss.; V. DI GREGORIO, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in *Danno resp.*, 2022, p. 51 ss.; J. KNETSCH, *Are existing tort theories ready for AR*, in L.A. DI MATTEO, C. PONCIBÒ e M. CANNARSA (a cura di), *The Cambridge Handbook of Artificial Intelligence. Global perspective on law and ethics*, Cambridge, 2022, p. 108 ss.; M.L. MONTAGNANI, *Liability and emerging digital technologies: an EU perspective*, in V. FALCE (a cura di), *Fairness e innovazione nel mercato unico digitale*, Torino, 2020, p. 85 ss.; L. DI DONNA, *Intelligenza artificiale e rimedi risarcitori*, Milano, 2022, p. 21 ss.; A. BERTOLINI, *Robot sas product: the case for a realistic analysis of robotic applications and liability rules*, in *Law Innovation and Technology*, 2013, p. 214 ss.

<sup>103</sup> A. AMIDEI, *Il (futuro)*, cit., p. 29.

<sup>104</sup> Sostenevano un tale approccio, U. RUFFOLO, *Per i fondamenti di diritto della robotica self-learning: dalla machinery produttiva all'auto driverless: verso una «responsabilità da algoritmo?»* in U. RUFFOLO (a cura di), *Intelligenza artificiale e responsabilità*, Milano, 2017, p.1 ss. Sul tema v. altresì, da ultimo, L. DI DONNA, *Intelligenza artificiale e rimedi risarcitori*, cit., p. 69 ss.

<sup>105</sup> Cfr. A. AMIDEI, *Il (futuro)*, cit., p. 29.

<sup>106</sup> Direttiva (UE) 2019/770 del Parlamento europeo e del Consiglio del 20 maggio 2019 relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali. Sul tema v. in particolare A. DE FRANCESCO, *La vendita di beni con elementi digitali*, Napoli, 2019, p. 70 ss.; C. CAMARDI, *Prime osservazioni sulla Direttiva (UE) 2019/770 sui contratti per la fornitura di contenuti e servizi digitali. Operazioni di consumo e circolazione di dati personali*, in *Giust. civ.*, 2019, p. 499 ss.; G. RESTA, *I dati personali oggetto del contratto. Riflessioni sul coordinamento tra la Direttiva (UE) 2019/770 e il Regolamento (UE) 2016/679*, in A.M. GAMBINO e A. STAZI (a cura di), *La circolazione dei dati. Titolarità, strumenti negoziali, diritti e tutele*, Pisa, 2020, p. 55 ss.

*Digital Services Act*<sup>107</sup>, il *draft* dell'*AI Act*<sup>108</sup>, il *draft* per un *Product Safety Regulation*<sup>109</sup> e il *draft* della direttiva *AI Liability*<sup>110</sup>.

Per quanto di nostro interesse, si evidenzieranno di séguito le modifiche principali.

Innanzitutto, può rilevarsi che il *draft* di direttiva *product liability* permetterebbe l'estensibilità della nozione di «prodotto» delineata dalla Direttiva 85/374/CEE, e quindi l'applicabilità del regime di responsabilità da prodotto, «alla realtà informatica e ai beni (interamente o parzialmente) dematerializzati che la contraddistinguono»<sup>111</sup>, compresi i sistemi di IA.

Benchè l'art. 4 del *draft* lasci sostanzialmente impregiudicate rispetto alla vigente direttiva sia la definizione di «fabbricante» sia quella di «prodotto», lo stesso però specifica che «[i]l termine “prodotto” comprende l'elettricità, i file di fabbricazione digitale e il software».

La chiarificazione è più che rilevante.

Da una parte, il riconoscimento dei *digital manufacturing file*<sup>112</sup> come prodotto è il riflesso della constatazione della sempre maggiore facilità che operatori non professionali producano beni di consumo<sup>113</sup> e che sia, dunque, necessaria anche in questo contesto una tutela maggiore del consumatore.

Dall'altra parte, il richiamo al *software* come prodotto offre una soluzione normativa all'estenuante dibattito sulla natura giuridica del *software*.

La possibilità di applicarsi il regime della *product liability*<sup>114</sup> è stata, infatti, da sempre messa in discussione in ragione della natura ibrida e non tangibile del

---

<sup>107</sup> Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali).

<sup>108</sup> Commissione UE, Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'Unione, 21 aprile 2021, COM(2021)206 final.

<sup>109</sup> Commissione UE, Proposta di Regolamento del Parlamento europeo e del Consiglio relativo alla sicurezza generale dei prodotti, 30 giugno 2021, COM(2021) 346 final.

<sup>110</sup> Commissione UE, Proposta di Direttiva del Parlamento europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale, 28 settembre 2022, COM(2022) 496 final.

<sup>111</sup> Cfr. A. AMIDEI, *Il futuro*, cit., p. 32.

<sup>112</sup> Si tratta di *file* che contengono le informazioni funzionali necessarie per produrre un oggetto tangibile consentendo il controllo automatizzato di macchinari o strumenti, come trapani, torni, frese e stampanti 3D (considerando 14).

<sup>113</sup> Caso eclatante è quello della stampa 3D. Qui è il consumatore assume anche la veste di produttore (si tratta di caso esemplificativo della nozione di «prosumer»). Sul punto si v. G. HOWELLS, C. TWIGG-FLESNER e C. WILLETT, *Protecting the values of consumer law in the digital economy the case of 3D-printing*, in A. DE FRANCESCHI e R. SCHULZE (a cura di), *Digital revolution – New challenges for Law*, Monaco, 2019, p. 224 ss.

<sup>114</sup> In senso favorevole all'estensione del regime di responsabilità da prodotto difettoso al *software* si v., senza pretesa di esaustività C. PERLINGIERI, *Diritto privato delle nuove tecnologie: contenuti e competenze*, in *Tecn. dir.*, 2021, p. 83 ss.; A. ALBANESE, *La responsabilità civile per i danni da circolazione di veicoli ad elevata automazione*, in *Contr. impr.*, 2020, p. 995 ss.; R. MONTINARO, *Responsabilità da prodotto difettoso e tecnologie digitali tra soft law e hard law*, in *Pers. merc.*, 2020, p. 370 ss.; A. ZACCARIA, *La responsabilità del “produttore” di software*, in *Contr. impr.*, 1993, p. 303 ss.; G. WAGNER, *Robot liability*, in S. LOHSSE, R. SCHULZE e D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things*, cit., p. 41.

In senso contrario, v., in particolare, K. ALHEIT, *The applicability Of the EU Product Liability Directive to software*, in *Comp. Int. L. Jour. South Africa*, 2001, p. 200 ss.

*software*, sulla scorta della quale veniva da molti negata la sua qualificabilità come «bene mobile» e, quindi, come «prodotto» ai sensi della Direttiva 85/374/CEE, preferendosi una lettura del *software* come «servizio», quindi esulante dall'ambito di applicazione della direttiva stessa<sup>115</sup>.

Il *draft PLD* neutralizzerebbe, perciò, il terreno di disputa affermando, sin dal considerando 12, che nell'era digitale i prodotti possono essere tangibili o intangibili. Il *software*, come sistemi operativi, *firmware*, programmi per *computer*, applicazioni o sistemi di IA, è sempre più diffuso sul mercato e svolge un ruolo sempre più importante per la sicurezza dei prodotti. Il *software* può essere immesso sul mercato come prodotto a sé stante e può essere successivamente integrato in altri prodotti come componente, ed è in grado di causare danni attraverso la sua esecuzione.

Nell'interesse della certezza del diritto, viene, quindi, chiarito che il *software* è un prodotto ai fini dell'applicazione della responsabilità senza colpa, a prescindere dalle modalità di fornitura o di utilizzo, e quindi a prescindere dal fatto che il *software* sia memorizzato su un dispositivo o che vi si acceda tramite tecnologie *cloud*. Eccezione fa il codice sorgente del *software* che, trattandosi di pura informazione, non può essere considerato prodotto.

Alla stregua del *software*, anche il sistema di IA è espressamente considerato prodotto dal *draft*.

In secondo luogo, la proposta chiarisce che possono essere ritenuti responsabili non solo i produttori di hardware, ma anche i fornitori di *software* e i fornitori di servizi digitali che influiscono sul funzionamento del prodotto (ad esempio un servizio di navigazione in un veicolo autonomo o più in generale al mondo dell'*Internet of things*).

L'ampliamento del novero dei soggetti responsabili è, certamente, realizzato attraverso la nuova di definizione di componente di cui all'art. 4 PDL in base al quale per essa si intende «qualsiasi elemento, materiale o immateriale, o qualsiasi servizio correlato, integrato o interconnesso con un prodotto dal fabbricante di tale prodotto o sotto il controllo di quest'ultimo».

È sempre più frequente, infatti, che i servizi digitali siano integrati o interconnessi con un prodotto in modo tale che l'assenza del servizio impedirebbe al prodotto di svolgere una delle sue funzioni (ad esempio la fornitura continua di dati sul traffico in un sistema di navigazione. Sebbene la presente direttiva non si applichi ai servizi in quanto tali, è necessario estendere la responsabilità senza colpa a tali servizi digitali, poiché essi determinano la sicurezza del prodotto tanto quanto i componenti fisici o digitali. Tali servizi connessi dovrebbero essere considerati come componenti del prodotto a cui sono interconnessi, quando sono sotto il controllo del produttore di tale prodotto, nel senso che sono forniti dal produttore stesso o che il produttore li raccomanda o influenza in altro modo la loro fornitura da parte di terzi.

In terzo luogo, la proposta garantisce che i produttori possano essere ritenuti responsabili per le modifiche apportate ai prodotti che hanno già immesso sul

---

<sup>115</sup> Cfr. A. AMIDEI, *Il (futuro)*, cit., p. 33.

mercato, anche quando tali modifiche sono innescate da aggiornamenti del *software* o dall'apprendimento automatico.

In altre parole rileverebbe il cd. difetto da autoapprendimento o di aggiornamento.

A tale risultato, però, la proposta, non arriva modificando l'attuale definizione di «difettosità». Piuttosto si propone di integrare i parametri alla stregua dei quali effettuare il cd. *consumer expectation test*.

Ad esempio, considerando la crescente diffusione di prodotti interconnessi, la valutazione della sicurezza di un prodotto dovrebbe tenere conto anche degli effetti di altri prodotti sul prodotto in questione.

Si dovrebbe, cosò, tenere conto dei «requisiti di sicurezza» di un prodotto<sup>116</sup>, di altri prodotti che ci si può ragionevolmente attendere siano utilizzati insieme al prodotto<sup>117</sup>, della sua capacità di apprendere dopo l'impiego, per riflettere la legittima aspettativa che il *software* di un prodotto e gli algoritmi sottostanti siano progettati in modo tale da prevenire comportamenti pericolosi del prodotto<sup>118</sup>, e del fatto che nell'era digitale molti prodotti rimangono sotto il controllo del produttore oltre il momento in cui vengono immessi sul mercato.

Tali due ultimi aspetti sono estremamente rilevanti.

Da un lato, considerare la capacità del prodotto di continuare ad imparare dopo la sua diffusione potrebbe superare la dimensione statica che aveva connotato la concezione del prodotto posta alla base della direttiva dell'85. In tal modo, «si prenderebbe, invece, atto della possibile sussistenza di una duplice dimensione dinamica dei prodotti a elevato contenuto tecnologico, ossia del fatto, da un lato, che essi sono suscettibili di modifiche del tempo, *in primis* perché possono esservi integrati nuovi elementi quali *software*, applicazioni, aggiornamenti e simili; e, dall'altro, che per taluni prodotti (quelli *AI-based*) particolarmente incidente può risultare il mutamento imposto dal *machine learning*, con la possibilità che un difetto del prodotto si generi e si palesi nel corso e per effetto del processo di autoapprendimento e autonoma evoluzione del bene «smart»<sup>119</sup>.

Dall'altro lato, prendere atto che, soprattutto in relazione ai beni altamente tecnologici, sussista per il produttore la possibilità di un controllo *post-market*, non solo giustifica forme di responsabilizzazione del fabbricante anche con riguardo a difettosità «sopravvenute» al momento di immissione in commercio, ma potrebbe consentire di riconoscere in capo al produttore, anche in via interpretativa, un implicito di «dovere di aggiornamento» tale per cui «il fabbricante che abbia mantenuto un collegamento con il prodotto idoneo a

---

<sup>116</sup> A. AMIDEI, *Il (futuro)*, cit., p. 41 afferma che «evidente è il riferimento ai requisiti di natura tecnica facenti parte del complesso di norme in materia di *product safety*, quale combinazione a seconda della natura del prodotto delle prescrizioni di cui alla Direttiva 2001/95/CE e delle varie norme europee di armonizzazione verticale *sector specific*, cui possono altresì sommarsi gli standard tecnici armonizzati adottati da specifici enti riconosciuti ai sensi del diritto unionale». Ad ogni modo il *draft* riporta che un prodotto può essere ritenuto difettoso anche a causa della sua vulnerabilità in termini di cybersecurity.

<sup>117</sup> Si pensi ai prodotti destinati a divenire parte di un *Internet of things*.

<sup>118</sup> Tale elemento è esplicitamente riferito ai prodotti *AI powered* ossia dotati di capacità di autoapprendimento e evoluzione autonoma.

<sup>119</sup> A. AMIDEI, *Il (futuro)*, cit., p. 42.

consentirne il controllo e che tuttavia non vi abbia apportato modifiche, *upgrade* o aggiornamenti necessari per assicurare la sicurezza, risponderebbe nei confronti del soggetto che dovesse essere lesa»<sup>120</sup> (considerando 37 e 38).

La questione, ovviamente, si riallaccia alla questione della esimente del rischio da sviluppo in riferimento a prodotti altamente tecnologici o *AI based* relativamente alla quale il maggior dubbio attiene il confinamento del livello di conoscenze che deve essere preso a riferimento ai fini dell'applicazione dell'esimente.

La proposta *PLD* preciserebbe al considerando 39 – per poi trasportarlo all'art. 10 – che lo stato delle conoscenze scientifiche e tecniche, deve essere determinato con riferimento al livello più avanzato di conoscenze oggettive accessibili, e non alle conoscenze effettive del produttore in questione, al momento dell'immissione del prodotto sul mercato, della messa in servizio o nel periodo in cui il prodotto era sotto il controllo del produttore.

Indirettamente, in tal momento, «al produttore che mantenga un certo livello di controllo sul prodotto anche dopo la sua messa in commercio sarebbe così imposto un dovere di monitoraggio non soltanto del funzionamento del prodotto stesso, ma altresì dello Stato delle conoscenze tecniche scientifiche rilevanti al fine di garantire la sicurezza, nonché di conseguente adozione di ogni misura che dovesse rivelarsi necessaria alla luce del progredire di tali conoscenze»<sup>121</sup>.

In quarto luogo, il *draft PLD* si occuperebbe di alleggerire l'onere della prova nei casi di particolare complessità tecnica o scientifica, tra cui quelli che coinvolgono sistemi di intelligenza artificiale, dal momento che le persone danneggiate si trovano spesso in una posizione di notevole svantaggio rispetto ai produttori in termini di accesso e comprensione delle informazioni sulle modalità di produzione e sul funzionamento di un prodotto (considerando 30).

Così, all'art. 9 vengono contemplati una serie di meccanismi presuntivi che permettono al giudice di basare l'esistenza della difettosità del bene o del nesso di causalità fra difetto e danno sulla presenza di un altro fatto che è stato provato<sup>122</sup>.

---

<sup>120</sup> A. AMIDEI, *o.n.c.* p. 44.

<sup>121</sup> A. AMIDEI, *o.n.c.*, p. 47.

<sup>122</sup> Art. 9 della proposta *PLD*, rubricato onere della prova, dispone: « 1. Gli Stati membri garantiscono che il ricorrente sia tenuto a provare la difettosità del prodotto, il danno subito e il nesso di causalità tra la difettosità e il danno.

2. La difettosità del prodotto si presume quando si verifica una delle seguenti condizioni:

(a) il convenuto è venuto meno all'obbligo di divulgare le prove pertinenti a sua disposizione ai sensi dell'articolo 8, paragrafo 1;

(b) il richiedente stabilisce che il prodotto non è conforme ai requisiti di sicurezza obbligatori stabiliti dal diritto dell'Unione o dal diritto nazionale che mirano a proteggere dal rischio del danno che si è verificato; oppure

(c) il richiedente dimostri che il danno è stato causato da un evidente malfunzionamento del prodotto durante il normale utilizzo o in circostanze ordinarie.

3. Il nesso causale tra la difettosità del prodotto e il danno è presunto, qualora sia stato accertato che il prodotto è difettoso e il danno causato è di tipo tipicamente coerente con il difetto in questione.

4. Qualora un tribunale nazionale ritenga che il ricorrente abbia eccessive difficoltà, a causa della complessità tecnica o scientifica, a dimostrare la difettosità del prodotto o il nesso causale tra la sua

Infine, ma non per importanza, deve sottolinearsi che la proposta di direttiva *PLD* si esprime anche con riguardo ai rapporti tra il regime di responsabilità da prodotto e altre norme di responsabilità previste dagli ordinamenti domestici.

Diversamente dall'art. 13 della direttiva dell'85, il *draft* all'art. 2 esplicita che la direttiva non pregiudica nessun diritto che una persona danneggiata possa avere in base alle norme nazionali sulla responsabilità contrattuale o sulla responsabilità extracontrattuale per motivi diversi dalla difettosità di un prodotto, comprese le norme nazionali che attuano il diritto dell'Unione, come la direttiva sulla responsabilità civile dell'IA (v. capitolo III).

In tal modo, «sarebbe consentito, finalmente, di superare le incertezze emerse, anche a fronte di orientamenti della Corte di giustizia UE<sup>123</sup> non sempre di univoca interpretazione con riferimento alla possibilità di considerare le norme unionali di responsabilità da prodotto come introduttive di un livello di tutela minima, e non invece una disciplina uniforme; e dunque come applicabili concorrentemente ad altre norme di responsabilità fondate su diversi presupposti, tra cui in particolare, con specifico riferimento all'ordinamento italiano, il regime di responsabilità da attività pericolosa ex art 2050 c.c.<sup>124</sup>».

---

difettosità e il danno, o entrambi, la difettosità del prodotto o il nesso causale tra la sua difettosità e il danno, o entrambi, si presume quando l'attore ha dimostrato, sulla base di prove sufficientemente rilevanti, che:

(a) il prodotto ha contribuito al danno; e

(b) è probabile che il prodotto sia difettoso o che la sua difettosità sia una probabile causa del danno, o entrambe le cose.

Il convenuto ha il diritto di contestare l'esistenza di difficoltà eccessive o la probabilità di cui al primo comma.

5. Il convenuto ha il diritto di confutare una qualsiasi delle presunzioni di cui ai paragrafi 2, 3 e 4.»

<sup>123</sup> Si fa riferimento alle note Corte giust., 25 aprile 2002, c. 154/00, Commissione c. Repubblica ellenica; Corte giust., 25 aprile 2002, c. 52/00, Commissione c. Repubblica francese; Corte giust., 25 aprile 2002, c. 183/00, Medicina Asturiana, tutte in *Resp. civ. prev.*, 2002, p. 997 ss., con nota di S. BASTIANON, *Responsabilità del produttore per prodotti difettosi: quale tutela per il consumatore?*.

<sup>124</sup> A. AMIDEI, *Il (futuro)*, cit., p. 52-53. L'a., circa il dibattito in questione, ricorda unicamente come questo si sia intrecciato con quello sulla natura oggettiva o meno dei due regimi dei due regimi di responsabilità in questione.

Coloro che escludevano il cumulo delle due responsabilità lo facevano ritenendo la *product liability*, prevista dalla Direttiva 85/374/CEE, una normativa di disciplina uniforme, e non di tutela minima; e affermando che le norme di responsabilità da prodotto potessero ritenersi cumulabili esclusivamente con regimi di responsabilità extracontrattuali fondati su presupposti differenti, individuando tale diversità di presupposti nella necessaria natura colposa, e non oggettiva, del regime concorrente.

Per contro, la giurisprudenza ha generalmente ammesso tale cumulo.

Se adottata, la proposta nuova formulazione contribuirebbe a chiarire l'irrilevanza, ai fini della questione, della natura oggettiva o meno dei regimi di responsabilità in parola: ad assumere rilievo ai fini della determinazione della applicabilità, a una determinata fattispecie, di norme suscettibili di cumularsi a quelle di responsabilità da prodotto sarebbe, infatti, la sola circostanza che le stesse abbiano presupposti applicativi diversi dalla «difettosità» del prodotto.

## CAPITOLO III

### Il problema della responsabilità civile

SOMMARIO: 1. Introduzione – 2. I paradigmi di responsabilità applicabili in base all'ordinamento nazionale – 2.1. La responsabilità per colpa. – 2.2. La responsabilità da attività pericolosa ex art. 2050 c.c. – 2.3. La responsabilità da cose in custodia ex art. 2051 c.c. – 3. Il paradigma previsto dalla proposta di direttiva *AI liability* (*AILD*). – 3.1. *AILD* e *PLD*: criticità, osservazioni e confronti.

#### 1. Introduzione.

Come apprezzato nel capitolo II, assumendo che anche i più autonomi sistemi di IA siano oggetti di diritto, in caso di danni dagli stessi cagionati sembra possibile applicare la disciplina della responsabilità da prodotto difettoso (v. *supra*) e, parallelamente, le altre responsabilità di natura contrattuale ed extracontrattuale che – come specificatamente preciserebbe il nuovo *draft* di direttiva *product liability* – non abbiano la loro genesi nella difettosità del prodotto.

Proprio su tali «altre responsabilità» è necessario focalizzare ora l'attenzione sottolineando che, ai fini del nostro interesse, si escluderà dall'analisi la responsabilità cd. contrattuale mentre si affronteranno nel dettaglio quei paradigmi di responsabilità extracontrattuale<sup>1</sup> che, in assenza di una

---

<sup>1</sup> La responsabilità extracontrattuale è la responsabilità che ha fonte nell'illecito civile cioè nel fatto lesivo di interessi giuridicamente tutelati nella vita di relazione. Così la definisce C.M. BIANCA, *Diritto civile*, V, *La responsabilità*, Milano, 2021, p. 517 ss. Sul punto si v. anche L. CORSARO, *Responsabilità civile I, Diritto civile*, in *Enc. giur. Treccani*, XXVI, 1991 «La responsabilità extracontrattuale è precisamente quindi una situazione giuridica soggettiva, ossia la condizione di chi è tenuto al risarcimento del danno ed assoggettato agli altri effetti sanzionatori dell'illecito. La responsabilità extracontrattuale è però intesa anche come istituto, ossia come il complesso unitario degli effetti che la legge pone a carico del responsabile a tutela degli interessi della vita di relazione. A questo significato si riconduce la definizione come della responsabilità civile come il “fenomeno” della reazione dell'ordinamento alla lesione degli interessi giuridicamente tutelati».

Sulla responsabilità civile in generale, fra molti, si v. trattazioni generali e sistematiche della materia: A. DE CUPIS, *Il danno. Teoria generale della responsabilità civile*, Milano, 1954; ID., *Dei fatti illeciti. Art. 2043-2059, IV, Delle obbligazioni* in *Comm. c.c. Scialoja e Branca*, 2<sup>ed.</sup>, Bologna-Roma, 1971; R. SCOGNAMIGLIO, *Responsabilità civile*, in *Noviss. Dig. it.*, XV, Torino, 1968; ID., *Illecito (diritto vigente)*, in VIII, 1962, p. 164; P. TRIMARCHI, *Illecito (diritto privato)*, in *Enc. dir.*, XX, Milano, 1970, p. 90; ID., *La responsabilità civile: atti illeciti, rischio, danno*, 3<sup>ed.</sup>, Milano, 2021; P. FORCHIELLI, *Responsabilità civile*, Padova, 1983; F.D. BUSNELLI, voce *Illecito civile*, in *Enc. giur. Treccani*, Roma, 1989, p. 33 ss.; L. CORSARO, *Responsabilità civile I) Diritto civile*, in *Enc. giur. Treccani*, XXVI, 1991; G. ALPA, M. BESSONE e V. ZENO ZENCOVICH, *Obbligazioni e contratti. VI. I fatti illeciti*, in *Tratt. Rescigno*, XIV, Torino, 1995; C. SALVI, *Responsabilità extracontrattuale (diritto vigente)*, in *Enc. dir.*, XXXIX, 1998, p. 1186, nonché ID., *La responsabilità civile*, in *Tratt. dir. priv. Indica e Zatti*, 3<sup>ed.</sup>, Milano, 2019; P.G. MONATERI, *Responsabilità civile nel diritto comparato*, in *Dig. disc. priv., Sez. civ.*, XVII, Torino, 1998; ID., *La responsabilità civile*, in *Le fonti delle obbligazioni*, III, in *Tratt. dir. civ. Sacco*, Torino, 1998; G. ALPA, *La responsabilità civile*, in *Tratt. dir. civ.*, IV, Milano, 1999; G. VISINTINI, *Trattato breve della responsabilità civile. Fatti illeciti. Inadempimento. Danno risarcibile*, XV, Padova, 2005; C. CASTRONOVO, *La nuova responsabilità civile*, 3<sup>ed.</sup>, Milano, 2006; A. PROCIDA MIRABELLI DI LAURO, *La responsabilità civile*, Torino, 2008; U. CARNEVALI (a cura di), *Commentario del Codice civile. Dei fatti illeciti, I, Art. 2043 - art. 96 c.p.c.*, Torino, 2011; M. BARCELLONA, *Trattato della responsabilità civile*, Torino, 2011; M. FRANZONI, *Dei fatti illeciti. Artt. 2043-2059*, in *Commentario c.c. Scialoja-Branca-Galgano*, 2<sup>a</sup> ed., Bologna, 2020; C.M. BIANCA, *Diritto civile*, V, cit., p. 517 ss.

Amplie raccolte e rassegne di dottrina e giurisprudenza: G. ALPA e M. BESSONE, *La responsabilità civile*, Milano, 2001; M. SELLA, *La responsabilità civile nei nuovi orientamenti giurisprudenziali*, Milano, 2007;

regolamentazione ad *hoc*, potrebbero, astrattamente e concretamente, richiamarsi per regolamentare le conseguenze pregiudizievoli derivanti dall'utilizzo di un sistema di IA<sup>2</sup> ferma la molteplicità e la equiordinazione dei criteri di imputazione<sup>3</sup>.

Di talché, si orienterà lo sguardo, innanzitutto, verso il panorama normativo nazionale approfondendo la responsabilità per colpa (art. 2043 c.c.), la responsabilità da attività pericolose (art. 2050 c.c.) e la responsabilità da cose in custodia (art. 2051 c.c.)<sup>4</sup>.

Ad ogni vaglio si indicheranno i pro e i contro della scelta di un determinato paradigma<sup>5</sup>.

In secondo luogo, si scandaglierà la recente proposta di direttiva europea sulla responsabilità da IA (*AI liability directive, AIDL*)<sup>6</sup> e sull'impianto di questa si tenteranno di offrire delle riflessioni.

---

F. DI MARZIO, *Codice della responsabilità civile commentato*, Milano, 2017; P. STANZIONE, *Trattato della responsabilità civile*, Padova, 2012.

Un'approfondita analisi storico-domatica degli istituti della responsabilità è svolta dal C. MAIORCA, *Responsabilità (teoria generale)*, in *Enc. dir.*, XXXX, 1988, p. 1004, e ID., *I fondamenti della responsabilità*, Milano, 1990 (recens. da C.M. BIANCA, in *Rass. dir. civ.*, 1991, I, p. 790).

<sup>2</sup> Non si farà riferimento qui ai casi di responsabilità per danni derivanti da trattamento automatizzato dei dati ritenendo che il tema meriti una trattazione autonoma.

<sup>3</sup> La ricostruzione che al carattere unitario di qualificazione del danno (l'ingiustizia) accosta la molteplicità dei criteri di imputazione è da ritenersi preferibile. Cfr. S. RODOTÀ, *Il problema della responsabilità civile*, Milano, 1967, *passim*; R. SCOGNAMIGLIO, *Responsabilità civile*, cit., *passim*.

*Contra*: l'orientamento tradizionale accoglie la concezione etica della responsabilità ripudiando l'idea che questa possa esistere senza colpa. Cfr. G. CIAN, *Antigiuridicità e colpevolezza. Saggio per una teoria dell'illecito civile*, Padova, 1966, pp. 62, 273, 351, il quale vede un atto di violazione della legge che non può prescindere dalla colpevolezza (giudizio di riprovevolezza soggettiva) in quanto il comando normativo può essere violato solo da chi lo abbia percepito in condizioni di libertà e capacità; A. DE CUPIS, *Dei fatti illeciti*, cit., p. 4, il quale afferma che la regola della colpa «è cardine e principio fondamentale della responsabilità civile. Di fronte a questa regola generale ne sussistono altre [...] ma [...] più limitate»; F.D. BUSNELLI, *voce Illecito civile*, cit., p. 6, laddove scrive che «la regola generale è imprescindibilmente legata alla regola della colpa».

Diversamente opinano coloro che abbracciano una concezione oggettiva della responsabilità, ritenendola strumento di riequilibrio economico del danno. Fra molti si v. G. VENEZIAN, *Studi delle obbligazioni*, in *Opere giuridiche di Giacomo Venezian*, I, Roma, 1919; R. SCOGNAMIGLIO, *Illecito*, cit. Gli aa. non ritengono che la responsabilità sia uno strumento repressivo che reprime un atto vietato, ma, piuttosto, una tecnica per riparare il danno. All'interno di tale orientamento emerge il filone di pensiero che fa ricorso all'analisi economica per guidare l'allocatione dei danni. Illuminante è stato G. CALABRESI con il suo *Costo degli incidenti e responsabilità civile, Analisi economico-giudiziaria* (tr.it), Milano, 1975.

L'orientamento eclettico, tuttavia, è quello che ha preso più piede. Esso ravvisa il fondamento della responsabilità civile in due principi: la colpa ed il rischio. Ad ognuno di questi corrisponderebbe un autonomo sistema di responsabilità. Rispettivamente, la responsabilità per colpa e la responsabilità per rischio d'impresa (secondo P. TRIMARCHI, *Rischio e responsabilità oggettiva*, Milano, 1961) o da esposizione al pericolo (secondo M. COMPORTI, *Esposizione al pericolo e responsabilità civile*, Napoli, 1965).

<sup>4</sup> Come noto, il principio generale della responsabilità extracontrattuale è sancito dall'art. 2043 c.c.: qualunque fatto doloso o colposo che cagiona ad altri un danno ingiusto, obbliga colui che ha commesso il fatto a risarcire il danno. Accanto alla figura generale dell'illecito si pongono, poi, figure speciali di illecito, caratterizzate dalla particolarità del fatto lesivo, in quanto compiuto nell'ambito di una determinata attività (es: esercizio di attività pericolose), o dalla particolarità del mezzo dannoso (es: animali, prodotti difettosi) o dalla particolarità del rapporto intercorrente tra il responsabile e l'autore del fatto (es: insegnante-allievo). In relazione alle speciali figure di illecito può essere varia l'incidenza della colpa, potendosi distinguere tra illeciti che prescindono dalla colpa (responsabilità oggettiva) ed illeciti in cui la colpa è presunta (responsabilità aggravata). Cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 518-519.

<sup>5</sup> Si consenta il rinvio a M.M. MOLlicONE, *Il rischio dell'intelligenza artificiale applicata. Modelli di allocatione a confronto*, in *AJI*, 2023, 18, p. 2108-2133.

<sup>6</sup> Proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496.

## 2. I paradigmi di responsabilità applicabili in base all'ordinamento nazionale.

«Il bello della responsabilità civile è che, pur non mutando spesso paradigma, se si fa eccezione di norme speciali settoriali e che peraltro emergono a valle di un processo evolutivo del sistema, riesce ad essere sempre moderna e a coprire le polluzioni innovative delle società. Essa di fronte a fenomeni sociali nuovi svolge per definizione il ruolo di cane da guardia del diritto, un *first responder* in emergenza se si vuole, in assenza cioè di una normativa ad *hoc*. È il luogo della sperimentazione regolatoria, il letto di Procuste su cui si temprano i ragionamenti sulla disciplina delle dinamiche della società e dei rapporti interindividuali. [...] Il riferimento a principi generali (*neminem laedere*) cristallizzati in norme di legge (es. 2043 c.c.) permette livelli di flessibilità che fanno del rimedio risarcitorio uno strumento polivalente e capace di coprire le aree più diverse delle relazioni sociali»<sup>7</sup>.

È proprio nella nota origine pretoria della responsabilità civile che si rinvengono le basi per considerare tale strumento sia un rimedio *ex post* teso alla compensazione dei danni, sia un rimedio *ex ante* con effetti regolatori di quelle ipotesi successive sussumibili sotto il principio emerso nel caso singolo, svolgendo, perciò, un ruolo di specchio del sentire delle società in cui il rimedio opera<sup>8</sup>.

In questa direzione spesso agisce l'azione suppletiva della giurisprudenza che «anticipa» – ed a volte riscrive – l'intervento del legislatore. «A queste dinamiche – regole generali vs regole settoriali; normazione legislativa e regolazione giurisprudenziale; interazione con istituti diversi dalla r[esponsabilità] c[ivile] che ne condizionano però la operatività - si sono nel tempo aggiunte dinamiche di normazione multilivello in cui la disciplina di derivazione sovranazionale ha anticipato aree di intervento della r[esponsabilità] c[ivile] o ha contribuito a svecchiare vecchi paradigmi mettendo il vino vecchio in otri nuovi»<sup>9</sup>.

---

<sup>7</sup> G. COMANDÈ, *Il "dividendo tecnologico" della responsabilità civile, prospettive sul contenzioso*, in *Danno Resp. civ.*, 2021, III, p. 269-282, afferma che il ruolo metaforico sopradetto viene svolto dalla responsabilità civile in fasi normalmente successive strettamente connesse al contenzioso. «La dialettica regole generali e regole specifiche (non necessariamente speciali) illustra una dinamica tutta interna alla evoluzione della r[esponsabilità] c[ivile] ma che spesso si è storicamente affiancata al matrimonio della r[esponsabilità] c[ivile] con altri istituti giuridici come, di volta in volta, varie formule assicurative fino alla sua sostituzione con meccanismi di sicurezza sociale. Esempio paradigmatico di ciò è la multiforme evoluzione del risarcimento del danno alla persona che, pur ruotando ancora oggi attorno alle regole di responsabilità civile, ha maturato soluzioni tecniche diverse graduando l'intervento delle regole di r[esponsabilità] c[ivile] (oscillando tra colpa e responsabilità oggettiva passando attraverso presunzioni di colpa, come già scolpito nel Titolo IX del libro IV del codice civile), sostituendole in parte (meccanismi indennitari in sede infortunistica, per esempio), collegandola ad un obbligo assicurativo (circolazione stradale), forzando l'internalizzazione dei costi a fronte di modulazioni degli obblighi risarcitori (ad esempio la r[esponsabilità] c[ivile] prodotti)».

Sulla funzione polivalente della responsabilità civile v. S. RODOTÀ, *Il problema della responsabilità civile*, Milano, 1967.

<sup>8</sup> G. COMANDÈ, *Il "dividendo tecnologico"*, cit., p. 269-282; ID., *Risarcimento del Danno alla Persona e Alternative Istituzionali*, Torino, 1999. Sulla funzione deterrente della responsabilità civile si v. anche G. CIAN, *Antigiuridicità e colpevolezza. Saggio per una teoria dell'illecito civile*, Padova, 1966, p. VI.

<sup>9</sup> G. COMANDÈ, *Il "dividendo tecnologico"*, cit., p. 270.

In brevi termini, la responsabilità civile, grazie alla sua flessibilità, probabilmente è lo strumento più immediato per dare una rapida risposta all'evoluzione della società tecnologica. Il cambiamento dei valori fondanti l'ordinamento trova, infatti, significativo riscontro nell'evoluzione della responsabilità civile<sup>10</sup>, la quale si propone sempre più in forme autonome dalla nozione di illecito civile<sup>11</sup>.

## 2.1. La responsabilità per colpa.

Tutti gli ordinamenti giuridici europei riconoscono la responsabilità per colpa<sup>12</sup> e per tutti la medesima rappresenta una clausola generale, applicabile di

---

<sup>10</sup> Sull'incidenza dell'evoluzione ordinamentale sugli istituti del diritto civile, tra i quali anche la responsabilità civile, cfr. P. PERLINGIERI, *Interpretazione ed evoluzione dell'ordinamento*, in *Riv. Dir. Priv.*, 2011, p. 159 ss.; N. LIPARI, *Le categorie del diritto civile*, Milano, 2013, p. 195 ss.

<sup>11</sup> C.M. BIANCA, *Diritto civile*, V, cit., p. 518-519, delinea «Il significato dell'illecito civile quale violazione di una norma giuridica sanzionata dall'obbligo del risarcimento del danno» il quale troverebbe «conferma puntuale nella previsione del codice sulle fonti dell'obbligazione, dove il fatto illecito è contrapposto alle fonti legali, ossia gli atti o fatti idonei a produrre effetti obbligatori «in conformità dell'ordinamento giuridico» (art. 1173). Dunque, chi commette un illecito non crea un semplice presupposto di determinati effetti obbligatori ma pone in essere un fatto in contrasto con l'ordinamento giuridico, ossia in violazione di una norma. La necessità giuridica di non violare la norma mediante comportamenti lesivi degli interessi altrui si esprime in termini del dovere. L'illecito civile può allora essere definito come violazione dell'obbligo di salvaguardia degli interessi altrui tutelati (*alterum non laedere*).

<sup>12</sup> Per il diritto francese: R. SAVATIER, *Traité de la responsabilité civile en droit français*, II, *Conséquences et aspects divers de la responsabilité*, Parigi, 1951; H. e L. MAZEAUD e A. TUNC, *Traité théorique et pratique de la responsabilité civile délictuelle et contractuelle*, 5 e 6<sup>ed.</sup>, Parigi, 1958-1965; A. TUNC, *La responsabilité civile*, 2<sup>ed.</sup>, Parigi, 1989; B. STARCK, H. ROLAND e L. BOYER, *Obligations. Responsabilité délictuelle*, 5<sup>ed.</sup>, Paris, 1996; G. VINEY, *Introduction à la responsabilité*, in *Traité Ghestin*, Paris, 1995; G. VINEY e P. JOURDAIN, *Les conditions de la responsabilité: dommage, fait générateur, régimes spéciaux, causalité*, Parigi, 2013; IDD. e S. CARVAL, *Les effets de la responsabilité* 4<sup>ed.</sup>, Parigi, 2017; P. LE TOURNEAU E L. CADIEU, *Droit de la responsabilité et des contrats*, 4<sup>ed.</sup>, Parigi, 2002.

Per la letteratura di lingua inglese: B.M.E. MCMAHON E W. BINCHY, *Irish Law of Torts*, 4<sup>ed.</sup>, Dublino, 1990; K. HORSEY A. E. RACKLEY, *Tort Law*, 7<sup>ed.</sup>, Oxford, 2021; B.S. MARKESINIS e S. DEAKIN, *Tort Law*, 7<sup>ed.</sup>, a cura di S. Deakin, A. Johnston e B.S. Markesinis, Oxford, 2012; J.G. FLEMING, *The Law of Torts*, 6<sup>ed.</sup>, Londra, 1983, e ID., *An Introduction to the Law of Torts*, Oxford, 1985; T. WEIR, *Tort Law*, Oxford, 2000; M. LUNNEY E K. OLIPHANT, *Tort Law*, Oxford, 2000; E. QUILL, *Torts in Ireland*, Dublin, 1999.

Per la letteratura di lingua tedesca: K. SCHÄFER, §§ 823-835, in J. VON STAUDINGER, *Kommentar zum Bürgerlichen Gesetzbuch*, II, *Recht der Schuldverhältnisse*, Berlino, 1978-1986; J. HAGER, §§ 823 E-I, 824, 825, *ivi*; H.J. MERTENS, §§ 823-838, 840-85, in *Münchener Kommentar zum BGB*, 3<sup>ed.</sup>, 1997; G. WAGNER, § 823 *Schadensersatzpflicht*, *ivi*, 5<sup>ed.</sup>, 2009; K. LARENZ, *Lehrbuch des Schuldrechts*, I, Berlino, 1960; H. KÖTZ, *Deliktsrecht*, Francoforte sul Meno, 1979.

Per la letteratura di lingua spagnola: J.L. DE LOS MOZOS e C.A. SOTO COAGUILA (a cura di), *Responsabilidad civil. Derecho de daños. Teoría de la responsabilidad civil*, Lima, 2006; L.F. REGLERO CAMPOS e J.M. BUSTO LAGO (a cura di), *Tratado de la responsabilidad civil*, 5<sup>ed.</sup>, Elcano (Navarra), 2014; F.J. GARCIA ROCA, *Derecho de daños, Textos y materiales*, 3<sup>ed.</sup>, Valencia, 2000; L. DIEZ-PICAZO e J. PONCE DE LEON, *Derecho de daños*, Madrid, 1999; M. YZQUIERDO TOLSADA, *Sistema de responsabilidad civil, contractual y extracontractual*, Madrid, 2001.

Copiosa la dottrina argentina: A.A. ALTERINI, *Estudios de derecho civil*, Buenos Aires, 2007; A.J. BUERES e A. KEMELMAJER DE CARLUCCI (a cura di), *Responsabilidad por daños en el tercer milenio, Homenaje al Profesor Atilio Anibal Alterini*, Buenos Aires, 1997; A.A. ALTERINI e R.M. LOPEZ CABANA, *Temas de responsabilidad contractual y extracontractual*, Buenos Aires, 1995; A.A. ALTERINI, O.J. AMEAL, e R.M. LOPEZ CABANA, *Derecho de obligaciones civiles y comerciales*, 4<sup>ed.</sup>, Buenos Aires, 1993; A.A. Alterini e R.M. Lopez Cabana sono anche i direttori di una *Enciclopedia de la responsabilidad civil*, iniziata a pubblicarsi a Buenos Aires nel 1997.

*default* in assenza di specifiche alternative<sup>13</sup>. Si tratta del paradigma generale della responsabilità extracontrattuale che a livello europeo è previsto dall'art. 1:101, comma 1 e comma 2 lett. a) dei Principi di diritto europeo della responsabilità civile (PETL), mentre in Italia è sancito dall'art. 2043 c.c. in base al quale «qualunque fatto doloso o colposo che cagiona ad altri un danno ingiusto obbliga colui che ha commesso il fatto a risarcire il danno».

La base per attribuire la responsabilità è una presunta condotta negligente della parte danneggiante. Più specificamente, gli elementi costitutivi della generale figura dell'illecito civile si riconoscono nel fatto illecito, nel collegamento soggettivo del fatto al soggetto agente a titolo di dolo o colpa, nel danno-evento (lesione di un interesse meritevole di tutela)<sup>14</sup>, nel danno- conseguenza (conseguenze pregiudizievoli di natura patrimoniale o non patrimoniale)<sup>15</sup> e nel doppio nesso di causalità, materiale (che collega il fatto al danno-evento) e giuridica (che collega il danno-evento al danno- conseguenza).

A questo punto, sembra opportuno analizzare i singoli elementi costitutivi, sottolineando sin da ora che lo si farà senza pretesa di esaustività in quanto, ai fini del nostro discorso, è utile, piuttosto, porre l'accento sugli aspetti più problematici in presenza di un danno da IA. Perciò, ci si soffermerà maggiormente sul profilo probatorio dell'elemento soggettivo e del nesso di causalità materiale.

Orbene, primo elemento costitutivo della responsabilità aquiliana è il fatto illecito il quale si presenta ben più ampio di quello considerato nel diritto penale per come declinato dal principio personalistico<sup>16</sup>. Sotto il profilo naturalistico, il fatto può realizzarsi attraverso un movimento (azione), o un'inerzia (omissione) la quale ultima, però, deve essere qualificata dalla concomitante violazione di un

---

In diritto portoghese va segnalato A. MENEZES CORDEIRO, *Tratado de direito civil português, Direito da obrigações*, III, Coimbra, 2010, p. 285 ss.; C. CHAVES de FARIAS, P. BRAGA NETTO e N. ROSENVALD, *Novo Tratado de Responsabilidade Civil*, 4<sup>ed.</sup>, São Paulo, 2019.

Per lo studio comparatistico della materia vanno ancora segnalati i due tomi sui *torts* (vol. XI) della *International Encyclopedia of Comparative Law*, curati dal A. TUNC, Tübingen London, 1983, nonché G. PONZANELLI, *La responsabilità civile, Profili di diritto comparato*, Bologna, 1992, e F. PARISI, *Liability for Negligence and Judicial Discretion*, 2<sup>ed.</sup>, Berkeley California, 1992. Per un raffronto dei regimi di responsabilità extracontrattuali vigenti negli ordinamenti inglese, francese e tedesco vedi C. VAN DAM, *European Tort Law*, 2<sup>ed.</sup>, Oxford, 2007.

<sup>13</sup> M. GEISTFELD, E. KARNER, B.A. KOCH e C. WENDERHOSTE, *Civil Liability for Artificial Intelligence and Software*, in *Tort and Insurance law*, vol. 37, Vienna, 2023, p. 44.

<sup>14</sup> Letteralmente la norma parla di danno ingiusto. Questo consiste nell'effetto del fatto illecito, ossia il pregiudizio arrecato sugli interessi giuridicamente rilevanti della controparte e meritevoli di tutela. Sull'evoluzione del cd. danno ingiusto si v. M. FRANZONI, *Dei fatti illeciti*, cit., p. 159 ss.; G. CHINÈ e A. ZOPPINI, *Manuale di diritto civile*, IX<sup>ed.</sup>, Molfetta, 2020, p. 1987-1903;

<sup>15</sup> Non tutti distinguono il danno-evento dal danno-conseguenza. Si tratta di una differenziazione per lo più giurisprudenziale ma si v. anche C.M. BIANCA, *Diritto civile*, V, cit., p. 558-559. Da ultimo, la Cass., Sez. Un., 15 novembre 2022, n. 33645, sul danno da occupazione abusiva hanno ritrattato la duplicità del danno ritenendo che in realtà siano due facce della stessa medaglia ossia del danno ingiusto.

<sup>16</sup> Nel diritto penale il fatto da cui deriva responsabilità penale, appunto, deve necessariamente essere umano, imputabile, anti-giuridico e colpevole. Nel diritto civile, invece, in virtù di una logica preminentemente riparatoria della responsabilità civile – e non quindi sanzionatoria – il fatto può anche abbracciare relazioni meramente fattuali, purché accedano al piano dei fatti giuridicamente rilevanti. Tuttavia, con specifico riferimento all'art. 2043 c.c. è osservazione comune e tradizionale che sia più corretto dire «atto», essendo rilevante l'elemento volitivo del comportamento. Sul punto v. M. FRANZONI, *Dei fatti illeciti*, cit., p. 29.

obbligo giuridico di attivarsi<sup>17</sup>. Sotto il profilo giuridico, lo stesso deve, invece, presentarsi «*contra ius*» e «*non iure*». Di fatti, non solo deve essere in grado di comportare la lesione di un interesse meritevole di protezione giuridica (cagionare un danno appunto), ma non deve essere posto in essere in presenza di una causa di giustificazione.

In ogni caso, il fatto deve essere imputabile<sup>18</sup> al suo autore. In altre parole si richiede quel minimo di attitudine psichica ad agire e a valutare le conseguenze del proprio comportamento. Tale attitudine, nel diritto civile, deve essere valutata dal giudice caso per caso<sup>19</sup>.

Trattandosi di un elemento costitutivo dell'illecito civile la prova del fatto lesivo è a carico del danneggiato.

Il secondo elemento costitutivo dell'illecito è quello soggettivo che, come dispone la lettera della legge, può consistere nel dolo o nella colpa.

Quanto al dolo, è opinione generalmente condivisa che questo si identifichi con la nozione penalistica del dolo generico (art. 43 c.p.), prescindendo da elementi specifici di intenzionalità o di frode, risolvendosi nella volontà di cagionare un danno<sup>20</sup>. Prevalde l'idea di escludere il dolo, per errore di fatto o di diritto, laddove manchi la consapevolezza dell'ingiustizia del danno e di accogliere la generale regola dell'equivalenza fra il dolo e la colpa, nel senso che l'*an* e il *quantum* restano immutati sia che il comportamento si qualifichi come doloso che colposo. Vi sono, tuttavia, delle ipotesi, sancite legislativamente<sup>21</sup> o giurisprudenzialmente<sup>22</sup>, nelle quali la responsabilità sorge solo in presenza di

---

<sup>17</sup> Sulla fonte dell'obbligo di attivazione, al di fuori della legge, si discute in ambito penale. Ad ogni modo, in ambito civile pacificamente questo può derivare dalla legge, dal contratto, da un contatto sociale qualificato o da un'autonoma assunzione del rischio. Nei casi di responsabilità oggettiva – come si vedrà – la distinzione fra comportamento attivo ed omissivo ha poco senso.

<sup>18</sup> L'ordinamento subordina la responsabilità civile, e il conseguente obbligo risarcitorio, alla circostanza che l'autore del fatto dannoso fosse capace di intendere e di volere al momento in cui l'ha commesso. Cfr. C. SALVI, *La responsabilità civile*, in *Tratt. dir. priv.* Iudica e Zatti, 3<sup>ed.</sup>, Milano, 2019, p. 165. Per la definizione dei caratteri dell'imputabilità l'interprete dispone, oltre all'art. 2046 c.c., anche degli artt. 428 c.c. e 85 c.p.. Si discute se le diverse norme si riferiscano ad un concetto unitario di imputabilità. Il prevalente orientamento sembra orientarsi nel senso di negare che, ai fini della determinazione della nozione di imputabilità rilevante ai sensi dell'art. 2046 c.c. si possa tener conto della disciplina penalistica e delle nozioni di incapacità legale e naturale di agire rilevanti in materia di autonomia negoziale. Si tende, dunque, ad escludere non solo l'applicabilità diretta, ma anche analogica, della disciplina penalistica. Cfr. R. SCOGNAMIGLIO, *Responsabilità civile*, cit., p. 639; G. VISINTINI, *Trattato breve della responsabilità civile: fatti illeciti, inadempimento, danno risarcibile*, XV, Padova, 2005, p. 591; M. FRANZONI, *Dei fatti illeciti*, cit., p. 1032; P.G. MONATERI, *La responsabilità civile*, cit., p. 261.

<sup>19</sup> «Il giudice non può limitarsi a tenere presente l'età dell'autore del fatto, ma deve anche considerare lo sviluppo intellettuale, quello fisico, l'assenza di eventuali malattie ritardanti, la capacità del minore di rendersi conto della illiceità della sua azione e la capacità del volere con riferimento all'attitudine di autodeterminarsi». Cfr. Cass., 26 giugno 2001, n. 8740 in *forplus.it*.

Il rapporto tra imputabilità e la colpevolezza di cui all'art. 2043 c.c. è comunque discusso. Per approfondimenti si v. C. SALVI, *La responsabilità civile*, cit., p. 165-168.

<sup>20</sup> ID., *o.u.c.*, p. 168-170.

<sup>21</sup> Si pensi, ad esempio, alla responsabilità del proprietario per atti emulativi ex art. 833 c.c.; del giudice ex art. 2, l. 13 aprile 1988, n. 117; del cancelliere e dell'ufficiale giudiziario ex art. 60 c.p.c.; degli impiegati civili dello Stato ex artt. 22 ss. t.u. imp. civ. St.; del terzo acquirente del bene ipotecato ex art. 2864 comma 1; la responsabilità per lite temeraria ex art. 96, comma 1 c.p.c.

<sup>22</sup> Si immagini le ipotesi di induzione all'inadempimento e di contratto in frode ai terzi.

una qualità peculiare della colpevolezza, in termini di colpa grave, o di dolo, ovvero di una particolare intensità e direzione del dolo<sup>23</sup>.

Più circoscritta appare, invece, la peculiarità del dolo in punto di *quantum* del risarcimento, dal momento che questo è governato, in via di principio, dal criterio della riparazione integrale.

Quanto alla colpa, deve sin da subito constatarsi che il suo accertamento appare più complesso rispetto al dolo: si tratta di individuare quali, fra gli eventi causati ma non voluti dall'agente, gli vadano imputati.

È noto che negli ordinamenti contemporanei sia largamente prevalente la concezione di colpa che deduce la colpa «non da un giudizio sulla riprovevolezza soggettiva del comportamento (e quindi dall'esame dell'atteggiamento psicologico dell'agente o delle sue doti personali di intelligenza o prudenza), ma dal rapporto tra il comportamento dannoso e quello richiesto dall'ordinamento, nelle stesse circostanze concrete, al fine di evitare le lesioni degli interessi altrui»<sup>24</sup>.

A quelle critiche mosse a tale modello di colpa, sottolineanti che, aderendo ad un modello «astratto» di colpa, si finirebbe per introdurre una responsabilità oggettiva «mascherata», ancorata a criteri standardizzati, senza tenere in considerazione tutte quelle circostanze concrete del fatto e le condizioni soggettive dell'agente che hanno un forte impatto sul giudizio di rimproverabilità, si è risposto evidenziando, non solo la netta distinzione dalla cd. responsabilità oggettiva, nella quale l'imputazione prescinde dall'apprezzamento valutativo della condotta, e talvolta dalla circostanza stessa che vi sia stato un «comportamento» del responsabile, ma anche che la connotazione soggettiva della colpa non si perderebbe completamente. Di fatti, la concezione più moderna di colpa si compone della duplice concezione oggettiva-soggettiva: oggettiva, perché esprime il disvalore della condotta posta in essere in violazione di una regola cautelare; soggettiva, perché il giudizio di imputazione del fatto illecito e di rimproverabilità del fatto all'agente deve trovare un aggancio di tipo soggettivo psicologico, verificando se le conseguenze delle proprie azioni potevano essere, in concreto, prevedute ed evitate da quel

---

<sup>23</sup> Ciò avviene nel cd. *animus nocendi* di cui all'art. 833 c.c., oppure in talune figure giurisprudenziali di illecito concorrenziale. Cfr. C. SALVI, *La responsabilità civile*, cit., p. 169.

<sup>24</sup> *Id.*, *o.u.c.*, p. 170. Si distingue la concezione soggettiva della colpa dalla concezione oggettiva.

In passato la colpa veniva intesa maggiormente in un'ottica soggettiva quante volte se ne mutuava la definizione dal diritto penale (artt. 42 e 43 c.p.). Così, da una parte si aveva la cd. colpa generica (negligenza, imprudenza, imperizia), dall'altra quella specifica (identificantesi nella posa in essere di comportamenti in violazione di leggi, regolamenti, ordini, discipline e precetti dell'autorità). Tale distinguo comportava anzitutto che, sul piano processuale, la colpa generica si accertasse secondo i criteri della prevedibilità, dovendo tener conto delle condizioni in cui si poteva prevedere e prevenire l'evento dannoso e dello sforzo esigibile dall'uomo medio per evitare il danno stesso. Invece, per l'accertamento della colpa specifica, sarebbe stato sufficiente l'accertamento (oggettivo) della violazione del precetto.

La progressiva presa di distanza tra l'illecito civile e l'illecito penale ha tuttavia portato la dottrina più moderna a preferire, nell'imputazione della responsabilità aquiliana, una nozione «oggettiva» di colpa, in termini di obiettiva difformità del comportamento concretamente tenuto rispetto parametro astrattamente predefinito da criteri codificati espressamente, o comunque desumibili dal dovere di diligenza che *l'homo eiusdem conditionis et professionis* è tenuto ad osservare.

determinato soggetto, tenendo conto di tutte le circostanze esistenti nel caso di specie<sup>25</sup>.

«Così come non è sufficiente, quindi, la mera violazione della regola di condotta, richiedendosi (ex art. 2046 c.c.) che il comportamento sia posto in essere da soggetti capaci di autodeterminarsi e di valutare le conseguenze delle proprie azioni, è altresì necessario che il comportamento diligente, prudente e perito non venga valutato in astratto, in base a standard di comportamento, ma venga calato e sussunto nelle circostanze del caso concreto. Pertanto, lo standard medio, costruito sul *diligens pater familias*, può essere superato (o meglio elevato) nei casi in cui si sia in presenza di soggetti agenti inferiori o superiori al “tipo medio di consociato”»<sup>26</sup>.

La colpa, tuttavia, si atteggia diversamente a seconda che venga in rilievo una condotta attiva, intesa come azione o *facere*, ovvero una condotta omissiva, statica o *non facere*. Ebbene, mentre il precetto generale del *neminem laedere* richiede l’astensione da ogni attività che possa recar danno ad altri, non può dirsi, invece, che sempre e necessariamente sussista colpa in relazione ad un non fare per il sol fatto che l’eventuale attività del soggetto avrebbe potuto impedire l’evento dannoso. Infatti, non la semplice inerzia può dar luogo a responsabilità per colpa (secondo il noto brocardo *ex nihilo, nihil fit*, «dal nulla non può derivare nulla»), ma soltanto quella inattività che si risolve in una vera e propria omissione, nel senso giuridico del termine, cioè nel mancato compimento di una attività specificamente dovuta e possibile nel caso concreto (*non facere quod debeat*)<sup>27</sup>.

L’onere di provare la colpa, così come tutti gli elementi dell’illecito aquiliano, grava sul danneggiato. Detto onere è alleggerito dalla possibilità di ricorrere a presunzioni semplici basate sull’*id quod plerumque accidit* il cui utilizzo – come conferma Bianca<sup>28</sup> – dimostra l’abbandono della concezione etica della responsabilità. In altri termini, il danneggiato non deve provare un riprovevole atteggiamento psichico quanto piuttosto l’obiettiva deviazione dagli standard sociali e normativi di condotta. La prova di questa deviazione molte volte può essere data già dalla prova del danno<sup>29</sup> quando il danno sia un risultato scaturito normalmente da un’azione colposa.

Gli elementi terzo e quarto sono rispettivamente costituiti dal danno-evento (o danno ingiusto) e dal danno-conseguenza. Danno ingiusto è la lesione di un

---

<sup>25</sup> Si fa riferimento al cd. giudizio di prevedibilità-evitabilità dell’evento.

<sup>26</sup> Cfr. G. CHINÈ e A. ZOPPINI, *Manuale di diritto civile*, cit., p. 1925. Un’opinione diffusa, invece, rifiuta di assegnare rilievo, sul piano aquiliano, alle variazioni soggettive dello standard, anche o proprio in virtù delle varie complessità probatorie eccessive che l’adozione di quel canone produrrebbe nello svolgersi dei giudizi. V. al riguardo, Cass., 12 settembre 2013, n. 20904, in *Guida dir.*, 2013, f. 43, p. 59; Cass., 30 maggio 1996, n. 5005, in *Giust. civ. mass.*, 1996, p. 797; Cass. 1° febbraio 1991, n. 977, in *Giur. it.*, 1991, I, I, c. 1379.

<sup>27</sup> *Ibidem*.

<sup>28</sup> C.M. BIANCA, *Diritto civile*, V, cit., p. 558.

<sup>29</sup> M. BUSSANI, *L’illecito civile*, in *Tratt. dir. civ. CNN*, diretto da P. Perlingieri, Napoli, 2020, p. 573 evidenzia che non vi è chi non veda «come in queste ipotesi l’onere della prova sia sostanzialmente ridotto all’allegazione dei fatti accaduti, mentre è lasciato al convenuto il compito di prospettare al giudice i fattori atti a smentire l’interpretazione degli eventi prospettata dalla vittima, oppure l’inoperatività, nella specie, delle regole di comune esperienza».

qualsiasi interesse tutelato nella vita di relazione<sup>30</sup>. Esso può manifestarsi anche sotto le spoglie di pericolo di danno ma in questo caso, piuttosto che la tutela risarcitoria interverrà la tutela preventiva dell'inibitoria.

Diversamente, il danno conseguenza consiste nelle conseguenze giuridiche patrimoniali<sup>31</sup> e non patrimoniali<sup>32</sup> che seguono la lesione di un interesse giuridico meritevole di tutela.

Entrambi i danni meriterebbero un approfondimento dedicato ma, in questa sede, ci si limiterà alle parole anzidette.

Ultimo elemento costitutivo dell'illecito civile è la causalità. È principio generale del torto quello per cui è possibile imporre l'obbligo risarcitorio in capo a taluno solo se, e nella misura in cui, il danno del quale si discute appare come il risultato della sua attività o inerzia colpevole (o del fatto per il quale egli si è chiamato a rispondere ad altro titolo). Per un verso deve accertarsi il collegamento fra l'evento di cui è rimasto vittima l'attore e la condotta o

---

<sup>30</sup> La clausola generale del danno ingiusto assolve alla funzione di selezionare qualitativamente la circolazione della ricchezza, mediante la cernita degli interessi meritevoli di tutela, la cui lesione può essere foriera di un danno risarcibile. La nozione di ingiustizia del danno è stata, dagli anni '60 in poi, il punto cruciale del dibattito sulla responsabilità extracontrattuale. Un orientamento ravvisando nell'ingiustizia del danno una clausola generale attributiva di poteri creativi all'interprete, descrive la norma come un sistema aperto, tale cioè da garantire l'ingresso di sempre nuove ipotesi di illecito e di pregiudizi risarcibili. Si consente in questo modo al giudice di adeguare il principio di responsabilità il significato di ingiustizia del danno ai valori di una società in evoluzione e di conferire rilievo ma di interessi emergenti dalla dinamica dei rapporti sociali ed economici. Cfr. ALPA, *Il problema dell'atipicità dell'illecito civile*, Napoli, 1979; G. VISINTINI, *Trattato breve*, cit., p. 421; M. FRANZONI, *Dei fatti illeciti*, cit., *passim*.

Per approfondimenti sugli attributi necessari del danno aquiliano si v. C. SALVI, *Danno*, in *Dig. Disc. Priv.*, Sez. civ., V, Torino, 1989, p. 63; D. MESSINETTI, *Danno giuridico*, in *Enc. dir.*, Agg. I, Milano, 1997, p. 469; A. DE CUPIS, *Il danno*, cit.; E. GIUSANA, *Il concetto di danno giuridico*, Milano, 1944.

Per approfondimenti sul concetto di ingiustizia cfr. S. RODOTÀ, *Il problema della responsabilità civile*, cit., p. 187; G. CATTANEO, *La responsabilità del professionista*, cit., p. 157 ss.; G. PACCHIONI, *Dei delitti e dei quasi delitti*, Padova, 1940, p. 84 ss.; P. SCHLESINGER, *La "ingiustizia" del danno nell'illecito civile*, in *JUS*, 1960, III, p. 339-340.

Nella definizione di danno ingiusto si utilizzano delle perifrasi scolari. Per il danno ingiusto come lesione di un interesse «meritevole di tutela», che mutua volutamente l'art. 1322 c.c., si v. F. GALGANO, *Le mobili frontiere del danno ingiusto*, in *Contr. impr.*, 1985, p. 1, 7; V. ZENO-ZENCOVICH, *I fatti illeciti*, in *Tratt. dir. priv.* Rescigno, 2<sup>ed.</sup>, Torino, 1995, p. 134 ss.; G. ALPA, *Il problema dell'atipicità dell'illecito*, cit., p. 127 ss.

Per danno ingiusto come lesione di un interesse «giuridicamente protetto» cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 587 ss.

<sup>31</sup> «Un danno si definisce patrimoniale quando l'illecito ha compromesso utilità suscettibili di essere agilmente valutate in denaro. A rilevare non è l'interesse colpito ma le conseguenze provocate di cui l'attore chiedi il ristoro» in M. BUSSANI, *L'illecito civile*, cit., p. 174. Cfr. F.D. BUSNELLI e S. PATTI, *Danno e responsabilità civile*, 3<sup>ed.</sup>, Torino, 2013, p. 3-7; C. SALVI, *La responsabilità civile*, cit., p. 54; M. FRANZONI, *Dei fatti illeciti*, cit., p. 808 ss.; F. MASTROPAOLO, *Danno. III) Risarcimento del danno*, in *Enc. giur.*, IX, Roma, 1988, p. 6.

<sup>32</sup> Anche il danno non patrimoniale è tale, non per qualità intrinseche delle posizioni violate, ma perché insuscettibili di immediata traduzione economica sono considerati i riverberi dannosi che all'interessato sono derivati da quella violazione. Così lo definisce M. BUSSANI, *L'illecito civile*, cit., p. 182.

Per le primissime indicazioni al riguardo, cfr. M. TAMPIERI, *Il danno non patrimoniale. La lesione di valori costituzionalmente rilevanti*, Padova, 2015; P. VIRGADAMO, *Danno non patrimoniale e "ingiustizia conformata"*, Torino, 2014; F.D. BUSNELLI e S. PATTI, *Danno e responsabilità civile*, cit., p. 59-95; A. ASTONE, *Art. 2059. Danni non patrimoniali*, in *Comm. del cod. civ.* Schlesinger e Busnelli, Milano 2012, p. 119-158; ID., *I danni non patrimoniali alla persona: la prova*, Milano 2011; C. BONA, *Studio sul danno non patrimoniale*, Milano 2012; S. PATTI e S. DELLE MONACHE (a cura di), *Responsabilità civile. Danno non patrimoniale*, Milano, 2010; E. NAVARRETTA (a cura di), *Il danno non patrimoniale: principi, regole e tabelle per la liquidazione*, Milano 2010; P. CENDON e P. ZIVIZ, *Danno esistenziale. Una nuova categoria della responsabilità civile*, Milano, 2000; G. BONILINI, *Danno morale*, in *Dig. civ.*, V, Torino, 1989, p. 83; R. SCOGNAMIGLIO, *Danno morale*, in *Noviss. dig. it.*, V, Torino 1960, p. 146.

l'omissione di chi ne è ritenuto l'autore e, per altro verso, è necessario segnare l'orizzonte delle conseguenze negative che possono costituire oggetto di risarcimento da parte di quest'ultimo. In altre parole, «il requisito della causalità vale a individuare il soggetto che ha contribuito all'occorrere dell'evento e selezionare, nel paniere delle perdite di cui la vittima prova di essere incorsa, le poste che possono essere traslate sulle spalle del responsabile»<sup>33</sup>.

Proprio tale manifestazione «bifronte» della causalità ha condotto la dottrina prevalente<sup>34</sup> a riconoscere ben due nessi di causalità: l'uno materiale (si riferisce al rapporto che lega la condotta di un soggetto all'evento dannoso, *Haftungsbergründende Kausalität*), l'altro giuridico (finalizzato a stabilire la rilevanza giuridica delle conseguenze economiche sfavorevoli derivanti dal fatto illecito, *Haftungsausfüllende Kausalität*). Il primo sarebbe interno alla fase dell'*an respondeatur*, il secondo sarebbe pertinente al *quantum respondeatur*<sup>35</sup>.

<sup>33</sup> M. BUSSANI, *L'illecito civile*, cit., p. 631.

<sup>34</sup> Tra molti v. G. GORLA, *Sulla cosiddetta causalità giuridica: "fatto dannoso e conseguenze"*, in *Riv. dir. comm.*, 1951, I-II, p. 405 ss.; e pure F. REALMONTE, *Il problema del rapporto di causalità nel risarcimento del danno*, Milano, 1967, p. 13; V. GERI, *Il rapporto di causalità in diritto civile negli illeciti civili*, in *Arch. resp. civ.*, 1962, p. 338; V. CARBONE, *Il rapporto di causalità*, in G. ALPA E M. BESSONE (a cura di), *La responsabilità civile. Una rassegna di dottrina e giurisprudenza*, II, 2, Torino, 1987, p. 148; M. FRANZONI, *Dei fatti illeciti*, cit., p. 773 ss.; A. BELVEDERE, *Causalità giuridica?*, in *Riv. dir. civ.*, 2006, I, p. 7 ss.; R. PUCELLA, *Causalità di fatto e causalità giuridica*, in U. CARNEVALI (a cura di), *Dei fatti illeciti*, I, cit., p. 405 ss.

In senso critico si esprime M. BUSSANI, *L'illecito civile*, cit., p. 637, il quale riscontra una «fragilità operativa» proprio nel presupposto della bipartizione del nesso causale. Fa riferimento all'idea «che il nesso causale possa essere oggetto di un accertamento di carattere 'meccanico', volto a sceverare e misurare, nell'ipotesi sottoposta all'attenzione del giudice, atti, fatti, eventi e conseguenze in modo sempre nitido, e sempre privo di incroci rispetto ai percorsi nei quali si sonda l'esistenza, o la consistenza delle componenti ulteriori della fattispecie – il danno, la sua ingiustizia, la colpa, e così via». Diversamente l'a. non ritiene possibile immaginare la generalizzazione astratta dell'imputazione di un fatto, un'azione, un evento o una conseguenza. Del medesimo avviso sono: M. INFANTINO, *La causalità nella responsabilità extracontrattuale. Studio di diritto comparato*, Napoli, 2012, p. 2937; P. CENDON, *Il dolo nella responsabilità extracontrattuale*, Torino, 1974, p. 368; v. pure ID., *Conseguenze volute e risarcimento del danno*, in Id., *Raccolta di scritti per il 50° Anniversario della Facoltà di Economia e Commercio di Trieste*, Udine, 1974, p. 61 ss.; P. TRIMARCHI, *La responsabilità civile*, cit., p. 536 ss.; ID., *Causalità e danno*, cit., p. 42 ss.; G. GORLA, *Sulla cosiddetta causalità giuridica*, cit., p. 410-412; F. CARNELUTTI, *Perseverare diabolicum*, cit.. Per riflessioni con varietà di accenti, C. MAIORCA, *I fondamenti della responsabilità*, Milano, 1990, p. 419; C. CASTRONOVO, *Il risarcimento e danno*, in *Riv. dir. civ.*, 2006, p. 81-97; M. CAPECCHI, *Il nesso di causalità. Da elemento della fattispecie "fatto illecito" a criterio di limitazione del risarcimento del danno*, 2<sup>ed.</sup>, Padova 2005, p. 43-49; R. SCOGNAMIGLIO, *Appunti sulla nozione di danno*, in *Riv. trim.*, 1969, p. 464 ss.; A. DE CUPIS, *Il problema giuridico del "quantum respondeatur"*, in *Riv. dir. civ.*, 1967, p. 516 ss.; L. DEVOTO, *La concezione analitica dell'illecito*, in *Riv. dir. civ.*, 1965, I, p. 498 ss., 510; S. RODOTÀ, *Il problema della responsabilità civile*, cit., p. 77; F.D. BUSNELLI, *La lesione del credito da parte di terzi*, Milano, 1963, pp. 120, nota 42; P. FORCHIELLI, *Il rapporto di causalità nell'illecito civile*, Bologna, 1960, p. 22; P. CALAMANDREI, *La condanna generica ai danni*, in *Studi sul processo civile*, III, Padova, 1934, p. 231, («quantità è misura di esistenza così come l'esistenza non può essere affermata se non entro certi limiti, il riconoscimento dei quali implica un giudizio di quantità»).

<sup>35</sup> Un paradigma normativo di tale distinzione, secondo la dottrina e la giurisprudenza prevalente, è ravvisabile rispettivamente nel I e nel II comma dell'art. 1227 c.c. Il comma I, partendo dalla disposizione (se il fatto colposo del creditore ha concorso a cagionare il danno, il risarcimento è diminuito secondo la gravità della colpa e l'entità delle conseguenze che ne sono derivate), lascia chiaramente intendere che il legislatore ha preso in esame l'ipotesi in cui il fatto del creditore-danneggiato interviene a spezzare il legame, a monte, tra comportamento del soggetto agente ed evento, escludendo così la totale imputabilità del fatto all'agente, e limitando di conseguenza la responsabilità di quest'ultimo. Il comma II, al contrario, (il risarcimento non è dovuto per i danni che il creditore avrebbe potuto evitare usando l'ordinaria diligenza), chiarisce in che modo il fatto del creditore possa influire, a valle, sul diverso rapporto evento-danno, e cioè rendendo non più risarcibili talune delle conseguenze immediate e dirette dell'evento, nonostante sia già stata accertata la piena responsabilità del danneggiante, e sia già stato determinato il risarcimento attraverso

Quanto all'accertamento del nesso di causalità giuridica, risulta tendenzialmente pacifico in giurisprudenza<sup>36</sup> il principio per cui ci si debba riferire, in base all'art. 1223 c.c., al criterio della cd. regolarità causale, alla stregua del quale sono risarcibili i danni immediati e diretti in quanto normale conseguenza dell'evento lesivo, secondo l'*id quod plerumque accidit* o secondo la comune esperienza.

Quanto all'accertamento del nesso di causalità materiale, quello che avvince la condotta all'evento, invece, si assiste ad un vivace dibattito dottrinale e giurisprudenziale dovuto al fatto che, l'unico accenno che il legislatore del '42 fa al nesso di causalità, è quello ricavabile dall'art. 2043 c.c. laddove dispone che il fatto deve essere addebitato a chi «cagiona» il danno. Diversamente, nel sistema penalistico, anche se laconicamente, la causalità materiale viene disciplinata dagli artt. 40 e 41 c.p. Di qui sorge il problema se, nel silenzio del codice civile, l'accertamento della causalità civile segua i medesimi criteri elaborati in tema di causalità penale<sup>37</sup> oppure sia sottoposta a regole autonome in ragione della

---

il filtro dell'art. 1223 c.c.. Così, da una lettura sistematica delle norme del codice civile vigente dettate in tema di risarcimento del danno, sembra emergere un plesso operativo razionale e coerente, articolato secondo un criterio di consequenzialità in virtù del quale il legislatore opera una netta separazione tra il momento (strutturale) dell'accertamento della responsabilità e quello (funzionale) dal contenuto della stessa. Cfr. G. CHINÈ e A. ZOPPINI, *Manuale di diritto civile*, cit., p. 1928.

<sup>36</sup> In dottrina la tematica è oggetto di dibattito.

<sup>37</sup> Dal punto di vista penalistico, hanno posto dei rilevanti *benchmarks* sull'accertamento della causalità le Sezioni Unite della Cassazione penale con la celebre Sentenza Franzese (Cass. pen., Sez. un., 11 settembre 2002, n. 30328 in *Danno resp.*, 2003, p. 195, con nota di S. CACACE, *L'omissione del medico e il rispetto della presunzione d'innocenza nell'accertamento del nesso causale*, *ivi*, p. 201; M. CAPECCHI, *Le sezioni unite penali sul nesso di causalità omissiva. Quali riflessioni per la responsabilità civile del medico?* in *Nuova giur. civ. comm.*, 2003, p. 257; F. STELLA, *Etica e razionalità del processo penale nella recente sentenza sulla causalità delle Sezioni Unite della Suprema Corte di cassazione*, in *Riv. it. dir. proc. pen.*, 2002, p. 767).

Gli artt. 40 e 41 c.p., pacificamente, devono essere letti alla luce della «teoria condizionalistica» o dell'«equivalenza causale», temperata, ma in realtà ribadita mediante il riferimento, specularmente e in negativo, alla «causalità umana», quanto alle serie causali sopravvenute, autonome e indipendenti, da sole sufficienti a determinare l'evento ex art. 41 comma 2. È, dunque, «causa» dell'evento la condotta umana, attiva o omissiva, che si pone come *conditio sine qua non* nella catena degli antecedenti che hanno concorso a produrre il risultato, senza la quale l'evento da cui dipende l'esistenza del reato non si sarebbe verificato.

«La verifica della causalità postula il ricorso al “giudizio controfattuale”, da compiersi ex post, cioè dopo e rispetto all'evento così come si è verificato, *hic et nunc*. Esso si sostanzia nella eliminazione mentale della condotta umana data per condizione tante per verificare se, senza questa, l'evento si sarebbe ugualmente prodotto.

A tal fine, si suole costruire il giudizio secondo la “doppia formula”:

- a. la condotta umana “è” condizione necessaria dell'evento se, eliminata mentalmente dal novero dei fatti realmente accaduti, l'evento non si sarebbe verificato (formula positiva);
- b. la condotta umana “non è” condizione necessaria dell'evento se, eliminata mentalmente mediante il medesimo procedimento, l'evento si sarebbe ugualmente verificato (formula negativa).

Lo statuto logico della causalità rimane lo stesso anche quando venga in rilievo la peculiare forma di «causalità omissiva» unica differenza è che l'accertamento di quest'ultima richiederà un giudizio doppiamente ipotetico perché la formula «dovrà rispondere al quesito se, “*mentalmente eliminato*” il mancato compimento dell'azione doverosa e “*sostituito alla componente statica un ipotetico processo dinamico corrispondente al comportamento doveroso, supposto come realizzato*” il singolo evento lesivo, *hic et nunc* verificatosi, sarebbe, o non, venuto meno.

Ciò posto, il giudizio controfattuale va condotto sulla base di una generalizzata regola di esperienza o di una legge scientifica di copertura. Di talché, un antecedente può essere configurato come condizione necessaria solo se esso rientri nel novero di quelli che, sulla base di una successione regolare conforme ad una generalizzata regola di esperienza o ad una legge dotata di validità scientifica, frutto della migliore scienza ed esperienza del momento storico, conducano a eventi del “tipo” di quello verificatosi in concreto.

diversità degli scopi e della funzione tra illecito civile e illecito penale. La questione non è meramente teorica, se si consideri che, a seconda dell'opzione prescelta, si ampliano o restringono le maglie della tutela risarcitoria del danneggiato.

L'opinione oggi prevalente in giurisprudenza ed in dottrina<sup>38</sup>, valorizzando le profonde differenze morfologiche e funzionali dei due sistemi di responsabilità, opta per la sostanziale autonomia dell'accertamento della responsabilità civile da quella penale. Questo si afferma lasciando invariata la definizione di «nesso di causalità» definita dalle Sezioni Unite con la sentenza Franzese<sup>39</sup> e sostenendone la divergenza da un punto di vista operativo, in sede di accertamento dello stesso. Fondamentali, in merito, sono state le pronunce delle Sezioni Unite civili del

---

Il sapere scientifico accessibile al giudice è costituito, a sua volta, sia da leggi “universalì” (invero assai rare, che asseriscono nella successione di determinati eventi invariabili regolarità senza eccezioni, sia da leggi “statistiche” che si limitano ad affermare che il verificarsi di un evento è accompagnato dal verificarsi di un altro evento in una certa percentuale di casi e con una frequenza relativa, con la conseguenza che quest'ultime [...] sono tanto più dotate di “alto grado di credibilità razionale” o “probabilità logica”, quanto più trovano applicazione in un numero sufficientemente elevato di casi e ricevono conferma mediante il ricorso a metodi di prova razionali ed empiricamente controllabili.

Il ricorso a dette ‘generalizzazioni scientificamente valide’ [...] è un dato irrinunciabile, perché consente di ancorare il giudizio controfattuale, altrimenti insidiato da ampi di discrezionalità e di indeterminazione, a parametri oggettivi, in conformità alle ragioni di determinatezza e legalità delle fattispecie di reato.

Di qui la questione, rimessa al vaglio delle Sezioni unite, se “*la sussistenza del nesso di causalità fra condotta omissiva ed evento, [...] debba essere ricondotta all'accertamento che con il comportamento dovuto ed omesso l'evento sarebbe stato impedito con elevato grado di probabilità “vicino alla certezza”, e cioè in una percentuale di casi “quasi prossima a cento”, ovvero siano sufficienti, a tal fine, soltanto “serie ed apprezzabili probabilità di successo” della condotta che avrebbero impedito l'evento?*”.

La soluzione offerta fuoriesce dalle secche della citata alternativa di tipo statistico-probabilistico (serie ed apprezzabili probabilità o percentuale di casi prossima a cento?), per abbracciare il diverso della «probabilità logica o certezza processuale». [...] In sostanza, il criterio empirico della probabilità statistica è sì necessario ai fini dell'accertamento del nesso causale, ma non sufficiente, imponendosi al giudice una ‘verifica aggiuntiva’ di tipo induttivo, dell'attendibilità dell'impiego dello stesso alla fattispecie concreta, alla stregua di tutto il materiale probatorio disponibile, anche al fine di escludere la ricorrenza ai processi eziologici alternativi.

In termini, il nesso causale deve essere accertato, così come tutti gli elementi costitutivi della penale responsabilità, con alto grado di credibilità razionale, ossia oltre ogni ragionevole dubbio, con un procedimento non diverso dalla sequenza inferenziale dettata in tema di prova indiziaria dall'art. 192, comma 2 c.p.p., con l'avvertenza naturalmente che la permanenza di un ragionevole dubbio, nonostante tutte le verifiche compiute, dovrà necessariamente comportare la neutralizzazione dell'ipotesi prospettata dall'accusa e l'esito assolutorio del giudizio». Cfr. G. CHINÈ e A. ZOPPINI, *Manuale di diritto civile*, cit., p. 1929-1931.

<sup>38</sup> Nel senso di osteggiare la trapiantabilità delle istruzioni penalistiche nel terreno civile si v., con varietà di argomenti, V. ZENO-ZENCOVICH, *Questioni in tema di responsabilità per colpa professionale* nota a Cass. pen., 12 luglio 1991, n. 371, in *Nuova giur. civ. comm.*, 1992, I, p. 358 ss.

P. CENDON, *Conseguenze volute e risarcimento del danno*, in ID., *Raccolta di scritti per il 50° Anniversario della Facoltà di Economia e Commercio di Trieste*, Udine, 1974, pp. 61, 76; S. RODOTÀ, *Il problema della responsabilità civile*, cit., p. 57; C. SALVI, *Responsabilità extracontrattuale*, cit., p. 1249; . C. CASTRONOVO, *L'obbligazione senza prestazione ai confini tra contratto e torto*, in *AA.VV.*, *Le ragioni e diritto. Scritti in onore di Luigi Mengoni*, I, Milano, 1995, pp. 191, 217; G. PONZANELLI, *La responsabilità civile*, cit., p. 15; V. ZENO ZENCOVICH, *La responsabilità civile da reato. Lineamenti e prospettive di un sottosistema giurisprudenziale*, Padova, 1989, p. 39; G. VALCAVI, *Intorno al rapporto di causalità nel torto civile*, in *Riv. dir. civ.*, 1995, II, pp. 482, 487; M.L. RUFFINI, *Concausa non imputabile e determinazione del quantum di responsabilità: il percorso giurisprudenziale*, in *Danno resp.*, 2014, p. 669.

<sup>39</sup> V. nota 37.

2007<sup>40</sup> e del 2008<sup>41</sup> grazie alle quali viene affermato che «l'applicazione dei principi generali di cui agli artt. 40 e 41 c.p., temperati dalla “regolarità causale”<sup>42</sup>, ai fini della ricostruzione del nesso eziologico va adeguata alle peculiarità delle singole fattispecie normative di responsabilità civile»<sup>43</sup>. Ciò significa che, se da una parte deve ammettersi, in sede civile, la presenza di plurimi criteri di imputazione del danno capaci di selezionare serie causali differenti, senza, però, eliminare l'eziologia in sé<sup>44</sup>, dall'altra parte la regola probatoria deve ritenersi meno rigorosa di quella penale ed appuntarsi su una «probabilità prevalente» ossia sul «più probabile che non».

Ciò vuol dire che il danno si considera derivante da un fatto se le probabilità che il danno sia la conseguenza di quel fatto siano maggiori delle probabilità che non lo sia<sup>45</sup>, tenendo a mente che deve sempre farsi riferimento ad una probabilità «relativamente» prevalente, calcolata alla luce del caso concreto<sup>46</sup>. La divaricazione dei criteri causali sarebbe giustificata dalla diversità dei «valori in gioco»<sup>47</sup>.

Orbene, se si accerta che il danno è causalmente imputabile al responsabile la sua risarcibilità deve avvenire per intero. Trova infatti applicazione il principio,

---

<sup>40</sup> Cass., 16 ottobre 2007, n. 21619, in *Resp. civ. prev.*, 2008, p. 323, nella quale per la prima volta, *expressis verbis*, si afferma che «il principio accolto in materia penale dalla sentenza Franzese [...] è difficilmente trasferibile nel campo della responsabilità civile, ove il rapporto causale può ritenersi sussistente anche in presenza di meno elevate soglie di probabilità, e cioè quando risulta che l'opera del sanitario, se correttamente e tempestivamente intervenuta, avrebbe avuto non già la certezza bensì soltanto serie ed apprezzabili probabilità di successo».

<sup>41</sup> Cass., Sez. un., 11 gennaio 2008, n. 581 in *Resp. civ. prev.*, 2008, p. 841, la quale fissa importanti principi in tema di *dies a quo* della prescrizione per il risarcimento dei danni lungolatenti e di responsabilità del Ministero della salute per danni da emotrasfusioni.

<sup>42</sup> Il criterio della regolarità causale ravvisa il danno causato dal fatto del responsabile nel danno che ne costituisce una conseguenza normale. V. Cass., 19 dicembre 2006, n. 27168 in *www.foroplus.it*, laddove si afferma che, ai fini del nesso di causalità occorre che il fatto costituisca un antecedente necessario dell'evento, nel senso che quest'ultimo rientri tra le conseguenze normali ed ordinarie.

Esistono, tuttavia, altri criteri o «formule» elaborate dalla tradizione scolare per accertare il nesso causale o temperare la teoria condizionalistica. Si fa riferimento al criterio dello «scopo della norma violata», alla «causalità adeguata», e alla «creazione di un rischio». Per approfondimenti cfr. M. BUSSANI, *L'illecito civile*, cit., p. 631 ss., il quale ritiene che non è dato immaginare la generalizzazione astratta dell'imputazione di un fatto, un'azione, un evento o una conseguenza. Di talchè, si può dire che la causalità si atteggi diversamente in relazione alle esigenze di volta in volta poste davanti all'interprete e da questi valorizzate. Di fatti, l'a. sostiene che il giudizio causale è nelle mani dell'interprete; C.M. BIANCA, *Diritto civile*, V, cit., p. 138-144.

<sup>43</sup> Cass., Sez. un., 11 gennaio 2008, n. 581, cit.

<sup>44</sup> Ci si riferisce al discorso in base al quale, mentre nella responsabilità penale il rapporto eziologico ha sempre come punto di riferimento iniziale la condotta dell'agente, «in tema di responsabilità civile extracontrattuale il punto di partenza del segmento causale rilevante può essere anche altro, se in questi termini la norma fissa il criterio di imputazione, ma le regole per ritenere sussistente, concorrente, insussistente o interrotto il nesso causale tra tale elemento e l'evento dannoso, in assenza di altre disposizioni normative, rimangono quelle fissate dagli artt. 40 e 41 c.p.»

<sup>45</sup> Cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 143.

<sup>46</sup> *Ibidem* in nota 93. V. Cass., 29 settembre 2015, n. 19213, in *www.foroplus.it*, nella quale si specifica che l'adozione del criterio della probabilità relativa si delinea in una analisi specifica e puntuale di tutte le risultanze probatorie del singolo processo, sicché la concorrenza di cause di diversa incidenza probabilistica deve essere attentamente valutata e valorizzata in ragione della specificità del caso concreto, senza limitarsi ad un meccanico e semplicistico ricorso alla regola del 51% ma facendosi luogo ad una compiuta valutazione dell'evidenza del probabile.

Cfr. anche M. FRANZONI, *Dei fatti illeciti*, cit., p. 51-52, spec. nota 63.

<sup>47</sup> Cass., Sez. un., 11 gennaio 2008, n. 581, cit.

accolto in giurisprudenza, della integralità del risarcimento in base al quale il danno causalmente imputabile al responsabile va risarcito interamente mentre non va risarcito il danno non causalmente imputabile al responsabile (principio anche detto del tutto niente o *all or nothing*). Una responsabilità «frazionabile» sembra ammissibile, oggi, unicamente in tema di risarcimento da perdita di *chance* ma sul punto non ci si sofferma.

Essendo il nesso causale elemento costitutivo dell'illecito, la prova dello stesso – ex art. 2697 c.c. in base al quale «chi vuol far valere un diritto in giudizio deve provare i fatti che ne costituiscono il fondamento» – incombe sul danneggiato. L'onere della prova è assolto dimostrando gli elementi di fatto e di diritto dai quali si possa desumere che il danno è causalmente ascrivibile al terzo responsabile in via extracontrattuale ed è agevolato, largamente, dall'onere, che incombe sul convenuto, di prendere posizione sui fatti addotti dall'attore, per cui i fatti non contestati s'intendono provati ex art. 167 c.p.c. La gravosità dell'onere è altrettanto diminuita da una molteplicità di strumenti – oltre quello già menzionato – quali il ricorso alle presunzioni o alle nozioni di comune esperienza circa lo scorrere delle vicende umane, il ricorso da parte del giudice o delle parti alla sapienza peritale.

La quantità di prova necessaria ad attestare l'esistenza – o l'assenza – del legame causale cambia al variare degli indizi che si sono saputi raccogliere con riferimento agli altri elementi costitutivi dell'illecito. Più numerosi sono questi ultimi, minore è l'esigenza di approfondire la prima. Di fatti, come attenta dottrina rileva<sup>48</sup>, i problemi di «fragilità» della prova del nesso causale trovano un puntello nell'accertamento della serietà del pregiudizio sofferto dalla vittima piuttosto che nelle «imperdonabili negligenze imputabili a chi è da questa indicato quale responsabile»<sup>49</sup>.

Orbene, alla luce di quanto sin qui analizzato si possono tentare delle riflessioni circa la *suitability* del criterio della responsabilità per colpa ai danni cagionati dall'IA «forte».

Nella sostanza, il danno causato da o nel corso del funzionamento di alcuni sistemi di IA non sembra mostrarsi di tipo diverso da qualsiasi altro danno rilevante in materia di responsabilità civile: la vittima potrebbe subire danni fisici o addirittura morire, potrebbero esserci danni alla proprietà o qualsiasi altro tipo di perdita. Perciò, il paradigma dell'art. 2043 c.c. potrebbe ben applicarsi anche in questo caso. Tra l'altro, l'imputazione per colpa, come molti sostengono, sarebbe il criterio preferibile per rispondere anche alla necessità «morale» di

---

<sup>48</sup> M. BUSSANI, *L'illecito civile*, cit., p. 685 -686.

<sup>49</sup> *Ibidem*. L'a. scrive che «Poco importa, in siffatte ipotesi, che non sia possibile offrire al giudice la certezza che un comportamento diverso da quello criticato sarebbe valso a impedire l'evento dannoso: gran parte dei dubbi circa l'eziologia del danno possono essere facilmente dissipati non appena divenga evidente l'entità della sciagura di cui si discute, o gli elevati tassi di repressibilità che connotano la posizione dell'autore del fatto. Cfr. anche P. CENDON, *Circostanze incerte e responsabilità civile*, in *Riv. trim.*, 1999, p.1237, 1240; P. TRIMARCHI, *Causalità e danno*, Milano, 1967, p. 13.

ascrivere il danno unicamente in capo a colui cui possa muoversi un rimprovero. E ciò sarebbe l'espressione di alto grado di civiltà<sup>50</sup>.

Tuttavia, esistono delle ragioni per le quali ritenere che il menzionato criterio di imputazione non sia il più opportuno in quanto aggrava eccessivamente la posizione del danneggiato sol si consideri la difficoltà di adempiere all'onere della prova nella sua totalità.

Si devono tenere a mente, infatti, le peculiarità della tecnologia che - come esposto nel capitolo I – richiedono delle nozioni altamente tecniche. Conseguentemente, il danneggiato dovrà fare i conti con quella che si è *supra* chiamata *black box* ossia l'inesplicabilità dell'algoritmo<sup>51</sup>, l'incapacità di conoscere, e quindi ripercorrere, il procedimento logico-matematico che ha portato ad un dato *output* che si è rivelato dannoso.

Orbene, tale inesplicabilità – diretto corollario dell'apprendimento automatico – si riflette sull'onere della prova del danneggiato in più modi.

Innanzitutto, il comportamento autonomo dall'IA ostacola la prova del nesso causale fra il risultato dannoso e l'azione (o l'omissione) della persona potenzialmente responsabile, mettendo alla prova il menzionato modello prevalentemente «antropocentrico e monocausale» di causazione del danno<sup>52</sup>.

Invero, il fatto che, ad esempio, la rete neurale profonda costituente l'IA operi seguendo regole non interamente predisposte da un umano, mutando la propria essenza costantemente, rende il funzionamento dell'IA poco trasparente e, conseguentemente, rende difficile individuare quali dati di *input* generino quello specifico *output* dannoso<sup>53</sup>. In altri termini, risulta complesso identificare una «causa specifica» – la *conditio sine qua non* – ossia risalire al comportamento lesivo e, quindi, al soggetto che ha agito (o non agito) «per mezzo dell'IA»<sup>54</sup>, sulla quale doveva operare un «controllo»<sup>55</sup>. Si parla, perciò, anche di «problema di dissipazione del controllo» umano che, per la verità, si atteggia più ad un «problema di dispersione del controllo tra più soggetti» considerato che, la vera difficoltà risiede nel riconoscere tra i numerosissimi soggetti che possono venire in gioco nella catena del valore di un'IA, quale o quali abbiano tenuto un

---

<sup>50</sup> M. BUSSANI, *L'illecito civile*, cit., p. 48. L'a. richiama il celebre argomento di R. VON JHERING, *Das Schuldmoment in römischen Privatrecht*, Giessen, 1867, p. 8-10, ripreso, a livello locale, da A. DE CUPIS, *Il danno*, cit., p.121, il quale scrive «il passaggio della responsabilità vincolata alla colpa costituisce un grande progresso nella storia del diritto, né è lecito tornare alle rozze concezioni dei tempi primitivi».

<sup>51</sup> Offrono una soluzione per affrontare l'inesplicabilità dell'IA: P.H. PADOVAN, C.M. MARTINS e C. REED, *Black is the new orange: how to determine AI liability*, in *A.I. and Law*, Berlino, 2023, p. 133-167. Gli a. suggeriscono di utilizzare tecniche di XAI (*explainable artificial intelligence*) per analizzare la causalità materiale e giuridica offrendo una comparazione della valutazione del nesso causa negli ordinamenti di *civil law* ed in quelli di *common law*, in particolare con riferimento all'ordinamento brasiliano e inglese.

<sup>52</sup> O. DHEU e J. DE BRUYNE, *Artificial Intelligence and Tort Law: A 'Multi-faceted' Reality*, in *Eu. Rev. Priv. Law*, 2023, II e III, p. 261-298.

<sup>53</sup> Si v. M. BUITEN, A. DE STREEL e M. PEITZ, *EU Liability Rules for the Age of Artificial Intelligence*, in *www.ssrn.com*, 2021, p. 34- 35, I quali sottolineano che la valutazione del nesso di causalità richiederà spesso la consulenza di esperti, il cui costo può scoraggiare le parti lese dall'intentare causa.

<sup>54</sup> Si potrebbe, così, parlare, più propriamente, di danno «mediato dall'IA».

<sup>55</sup> Si fa riferimento ad un qualsiasi tipo di controllo dal momento che, per il momento, non esistono specificazioni.

comportamento «realmente efficiente» alla causazione del danno<sup>56</sup>. A nulla o a poco gioverebbero sia gli strumenti di alleggerimento della prova già menzionati, sia le norme sulla solidarietà nella responsabilità extracontrattuale, considerato che non sono dotate di alcuna funzione selettiva delle serie causali rilevanti. Ebbene, come è stato osservato da Schütte<sup>57</sup>, tale quadro rischierebbe di alimentare la «cultura del reclamo», ossia di generare un «contenzioso eccessivo o abusivo» in cui i danneggiati potrebbero indirizzare le richieste di risarcimento contro «tutti i possibili convenuti», nonostante solo uno di essi sia la vera causa del danno.

In secondo luogo, seppur il danneggiato riuscisse ad individuare il «comportamento» cagionante il danno e, dunque, il soggetto imputabile, molto ardua sarebbe la dimostrazione della colpevolezza. Se per il dolo, probabilmente, potrebbero raccogliersi degli elementi indiziari inequivocabili – vista l'intensità dell'elemento soggettivo –, per la colpa si richiederebbe da parte del danneggiato un comportamento che va oltre la ragionevole esigibilità. Invero, egli non solo dovrebbe individuare la norma cautelare violata, tesa ad evitare proprio il danno prodotto attraverso l'impartizione della condotta alternativa lecita, ma dovrebbe, altresì, dimostrare che il danno, tenendo quel comportamento, fosse prevedibile. Il tutto navigando a tentoni fra l'ancóra parziale comprensione e valutazione dei rischi legati alla tecnologia IA, normative quasi inesistenti, codici di condotta e *standard* di diligenza in corso di formazione. Oltre al fatto che, talvolta, una colpa vera e propria potrebbe anche non sussistere laddove l'IA non sia più sotto il «controllo» di nessuno o detto controllo sia stato effettuato correttamente.

Alla luce di tutto quanto sin qui riportato, non vi è chi non veda che, forse, il criterio della colpa non sia il preferibile di fronte ad un danno «mediato» dalla IA.

## 2.2. La responsabilità da attività pericolosa ex art. 2050 c.c.

Alla luce delle peculiari caratteristiche delle IA ed in virtù dell'orientamento ormai acquisito, secondo cui i criteri di imputazione sfuggono ad ogni

---

<sup>56</sup> Di tale avviso è anche D. KIM, *Artificial Intelligence Should not Become a "Black Hole" for Human Agency in Tort Law*, in *Tort Law Rev.*, 2023, 29, p. 152 ss.

Le parti coinvolte vanno dal programmatore di *software*, ai fornitori di rete, ai fornitori di dati fino alle autorità pubbliche, ad esempio in qualità di gestori di infrastrutture.

Possono concorrere ad aumentare i soggetti coinvolti e, quindi, le difficoltà di individuazione della «causa specifica» diversi fattori. Fra questi vi rientra la presenza di diverse componenti nel prodotto *AI based* e la possibilità che ognuna di queste componenti comunichi tra loro (dispositivi *IoT*). La loro interazione può influire sul nesso di causalità tra la condotta delle parti e il danno. Si consideri, ad esempio, un termostato intelligente, composto da sensori di temperatura e umidità distribuiti in tutta la casa e da un'unità termostatica centrale: eventuali errori nell'impostazione della temperatura del termostato potrebbero essere la conseguenza di un'errata raccolta di dati nel sensore, di un difetto nei componenti fisici dei sensori o del software che ottimizza e analizza tutti i dati raccolti. Se il termostato è collegato anche ad altri dispositivi *IoT* presenti nell'abitazione, gli errori nella raccolta e nell'analisi dei dati raccolti da questi altri dispositivi, o i difetti nei loro componenti, potrebbero influire anche sul termostato intelligente. L'interazione di diversi componenti e dispositivi può dare risultati inaspettati, rendendo difficile l'identificazione dell'origine del danno. Inoltre, moltiplica gli attori coinvolti.

<sup>57</sup> B. SCHÜTTE, L. MAJEWSKI e K. HAVU, *Damages Liability for Harm Caused by Artificial Intelligence – EU Law in Flux*, in *Legal Studies Research Paper Series*, 2021, 69, p. 26.

classificazione d'ordine gerarchico che non sia fondata su ragioni di ordine quantitativo o statistico (maggiore frequenza d'applicazione) o su motivi, verbalizzati o meno, di politica del diritto (maggiore rilevanza socio-economico delle fattispecie in relazione alle quali uno dei modelli è chiamato a operare)<sup>58</sup>, è opportuno verificare se, in realtà, i danni cagionati per il tramite di una IA possano sussumersi sotto l'ombrello regolamentare di un altro paradigma normativo di responsabilità. D'altronde, come attenta dottrina nota<sup>59</sup>, «la forma di responsabilità in effetti applicata dipende [...] dalle modalità tramite le quali si combinano sul piano operativo gli ingredienti a quella regola essenziali».

A séguito della seconda rivoluzione industriale del secolo scorso, la supremazia del dogma «nessuna responsabilità senza colpa» iniziò a vacillare di fronte alla massificazione dei danni<sup>60</sup> e presero vita due tendenze ancorá attuali. La prima è volta a garantire l'indennizzo di certi danni di piú spiccata rilevanza sociale mediante assicurazioni obbligatorie<sup>61</sup>. La seconda tendenza è volta a sancire in piú ampia misura regole di responsabilità aggravata o, senz'altro, di responsabilità oggettiva a carico di soggetti che creano situazioni di pericolosità per i terzi<sup>62</sup>.

---

<sup>58</sup> M. BUSSANI, *L'illecito civile*, cit., p. 49. La questione è ben approfondita da S. RODOTÁ, *Il problema della responsabilità civile*, cit., p.158.

<sup>59</sup> M. BUSSANI, *L'illecito civile*, cit., p. 49. In argomento, per considerazioni di ordine generale cfr. V. ROPPO, *Spigolature di responsabilità oggettiva e funzione deterrente*, in *Liber Amicorum per F.D. Busnelli*, II, Milano, 2008, p. 453, 460- 463, 468-469; C. SALVI, *La responsabilità civile*, cit., p. 153 ss.

<sup>60</sup> La massificazione dei danni è data, precisamente, dal moltiplicarsi delle cause di pregiudizio a séguito del moltiplicarsi di prestazioni, prodotti e attività fatalmente rischiose, che coinvolgono il grosso pubblico.

A fronte dei danni di massa la regola della illiceità per colpa è risultata inadeguata in quanto le vittime rimangono normalmente indifese già per la difficoltà di provare la «colpa» di un responsabile. Da qui l'esigenza di una piú intensa tutela di persone e cose nella vita di relazione.

<sup>61</sup> Significativo, al riguardo, è il settore degli infortuni sul lavoro.

<sup>62</sup> L'incidenza del progresso tecnologico sulla responsabilità civile è rilevata da F. DI CIOMMO, *Evoluzione tecnologica e regole di responsabilità civile*, Napoli, 2003, p. 161, che ravvisa una perdita di centralità della norma generale dell'art. 2043.

Nell'ambito di questa tendenza una moderna corrente di pensiero ha contrapposto alla concezione etica della responsabilità extracontrattuale la concezione che ravvisa nella responsabilità extracontrattuale una tecnica di allocazione dei danni. Nell'ambito di questa concezione possono poi distinguersi due filoni. I dati che caratterizzano la concezione oggettiva sono il ripudio della colpa come fondamento della responsabilità civile e il rigetto della nozione di illecito quale fatto antiggiuridico.

Secondo la concezione della responsabilità extracontrattuale quale tecnica di allocazione dei danni la responsabilità extracontrattuale non deve intendersi come un sistema repressivo che colpisce un atto vietato ma come uno strumento di riequilibrio economico del danno: compito della responsabilità extracontrattuale non è quello di punire i colpevoli ma di far sí che i danni siano riparati. Secondo questa concezione vari sono i criteri mediante i quali deve procedere l'allocazione dei danni ma sono tutti criteri di obiettiva imputazione del danno che prescindono dalla violazione di un dovere. La colpa, si ammette, può avere rilevanza come uno dei criteri di imputazione del danno ma si tratterebbe di un criterio tecnico che prescinde dal fatto e che riguarderebbe una fascia ristretta di casi. La colpa avrebbe quindi ormai un'importanza marginale, destinata a sparire. Cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 524-526.

Per l'appunto, alcuni studiosi<sup>63</sup>, ritengono più appropriato all'ipotesi di nostro interesse applicare il paradigma della responsabilità da attività pericolosa<sup>64</sup>.

L'art. 2050 c.c. prevede che «chiunque cagiona un danno ad altri nell'esercizio di un'attività pericolosa, per sua natura o per la natura dei mezzi adoperati, è tenuto al risarcimento, se non prova di aver adottato tutte le misure idonee ad evitare il danno».

La norma, che al momento dell'entrata in vigore era priva di precedenti nei sistemi codicistici, costituisce la generalizzazione di fattispecie tipiche di responsabilità che erano state introdotte dalla legislazione speciale – in particolare da quella sulla circolazione dei veicoli – per far fronte alle conseguenze di attività pericolose non riconducibili alle tradizionali figure di danno da cosa o da animale. Il risultato conseguito è stato ampiamente criticato dalla dottrina sia per la formulazione tecnica della prova liberatoria, sia per la mancata adozione di una chiara regola di responsabilità oggettiva, sia infine per la carenza di coordinamento con le altre ipotesi di responsabilità diverse dalla colpa e in particolare con gli artt. 2051 e 2052 c.c.<sup>65</sup>.

In particolare, della responsabilità da attività pericolosa si contesta la natura giuridica. Alcuni studiosi, così come il diritto vivente, ne sostengono la natura di responsabilità oggettiva<sup>66</sup>. Di contrario avviso, invece, il filone prevalente – soprattutto in dottrina<sup>67</sup> – secondo cui detta responsabilità non si emancipa dalla imputazione per colpa. L'effetto risarcitorio, infatti, non si ricollegerebbe automaticamente alla creazione di un pericolo, essendo previsto l'esonero attraverso la prova di un elemento che attiene sempre le modalità di svolgimento dell'attività pericolosa ossia «tutte le misure idonee a evitare il danno». La norma perciò, come definito da autorevole dottrina, dimostrerebbe «l'incapacità del legislatore italiano di superare il muro della colpa»<sup>68</sup>.

---

<sup>63</sup> U. RUFFOLO, *Intelligenza artificiale e responsabilità. Responsabilità da "algoritmo"? A.I. e automobili self-driving, automazione produttiva, robotizzazione medico-farmaceutica. A.I. e attività contrattuali. Le tendenze e discipline unionali*. Convegno del 29 novembre 2017, Università degli stranieri di Perugia, Milano, 2017; ID., *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 2019, p. 1689 ss.; ID., *Responsabilità da produzione e gestione di A.I. self-Learning*, in S. GIOVA e I. PRISCO (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*, Atti del 15° Convegno Nazionale Sisdic, Napoli 2020.

<sup>64</sup> Sulla disciplina della responsabilità da attività pericolosa in generale si v. M. FRANZONI, *Responsabilità per l'esercizio di attività pericolose*, in G. ALPA e M. BESSONE (a cura di), *La responsabilità civile*, cit.; M. COMPORI, *Responsabilità per l'esercizio di attività pericolose*, in *Comm. c.c.* Busnelli, Milano, 2009; E. AL MUREDEN, *La responsabilità per esercizio di attività pericolose a quarant'anni dal caso Seveso*, in *Contr. impr.*, 2016, 3, p. 647 ss. La dottrina nell'analizzare la responsabilità derivante dall'utilizzo dell'intelligenza artificiale in campo medico, ha fatto espreso riferimento a questa norma. Sul punto si veda AR. FUSARO, *Attività pericolose e dintorni*, in *Riv. dir. civ.*, 2013, 6, p. 1337 ss.

<sup>65</sup> Cfr. C. SALVI, *La responsabilità civile*, cit., p. 184.

<sup>66</sup> Fra molti si v. P. TRIMARCHI, *Rischio e responsabilità oggettiva*, cit.; M. COMPORI, *Esposizione al pericolo*, cit.; ID., *Fatti illeciti: le responsabilità oggettive*, in *Cod. civ. Comm.* Schlesinger Busnelli, Milano, 2009, p. 156.

In senso conforme v. Cass., 2 marzo 2011, n. 3022, che afferma la «presunzione di responsabilità» nel caso di adozione di «misure diverse da quelle prescritte da norme legislative o regolamentari senza che vi sia alcuna possibilità di valutarne l'idoneità»; mentre può scagionare l'autore dell'attività pericolosa solo la prova dei fatti che non rientrano tra le dette misure ma ammontano a caso fortuito o a una causa esterna.

<sup>67</sup> Cfr. C. SALVI, *La responsabilità civile*, p. 185. In senso conforme C. CASTRONOVO, *La responsabilità civile*, Milano, 2018, p. 454.

<sup>68</sup> C. CASTRONOVO, *o.u.c.*, p. 458.

Ad ogni buon conto, l'applicazione dell'art. 2050 c.c. richiede, innanzitutto, che il fatto dannoso sia inserito in una «attività», il che presuppone un minimo di continuità ed organizzazione dei mezzi. L'atto isolato benché estremamente pericoloso sembrerebbe meglio sussumibile sotto l'art. 2043 c.c.<sup>69</sup> L'utilizzo del termine «attività» porterebbe ad escludere che il danno debba derivare da un comportamento personale del responsabile, potendo l'attività essere svolta da altri soggetti, benché sotto il suo controllo<sup>70</sup>.

L'attività deve, poi, essere «pericolosa»<sup>71</sup>. Il legislatore non ha tipizzato tutte le attività pericolose. Conseguentemente la valutazione è rimessa, di volta in volta, al giudice di merito<sup>72</sup>. La giurisprudenza, di base, definisce come pericolose quelle attività che per la loro stessa natura o per le caratteristiche dei mezzi adoperati comportano la rilevante possibilità del verificarsi di un danno per la loro spiccata potenzialità lesiva. Il criterio da adottare per la valutazione della pericolosità è di tipo quantitativo o statistico, da accertare secondo una valutazione *ex ante*, diretta cioè ad accertare se una determinata attività aveva, al momento del suo esercizio, una rilevante probabilità di danno<sup>73</sup>. Ove si tratti di attività nuove, per le quali manchino dati statistici, «il giudizio di pericolosità potrà essere desunto dalla normale prevedibilità, basata sulle regole offerte dalla tecnica e su ogni altro utile elemento»<sup>74</sup>.

È di tutta evidenza, quindi, che *in vi* il cd. rischio da sviluppo, costituente limite esterno alla responsabilità per danno da prodotto (v. capitolo II), se particolarmente incidente, potrebbe costituire indice, se non addirittura prova, della pericolosità dell'attività connotata da tale rischio.

Infine, la norma presuppone che i danni siano derivati dallo «svolgimento dell'attività». Perciò, indispensabile requisito è la prova del nesso causale tra lo svolgimento dell'attività ed il danno. Il concetto di svolgimento di attività pericolosa non deve essere utilizzato per restringere le ipotesi di applicabilità dell'art. 2050 c.c., ma al solo fine di qualificare correttamente la natura dell'illecito. Di talché, il nesso si ritiene sussistente anche qualora il danno si

---

<sup>69</sup> Tutte le attività umane contengono in sé un grado più o meno elevato di pericolosità per coloro che direttamente le esercitano o per i terzi estranei. Ne consegue che, interpretando in senso ampio la locuzione «attività pericolose», l'art. 2043 c.c. diventerebbe completamente superfluo, in quanto inutile duplicato (in termini Cass., 21 ottobre 2005, n. 20359, in *Foropus.it*, 2005).

<sup>70</sup> C. SALVI, *u.o.c.*, cit., p. 185.

<sup>71</sup> È bene tenere a mente la differenza tra pericolosità della condotta e pericolosità dell'attività. La prima riguarda un'attività normalmente innocua, che assume i caratteri della pericolosità a causa della condotta imprudente o negligente dell'operatore, ed è elemento costitutivo della responsabilità ai sensi dell'art. 2043 c.c.; la seconda concerne un'attività che, invece, è potenzialmente dannosa di per sé per l'alta percentuale di danni che può provocare in ragione della sua natura o della tipologia dei mezzi adoperati e rappresenta una componente della responsabilità disciplinata dall'art. 2050 c.c. cfr. M. FRANZONI, *Dei fatti illeciti*, cit., p. 385-386.

<sup>72</sup> Per una casistica si v. M. FRANZONI, *u.o.c.*, cit., p. 388 ss.

<sup>73</sup> Così ritiene C. SALVI, *La responsabilità civile*, cit., p. 184. M. FRANZONI, *Dei fatti illeciti*, cit., p. 386, aggiunge che la valutazione *ex ante* dovrebbe essere immediatamente integrata da un criterio di tipo qualitativo, che tenga conto dell'entità e della gravità dei danni causati.

In senso contrario si esprime la giurisprudenza che parla di «prognosi postuma». Fra altre v. Cass., 20 maggio 2015, n. 10268, in *Foropus.it*, 2015.

<sup>74</sup> M. FRANZONI, *Dei fatti illeciti*, cit., p. 386; M. COMPORI, *Esposizione al pericolo*, cit., p. 291.

produca in una fase temporalmente successiva all'immediatezza dell'attività. Questo vale purchè il danno ne dipenda in modo «sufficientemente mediato».

Chi cagiona danno nello svolgimento di un'attività pericolosa può, poi, liberarsi dalla responsabilità solo provando «di avere adottato tutte le misure idonee a evitare il danno».

Secondo tale fattispecie, quindi, l'esercente l'attività potrebbe andare esente unicamente dimostrando di aver agito coerentemente alla diligenza professionale richiesta, per questa intendendosi l'adozione di tutte le cautele normalmente adeguate in termini di perizia e prudenza<sup>75</sup>. Una lettura esclusivamente letterale della norma non consentirebbe mai la prova liberatoria sol si consideri che l'adozione di davvero tutte le misure astrattamente possibili certamente eviterebbe il danno. Ma è concretamente impossibile stabilire a priori, in via assoluta, quali siano gli accorgimenti da tenere per evitare un danno, specie con riferimento ad attività nuove o innovative. Indi per cui, si starebbe di fronte ad una *probatio diabolica* a carico dell'esercente sol si consideri che «la totalità di ciò che non è quantitativamente determinato è in se inattuabile»<sup>76</sup>.

Così, la giurisprudenza maggioritaria interpreta la norma in virtù della sua funzione che è quella di fissare un limite di responsabilità a carico di chi esercita un'attività pericolosa. Questo limite è segnato dalla diligenza professionale<sup>77</sup>. Tuttavia, secondo l'interpretazione giurisprudenziale corrente, la prova è molto rigorosa. Di fatti, si ritiene che non si libera da responsabilità l'esercente perito, prudente e diligente, ma solo quello che dimostri che l'evento dannoso «non si sarebbe potuto evitare mediante l'adozione delle misure di prevenzione che le leggi dell'arte o la comune diligenza imponevano»<sup>78</sup>. E si aggiunge che «tra le misure idonee che l'esercente di attività pericolose deve adottare, per sottrarsi alla responsabilità, vanno comprese tutte le misure previste da leggi o regolamenti che disciplinano esercizio di quelle attività»<sup>79</sup>. «Ciò significa che l'aver omesso misure non espressamente previste per quello specifico settore di attività pericolose è considerato un comportamento di per sé riprovevole, idoneo a far insorgere la presunzione semplice che, avendone adottate altre, il danno non si sarebbe prodotto. È vero che all'esercente non è precluso il potere discrezionale di scelta sulle misure preventive, lasciando così intendere che non tutte debbano essere adottate. Ciononostante, qualora il danno accada per l'omissione di cautele prescritte da una norma legislativa o regolamentare, l'agente non può essere esonerato dalla responsabilità in questione, dovendo ritenersi che, se fossero state adottate, il danno non si sarebbe verificato»<sup>80</sup>. Non vi è chi non veda come la prova sia così rigorosa da essere nella sostanza

---

<sup>75</sup> Cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 682.

<sup>76</sup> C. CASTRONOVO, *La responsabilità civile*, cit., p. 461.

<sup>77</sup> V. la Relazione al Re (n. 795) che, nell'attribuire anche in questo paradigma di responsabilità un ruolo alla colpa benchè in sede di prova liberatoria, ritiene che qui il dovere di diligenza posto a carico dell'esercente sia ampliato.

<sup>78</sup> Sul punto Cass., 5 luglio 2017, n. 16637, in *ForoPlus.it*, 2017. Cfr. anche M. FRANZONI, *Dei fatti illeciti*, cit., p. 399-400.

<sup>79</sup> Cass., 26 novembre 1977, n. 5155, in *ForoPlus.it*, 1977.

<sup>80</sup> Cfr. M. FRANZONI, *o.n.c.*, cit., p. 400.

equiparabile al caso fortuito. Invero, «l'esonero di responsabilità potrebbe aversi solo se dalla prova addotta possano ricavarsi «elementi presuntivi» circa l'identificazione di una causa non imputabile che abbia reso oggettivamente impossibile l'adempimento del dovere di diligenza»<sup>81</sup>.

E si badi, la valutazione delle idoneità delle misure preventive non deve essere valutata in riferimento allo specifico atto che nella catena delle attività ha determinato il danno – differenza importante con la responsabilità ex art. 2043 c.c. – quanto, invece, alle complessive modalità di organizzazione dell'attività.

Orbene, considerato quanto sopra, può riflettersi ora sull'applicazione dell'art. 2050 c.c. con riferimento all'IA.

Certamente, dovendo il danneggiato, in questo caso provare unicamente il danno ed il nesso causale tra questo e lo svolgimento dell'attività, il suo onere probatorio sembra sensibilmente alleggerito rispetto all'utilizzo del paradigma dell'imputazione per colpa ex art. 2043 c.c. D'altro canto, rappresentando l'art. 2050 c.c. una responsabilità «intermedia» – almeno accogliendo l'interpretazione prevalente – non si escluderebbe nemmeno un ancoraggio soggettivo ai fini dell'imputabilità, ritrovando la colpa nella prova liberatoria sotto le spoglie della diligenza tecnica. Probabilmente, tale paradigma è quello da ritenersi preferibile.

Eppure, anche l'art. 2050 c.c. scricchiola se proiettato su uno scenario in cui è protagonista un'IA.

Innanzitutto, si dubita della opportunità di accostare l'utilizzo di un'IA, seppur autonoma, ad una attività di spiccata potenzialità lesiva per natura o per proprie caratteristiche. D'altronde, si tratta pur sempre di una funzione logico-matematica a dati variabili. Di talché, risulta difficile pensare che sia innatamente pericolosa<sup>82</sup>. L'IA è pur sempre simbolo della tecnica e, per ciò stesso, resta più affidabile dell'uomo, rappresentando, d'altronde, un mezzo correttivo o integrativo delle imprecisioni umane<sup>83</sup>. «Al contrario, potrebbe accadere che la pericolosità non concerna l'attività svolta dai sistemi intelligenti in sé, quanto, piuttosto, le interazioni del sistema con il mondo esterno. Pertanto potrebbe succedere che un'attività produttiva, ritenuta fino ad oggi non pericolosa, lo divenga con l'ingresso dell'intelligenza artificiale nel processo di produzione o direttamente nel prodotto»<sup>84</sup>.

In secondo luogo, tale meccanismo non appare pienamente soddisfacente in punto di prova liberatoria. Presupponendo, infatti, che l'esercente l'attività si possa liberare solo dimostrando di aver adottato tutte le misure idonee ad evitare

---

<sup>81</sup> C. SALVI, *La responsabilità civile*, cit., p. 190.

<sup>82</sup> Sul tema si consenta il rinvio a M.M. MOLLICONE, *Il rischio dell'intelligenza artificiale*, cit., p. 2108.

Di avviso contrario a ritenere l'applicazione dell'intelligenza artificiale pericolosa è M. COSTANZA, *L'intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Giur. it.*, 2019, VII, p. 1688: «all'intelligenza artificiale come emblema della tecnica e perciò come ente ritenuto più affidabile dell'uomo non si addice allora l'attributo di pericolosa. [...] La intelligenza artificiale come mezzo correttivo o integrativo delle umane carenze non tollerebbe alcun attributo che la qualifici come rischiosa; anzi, la intelligenza artificiale sarebbe un ente non pericoloso perché capace di evitare gli inconvenienti che senza il suo intervento possono generarsi con lo svolgimento di certe attività».

<sup>83</sup> G. D'ALFONSO, *Il regime di responsabilità da cose in custodia tra questioni tradizionali e "responsabilità da algoritmo"*, in *EJPLT*, 2022, I, p. 1 ss.

<sup>84</sup> *Ibidem*.

il danno, di fatto la prova sembra difficilmente raggiungibile sol si consideri la totale imprevedibilità dell'*output* e, quindi, l'impossibilità di conoscere quali siano le misure idonee ad evitare l'imprevedibile. Sicuramente, soccorrerebbero nella valutazione della idoneità delle misure «la precisione della stesura degli obiettivi da raggiungere con l'impiego dell'algoritmo nella procedura, la purezza dei dati inseriti e la considerazione di un approccio etico allo sviluppo dei sistemi di IA da parte dei programmatori»<sup>85</sup>. Ma ciò non sembra essere sufficiente ad una piena e consapevole valutazione dei rischi *ex ante* e, quindi, alla predisposizione di misure preventive in quanto non esistono ancora abbastanza dati statistici per sintetizzare l'ignoto tecnologico.

Infine, non è scontato che il danno derivante dall'utilizzo-applicazione di un'IA intervenga nell'esercizio di un'attività. Ciò accade il più delle volte ma ciò non toglie che non possa avvenire diversamente. Ebbene, in quel remoto caso la responsabilità da attività pericolose sembrerebbe completamente inadatta.

### 2.3. La responsabilità da cose in custodia ex art. 2051 c.c.

Per i motivi suesposti, vi è chi<sup>86</sup> prende in considerazione il fatto di applicare il paradigma della responsabilità da cose in custodia ex art. 2051 c.c.<sup>87</sup>

L'art. 2051 c.c. regola la responsabilità civile del custode della cosa per i danni da questa provocati, statuendo che «ciascuno è responsabile del danno cagionato dalle cose che ha in custodia, salvo che non provi il caso fortuito»<sup>88</sup>.

Sia rispetto al criterio di imputazione per colpa (art. 2043 c.c.) che a quello per esercizio dell'attività pericolosa (art. 2050 c.c.), il danneggiato, che richieda la tutela giurisdizionale ex art. 2051 c.c., ha accesso ad un «regime probatorio semplificato», in quanto legittimato a richiedere il risarcimento del danno patito, in ragione del mero rapporto sussistente tra la *res* ed il soggetto che lo custodisca.

I presupposti richiesti per l'insorgenza della responsabilità sono tre.

---

<sup>85</sup> Così afferma M. FERRARI, *Il vantaggio della responsabilità concorsuale da uso "organizzato" di algoritmi*, in *Il Mulino*, 2021, II, p. 424 la quale riprende G. FASANO, *L'intelligenza artificiale nella cura dell'interesse generale*, in *Giornale dir. amm.*, 2020, VI, p. 724 che sottolinea l'importanza del rapporto fra neutralità dell'algoritmo e imparzialità dei suoi risultati: «La neutralità dell'algoritmo (rappresentata dal fatto che il meccanismo di funzionamento non parteggia per nessuna delle possibili scelte di *output*) non può tradursi in una scelta imparziale se la sua fase di progettazione non è governata da criteri improntati alla massima imparzialità», perciò l'imparzialità si anniderebbe nelle scelte umane del programmatore.

<sup>86</sup> U. RUFFOLO, *Responsabilità da produzione e gestione di ai self Learning*, in ID., S. GIOVA e I. PRISCO (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*, Atti del 15° Convegno Nazionale Sisidic, Napoli 2020, p. 233 ss.; M. RATTI, *Riflessioni in materia di responsabilità civile e danno cagionato da dispositivo intelligente alla luce dell'attuale scenario normativo*, in *Contr. impr.*, 2020, III, p. 1174-1191; A. SANTOSUOSSO, C. BOSCARATO e F. CAROLEO, *Robot e diritto. Una prima ricognizione*, in *Nuova giur. civ. comm.*, 2012, II, p. 494-516, i quali sottolineano come, accedendo alla tesi secondo cui i robot sono meri oggetti, l'art. 2051 c.c. sarebbe l'unica norma certamente applicabile al caso dei danni da questi prodotti.

In senso contrario a questa potenziale applicazione dell'art. 2051 c.c. è M. COSTANZA, *L'intelligenza*, cit., p. 1687, la quale ritiene che la norma si riferisca unicamente alle cose inanimate. Di talché, la logica codicistica sottesa sarebbe esageratamente semplice per riferirsi alla IA.

<sup>87</sup> In generale sulla responsabilità per danno da cose in custodia v. tra molti L. CORSARO, *Responsabilità da cose*, in *Dig. disc. priv., Sez. civ.*, XVII, 1998, Torino, p. 103 ss; M. FRANZONI, *La responsabilità oggettiva. Il danno da cose e da animali*, Padova, 1988.

<sup>88</sup> G. D'ALFONSO, *Il regime di responsabilità da cose in custodia*, cit., p. 1 ss.

Innanzitutto, occorre che ad aver causato il danno sia stata una cosa<sup>89</sup>. A differenza di quanto previsto dall'art. 2050 c.c., le cose non devono essere di una particolare natura, o meglio intrinsecamente pericolose. Quindi si comprendono cose inerti e cose in movimento<sup>90</sup>, cose mobili ed immobili<sup>91</sup>, liquidi, solidi ed universalità di mobili<sup>92</sup>. Nonostante queste precisazioni, «secondo un orientamento largamente seguito, il danno deve essere causato dal dinamismo connaturato alla cosa o dall'insorgere in essa di un agente dannoso; ed è su questa premessa che si fonda l'opinione per la quale vi sarebbe un criterio di imputazione, comune a tutte le ipotesi di responsabilità oggettiva, fondato sul concetto di "cose seagenti"<sup>93</sup>. Ma il ragionamento non è convincente giacché attribuisce pregio alla distinzione tra cose pericolose e cose non pericolose, distinzione che, invece, è priva di fondamento, poiché la pericolosità non è una caratteristica immanente della cosa, ma una caratteristica che possono assumere tutte le cose in date circostanze di fatto»<sup>94</sup>.

In secondo luogo, deve sussistere relazione qualificata con la *res* ossia un «rapporto di custodia»<sup>95</sup>. Tale rapporto si concretizza nell'esercizio di un «potere di governo» che il custode esercita sulla cosa e che gli consente di controllarla, di intervenire sulla stessa, al fine di eliminare o quantomeno attenuare eventuali situazioni di pericolo e di escludere l'ingerenza di terzi. La norma non implica, perciò, neppure indirettamente uno specifico obbligo di custodire la cosa quale, ad esempio, quello previsto nel contratto di deposito<sup>96</sup>. La norma ha la funzione di imputare la responsabilità a chi si trovi nelle condizioni di controllare i rischi inerenti alla cosa<sup>97</sup>. Quindi, è custode il soggetto che di fatto controlli le modalità di uso e di conservazione della cosa<sup>98</sup>, indipendentemente dal fatto che questi sia il proprietario, possessore o detentore<sup>99</sup>.

In terzo luogo, deve esistere il nesso di causalità tra la cosa e l'evento dannoso.

Sul danneggiato incombe, quindi, l'onere di provare la derivazione del danno dal fatto della cosa in custodia. In altri termini, egli deve dimostrare l'esistenza

---

<sup>89</sup> E non l'uomo per mezzo della stessa.

<sup>90</sup> Fra tante si v. Cass., 11 giugno 1998, n. 5796, in *Foropius.it*.

<sup>91</sup> Per approfondimenti si v. V. GERI, *Responsabilità civile da cose in custodia animali e rovina di edificio*, Milano, 1974, p. 102 ss., il quale afferma che la categoria delle cose inanimate e comprensiva di tutte le cose inerti virgola in movimento, pericolose, non pericolose ecc. Anche F. MESSINEO, *Manuale di diritto civile e commerciale*, V, 1947, p. 586 che confronta l'art. 2051 e l'art. 2052, il quale ultimo vede proprio una figura di responsabilità derivante da una cosa animata.

<sup>92</sup> Cfr. Cass., 8 aprile 1997, n. 3041, in *Vita notar.*, 1997, p. 868, con nota di F. PIZZETTI, *Nuovi profili della responsabilità per danno da cose in custodia ex art. 2051 c.c.*, in *Arch. Civ.*, 1997, p. 1108.

<sup>93</sup> Cfr. C. SALVI, *La responsabilità civile*, cit., p. 176; D. DE MARTINI, *I fatti produttivi di danno risarcibile*, Padova, 1983, p. 196 ss.

<sup>94</sup> M. FRANZONI, *Dei fatti illeciti*, cit., p. 408-409.

<sup>95</sup> R. SCOGNAMIGLIO, voce «Responsabilità civile», cit., p. 644; E. BONVICINI, *La responsabilità civile per fatto altrui*, Milano, 1976, p. 282; C. SALVI, voce «Responsabilità extracontrattuale» (dir.vig.), in *Enc.dir.*, XXXIX, Milano, 1988, p. 1227.

<sup>96</sup> C. SALVI, *La responsabilità civile*, cit., p. 177.

<sup>97</sup> Richiamando TRIMARCHI, *Rischio e responsabilità oggettiva*, cit. p. 244.

<sup>98</sup> Cass., 1 febbraio 2018, n. 2482, in *foropius.it*.

<sup>99</sup> Già G. BRANCA, *Profili storici della responsabilità per cose in custodia nel nuovo codice delle obbligazioni*, in *Resp. civ.*, 1941, p. 169.

del rapporto di custodia tra il danneggiante e la cosa e il nesso causale tra la cosa in custodia e l'evento lesivo.

Dall'altro lato, il custode, per andare esente da responsabilità, dovrà provare l'esistenza del «caso fortuito» idoneo ad elidere tale rapporto causale<sup>100</sup>. Esso deve intendersi in senso ampio, dunque, comprensivo del fatto naturale, del fatto del terzo, del fatto del danneggiato, del *factum principis*. In altre parole, «per “caso fortuito” deve intendersi tutto ciò che costituisca un'eccezione alla normale sequenza causale e che, ricondotto ad una valutazione ex ante ovvero in astratto, non sia oggettivamente prevedibile, con il risultato che all'imprevedibilità consegua l'inevitabilità dell'evento»<sup>101</sup>.

Orbene, riportando il tutto alla tematica dell'IA, non parrebbe per molti astruso utilizzare il paradigma di cui si è sin qui parlato sol si consideri che, confrontandolo con l'art. 2050 c.c., il regime della responsabilità della cosa in custodia ha il pregio di prescindere dalla prova della pericolosità dell'attività.

Si riconoscerebbe, quindi, la responsabilità in capo, non solo all'utilizzatore dell'IA o al gestore, ma anche all'addestratore (colui che più propriamente immette i dati) dato che è proprio costui che ha il «potere di governo», ossia che espone la IA a nuove «esperienze» «alimentando» l'algoritmo e, così, influenzando continuamente il suo comportamento.

Di fatti, secondo tale visione, le macchine e i sistemi basati sull'IA sarebbero tutti indubbiamente cose<sup>102</sup>.

Il *facere* della IA, sia simbolica-logico-semanticamente, sia sub- simbolica- biologico-empirica, rientrerebbe nel novero delle *res*, sviluppandosi, attualmente, in uno *status* di «inconsapevolezza». In altre parole, la macchina funzionerebbe ma non si renderebbe conto di funzionare.

Sarebbe, quindi, solo un «bene intelligente». E proprio tale locuzione suggerirebbe ancor di più l'inclusione di tale fenomeno nel novero dell'art. 2051 c.c..

Trattasi, invero, questa di previsione simmetrica all'art. 2052 c.c. che detta una disciplina identica per la responsabilità per fatto dell'animale che, in altre parole, potrebbe riformularsi in responsabilità per fatto dell'intelligenza naturale animale, anch'essa capace di autoapprendere.

---

<sup>100</sup> L'interpretazione del caso fortuito ha dato luogo a non poche controversie dal momento che esso rappresenta il punto nodale del dibattito concernente la natura della responsabilità da cose in custodia. La prevalente dottrina e giurisprudenza parla della responsabilità da cose in custodia come una responsabilità oggettiva. La pronuncia che ha dato avvio alla moderna concezione dell'art. 2051 c.c. è la sentenza Cass., Sez.un., 11 novembre 1991, n. 12019, in *Nuova giur. civ. comm.*, 1992, p. 81, con nota di G. ALPA, *Responsabilità per rischio da cose in custodia*.

Di avviso contrario, sostenendo la natura di responsabilità aggravata ed il caso fortuito come elemento soggettivo connesso alla colpa, tra altri, cfr. S. PATTI, *Prove*, in *Comm. c.c. Scialoja-Branca-Galgano-De Nova*, sub art. 2697 c.c., Bologna, 2015, pp. 163 ss.

<sup>101</sup> G. D'ALFONSO, *Il regime di responsabilità da cose in custodia*, cit., p. 7.

<sup>102</sup> A. BERTOLINI, *Artificial intelligence does not exist! Defying the technology-neutrality narrative in the regulation of civil liability for advanced technologies*, in *Europa dir. priv.*, 2022, II, p. 370 afferma che non solo è inutile sanzionare qualcosa che è oggettivamente incapace di temere la sanzione stessa - perché non riuscirebbe comunque a indurre il rispetto delle regole -, ma è anche inutile, perché alla fine le conseguenze legali saranno a carico dell'essere umano che sta dietro, chiamato a fornire i mezzi necessari.

Di talché, non sembrerebbe troppo difficile inquadrare nell'art. 2051 c.c. la responsabilità per il fatto della cosa intelligente e quindi della intelligenza artificiale.

A parere di chi scrive, tuttavia, detto paradigma non sembra il migliore.

Innanzitutto, appare assai rigido sol si consideri la talmente stretta prova liberatoria del caso fortuito tale da essere motivo di disincentivazione dell'innovazione<sup>103</sup>.

Inoltre, «il parametro della custodia potrebbe risultare inidoneo, in considerazione della circostanza che la custodia di un dispositivo intelligente, soprattutto se autonomo, sarebbe eccessivamente complessa e squilibrata per il custode che potrebbe non essere in grado di controllarlo»<sup>104</sup>.

In aggiunta, il potere di governo sulle «cose intelligenti» non sarebbe riscontrabile nei cd. «sistemi chiusi» ossia quei sistemi in cui il produttore dell'hardware non permetta l'impiego di software di aziende che non siano specificamente autorizzate<sup>105</sup>. Di talché, in tali ipotesi, il titolare-utilizzatore-custode non avrebbe alcun controllo sulle condizioni di rischio del dispositivo.

Ne discenderebbe che, nel sistema chiuso, l'eventuale alleggerimento di responsabilità di tali soggetti «potrebbe essere compensata dall'aggravamento, in termini di responsabilità, della posizione del produttore del dispositivo, perché si tratta del soggetto idoneo, in maggior misura, a minimizzare il rischio: costui è, infatti, in grado di controllarlo integralmente e di gestirlo, anche per mezzo di strumenti assicurativi oppure traslando sul prezzo del bene il costo del rischio»<sup>106</sup>.

### 3. Il paradigma previsto dalla proposta di direttiva *AI liability* (AILD).

Il problema della responsabilità nell'universo dell'IA, come accennato, non è stato percepito come rilevantissimo unicamente dall'ordinamento domestico, bensì anche dall'Unione europea che, pressoché contemporaneamente con l'insorgenza del problema, si è proposta di divenire un *leader* mondiale, sia nello sviluppo di sistemi di IA sicuri e affidabili<sup>107</sup> che nella regolamentazione del loro intero corso di «vita» – fisiologico e patologico –, per mezzo di un importante *iter* regolamentare.

---

<sup>103</sup> Cfr. L. FORT e V. IEVA, *Intelligenza artificiale, responsabilità civile e interpretazione analogica*, in *www.biodiritto.org*, 2020, VIII, p. 2.

<sup>104</sup> G. D'ALFONSO, *Il regime di responsabilità da cose in custodia*, cit., p. 25.

<sup>105</sup> Di contro, la disciplina della responsabilità da cosa in custodia sarebbe applicabile, quando ci si trovi di fronte ad un sistema aperto, nel quale l'utilizzatore, il custode o il titolare del prodotto eserciteranno un certo grado di controllo e potranno modificare le condizioni di rischio, accedendo a servizi o prodotti software che siano offerti da altri imprenditori.

<sup>106</sup> *Ibidem*.

<sup>107</sup> Considerando n. 5 della proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (COM 2021/206) nel suo ultimo testo di compromesso interistituzionale approvato definitivamente il 2 febbraio 2024 (*AI Act*). L'ultimo trilogo del 6 dicembre 2023 ha, dopo ampi dibattiti, portato ad un compromesso sui «*Foundation Models*», sull'IA «*General Purpose*» e su obbligazioni più stringenti in tema di *copyright*.

Sull'affidabilità dell'IA come obiettivo di lungo termine v. HLEG AI, *Ethics Guidelines for Trustworthy AI*, 2019; J. LAUX, S. WACHTER e B. MITTELSTADT, *Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk*, in *ssrn.com*, 2022,

Infatti, per promuovere la diffusione di un'IA antropocentrica, affidabile ed etica<sup>108</sup> e garantire un livello elevato di protezione della salute, della sicurezza, dei diritti fondamentali, della democrazia e dello Stato di diritto, nonché dell'ambiente, dagli effetti nocivi dei sistemi di IA nell'Unione, sostenendo al contempo l'innovazione e migliorando il funzionamento del mercato interno, l'Unione europea ha ritenuto necessario predisporre un quadro giuridico uniforme che stabilisca norme armonizzate sull'IA e che allo stesso tempo soddisfi un elevato livello di protezione degli interessi pubblici.

Per raggiungere tale scopo, è stato ritenuto opportuno, da parte della Commissione europea, predisporre un pacchetto di riforme<sup>109</sup>.

Da una parte, uno schema di regolamento, il cd. *AI Act*<sup>110</sup>, diretto a stabilire norme che regolino l'immissione sul mercato, la messa in servizio e l'uso di determinati sistemi di IA, garantendo così il buon funzionamento del mercato interno e consentendo a tali sistemi di beneficiare del principio della libera circolazione delle merci e dei servizi. Il Regolamento proposto è volto ad introdurre nuovi *standards*, limitandosi a fissare regole armonizzate per lo sviluppo, l'immissione sul mercato e l'utilizzo di sistemi di IA nell'Unione, contribuendo, così, all'obiettivo di promuovere l'IA «*Made in Europe*», seguendo un approccio orizzontale proporzionato, basato sul rischio, differenziando tra gli usi dell'IA che creano: a) un rischio inaccettabile<sup>111</sup>; b) un rischio alto c) un rischio bassissimo o minimo.

---

<sup>108</sup> Così è stato richiesto dal Parlamento europeo con la sua Risoluzione del 20 ottobre 2020 recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate (2020/2012(INL)).

<sup>109</sup> Le riforme di cui si parla sono in linea con gli obiettivi del Libro Bianco sull'Intelligenza Artificiale (COM(2020) 65 definitivo), in cui la Commissione si è impegnata a promuovere l'adozione dell'Intelligenza Artificiale e ad affrontare i rischi associati ad alcuni dei suoi usi

<sup>110</sup> Proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (COM 2021/206).

<sup>111</sup> L'art 5 *AI Act* dichiara vietate le pratiche che riguardano (a) l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che impiega tecniche subliminali al di là della coscienza di una persona o tecniche manipolative o ingannevoli, con l'obiettivo o l'effetto di distorcere materialmente il comportamento di una persona o di un gruppo di persone, compromettendo in modo significativo la capacità della persona di prendere una decisione informata, inducendo così la persona a prendere una decisione che non avrebbe altrimenti preso, in un modo che causa o è probabile che causi a quella persona, a un'altra persona o a un gruppo di persone un danno significativo; (b) l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che sfrutta una qualsiasi vulnerabilità di una persona o di un gruppo specifico di persone a causa della loro età, della loro disabilità o di una specifica situazione sociale o economica, con l'obiettivo o l'effetto di distorcere materialmente il comportamento di quella persona o di una persona appartenente a quel gruppo in un modo che causa o è ragionevolmente probabile che causi a quella persona o a un'altra persona un danno significativo; (ba) l'immissione sul mercato o la messa in servizio per questo scopo specifico, o l'uso di sistemi di categorizzazione biometrica che categorizzano individualmente le persone fisiche sulla base dei loro dati biometrici per dedurre o inferire la loro razza, le opinioni politiche, l'appartenenza sindacale, le convinzioni religiose o filosofiche, la vita sessuale o l'orientamento sessuale. Questo divieto non copre l'etichettatura o il filtraggio di insiemi di dati biometrici acquisiti legalmente, come le immagini, sulla base di dati biometrici o la categorizzazione di dati biometrici nell'ambito dell'applicazione della legge; (c) l'immissione sul mercato, la messa in servizio o l'uso di sistemi di IA per la valutazione o la classificazione di persone fisiche o di gruppi di esse per un certo periodo di tempo, sulla base del loro comportamento sociale o di caratteristiche personali o di personalità conosciute, dedotte o previste, con il punteggio sociale che porta a uno o a entrambi i seguenti elementi: (i) trattamento svantaggioso o sfavorevole di determinate persone fisiche o di interi gruppi di esse in contesti sociali che non sono correlati ai contesti in cui i dati sono stati originariamente generati o raccolti; (ii) trattamento

Dall'altra parte, due schemi di direttive. L'una – di cui si è già trattato al capitolo II – che offre una revisione delle norme settoriali e orizzontali in materia di sicurezza dei prodotti (cd. *PLD*, v. *supra*)<sup>112</sup>; l'altra che predispose norme di livello europeo per affrontare le questioni in materia di responsabilità relative ai sistemi di IA (cd. *AILD*, *artificial intelligence liability directive*)<sup>113</sup>.

Proprio su quest'ultima deve soffermarsi ora la nostra attenzione.

Ebbene, l'Unione europea si è concentrata sulla risoluzione del problema della responsabilità da applicazione di IA, consapevole che un mancato suo intervento comporterebbe un'ulteriore frammentazione normativa e, conseguentemente, un aumento dei costi per le imprese attive in tutta l'Unione in quanto, riscontrando difficoltà nel prevedere le modalità di applicazione delle norme vigenti in materia di responsabilità, si mostrerebbero incapaci di valutare la loro esposizione alla responsabilità e a stipulare la relativa copertura assicurativa<sup>114</sup>.

---

svantaggioso o sfavorevole di determinate persone fisiche o di gruppi di esse che è ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla loro gravità; (d) l'uso di sistemi di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico ai fini dell'applicazione della legge, a meno che e nella misura in cui tale uso sia strettamente necessario per uno dei seguenti obiettivi: (i) la ricerca mirata di vittime specifiche di rapimento, traffico di esseri umani e sfruttamento sessuale di esseri umani, nonché la ricerca di persone scomparse; (ii) la prevenzione di una minaccia specifica, sostanziale e imminente alla vita o all'incolumità fisica di persone fisiche o di una minaccia reale e presente 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 della conduzione di un'indagine penale, dell'esercizio dell'azione penale o dell'esecuzione di una sanzione penale per i reati di cui all'Allegato IIa e punibili nello Stato membro interessato con una pena detentiva o una misura di sicurezza per un periodo massimo di almeno quattro anni. Questo paragrafo non pregiudica le disposizioni di cui all'Articolo 9 del GDPR per il trattamento dei dati biometrici per scopi diversi dall'applicazione della legge; (da) l'immissione sul mercato, la messa in servizio per questo scopo specifico, o l'uso di un sistema di IA per effettuare valutazioni del rischio di persone fisiche al fine di valutare o prevedere il rischio di una persona fisica di commettere un reato penale, basandosi esclusivamente sul profilo di una persona fisica o sulla valutazione dei tratti e delle caratteristiche della sua personalità. Questo divieto non si applica ai sistemi di IA utilizzati per supportare la valutazione umana del coinvolgimento di una persona in un'attività criminale, che è già basata su fatti oggettivi e verificabili direttamente collegati a un'attività criminale; (db) l'immissione sul mercato, la messa in servizio per questo scopo specifico o l'utilizzo di sistemi di IA che creano o ampliano i database di riconoscimento facciale attraverso lo *scraping* non mirato di immagini facciali da internet o da filmati CCTV; (dc) l'immissione sul mercato, la messa in servizio per questo scopo specifico o l'uso di sistemi di intelligenza artificiale per dedurre le emozioni di una persona fisica nei settori del lavoro e delle istituzioni educative, ad eccezione dei casi in cui l'uso del sistema di intelligenza artificiale è destinato a essere immesso sul mercato per motivi medici o di sicurezza.

<sup>112</sup> Le due Proposte sono strettamente collegate e si completano a vicenda. La Proposta *PLD* riguarda la responsabilità senza colpa del produttore per i prodotti difettosi e aggiornerebbe l'attuale quadro giuridico includendo il *software* e i file di fabbricazione digitale nella definizione di prodotto e chiarendo quando un servizio correlato deve essere trattato come un componente di un prodotto. Consentirebbe il risarcimento delle perdite materiali derivanti da morte, lesioni personali, danni alla proprietà e perdita o corruzione di dati. La Proposta *AILD* riguarda le richieste di responsabilità perseguite nell'ambito dei regimi nazionali di responsabilità per colpa, nei casi in cui alcuni sistemi di intelligenza artificiale siano coinvolti nel causare danni.

<sup>113</sup> Proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496.

<sup>114</sup> Così in M.M. MOLLICONE, *Il rischio dell'intelligenza artificiale*, cit., p. 2127. Il problema si pone, soprattutto, per le imprese attive a livello transfrontaliero. In assenza di norme armonizzate a livello di Ue in materia di risarcimento dei danni causati dai sistemi di IA, i fornitori, gli operatori e gli utenti dei sistemi di IA, da un lato, e i danneggiati, dall'altro, si troverebbero di fronte a 27 regimi di responsabilità differenti, che comporterebbero livelli di protezione diseguali e una distorsione della concorrenza tra le imprese di Stati membri diversi.

La recentissima *AILD*, perciò, che si proclama diretta a garantire a coloro che hanno subito danni causati dall'IA una protezione equivalente a quella di cui beneficiano quanti subiscono danni causati da prodotti di altro tipo, si volge a perseguire i suoi obiettivi con un testo magro (33 considerando e solo 9 articoli) e con un approccio in più fasi.

La prima fase, meno invasiva, si limiterebbe a specificare misure relative all'onere della prova<sup>115</sup>; la seconda fase, invece, prevederebbe il riesame della necessità di misure più rigorose o più ampie come l'utilizzo del paradigma di responsabilità oggettiva per le IA ad alto rischio.

Perciò, ad oggi (I fase), la proposta direttiva provvederebbe ad alleggerire l'onere della prova in modo molto mirato e proporzionato attraverso il ricorso alla «divulgazione»<sup>116</sup> e a «presunzioni relative»<sup>117</sup>.

La direttiva prevede, quindi, per coloro che chiedono il risarcimento del danno, la possibilità di ottenere informazioni sui sistemi di IA ad alto rischio che devono essere registrate/ documentate a norma dell'*AI Act*. Oltre a quanto indicato, le presunzioni relative garantiranno a coloro che chiedono il risarcimento dei danni causati dai sistemi di IA un onere della prova più ragionevole e la possibilità che le azioni di responsabilità giustificate abbiano esito positivo. La direttiva si applicherebbe, perciò, alle domande di risarcimento del danno causato da un sistema di IA nel quadro di azioni civili di responsabilità extracontrattuale, qualora tali azioni siano intentate nell'ambito di regimi di responsabilità per colpa, ossia, in particolare, regimi che prevedono la responsabilità legale di risarcire i danni causati da un'azione o un'omissione intenzionalmente lesiva o colposa.

È l'art. 1 a definire l'oggetto e l'ambito di applicazione, statuendo anche che gli Stati membri possono adottare o mantenere in vigore norme nazionali più favorevoli all'attore in relazione all'esigenza di motivare una domanda presentata nel quadro di un'azione civile di responsabilità extracontrattuale per ottenere il risarcimento danno causato da un sistema di IA, nel rispetto del diritto dell'Unione.

In altri termini, con la presente direttiva non si imporrebbe un particolare regime di responsabilità per i casi di danni da utilizzo di IA quanto piuttosto si offrirebbero degli strumenti di semplificazione dell'onere della prova tutte le volte in cui gli Stati scelgano di utilizzare il paradigma di responsabilità per colpa.

Quanto, più specificatamente, alla divulgazione degli elementi di prova e presunzione relativa di non conformità, deve richiamarsi l'art. 3 ove si prevede che un organo giurisdizionale possa ordinare la divulgazione-conservazione degli

---

<sup>115</sup> I cittadini dell'UE, le organizzazioni dei consumatori e le istituzioni accademiche hanno espresso un forte sostegno per le misure sull'onere della prova e la responsabilità senza colpa - denominata «responsabilità oggettiva» - abbinata alla copertura assicurativa obbligatoria. Di contro, la responsabilità oggettiva è stata ritenuta sproporzionata dalla maggior parte delle imprese. Perciò, l'armonizzazione dell'alleggerimento dell'onere della prova ha ottenuto maggiore sostegno, in particolare tra le PMI.

<sup>116</sup> Art. 3 proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496.

<sup>117</sup> Art. 4 proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496.

elementi di prova rilevanti in relazione a specifici sistemi di IA ad alto rischio<sup>118</sup> che si sospetta abbiano cagionato danni.

L'ordine di divulgazione, però, potrebbe sollecitarsi unicamente quando l'attore potenziale si sia visto opporre un rifiuto a séguito della richiesta di elementi di prova effettuata al fornitore, ad una persona soggetta agli obblighi del fornitore a norma dell'*AI Act* o ad un utente.

Ciò significa che l'attore potenziale, per poter istruire la sua causa e sperare di ottenere il risarcimento del danno, dovrà, innanzitutto, interpellare i soggetti che esercitano un «controllo» sui sistemi di IA, impiegando uno «sforzo proporzionato»<sup>119</sup>. In secondo luogo, in caso di rifiuto di informazioni da parte di quelli, dovrà fare richiesta al giudice di ordinarne la divulgazione, sempre che dia «prove sufficienti» a sostenere la «plausibilità» della domanda di risarcimento del danno.

A questo punto, i convenuti hanno due opzioni. O ottemperare all'ordine del giudice o disobbedire all'ingiunzione di divulgazione o conservazione degli elementi di prova.

In tale ultimo caso, anche al fine di ridurre i tempi di durata del contenzioso, scatterà una presunzione relativa di non conformità al pertinente obbligo di diligenza da parte del convenuto (quindi una «presunzione della colpa» del convenuto<sup>120</sup>), il quale avrà facoltà di dare prova contraria.

L'art. 4, poi, fa riferimento alla presunzione relativa del nesso causale in caso di colpa, diretta a porre rimedio alla difficoltà di dimostrare che un determinato *input*, di cui è responsabile la persona potenzialmente tenuta a rispondere del danno, ha provocato un determinato *output* del sistema di IA, che a sua volta ha causato il danno in questione.

Detta presunzione opererebbe unicamente laddove sia stata già accertata la colpa<sup>121</sup>, oltrechè la «ragionevole probabilità» che il comportamento colposo del convenuto (o di altri di cui il convenuto è responsabile) abbia influito sull'*output*

---

<sup>118</sup> Il fatto che la divulgazione degli elementi di prova sia limitata ai sistemi di IA ad alto rischio è coerente con l'*AI Act*, che prevede determinati obblighi specifici di documentazione, informazione e conservazione delle registrazioni per gli operatori coinvolti nella progettazione, nello sviluppo e nell'impiego di sistemi di IA ad alto rischio. Tale coerenza garantisce inoltre la necessaria proporzionalità evitando che gli operatori di sistemi di IA con un livello di rischio basso o nullo siano soggetti a un obbligo di documentazione delle informazioni simile a quello imposto dall'*AI Act* per i sistemi di a ad alto rischio. Cfr. considerando n. 18 della proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496.

<sup>119</sup> Letteralmente viene richiesto dall'art. 3, punto 2 della proposta di Direttiva sulla responsabilità da IA, Bruxelles, 28 settembre 2022, COM(2022)496. Nel valutare la proporzionalità di un ordine di divulgazione o di conservazione degli elementi di prova, gli organi giurisdizionali nazionali tengono conto dei legittimi interessi di tutte le parti, compresi i terzi interessati, specialmente in relazione alla protezione dei segreti commerciali (ex art. 2, § 1, della direttiva UE 2016/943) e delle informazioni riservate, come le informazioni relative alla sicurezza pubblica o nazionale.

<sup>120</sup> La disposizione deve leggersi in combinato disposto con l'art. 4, § 2 della medesima direttiva, il quale specifica che l'accertamento della colpa – in caso di IA ad alto rischio e domanda azionata verso fornitori *et similia* – è soddisfatto solo se parallelamente si dimostra che il convenuto ha violato le misure di cui agli artt. 24-28 *AI Act*.

<sup>121</sup> La direttiva di cui si parla non armonizza i criteri in base ai quali gli organi giurisdizionali accertano l'esistenza della colpa. Detti criteri restano disciplinati dal diritto nazionale e, ove siano armonizzati, dal diritto dell'Unione applicabile. Allo stesso modo, la direttiva non armonizza i criteri relativi ai tipi di danno risarcibile.

e che questo, o la sua assenza abbia cagionato il danno, facendo un distinguo a seconda che il sistema di IA sia ad alto o basso rischio o se sia impiegato, o meno, in un'attività professionale.

Con riferimento ai sistemi di IA ad alto rischio, in primo luogo, occorre ulteriormente distinguere a seconda che la domanda di risarcimento del danno venga intentata nei confronti di un fornitore<sup>122</sup> di sistemi di IA o di altra persona soggetta ai medesimi obblighi previsti dall'*AI Act*<sup>123</sup>, oppure di un utente.

Nel primo caso, si deve premettere che, la proposta direttiva limita l'accertamento della colpa richiedendo che si possa ottenere solamente qualora l'attore abbia dimostrato che i soggetti menzionati non abbiano rispettato uno dei requisiti previsti dai capi 2 e 3, Titolo III dell'*AI Act*<sup>124</sup>, tenendo conto delle misure adottate<sup>125</sup> nel quadro del sistema di gestione dei rischi e dei relativi risultati a norma dell'art. 9 e 16 *AI Act*.

Nel caso di domanda intentata contro un utente<sup>126</sup> di un sistema di IA, invece, la sua colpa si intende provata laddove l'attore dimostri o che l'utente non abbia rispettato l'obbligo di utilizzare il sistema di IA o di monitorarne il funzionamento conformemente alle istruzioni per l'uso che accompagnano il sistema o, se del caso, l'obbligo di sospenderne o interromperne l'uso a norma dell'art. 29 dell'*AI Act*; oppure che questi abbia esposto il sistema di IA a dati di *input* sotto il suo controllo che non sono pertinenti e sufficientemente rappresentativi alla luce della finalità prevista del sistema a norma dell'art. 29, § 3 dell'*AI Act*.

Ad ogni modo, indipendentemente dalla qualifica, il convenuto ha una prova contraria diretta a bloccare l'operatività della presunzione di causalità. Questa consiste nella dimostrazione che l'attore possa ragionevolmente accedere ad elementi di prova e competenze sufficienti per dimostrare l'esistenza del nesso di causalità di cui all'art. 4, § 1.

Con riferimento, poi, alle IA non ad alto rischio, la presunzione di causalità potrebbe operare unicamente in determinati casi ossia quando l'organo giurisdizionale nazionale ritiene eccessivamente difficile per l'attore dimostrare l'esistenza del nesso di condizionamento<sup>127</sup>.

---

<sup>122</sup> Il fornitore è l'entità che sviluppa il sistema di IA e lo immette sul mercato/lo mette in servizio con il proprio nome o marchio.

<sup>123</sup> Si fa riferimento agli obblighi dei fabbricanti ed altri soggetti, individuati, oggi, agli artt. 24-28 dell'*AI Act* in base all'ultimo testo di compromesso.

<sup>124</sup> *Ibidem*.

<sup>125</sup> La presunzione scatta solo se vengono violati i requisiti specifici e tassativamente elencati dall'*AI Act*. Questi comprendono: il regime di governance dei dati, compresi i dati di formazione, di cui all'art. 10, §§ da 2 a 4 *AI Act*; i requisiti di trasparenza di cui all'art. 13 *AI Act*; la disposizione sulla supervisione umana di cui all'art. 14 *AI Act*; i requisiti di performance, robustezza e cybersecurity di cui all'art. 15 *AI Act*; gli obblighi di azioni correttive, compresi il ritiro e il richiamo, e dovere di informazione di cui all'art. 21 *AI Act*.

<sup>126</sup> La definizione di «utente» è ancora una volta relegata all'*AI Act*, che nella sua versione attuale specifica che si tratta di «qualsiasi persona fisica o giuridica, autorità pubblica, agenzia o altro organismo che utilizza un sistema di IA sotto la sua autorità». L'aspetto degno di nota di questa definizione tautologica è che ora include i consumatori che utilizzano il sistema di IA per scopi non professionali.

<sup>127</sup> In questo caso, la presunzione di causalità tra la colpa (mancato monitoraggio) e la produzione dell'*output* da parte dell'IA non è legata alla violazione di alcun dovere specifico dell'*AI Act* (che in questi casi sono molto pochi, cfr. art. 52 *AI Act*). Piuttosto, si applica la disposizione generale di cui all'art. 4, § 1, lett. a) *AILD*, che richiede solo la violazione di un obbligo di diligenza previsto dal diritto dell'Unione o

Nel caso ulteriore, infine, in cui il convenuto sia un soggetto che ha utilizzato il sistema di IA nel corso di un'attività personale non professionale ossia un consumatore, la presunzione di causalità potrebbe operare solo nel caso in cui il convenuto abbia interferito materialmente con le condizioni di funzionamento del sistema di IA<sup>128</sup> o se lo stesso non abbia adempiuto all'obbligo di determinare, essendone in grado, le condizioni di funzionamento del sistema di IA<sup>129</sup>.

Orbene, al netto dell'analisi effettuata la proposta direttiva, benché apprezzabile, genera perplessità e non resta esente da critiche.

Diverse iniziative europee si sono già basate sull'inversione dell'onere della prova o su presunzioni confutabili<sup>130</sup>. Tuttavia, tali meccanismi meritano sempre di essere affrontati con attenzione dal momento che possono avere un impatto sull'innovazione modificando in modo significativo la distribuzione dei rischi a scapito dei produttori/fornitori *et alia*.

Considerato quanto sopra, non vi è chi non veda il pregio di tale iniziativa sovranazionale sol si consideri che questa voluta armonizzazione contribuisce a garantire lo stesso livello di protezione per tutti gli utenti in Europa e a salvaguardare la parità di condizioni per gli attori coinvolti, come produttori, fornitori o operatori.

D'altra parte, l'approccio mirato della proposta *AILD*, che copre solo un aspetto specifico della richiesta di risarcimento per responsabilità civile, dimostra chiaramente la sua insufficienza attraverso la riluttanza a sviluppare un nuovo regime di responsabilità civile che comprenda tutti gli aspetti del diritto civile<sup>131</sup>, preferendo, invece, un'armonizzazione minima.

Invero, lo sviluppo di un approccio uniforme a livello europeo resta, comunque, impegnativo, considerate le varie divergenze tra gli Stati membri su istituti sostanziali e procedurali della responsabilità. Si pensi, ad esempio, ai concetti di «ragionevolmente probabile», «prove sufficienti», «diligenza dovuta», «plausibilità della domanda» o «sforzo proporzionato». Il contenuto di questi dovrà essere applicato sempre in conformità al diritto nazionale e a seconda dell'interpretazione che ne darà il giudice. Anche sull'essenza del «nesso di causalità», che è al centro della proposta – e che non lo definisce – non vi è

---

nazionale e destinato a proteggere dal danno specifico. Tuttavia, questa presunzione si applica solo se il giudice nazionale ritiene che sia eccessivamente difficile per il ricorrente soddisfare l'onere della prova.

<sup>128</sup> Questo potrebbe essere il caso in cui un tosaerba basato sull'IA sia stato posizionato dal consumatore su un sentiero di ghiaia e abbia quindi sollevato un sasso che ha colpito un pedone.

<sup>129</sup> Ad esempio, il consumatore potrebbe sapere ed essere obbligato per contratto a utilizzare un sistema di IA solo in determinate condizioni atmosferiche, ma decidere di utilizzarlo comunque in condizioni meteorologiche avverse. Ciò può essere rilevante, ad esempio, per i veicoli a guida automatica utilizzati in modalità assistita in condizioni di guida difficili.

<sup>130</sup> Si v. la Direttiva sui contenuti digitali (*DCD* - Direttiva (UE) 2019/770 del Parlamento europeo e del Consiglio del 20 maggio 2019 relativa a determinati aspetti dei contratti di fornitura di contenuto digitale e di servizi digitali) o il Regolamento (CE) n. 261/2004 del Parlamento europeo e del Consiglio dell'11 febbraio 2004 che istituisce regole comuni in materia di compensazione ed assistenza ai passeggeri in caso di negato imbarco, di cancellazione del volo o di ritardo prolungato e che abroga il Regolamento (CEE) 295/91.

<sup>131</sup> O. DHEU, J. DE BRUYNE e C. DUCUING, *The European Commission's approach to extra-contractual liability and AI—An evaluation of the AI liability directive and the revised product liability directive*, in *Comp. L. Sec. Rev.*, 2023, I, p. 51 ss.; Idd., *Liability for Damage Caused by Artificial Intelligence—Some Food for Thought and Current Proposals, in Tort Liability and Autonomous Systems Accidents*, Cheltenham, 2023, p. 27-62.

convergenza di definizione tra gli Stati membri. Al diritto nazionale sono, altresì, lasciate le norme che stabiliscono quale parte abbia l'onere della prova, quale grado di certezza sia richiesto per lo *standard* di prova, come venga definita la colpa o cosa si intenda per «danno prevedibile ed evitabile»<sup>132</sup>.

Orbene, se da una parte, l'approccio nazionale consente alle giurisdizioni di basarsi sul *corpus* normativo esistente, compresa la giurisprudenza, e di adattarlo al contesto locale, dall'altra parte, comporta una frammentazione giuridica tra le legislazioni degli Stati membri, che potrebbe essere dannosa per la sicurezza giuridica degli utenti di prodotti o servizi guidati IA.

Inoltre, nei casi di torti con elementi di transnazionalità, applicandosi il regolamento Roma II<sup>133</sup> (artt. 4-9), si creerebbe un problema di *forum shopping* ogni qualvolta le leggi dei due Stati divergono sensibilmente, ad esempio, prevedendo una responsabilità per colpa, l'altro una responsabilità oggettiva.

### 3.1. AILD e PLD: criticità, osservazioni e confronti.

Alla luce dell'analisi appena effettuata, si può ora tentare qualche considerazione mettendo a sistema la proposta *AILD* con la proposta *PLD* (v. capitolo II, § 3.2.2.) dal momento che queste, insieme con l'*AI Act*, si propongono di esaurire il problema della responsabilità da applicazione e/o utilizzo di un sistema di IA<sup>134</sup> e di innescare il cd. «effetto Bruxelles» nella regolamentazione dell'IA<sup>135</sup>.

Sicuramente, le proposte, sebbene elaborate contestualmente e tese, in parte, ad applicarsi alle medesime fattispecie completandosi a vicenda, divergono per scopi, struttura e contenuto. Si delinea, così, un doppio binario di tutela: una responsabilità ampiamente armonizzata secondo la direttiva *PLD* (nella misura in cui si applica al *software*) e una responsabilità ampiamente non armonizzata secondo le norme del diritto civile nazionale (coperta dalla proposta *AILD*).

Invero, da un lato, la proposta *AILD* cerca di armonizzare solo le questioni procedurali, come la divulgazione delle prove e l'onere della prova, nei regimi nazionali di responsabilità degli Stati membri ai fini della responsabilità per l'IA,

---

<sup>132</sup> Qui entra in gioco l'estensione sia del concetto di «lontananza del danno» dal soggetto agente, sia del concetto di «danno ricadente nella sfera della vittima» i quali sono strettamente legati alla percezione che i soggetti hanno della tecnologia.

<sup>133</sup> Regolamento (CE) n. 864/2007 del Parlamento europeo e del Consiglio dell'11 luglio 2007 sulla legge applicabile alle obbligazioni extracontrattuali («Roma II»).

<sup>134</sup> IPSOS, *Indagine sulle imprese europee sull'uso delle tecnologie basate sull'IA*, 2020. La Commissione osserva che nella sua indagine del 2020, la responsabilità civile ha costituito il principale ostacolo esterno all'adozione dell'IA da parte delle imprese nell'UE (43%).

<sup>135</sup> M. HILDEBRANDT, *Global competition and convergence of AI Law*, in JAN. M. SMITS, J. HUSA, C. VALCKE, M. NARCISO (a cura di), *Elgar Encyclopedia for Comparative Law*, Cheltenham, 2023, p. 12-13; C. WENDEHORST, *AI liability in Europe: anticipating the EU AI Liability Directive*, in *adalovelaceinstitute.org*, 2022, p. 9.

Per «effetto Bruxelles» si fa riferimento alla capacità catalizzatrice che questa normativa potrebbe avere anche nei confronti di Paesi terzi. Il governo Statunitense, ad esempio, sta commentando e seguendo da vicino gli sviluppi. Molte norme dell'*AI Act* e del pacchetto sulla responsabilità civile si applicheranno all'IA sviluppata negli Stati Uniti e in altri luoghi del pianeta, se impiegata o utilizzata nell'UE. L'approccio normativo dell'UE all'IA avrà quindi ripercussioni globali e potrà essere emulato in varia misura in altre giurisdizioni.

legando in larga misura questi strumenti alle violazioni dell'*AI Act*. Dall'altro lato, e indipendentemente dall'*AI Act*, la proposta di *PLD* suggerisce un aggiornamento generale della classica responsabilità per danno da prodotto difettoso, con uno sguardo specifico, tuttavia, ai prodotti digitali più in generale e all'IA più in particolare.

L'aggiornamento specifico per l'IA della direttiva *PLD* e le nuove regole della proposta *AILD* appaiono, comunque, relativamente modeste in termini di portata e ambizione rispetto alla proposta molto più ampia presentata dal Parlamento europeo nell'ottobre 2020<sup>136</sup>. Mentre quest'ultima prevedeva un vero e proprio sistema di responsabilità oggettiva per i sistemi di IA ad alto rischio e un'inversione generale dell'onere della prova della colpa per i sistemi di IA non ad alto rischio, le proposte della Commissione consistono principalmente in regimi di divulgazione e alleggerimento dell'onere della prova per casi specifici. Tuttavia, esse continuano a distinguere tra sistemi ad alto rischio e sistemi non ad alto rischio, come previsto dall'*AI Act*, istituendo un intricato sistema di obblighi per i diversi modi di utilizzo e per le diverse tipologie di destinatari.

La regolamentazione attualmente proposta della responsabilità per l'IA affonda le sue basi giuridiche nell'art. 114 TFUE<sup>137</sup> e ciò è, effettivamente,

---

<sup>136</sup> Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale (2020/2014(INL))

<sup>137</sup> Art. 114 TFUE: «1. Salvo che i trattati non dispongano diversamente, si applicano le disposizioni seguenti per la realizzazione degli obiettivi dell'art. 26. Il Parlamento europeo e il Consiglio, deliberando secondo la procedura legislativa ordinaria e previa consultazione del Comitato economico e sociale, adottano le misure relative al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati membri che hanno per oggetto l'instaurazione ed il funzionamento del mercato interno.

2. Il paragrafo 1 non si applica alle disposizioni fiscali, a quelle relative alla libera circolazione delle persone e a quelle relative ai diritti ed interessi dei lavoratori dipendenti.

3. La Commissione, nelle sue proposte di cui al paragrafo 1 in materia di sanità, sicurezza, protezione dell'ambiente e protezione dei consumatori, si basa su un livello di protezione elevato, tenuto conto, in particolare, degli eventuali nuovi sviluppi fondati su riscontri scientifici. Anche il Parlamento europeo ed il Consiglio, nell'ambito delle rispettive competenze, cercheranno di conseguire tale obiettivo.

4. Allorché, dopo l'adozione di una misura di armonizzazione da parte del Parlamento europeo e del Consiglio, da parte del Consiglio o da parte della Commissione, uno Stato membro ritenga necessario mantenere disposizioni nazionali giustificate da esigenze importanti di cui all'art. 36 o relative alla protezione dell'ambiente o dell'ambiente di lavoro, esso notifica tali disposizioni alla Commissione precisando i motivi del mantenimento delle stesse.

5. Inoltre, fatto salvo il paragrafo 4, allorché, dopo l'adozione di una misura di armonizzazione da parte del Parlamento europeo e del Consiglio, da parte del Consiglio o da parte della Commissione, uno Stato membro ritenga necessario introdurre disposizioni nazionali fondate su nuove prove scientifiche inerenti alla protezione dell'ambiente o dell'ambiente di lavoro, giustificate da un problema specifico a detto Stato membro insorto dopo l'adozione della misura di armonizzazione, esso notifica le disposizioni previste alla Commissione precisando i motivi dell'introduzione delle stesse.

6. La Commissione, entro sei mesi dalle notifiche di cui ai paragrafi 4 e 5, approva o respinge le disposizioni nazionali in questione dopo aver verificato se esse costituiscano o no uno strumento di discriminazione arbitraria o una restrizione dissimulata nel commercio tra gli Stati membri e se rappresentino o no un ostacolo al funzionamento del mercato interno. In mancanza di decisione della Commissione entro detto periodo, le disposizioni nazionali di cui ai paragrafi 4 e 5 sono considerate approvate. Se giustificato dalla complessità della questione e in assenza di pericolo per la salute umana, la Commissione può notificare allo Stato membro interessato che il periodo di cui al presente paragrafo può essere prolungato per un ulteriore periodo di massimo sei mesi.

7. Quando uno Stato membro è autorizzato, a norma del paragrafo 6, a mantenere o a introdurre disposizioni nazionali che derogano a una misura di armonizzazione, la Commissione esamina immediatamente l'opportunità di proporre un adeguamento di detta misura.

giustificato dall'obiettivo di evitare la frammentazione del mercato unico digitale, i costi legali proibitivi per le aziende che forniscono IA in tutta l'UE e gli ostacoli per le persone danneggiate che intendono citare in giudizio i fornitori in contesti transfrontalieri. Tuttavia, le medesime ragioni addotte per l'utilizzo dell'art. 114 TFUE, probabilmente, deporrebbero meglio a favore dell'istituzione di un unico regime europeo armonizzato di responsabilità civile per i danni causati da prodotti (non solo per difettosità), compresi i sistemi di IA, da formalizzare attraverso l'emanazione di un unico regolamento globale. Un quadro unificato e coerente non andrebbe solo a vantaggio dell'industria dell'IA ma diminuirebbe drasticamente il rischio di esacerbare, in presenza di regimi frammentati, le difficoltà dei consumatori nel sapere quale regime applicare e come citare in giudizio i fornitori situati in Stati membri diversi<sup>138</sup>. Ciononostante, al momento una simile iniziativa è politicamente improbabile.

Entrambe le direttive si applicano in parallelo, ma coprono aspetti diversi e, secondo la Commissione, reciprocamente esclusivi. I danneggiati possono, quindi, scegliere se procedere in base alla *PLD*, al quadro della proposta *AILD* o a entrambi.

Si riportano di seguito le principali differenze fra le due proposte:

<b>Proposta <i>PLD</i></b>	<b>Proposta <i>AILD</i></b>
Pretesa radicata nel diritto dell'UE	Pretesa radicata nel diritto degli Stati membri
Aspetti materiali e procedurali della responsabilità da prodotto, anche IA	Aspetti procedurali della responsabilità per i sistemi di IA
Applicabile a prodotti <i>hardware</i> e <i>software</i> , compresi i sistemi di IA	Applicabile solo ai sistemi di IA
Presumibilmente responsabilità oggettiva	Responsabilità per colpa
Richieste di risarcimento contro produttori e altri soggetti della catena di fornitura	Richieste di risarcimento contro produttori, utenti professionali e consumatori
Danni ammissibili: proprietà, morte o lesioni personali e perdita di dati.	Danni ammissibili:

8. Quando uno Stato membro solleva un problema specifico di pubblica sanità in un settore che è stato precedentemente oggetto di misure di armonizzazione, esso lo sottopone alla Commissione che esamina immediatamente l'opportunità di proporre misure appropriate al Consiglio.

9. In deroga alla procedura di cui agli articoli 258 e 259, la Commissione o qualsiasi Stato membro può adire direttamente la Corte di giustizia dell'Unione europea ove ritenga che un altro Stato membro faccia un uso abusivo dei poteri contemplati dal presente art..

10. Le misure di armonizzazione di cui sopra comportano, nei casi opportuni, una clausola di salvaguardia che autorizza gli Stati membri ad adottare, per uno o più dei motivi di carattere non economico di cui all'art. 36, misure provvisorie soggette ad una procedura di controllo dell'Unione.

<sup>138</sup> La direttiva sulla protezione dei dati personali ne è un esempio lampante. La frammentazione del panorama della protezione dei dati negli Stati membri è stata una delle ragioni principali della scelta di un regolamento come strumento per il GDPR.

	potenzialmente anche i diritti fondamentali e il danno economico primario
Armonizzazione completa	Armonizzazione minima

Andando ad evidenziare nel dettaglio cosa delle due direttive preoccupa di più, deve cominciarsi dall'ambito di applicabilità.

Per delineare il suo ambito di applicazione, la proposta *AILD* fa molto affidamento all'*AI Act* attraverso continui riferimenti. Questo non solo implica che la sua applicabilità non possa ancora essere delineata in modo definitivo, in quanto l'*AI Act* è ancora oggetto di pendente compromesso fra il Parlamento europeo ed il Consiglio, ma anche che la sua operatività, applicandosi unicamente nei casi che riguardano i sistemi di IA, dipende strettamente dalla definizione che si da di questo termine. Termine che la proposta *AILD* mutua proprio dall'*AI Act*. Termine che, purtroppo, continua a perseguire i legislatori dell'*AI Act*<sup>139</sup> in quanto incapaci di trovare una quadra. La sfida, in effetti, è ardua. Anche nell'ambito dell'informatica, gli studiosi non sono molto d'accordo su una definizione di IA<sup>140</sup> (v. capitolo I, § 2). In genere, i concetti di IA si distinguono tra due sottotipi principali, l'apprendimento automatico da un lato e gli approcci logici o i sistemi esperti dall'altro. Questa differenza è ora incorporata nel nuovo considerando 6, secondo capoverso<sup>141</sup> che specifica la definizione più generale di IA contenuta nell'art. 3, § 1 dell'*AI Act* nella versione ultima<sup>142</sup>.

Il problema principale rimane quello di distinguere un sistema di IA da un *software* complesso ma più tradizionale, in particolare quando utilizza metodi statistici o regole artigianali.

Nella versione finale del testo di compromesso interistituzionale dell'*AI Act*, i metodi statistici sono richiamati dal considerando 6. Pertanto, la definizione di IA comprenderà probabilmente strumenti *software* piuttosto semplici, come la regressione lineare<sup>143</sup>.

Inoltre, il considerando 6 dell'*AI Act* comprende gli approcci basati sulla logica e sulla conoscenza, che includono i sistemi esperti e altri modelli basati su regole. A meno che il criterio di «autonomia» di cui all'art. 3, § 1 dell'*AI Act* non

<sup>139</sup> L. BERTUZZI, *Intelligenza Artificiale intelligenza artificiale definizione, governance*, in *euractiv.com*, 2022.

<sup>140</sup> Cfr. S. RUSSEL e P. NORVIG, *Intelligenza artificiale*, cit., p. 5; D.L. POOLE e A.K. MACKWORTH, *Artificial Intelligence: Foundations of Computational Agents*, 2<sup>ed.</sup>, Cambridge, 2017, p. 3-7.

<sup>141</sup> Considerando 6, secondo capoverso *AI Act*: «[...] The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives; and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved. The capacity of an AI system to infer goes beyond basic data processing, enable learning, reasoning or modelling [...]».

<sup>142</sup> Art. 3, § 1 *AI Act*: «“AI system” is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments».

Il considerando 6, secondo capoverso *AI Act* aggiunge che «The term “machine-based” refers to the fact that AI systems run on machines».

<sup>143</sup> Detti *software* sono stati utilizzati per decenni in varie tipologie di industrie, ricerca e istruzione.

sia dotato di un significato che giustifichi una certa capacità di apprendimento o una soglia di complessità, il considerando 6 rischia di far rientrare nell'ambito dell'*AI Act* anche sistemi di inferenza relativamente semplici e rudimentali<sup>144</sup>. Or dunque, se così fosse, il campo di applicazione così ampio suggerirebbe che sia l'*AI Act* che, di conseguenza, la proposta *AILD*, potrebbero essere considerati una «regolamentazione del *software* avanzato» piuttosto che una specifica «regolamentazione dell'IA».

Questo, forse inavvertitamente, avvicina in modo significativo l'*AILD* alla *PLD*<sup>145</sup>. Quest'ultima, infatti, si applica al *software* in generale, riservando alcune inversioni dell'onere della prova ai sistemi *software* particolarmente complessi<sup>146</sup>.

Ad ogni modo, dal momento che i confini mobili della definizione di «sistema di IA» o di «*software* complesso» aprono delle significative zone grigie, si suggerisce, al fine di una maggiore certezza del diritto, la redazione di un elenco esaustivo di esempi concreti di *software* che non rientrano nell'ambito di applicazione dell'*AI Act* e, quindi, della *AILD*.

Tuttavia, i problemi in tema di applicabilità della *AILD* non sono terminati. Sebbene la proposta si allinei all'*AI Act* per quanto riguarda la definizione di sistema di IA, si ritiene che la stessa debba essere da quello disaccoppiata per quanto riguarda la classificazione del rischio: la categoria di sistemi di IA ad alto rischio dell'*AI Act* sembra sia eccessiva che poco inclusiva, specialmente se trapiantata direttamente in un quadro di responsabilità per l'IA.

Se, da una parte, trapiantare semplicemente le categorie di rischio dall'*AI Act* garantisce la coerenza tra gli strumenti, dall'altra parte, pecca nel non considerare le peculiarità dell'istituto della responsabilità civile<sup>147</sup>.

La concretizzazione dei rischi può essere distribuita in modo molto diseguale tra gli individui. Se, in una società di un milione di cittadini, solo cinque persone rischiano di essere danneggiate in modo significativo da una certa applicazione dell'IA, questo potrebbe non essere sufficiente per classificare l'applicazione

---

<sup>144</sup> Cfr. P. HACKER, *The European AI liability directives – Critique of a half-hearted approach and lessons for the future*, in *Comp. law sec. rev.*, 2023, 51, p. 1 ss.

<sup>145</sup> *Ibidem*.

<sup>146</sup> Si potrebbe pensare che, se un *software* supera una certa soglia di complessità e i suoi risultati diventano difficili da prevedere, controllare o spiegare (senza ulteriori strumenti di analisi), non dovrebbe essere importante, dal punto di vista del rischio, se il modello sottostante si qualifica o meno come IA. Ciò che dovrebbe davvero rilevare sarebbe l'individuazione dell'inesco degli stessi rischi di imprevedibilità e opacità spesso attribuiti all'IA.

P. HACKER, *The European AI liability directives*, cit. p. 8, ritiene che, probabilmente, questi *software* «convenzionali» ma complessi, potenzialmente imprevedibili e opachi, dovrebbero seguire le stesse regole dei modelli di apprendimento automatico potenzialmente imprevedibili e opachi. Un approccio ancora più radicale abbandonerebbe del tutto la definizione di IA a favore di una vera e propria regolamentazione del rischio, che copra tutti i *software* che innescano rischi specifici come l'opacità, la discriminazione o l'imprevedibilità.

Cfr. Z.C. LIPTON, *Il mito dell'interpretabilità dei modelli: In machine learning, the concept of interpretability is both important and slippery*, in *researchgate.net*.

<sup>147</sup> In particolare, pecca nel non considerare la funzione della responsabilità civile. Secondo l'opinione prevalente, quest'ultima non è quella di distribuire i costi sociali, come voleva intenderla G. CALABRESI, *Costo degli incidenti*, cit., quanto piuttosto quella di riparare il danno subito individualmente. In altri termini, funzione precipua della responsabilità extracontrattuale sarebbe quella di porre il soggetto danneggiato nella medesima situazione di indifferenza nella quale si trovava prima della verifica del danno.

come ad alto rischio ai sensi dell'*AI Act*. Tuttavia, se il rischio è molto pronunciato a livello individuale, lo stesso deve essere affrontato con un sistema efficace di responsabilità e risarcimento<sup>148</sup>. Pertanto, nel caso di un danno totale atteso basso ma con una significativa variazione del rischio tra gli individui colpiti, la classificazione del rischio individuale e quella sociale possono divergere<sup>149</sup>. Si tratta di un fenomeno ben noto alla teoria delle scelte sociali in generale.

Questa differenza teorica dovrebbe scoraggiare dall'equiparare le categorie di rischio dell'*AI Act* a quelle della proposta *AILD*. Differentemente opinando, si confonderebbe il rischio sociale con quello individuale e si negherebbe il diritto a un rimedio efficace a coloro che hanno subito danni a séguito dell'applicazione di sistemi di IA a basso rischio sociale seppur ad alto rischio individuale<sup>150</sup>.

Perciò, non vi è chi non veda che, quantomeno nell'ottica della responsabilità civile, il Legislatore europeo dovrebbe riconoscere che rischi significativi possono incorrere anche con riferimento ad attività il cui rischio aggregato non merita la qualifica di rischio elevato ai sensi dell'*AI Act*. Si pensi all'*automotive*, ai sistemi di riconoscimento delle emozioni e modelli di tariffazione assicurativa. Pertanto, tali applicazioni dovrebbero almeno essere qualificate come ad alto rischio ai sensi della proposta *AILD*<sup>151</sup>.

A differenza della *AILD*, la proposta *PLD*, richiama molto meno l'*AI Act* per delineare il suo perimetro di applicazione. Rivoluzionando il regime *PLD*, ora copre esplicitamente il *software* di tutti i tipi escludendo dal suo ombrello

---

<sup>148</sup> P. HACKER, *The European AI liability directives*, cit. p. 9

<sup>149</sup> *Ibidem*.

<sup>150</sup> Tale questione preoccupa anche il Garante europeo per la protezione dei dati (*GEPD* o *EDPS*). Si v. *EPDS, Opinion 42/2023 on the Proposals for two Directives on AI liability rules*, in [www.edps.europa.eu](http://www.edps.europa.eu), 11 ottobre 2023. L'Autorità afferma che i danni effettivi causati dai sistemi di IA non etichettati come ad alto rischio potrebbero essere ancora piuttosto significativi nella pratica. Potrebbe essere il caso, ad esempio, dei sistemi di IA destinati ad essere utilizzati per prendere decisioni sull'idoneità delle persone fisiche all'assicurazione sulla casa o sulla responsabilità civile. Le vittime di danni causati da sistemi di IA non etichettati come ad alto rischio potrebbero, quindi, incontrare difficoltà nell'ottenere prove a sostegno della loro richiesta di risarcimento simili a quelle che si incontrano davanti a sistemi di IA etichettati come ad alto rischio. Alla luce dell'obiettivo di creare parità di condizioni, non bisognerebbe differenziare le persone colpite dai sistemi di IA in base alla loro classificazione come ad alto rischio o non.

<sup>151</sup> Problema inverso è, invece, quello che riguarda il trattamento del sistema di IA cd. «*for general purpose*» (*AGI* o *GPAIS*) e, quindi, dei cd. «*Foundation Models*» e «*Large Generative Models*» – come *ChatGPT*, *GPT-4*, *DALL-E 2* e *Stable Diffusion*. Le disposizioni sul tema contenute nell'*AI Act* sono state, durante la stesura di questo scritto, oggetto di un acceso dibattito. La questione si è appuntata sulla opportunità o meno di sottoporre tutti i *GPAIS* all'intera gamma di obblighi per i sistemi di IA di rischio elevato previsti dall'*AI Act* e, di conseguenza, alle norme di responsabilità più severe previste dalla proposta *AILD* (e dalla *PLD*). Molti di questi modelli, infatti, sono attualmente offerti come *software open-source* e ampiamente utilizzati dagli sviluppatori. Il Parlamento europeo, nella sua posizione del 14 giugno 2023, ha optato per un approccio scaglionato – preferibile – che prevede tre livelli di obblighi per i fornitori di *Foundation Models*. Ad ogni modo, tali obblighi implicano che i fornitori dei *Foundation Models* debbano tracciare, descrivere meticolosamente e stabilire sistemi di monitoraggio e mitigazione integrativi per migliaia di potenziali casi d'uso dei loro sistemi di IA generale ad alto rischio, anche se solo una frazione di questi potrebbe essere implementata nella pratica. Un obbligo del genere non sembra utile, né tantomeno sensato, dato l'elevatissimo numero di scenari puramente ipotetici. In definitiva, si tratterebbe di un esercizio di conformità inefficiente. Ad ogni modo il testo definitivo di compromesso interistituzionale del 2 febbraio 2024 prevede l'inserimento di un intero Titolo (VIII A) dedicato ai *General Purpose AI Models*.

rimediale unicamente il *software* libero e *open-source*<sup>152</sup> sviluppato o fornito al di fuori di un'attività commerciale, al fine di facilitare l'innovazione e l'accesso al *software*.

Tuttavia, deve notarsi che anche molti strumenti *software* commerciali si basano, in parte, su *software open-source*. Ciò solleva immediatamente la questione dell'applicabilità della *PLD* nel caso in cui un componente *open-source* di un *software* commerciale contenga un *bug* (difetto). Rendere possibile che i fornitori del *software* commerciale chiedano il danno agli sviluppatori del *software open-source* renderebbe vana l'esenzione di responsabilità prevista dalla *PLD*. D'altro canto, non può non considerarsi che spesso non esistono alternative economicamente valide al *software open-source* per cui i loro programmatori dovrebbero essere in una posizione di forza per escludere l'indennizzo nei loro termini di servizio. Sembra, quindi, più importante incentivare i fornitori commerciali a controllare accuratamente il *software open-source* che includono nei loro prodotti *software*.

Dal punto di vista dell'applicabilità soggettiva, invece, la proposta *PLD* suggerisce di applicare il regime di responsabilità agli «operatori economici» (artt. 1 e 7), sostituendo il concetto di «produttori». Quest'ultimo termine comprende, infatti, i fabbricanti del prodotto finito o di componenti, i produttori di materie prime, gli importatori e le persone che si rappresentano come produttori<sup>153</sup>.

Il termine «operatore economico», appare, invero, più ampio, comprendendo tutti i soggetti coperti dalla *PLD*, ma aggiungendo il fornitore di un servizio connesso, il rappresentante autorizzato e il fornitore di servizi di adempimento<sup>154</sup>. Ciò è sintomo del riconoscimento dell'importanza dei canali logistici e dei prodotti ibridi<sup>155</sup>.

A differenza, poi, della proposta *AILD*, la *PLD* non si applica agli «utenti» come definiti nell'*AI Act*<sup>156</sup>.

Un altro aspetto su cui riflettere sono i meccanismi di innesco della responsabilità in base alle direttive, la colpa per l'*AILD* e la difettosità del prodotto per la *PLD*.

Relativamente all'*AILD*, il considerando 15 della proposta *AILD* contiene una notevole restrizione dei tipi di colpa coperti dalla direttiva. In poche parole,

---

<sup>152</sup> Questa eccezione dovrebbe essere accolta con favore, in quanto chi contribuisce al *software open-source* spesso lavora al codice nel tempo libero e potrebbe essere notevolmente scoraggiato dal contribuire se fosse soggetto alla responsabilità della *PLD*.

<sup>153</sup> Ad esempio apponendo un marchio (art. 3, §§ 1 e 2, attuale *PLD*). La responsabilità sussidiaria dei fornitori è stabilita dall'art. 3, § 3 attuale *PLD*.

<sup>154</sup> L'importatore e il rappresentante autorizzato del produttore, nonché il fornitore di servizi di adempimento in via sussidiaria, possono essere ritenuti responsabili se il produttore stesso è stabilito al di fuori dell'UE (art. 7, §§ 2 e 3 proposta *PLD*). Se nessuno dei due può essere identificato, possono essere citati in giudizio i distributori (= entità lungo la catena di fornitura, art. 4, § 15 della proposta *PLD*).

<sup>155</sup> Inoltre, come riconoscimento dell'importanza del commercio elettronico, le piattaforme online possono essere ritenute responsabili alle stesse condizioni dei distributori, se non sono protette dall'eccezione di *hosting* prevista dalla legge sui servizi digitali.

<sup>156</sup> Nell'ultima versione dell'*AI Act* del 2 febbraio 2024, tra l'altro, il termine «users/utenti» è stato modificato con quello di «deployers/applicatori», sembrando ampliare l'ambito soggettivo di applicazione del Regolamento. Resta ferma la non applicazione del Regolamento, a norma dell'art. 5, § 5c *AI Act*, ai «deployers who are natural persons using AI systems in the course of a purely personal non-professional activity».

le agevolazioni (divulgazione e inversione dell'onere) non dovrebbero applicarsi se un agente umano si interpone tra l'*output* dell'IA e il danno. Ciò riguarda in particolare l'uso dell'IA come sistema di raccomandazione o di supporto alle decisioni.

Ad una più attenta lettura, questa restrizione sembra presentare due problemi.

In primo luogo, una simile formulazione potrebbe invitare ad aggirare la norma facendo sì che i risultati dell'IA vengano «timbrati» da attori umani che seguono ciecamente i consigli della macchina<sup>157</sup>.

In secondo luogo, l'art. 14 dell'*AI Act* obbligherebbe i fornitori a consentire la supervisione umana dei sistemi di IA ad alto rischio e l'art. 29, § 1, 1a e 4, dell'*AI Act* imporrebbe agli utenti professionali di attuare tale supervisione, rendendo gran parte della proposta *AILD* potenzialmente inapplicabile e obsoleta.

Relativamente alla *PLD*, invece, la nuova disposizione sulla difettosità rispecchierebbe solo parzialmente le considerazioni tecniche dell'IA. Il *test* dello *standard* di aspettativa del consumatore nel migliore dei casi appare vago, mentre nel peggiore dei casi rasenta un concetto privo di qualsiasi contenuto.

La *PLD*, poi, non affronta un aspetto fondamentale relativo alla difettosità dei sistemi di IA. Come è noto, i modelli di IA commettono errori che gli esseri umani non avrebbero commesso. Al contrario, possono evitare errori che l'uomo avrebbe commesso. L'incidente della Tesla, avvenuto nel maggio 2016, è un esempio di questa tendenza. Il pilota automatico ha scambiato un grosso camion che attraversava un incrocio per un ponte lontano e ha guidato il veicolo dritto contro il camion, uccidendo il passeggero della Tesla. Chiaramente, un conducente umano in condizioni normali non avrebbe scambiato il camion per un ponte. Ciò solleva la questione se questo bug costituisca un difetto ai sensi della *PLD*.

In base alla normativa *PLD* si possono individuare tre diversi tipi di difetti: di fabbricazione, di progettazione e di istruzione (avvertimento). Con riferimento ai sistemi di IA il problema sussiste nella identificazione del difetto. Esistono, perciò, due approcci diversi per la rilevazione del difetto.

Secondo l'approccio antropomorfo, un giudice dovrebbe riscontrare il difetto se un essere umano medio non avrebbe commesso l'errore (e i consumatori si sarebbero aspettati che l'uomo non lo commettesse). Ciò offrirebbe, da una parte, il vantaggio di incentivare gli operatori a trovare ed eliminare in modo proattivo gli errori (ad esempio del camion come ponte e casi limite simili). D'altra parte, ciò potrebbe portare a un eccesso di deterrenza. Di fatti, potrebbe accadere che un sistema di IA (come il pilota automatico) eviti tutti i possibili errori, compresi quelli commessi dagli esseri umani, ma continui a confondere binari e ponti in alcune condizioni molto rare e imprevedibili. In

---

<sup>157</sup> GRUPPO DI LAVORO ART. 29, Linee guida sul processo decisionale individuale automatizzato e sulla profilazione ai fini del regolamento 2016/679 (WP 251), punta, infatti, ad una comprensione più ampia del processo decisionale automatizzato, che sarà invalidato solo da un intervento umano sostanziale.

Del medesimo avviso anche recenti sentenze dei tribunali. Sul caso del *robo-firing* di Uber deciso dalla Corte d'Appello di Amsterdam, si veda J. TURNER, *Amsterdam Court Upholds Appeal in Algorithmic Decision-Making Test Case: Drivers v Uber and Ola*, in *Fountain Court Blog*, 2023, in [fountaincourt.co.uk](https://fountaincourt.co.uk).

una situazione del genere, in base allo *standard* antropomorfo, si troverebbe effettivamente un difetto.

Secondo un approccio più orientato al sistema di IA, invece, il giudice dovrebbe astrarre dal singolo caso e valutare il livello di rischio dell'intero sistema di IA inserito in tutti i dispositivi (nel caso citato le automobili). Solo se venisse rilevato un malfunzionamento a livello sistemico, verrebbe dichiarata la presenza di un difetto<sup>158</sup>.

Ulteriore tematica su cui soffermarsi attiene la divulgazione delle prove e l'accesso alle informazioni. In entrambe le proposte, nella loro forma attuale, permane una preoccupazione centrale: i fattori che fanno scattare la responsabilità (colpa; difettosità) devono essere provati, almeno in generale, senza alleggerimento dell'onere della prova, dovendo gli attori presentare fatti e prove per la «plausibilità della richiesta».

Una pietra miliare della proposta *AILD*, finalizzata a garantire un risarcimento efficace, è l'art. 3 che stabilisce specifici obblighi di divulgazione per i potenziali convenuti<sup>159</sup>. Questo è limitato alle controversie che coinvolgono sistemi di IA ad alto rischio.

La proposta di direttiva *PLD* contiene un meccanismo sostanzialmente simile per la divulgazione delle prove all'art. 8. Il suo ambito di applicazione è tuttavia significativamente più ampio di quello dell'equivalente proposta *AILD*. L'art. 8 della proposta *PLD*, infatti, può essere invocato per ottenere informazioni anche su sistemi di IA non ad alto rischio.

Tuttavia, esistono alcune piccole differenze tra i due sistemi di divulgazione.

L'*AILD* è più dettagliata nei suoi requisiti in diversi casi. Innanzitutto, a differenza di quanto richiesto dalla *PLD*, il potenziale attore deve aver intrapreso tutti i tentativi proporzionati di raccogliere le prove dal convenuto prima di ottenere un ordine di divulgazione da parte del Tribunale (art. 3, § 1, della proposta *AILD*).

In secondo luogo, contrariamente alla proposta di *PLD*, la proposta *AILD* contiene una disposizione specifica che autorizza i Tribunali a ordinare misure di conservazione delle prove.

Infine, ma forse è la cosa più importante, il meccanismo di divulgazione contenuto nell'*AILD* può essere invocato sia prima che durante un vero e proprio processo<sup>160</sup>. Al contrario, l'equivalente della *PLD* sembra applicarsi solo una volta che la domanda sia stata depositata in Tribunale.

La proposta *AILD*, poi, in un'ottica di continuità normativa, prevede che le informazioni potenzialmente richieste corrispondano in larga misura agli

---

<sup>158</sup> Molti studiosi suggeriscono di riferire la difettosità alla mancata *compliance* agli *standard* tecnici che regolano specifici settori industriali. In questo modo la responsabilità della *PLD* si allineerebbe con la *P'AI Act*, garantendo, così, una certa coesione e certezza giuridica. Ad oggi il problema risiede proprio nella creazione di tali *standard* sol si consideri l'assenza di esperienza stratificata che possa permettere un'adeguata valutazione dei rischi.

<sup>159</sup> Si noti che, secondo alcuni, a differenza della presunzione di causalità di cui all'art. 4, il meccanismo di divulgazione non sarebbe limitato alle cause di responsabilità per colpa e quindi si applicherebbe anche alle cause di responsabilità oggettiva.

<sup>160</sup> L'art. 3, § 1, della proposta *AILD* parla di «richiesta di un potenziale ricorrente [...] o di un ricorrente».

obblighi di documentazione, conservazione dei registri e informazione già previsti dall'*AI Act*. D'altro canto, però, l'allineamento delle informazioni della proposta *AILD* con l'*AI Act* può anche costituire un difetto significativo. Invero, i destinatari della divulgazione delle prove sono, generalmente, molto diversi da quelli della divulgazione tipica dell'*AI Act*. Quest'ultima, infatti, è rivolta ad esperti con conoscenze tecnologiche significative, come sviluppatori, consulenti tecnologici, osservatori dell'IA o agenzie di regolamentazione. Al contrario, in base ad entrambe le proposte direttive, le informazioni divulgate nei procedimenti di responsabilità saranno principalmente richieste da ricorrenti consumatori, o da consulenti legali senza una forte preparazione tecnica, e dovranno essere interpretate da giudici esperti di diritto, ma inesperti di IA. È vero che si potrebbe, comunque, ricorrere ad esperti per interpretare le informazioni tecniche, tuttavia, così facendo, non si fuggirebbe dagli alti costi legati alla «manutenzione» della tecnologia né si eviterebbe l'innalzamento delle barriere all'applicazione dei sistemi di IA.

Ciò mette subito in evidenza un altro punto debole delle proposte direttive ovvero l'imposto obbligo di fornire fatti e prove a sostegno della plausibilità della richiesta di risarcimento danni<sup>161</sup>. Se da una parte, una soglia di questo tipo potrebbe rivelarsi effettivamente necessaria per proteggere aziende e sviluppatori da cause vessatorie o dall'utilizzo abusivo del sistema di divulgazione, magari per ottenere segreti commerciali, tuttavia, dall'altra parte, va notato che difficilmente faciliterebbe l'azione del ricorrente<sup>162</sup>. Le regole di divulgazione dell'IA dovrebbero cercare di evitare un destino simile. Pertanto, per far scattare le norme sulla divulgazione delle prove dovrebbe essere sufficiente che i potenziali ricorrenti forniscano fatti e prove del danno e del coinvolgimento di un sistema di IA.

Ancóra, entrambe le proposte prevedono come condizione necessaria, per la divulgazione delle prove, che le informazioni siano a disposizione dell'imputato. Questo, a sua volta, può essere difficile da valutare dall'esterno. Sebbene sia chiaro il tipo di informazioni che un fornitore o un utente dovrebbe possedere in osservanza delle norme dell'*AI Act*, non è garantito che le prove siano effettivamente a disposizione dell'imputato. Inoltre, l'art. 8 della *PLD* si applica anche ai sistemi di IA non ad alto rischio, per i quali l'*AI Act* non prescrive alcun requisito specifico in materia di informazioni, registrazione o conservazione dei dati. Pertanto, si dovrebbe chiarire che entrambe le disposizioni si applichino ogni qualvolta le informazioni siano, o dovrebbero essere legalmente a

---

<sup>161</sup> Anche il GEPD con la sua *Opinion 42/2023 on the Proposals for two Directives on AI liability rules*, in [www.edps.europa.eu](http://www.edps.europa.eu), 11 ottobre 2023, sostiene che, soddisfare tale requisito potrebbe essere particolarmente difficile nel contesto dei sistemi di IA, dove i rischi di manipolazione, discriminazione e decisioni arbitrarie saranno certamente presenti, anche quando i fornitori, gli operatori e gli utenti abbiano *prima facie* rispettato il loro dovere di diligenza, come definito dal *draft* dell'*AI Act*.

<sup>162</sup> Un caso simile è previsto dall'art. 9 della Direttiva 2004/113/CE del Consiglio del 13 dicembre 2004 che attua il principio della parità di trattamento tra uomini e donne per quanto riguarda l'accesso a beni e servizi e la loro fornitura. I ricorrenti qui devono fornire fatti che suggeriscano un caso *prima facie* di discriminazione. Di conseguenza, l'onere della prova è invertito e il convenuto deve dimostrare che, in realtà, la decisione non è stata influenzata dalla discriminazione.

disposizione del convenuto (ad esempio ai sensi dell'*AI Act*<sup>163</sup>). Perciò, in realtà, la norma potrebbe coprire unicamente i casi in cui il convenuto, in violazione dell'*AI Act* o di altri obblighi di legge o contrattuali, non ha prodotto o conservato determinati tipi di informazioni, registri o documenti.

Ad ogni buon conto, tali violazioni non dovrebbero escludere l'obbligo del convenuto di divulgare le prove pertinenti, anche se ciò implica la creazione *ex post* delle informazioni ai fini della causa di risarcimento danni. Ad esempio, se il convenuto non ha divulgato i dati metrici sull'accuratezza e sugli impatti discriminatori richiesti dall'art. 11, allegato IV, § 2, lett. g), *AI Act*, egli dovrebbe produrre tali documenti nell'ambito della procedura di divulgazione.

Infine, probabilmente, l'ultimo e il più grande problema della procedura di divulgazione della *AILD* risiede nella sua limitazione ai sistemi di IA ad alto rischio ai sensi dell'*AI Act*<sup>164</sup>. L'idea è quella di risparmiare ai fornitori di sistemi di IA non ad alto rischio l'onere aggiuntivo di raccogliere informazioni che non sono tenuti a conservare o generare ai sensi dell'*AI Act*.

Ora, sebbene sia generalmente comprensibile dal punto di vista della proporzionalità, una simile restrizione è vanificata dall'applicabilità della parallela *PLD* che prevede che la divulgazione operi anche per gli sviluppatori di IA non ad alto rischio. Tra l'altro, l'intera impostazione normativa dipende, in via cruciale, dalla classificazione dei sistemi di IA come ad alto rischio o a basso rischio e, come noto, molto spesso, il confine appare sfumato in assenza di *standard* tecnici sufficienti e di valutazioni del rischio tecnologico affidabili. Come sopra detto, alcuni sistemi di IA non considerati ad alto rischio ai sensi dell'*AI Act* (al momento) possono certamente danneggiare sensibilmente i singoli soggetti in modo materiale o immateriale, lasciandoli senza opzioni di ricorso veramente efficaci ai sensi della proposta *AILD*<sup>165</sup>.

Con buona probabilità, sarebbe opportuno armonizzare i meccanismi di divulgazione prevedendo gli stessi prerequisiti indipendentemente dal fatto che la richiesta di risarcimento sia basata sulla *PLD* o sulle norme nazionali in materia di responsabilità.

Oltre al meccanismo di divulgazione, il secondo elemento centrale delle proposte direttive è l'alleggerimento, a favore dei ricorrenti, dell'onere di provare elementi specifici costitutivi la richiesta di risarcimento. Mentre l'*AILD* permette

---

<sup>163</sup> P. HACKER, *The European AI liability directives*, cit. p. 19.

<sup>164</sup> Anche il GEPD con la sua *Opinion 42/2023*, cit., in *edps.europa.eu*, 11 ottobre 2023, raccomanda che sia il meccanismo di divulgazione delle prove di cui all'art. 3 della Proposta *AILD* che la presunzione confutabile di nesso causale in caso di colpa di cui all'art. 4 (v. *infra*) siano scollegati dalla definizione di sistemi «ad alto rischio» come definiti dall'*AI Act* e si applichino a tutti i sistemi di IA.

Il GEPD ricorda che la Proposta di legge sull'IA cerca di affrontare il problema introducendo all'articolo 13, paragrafo 2, l'obbligo per i fornitori di sistemi di IA ad alto rischio di accompagnarli con istruzioni per l'uso che includano «informazioni concise, complete, corrette e chiare, pertinenti, accessibili e comprensibili agli utenti». Pertanto, affinché le informazioni divulgate ai sensi dell'articolo 3 della Proposta *AILD* abbiano un effettivo valore probatorio, il GEPD raccomanda di introdurre nell'*AILD* un requisito analogo in base al quale le informazioni divulgate non si limitano alla divulgazione della sola documentazione tecnica, ma devono essere integrate da spiegazioni chiare e comprensibili.

<sup>165</sup> Di conseguenza, questi sistemi o dovrebbero essere inclusi nella categoria ad alto rischio dell'*AI Act* o, in mancanza di questa, nell'elenco dei sistemi a cui si estende la procedura di divulgazione preprocessuale secondo la proposta *AILD*.

di presumere l'inosservanza di un obbligo di diligenza, la *PLD* si concentra sulla corrispondente presunzione di difettosità del prodotto. Altresì, entrambe le proposte offrono un'inversione dell'onere della prova per il nesso di condizionamento, l'una (proposta *AILD*) tra la violazione di un dovere di diligenza («colpa») e l'*output* di un sistema di IA, l'altra (proposta *PLD*) tra la difettosità di un sistema di IA e la lesione di un interesse meritevole di tutela (danno evento).

Proprio sulla presunzione di causalità prevista dalla *AILD* occorre spendere delle parole.

Con riguardo alla richiesta di risarcimento danni azionata nei confronti dell'utente di una IA ad alto rischio, sorgono dei dubbi circa l'utilità del meccanismo presuntivo, per come formulato, tra la violazione del dovere di diligenza da parte del soggetto tenuto al «controllo» dell'IA e l'*output*. Ai sensi dell'art. 4, § 1 *AILD*, si presume solo il collegamento tra il comportamento colposo umano (a monte o nel corso dell'applicazione di un sistema di IA) e l'*output*. Tuttavia, la violazione del dovere dell'utente riguarda proprio un'azione a valle dell'applicazione di un'IA (ad esempio la supervisione umana), la quale, quindi, ha luogo solo dopo che l'*output* dell'IA è stato prodotto. Pertanto, l'utente potrebbe facilmente confutare la presunzione di causalità affermando che, a prescindere da qualsiasi svista, il risultato del sistema di IA sarebbe stato lo stesso.

Pertanto, si ritiene che, in tal caso, sarebbe più utile che la presunzione avvenga il danno all'azione/omissione dell'utente, cioè la violazione del requisito di supervisione umana dell'art. 29 *AI Act*. Per vero, il punto cruciale dell'instaurazione di una causa contro l'utente, in termini di causalità, sta proprio nel dimostrare che, con la supervisione umana, l'*output* dell'IA sarebbe stato corretto dall'uomo nel *loop* (post-elaborazione), e non che l'*output* sarebbe stato diverso. Pertanto, l'art. 4, § 3, *AILD*, nella sua forma attuale, non raggiungerebbe l'obiettivo dichiarato.

La stessa critica si applica, *mutatis mutandis*, all'obbligo del fornitore di consentire una sorveglianza efficace durante l'uso, menzionato dall'art. 4, § 2, lett. c) *AILD* e originariamente contenuto nell'art. 14 *AI Act*.

Dubbia resta la formulazione anche con riferimento alle richieste azionate di fronte ad IA non ad alto rischio. Di fatti, la regola secondo cui la decisione dell'operatività della presunzione di causalità viene lasciata alla discrezionalità del giudice, non sembra giustificata quando i sistemi non ad alto rischio siano tanto opachi ed individualmente rischiosi (v. *supra*) quanto i sistemi ad alto rischio, come specificato dall'*AI Act*. Anche in questo caso, perciò, si evidenzia la necessità o di includere alcuni sistemi di IA individualmente rischiosi<sup>166</sup> nella categoria dei sistemi ad alto rischio, almeno ai fini dell'azione di responsabilità, o di estendere la presunzione di causalità di cui all'art. 4, §§ 2 e 3 *AILD* a tali sistemi.

---

<sup>166</sup> Come, ad esempio, i sistemi *automotive* o di riconoscimento delle emozioni.

Infine, la proposta *AILD* dovrebbe essere aggiornata per coprire i danni inflitti dall'uso di un sistema di IA vietato (art. 5 *AI Act*).

Parallelamente, anche con riguardo alla presunzione di difettosità contenuta nella proposta *PLD* si segnalano alcuni punti deboli.

In primo luogo, si ritiene che la presunzione di difettosità, ai sensi dell'art. 9, § 2, lett. a), *PLD*, dovrebbe scattare, non solo in caso di mancata ottemperanza ad un ordine di divulgazione delle prove ma anche in caso di mancata ottemperanza ad un ordine di conservazione delle stesse, come stabilito dall'art. 3, § 5, della proposta *AILD*.

In secondo luogo, si ritiene che il soddisfacimento da parte del ricorrente, ai fini della presunzione di difettosità, degli ultimi due criteri di cui all'art. 9, § 2 del *draft PLD* – violazione dei requisiti di sicurezza obbligatori o evidente malfunzionamento –, resta estremamente costoso. Di fatti, il convenuto potrebbe ben preferire di ostendere una qualsiasi informazione piuttosto che incorrere nella presunzione inversa susseguente il rifiuto di ostensione. I convenuti sarebbero, perciò, incentivati a sommergere i ricorrenti con grandi quantità di dati indipendentemente dalla loro pertinenza e conclusività rispetto al problema. Se così fosse, i ricorrenti, senza l'accesso a documenti e dati specifici, dovrebbero affrontare alti costi per assumere esperti di IA che possano decifrare e selezionare i dati rilevanti. Appare evidente, quindi, la necessità che la proposta direttiva (ma è un suggerimento che vale per entrambe) sia accompagnata da strategie funzionali a facilitare l'accesso al finanziamento delle controversie per i ricorrenti che non sono in grado di sostenere personalmente i costi del ricorso ad esperti di IA<sup>167</sup>. Di conseguenza, sarebbe più utile sviluppare un regime di responsabilità alternativo che tenga conto anche delle ipotetiche strategie difensive dei soggetti interessati.

In conclusione, alla luce di tutte le osservazioni presentate, apparrebbe, quindi, preferibile, delineare un regime di responsabilità maggiormente armonizzato e graduato<sup>168</sup>, non solo sulla base del rischio, ma anche in considerazione dei destinatari della richiesta di risarcimento.

Si potrebbe pensare, perciò, di riservare la responsabilità oggettiva agli operatori economici e agli utenti professionali di sistemi di IA ad alto rischio, con la particolare attenzione, con riguardo ai *Foundation Model (GPAIS)*, di serbarla solo per le loro applicazioni ad alto rischio. Diversamente, le PMI, gli operatori economici e gli utenti professionali di sistemi di IA non ad alto rischio dovrebbero essere soggetti solo a presunzioni confutabili di colpa e causalità, al fine di evitare un eccesso di deterrenza. Infine, nei confronti dei consumatori

---

<sup>167</sup> A. BERTOLINI, *Artificial Intelligence and Civil Liability*, in *Publications Office of the European Union*, 2020, p. 84.

<sup>168</sup> Ci si potrebbe anche ispirare alla precedente risoluzione del Parlamento europeo (v. PARLAMENTO EUROPEO, Risoluzione del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale (2020)), purchè non ci si dimentichi di ampliare l'ambito soggettivo di applicazione per includere, oltre agli utenti professionali, anche i produttori. Allo stesso tempo, le regole dovrebbero essere attuabili per gli sviluppatori di IA grandi e piccoli e, per questo motivo, i massimali di responsabilità, le assicurazioni agevolate per le PMI, le restrizioni a casi specifici di danni «illeghi» e altre misure dovrebbero essere revisionate per rendere la responsabilità oggettiva accessibile alle aziende più «innovative».

utenti di un sistema di IA si potrebbe applicare un regime di responsabilità basato sulla colpa con, al più, una presunzione di causalità in determinati casi.

In questo modo si modellerebbe la natura della responsabilità a seconda dell'esistenza di un modello di diligenza che esoneri dalla responsabilità, «cioè l'esistenza di un livello di prevenzione che esonera dal risarcimento del danno»<sup>169</sup>.

In questa prospettiva, poi, la futura regolamentazione dell'IA non dovrebbe incappare nella dimenticanza del significato sociale che l'istituto della responsabilità assume nel contesto della crisi climatica. La sostenibilità ambientale<sup>170</sup> dovrebbe essere maggiormente considerata nella normazione dell'IA, specie in qualità di *standard* il cui mancato raggiungimento comporti l'assunzione di responsabilità.

L'impronta di carbonio dell'apprendimento automatico, in particolare, è aumentata drasticamente negli ultimi anni<sup>171</sup>. Dato l'attuale ritmo di sviluppo e adozione dell'IA, il contributo dell'apprendimento automatico al riscaldamento globale è reale, evidente e in aumento. È inevitabile, quindi, introdurre incentivi chiari e convincenti per rendere l'IA più sostenibile. La proposta originaria della Commissione dell'*AI Act* era rimasta in gran parte silente sulle questioni di sostenibilità ambientale.

Il Parlamento europeo ha aggiunto diverse disposizioni in tal senso, solo, nella bozza di *AI Act* nella sua posizione adottata il 14 giugno 2023, poi consolidata nella versione del 2 febbraio 2024<sup>172</sup>. D'altro canto, la Commissione sostiene che la proposta *AILD* si collega e promuove il *Green Deal* europeo. Tuttavia, questa interazione è strettamente incentrata sulla facilitazione della diffusione dell'IA. La Commissione conclude che l'adozione di sistemi di IA sia benefica per l'ambiente, ma, questa affermazione sembra pericolosamente unilaterale<sup>173</sup>.

Pertanto, si ritiene che, per sistematicamente meglio operare, l'*AI Act* dovrebbe prevedere degli strumenti che dimostrino come l'utilizzo di sistemi di IA combattano effettivamente il cambiamento climatico. Potrebbe, ad esempio, imporsi una «valutazione d'impatto sulla sostenibilità» per i sistemi di IA. Tale valutazione potrebbe ispirarsi alle numerose proposte relative alle valutazioni d'impatto dei sistemi di IA in generale. Tale obbligo potrebbe applicarsi agli sviluppatori di sistemi di IA ad alto rischio e non dal momento che l'impronta di carbonio dei sistemi di IA non è correlata al loro livello di rischio per la salute, la sicurezza o i diritti fondamentali.

---

<sup>169</sup> P.G. MONATERI, *La responsabilità civile*, cit. p. 37. L'a. afferma che il criterio della colpa sia efficiente quando sia la vittima che il danneggiante possono fare qualcosa per impedire il danno. Cioè quando esiste la possibilità di prevenzioni bilaterali. Diversamente la responsabilità oggettiva è uno strumento efficiente quando la tecnologia di prevenzione è unilaterale, il risarcimento per la vittima è perfetto, ed è chiaro *ex ante* chi è il danneggiante potenziale e chi è la vittima potenziale.

<sup>170</sup> Si v., ad esempio, P. HACKER, *Sustainable AI Regulation, Working Paper*, presentato alla Privacy Law Scholars Conference, 2023, in *arxiv.org*.

<sup>171</sup> ACM TECHNOLOGY POLICY COUNCIL, *Computing and Climate Change*, in *TechBrief*, 2021, I, p. 1-2.

<sup>172</sup> Si veda Part. 69, § 2, lett. b) *AI Act* (codici di condotta volontari); Part. 84, § 4 *AI Act* (incoraggiamento alla Commissione a includere la sostenibilità ambientale dei sistemi di IA nei codici di condotta volontari nella sua prima revisione tre anni dopo la data di applicazione dell'*AI Act*).

<sup>173</sup> J. COWLS, A. TSAMADOS, M. TADDEO e L. FLORIDI, *The AI gambit: leveraging artificial intelligence to combat climate change- opportunities, challenges, and recommendations*, in *AI & Society*, 2023, 38, I, p. 283-307.

In particolare, durante la fase di modellazione, gli sviluppatori potrebbero dover confrontare diversi tipi di modelli (ad esempio, regressione lineare o apprendimento profondo) non solo in relazione alle loro prestazioni, ma anche all'impronta climatica stimata. Perciò, davanti a modelli che presentano prestazioni simili, gli sviluppatori sarebbero obbligati, in base alla nuova disposizione, a scegliere il modello più sostenibile (una sorta di *sustainability by design*).

## Conclusioni

In conclusione, dopo aver fatto un giro di boa, può dirsi che, ad oggi, i sistemi di IA, anche quelli denominati forti o ad alto rischio, sono qualificabili come «oggetto del diritto» e, perciò, la responsabilità per danno da loro applicazione non può che appuntarsi sull'essere umano, sia che le abbia offerto in dono la scintilla vitale, sia che la segua nel suo sviluppo, distribuzione, utilizzo.

Si sono esaminati i potenziali paradigmi applicabili di matrice domestica e la normativa europea *incoming* che, a latere della responsabilità dell'operatore economico per difettosità del prodotto, sembra optare per una armonizzazione minima di strumenti procedurali che allevino l'onere della prova in un contesto di responsabilità per colpa del soggetto che, in vario modo e sotto vari ruoli, cagiona un danno «per mano» di un sistema di IA.

L'Europa, tuttavia, al momento – in vista della futura esecuzione della I fase della direttiva – lascia agli Stati la spinosa possibilità di scegliere qualsivoglia paradigma di responsabilità si ritenga migliore per la fattispecie in discussione. Conseguentemente, spetterà nuovamente al nostro ordinamento italiano farsi carico della scelta ottimale.

Di talchè, ove ritenesse più appropriato, nonostante le varie criticità evidenziate, far operare, indipendentemente dal livello di rischio del sistema di IA, la clausola generale della responsabilità extracontrattuale per colpa ex art. 2043 c.c., si potrebbe attingere all'intero bacino di presunzioni relative predisposto dalla proposta *AILD*, nella speranza che, così come delineate, possano agevolare la posizione del danneggiato in via effettiva.

Diversamente sarebbe qualora si ritenesse più congruo, almeno per le IA più complesse, l'utilizzo di uno «schema» di responsabilità oggettiva o, comunque, aggravata per presunzione di colpa. Di fatti, in tal casi non vi sarebbe modo di ricorrere alle agevolazioni probatorie della *AILD*, applicandosi queste unicamente nei regimi di responsabilità per colpa. Unica eccezione – forse e secondo l'interpretazione di solo alcuni studiosi<sup>1</sup> – sarebbe il ricorso, di fronte ai sistemi di IA ad alto rischio, all'ordine di divulgazione da parte del giudice di cui all'art. 3, § 1 *AILD*.

L'utilizzo di regole di responsabilità più «stretta» per le applicazioni di IA forti evidenzerebbe anche una scelta da parte della società in termini di scopo attribuito all'istituto della responsabilità civile. Così optando, prevarrebbe, probabilmente, la convinzione che la responsabilità sia uno strumento di controllo del rischio e che, per i nuovi rischi, si debba preferire la sua versione oggettiva sol si consideri che la stessa, non solo influenzerebbe il livello di diligenza, ma anche il livello di attività, internalizzando completamente i rischi

---

<sup>1</sup> P. HACKER, *The European AI liability directives*, cit. p. 10.

economici dell'IA<sup>2</sup>. Specialmente se combinata con l'assicurazione obbligatoria<sup>3</sup> per la responsabilità civile, la responsabilità oggettiva opererebbe, quindi, come strumento di distribuzione del rischio.

Resterebbe fermo, tuttavia, il cd. «effetto agghiacciante» di tale regime dal momento che la prospettiva di essere ritenuti strettamente responsabili (si pensi ai produttori di *software*) potrebbe rallentare o addirittura bloccare lo sviluppo tecnologico e l'innovazione.

Perciò, apparirebbe preferibile – come anticipato nel capitolo III, § 3.1 – un quadro di responsabilità più bilanciato, differenziato, graduato sulla base del rischio, nonché settorializzato. Dovrebbero, così, convivere due approcci regolamentari: quello orizzontale, in grado di normare, da un punto di vista generale, tutti i sistemi di IA raggruppandoli per livello di rischio<sup>4</sup> (sociale ed individuale); e quello verticale, predisponente norme ad *hoc* per i più rilevanti settori di impiego di sistemi di IA<sup>5</sup>

Per i sistemi di IA ad alto rischio – come già affermato sempre nel capitolo III – si riterrebbe confacente il paradigma di responsabilità per attività pericolose che combina la rigidità maggiore del criterio di imputabilità ad una valutazione di rimproverabilità soggettiva sotto le spoglie della prova contraria. Diversamente dicasi per i sistemi di IA a rischio inaccettabile il cui utilizzo farebbe «meritare» ad i suoi sviluppatori, fornitori e distributori, una responsabilità oggettiva *strictu sensu*.

---

<sup>2</sup> H. ZECH, *Liability for AI: Public Policy Considerations*, in *ERA Forum*, 2021, 22, p. 152. L'a. ritiene che la responsabilità oggettiva incentiva anche l'ulteriore sviluppo delle tecnologie esistenti e, probabilmente, ne favorisce l'accettazione da parte del pubblico.

<sup>3</sup> L'idea è stata, tra l'altro, accolta dal considerando 21 della Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale (2020/2014(INL)) ove si legge che «le assicurazioni possono contribuire a garantire che le vittime ricevano un risarcimento congruo e che i rischi di tutti gli assicurati siano condivisi. Uno dei fattori su cui le compagnie di assicurazioni basano l'offerta di prodotti e servizi assicurativi è la valutazione del rischio, che è possibile effettuare attraverso l'accesso a dati storici sufficienti sulle richieste di risarcimento. L'impossibilità di accedere a dati di alta qualità oppure una quantità insufficiente di tali dati potrebbe essere uno dei motivi per cui, all'inizio, la creazione di prodotti assicurativi per le tecnologie nuove ed emergenti risulta difficoltosa. Tuttavia, l'ottimizzazione dei dati generati dalle nuove tecnologie e un migliore accesso ad essi, insieme all'obbligo di fornire informazioni ben documentate, rafforzerebbero la capacità degli assicuratori di modellare il rischio emergente e di promuovere lo sviluppo di coperture assicurative più innovative». L'intenzione è stata, poi, confermata anche dalla Proposta di Direttiva sulla responsabilità da intelligenza artificiale nel considerando 31 ove si ribadisce essere necessario, in quella che viene annunciata come 'seconda fase della direttiva', l'introduzione di «norme in materia di responsabilità oggettiva per le azioni avviate contro l'operatore di un sistema di IA, purché non siano già disciplinate da altre norme dell'Unione in materia di responsabilità, in particolare dalla direttiva 85/374/CEE, in combinazione con un'assicurazione obbligatoria per il funzionamento di determinati sistemi di IA, come suggerito dal Parlamento europeo [...]».

<sup>4</sup> Approccio seguito dalle riforme europee.

<sup>5</sup> L'IA è e sarà utilizzata in molti campi, tra cui i trasporti, la diagnosi medica, i mercati dei capitali, i prodotti e i servizi di consumo, la produzione industriale, la produzione e la distribuzione di energia. Considerando che anche gli aspetti della responsabilità civile sono in gran parte regolati separatamente, dovrebbero continuare a essere regolati separatamente quando vengono implementate soluzioni basate sull'IA. Un tale approccio alla regolamentazione dell'IA, specifico per tecnologia e applicazione, sarebbe anche più conforme ai principi di proporzionalità e sussidiarietà, riducendo al minimo i rischi di interferenze indesiderate con gli ordinamenti giuridici degli Stati membri.

Infine, per i sistemi a basso rischio si riterrebbe sufficiente l'utilizzo del paradigma di responsabilità per colpa accompagnato da agevolazioni probatorie, differenziate a seconda dei convenuti, per far fronte alle difficoltà generate dalla tecnica.

Ad ogni modo, qualsiasi paradigma lo Stato italiano opzionerà, per il tramite del Legislatore o della magistratura che affronterà per prima i casi pratici, ciò che qui preme sottolineare è il consolidamento di una tendenza che si può ravvisare se si sposta il fuoco dell'attenzione sulla funzione della dell'istituto della responsabilità civile<sup>6</sup>.

Come noto, individuare rigidamente la funzione della responsabilità civile è atteggiamento aprioristico e dommatico.

Come la migliore dottrina afferma<sup>7</sup> «la responsabilità civile ha – secondo il principio di differenziazione del quale v'è traccia nel riformato art. 118 cost – non un'unica funzione, ma una pluralità di funzioni (preventiva, compensativa, sanzionatoria, punitiva) che possono anche coesistere».

Di conseguenza, natura e funzione del rimedio risarcitorio mutano non solo nelle singole disposizioni, ma anche, problematicamente, da settore a settore. Ai diversi criteri di collegamento della responsabilità, legati alle funzioni delle singole fattispecie, perciò, corrisponde anche una pluralità di soluzioni rimediali.

---

<sup>6</sup> Secondo la concezione che ravvisa nella responsabilità extracontrattuale una tecnica di allocazione dei danni tale istituto non deve essere considerato come uno strumento repressivo che colpisce un atto vietato, ma uno strumento di riequilibrio economico del danno. Cfr. C.M. BIANCA, *Diritto civile*, V, cit., p. 524-526. Già il G. VENEZIAN, *Studi delle obbligazioni*, in ID., *Opere giuridiche* (a cura della famiglia e R. ACCADEMIA DELLE SCIENZE DI BOLOGNA) I, Roma, 1919, p. 16 si era occupato del tema e aveva contestato la validità del principio della colpa, quale principio di imputabilità morale del danno (concezione etica della colpa), sottolineando la necessità che il torto avesse comunque una sanzione a prescindere dalle condizioni soggettive del suo autore. Si v., altresì, R. SCOGNAMIGLIO, *Illecito (diritto vigente)*, NDI, VIII, Torino, 1962, p. 164 il quale, respinta la teoria che riconduce la responsabilità civile alla prospettiva dell'illecito, afferma che «la concezione più attendibile [...] è quella che fa capo all'idea della reazione al danno ingiusto». Il legislatore affronterebbe «il tema della ripartizione [...] dei danni che nell'ambito della vita di relazione [...] si possono verificare [...] attraverso l'obbligo di risarcimento [...] alla stregua dei criteri di volta in volta accolti: ed a prescindere [...] dalla finalità di proibire e colpire, attraverso un tale rimedio, l'attività che genera l'evento dannoso».

All'interno di tale concezione rileva il filone che fonda i criteri di allocazione dei danni sull'analisi economica del diritto. Si fa riferimento, innanzitutto, a G. CALABRESI, *Costo degli incidenti e responsabilità civile. Analisi economico-giudiziaria* (tr.it.), Milano, 1975, p. 47 il quale ravvisa la preminente funzione della responsabilità civile nella riduzione dei costi, funzione che viene distinta in tre fini: «il primo consiste nella riduzione del numero e della gravità dei sinistri [...] il secondo fine [...] mira [...] a ridurre il costo che gli incidenti causano alla società. Il terzo infine [...] mira a ridurre i costi di esercizio di un qualunque sistema di responsabilità civile». Al perseguimento di tali fini vengono, unicamente, ritenuti adeguati i metodi fondati sul controllo di mercato e sulla ricerca di soluzioni economiche ottimali. Per quanto attiene, in particolare, alla rimozione dei costi secondari (il secondo fine), l'a. prospetta il metodo del frazionamento dei danni (che è alla base dello strumento assicurativo) e della sua variante espressa dalla concezione della «tasca profonda»: «il metodo migliore per ridurre i costi secondari dei sinistri è di trasferirli su quella categoria di persone, la cui posizione sociale ed economica meno ne risentirebbe, su quelli, cioè, che [...] si sogliono definire "ricchi"».

<sup>7</sup> G. CIAN, *Antigiuridicità e colpevolezza. Saggio per una teoria dell'illecito civile*, Padova, 1966, p. VI. L'a. ammonisce il lettore di non dimenticare che un sistema, il quale si fondi sul presupposto della colpevolezza, e guardi all'obbligo del risarcimento non soltanto come a un mezzo di riparazione ma anche come a un male minacciato al trasgressore, costituisce, con il suo trattamento differenziato dei soggetti colpevoli e non colpevoli, un valido strumento di giustizia per indurre i singoli quella condotta accorta e prudente che può esimerli da responsabilità, e rappresenta perciò sempre un forte mezzo di tutela preventiva.

Sta alla dottrina fornire indicazioni, soffermandosi sulla «gravità» della lesione dei casi concreti, circa la funzione che l'istituto della responsabilità assume nel caso specifico e nell'ottica dell'intero sistema ordinamentale<sup>8</sup>.

Orbene, quanto detto vale anche nel campo dell'intelligenza artificiale. Tuttavia, nel settore specifico sembra di star assistendo ad un consolidamento della funzione deterrente dell'istituto della responsabilità in virtù di una profonda risemantizzazione della parola «responsabilità», sollecitata dal proposto *AI Act* e dalle collegate direttive<sup>9</sup>.

Accogliendo un'idea larga di «rispondere», pare ci si stia spostando, infatti, dalla «responsabilità» alla «responsabilizzazione» (o *accountability*) dei soggetti della catena del valore dei sistemi di IA dal momento che il *focus* dell'intera disciplina si poggia più sull'aspettativa (lesa) di un certo comportamento che sulla rilevanza delle conseguenze in termini di danno.

È proprio il principio di «accountability»<sup>10</sup> che genera una serie di aspettative rispetto all'adozione di determinate condotte. I soggetti coinvolti nell'applicazione dell'IA – e, quindi, potenzialmente chiamati a rispondere – devono sempre adottare di *default*, infatti, tutta una serie di misure giuridiche, organizzative, tecniche e di sicurezza tese ad evitare danni, tipiche ed atipiche. Perciò, «il punto prospettico si colloca “ex ante”»<sup>11</sup> e viene in luce una responsabilità di tipo organizzativo fisiologica<sup>12</sup>. In altre parole, è come se venisse riposizionata la linea del «rispondere» ad un momento anteriore.

Tale trasformazione della responsabilità civile è stata già oggetto di osservazioni da parte degli studiosi sin dal momento in cui l'idea di «rischio» è stata ritenuta centrale. Tuttavia, oggi, le riflessioni su una diversa impostazione di responsabilità – che secondo alcuni<sup>13</sup> richiama le logiche della responsabilità contrattuale (si pensi alla costruzione della prova liberatoria nei termini di «adozione di tutte le misure idonee a evitare il danno») – sono suffragate da norme, per lo più tutte europee.

In definitiva, oggi, il tema della gestione delle conseguenze dannose da applicazione di un sistema di IA sembra essere affrontato attingendo ad un modello di responsabilizzazione basato sulla «*compliance*», il quale trova la sua

---

<sup>8</sup> P. PERLINGIERI, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, IV, 4<sup>ed.</sup>, *Attività e responsabilità*, Napoli, 2020, p. 324-326.

<sup>9</sup> Tali sono le parole di T. DELLA MASSARA nell'intervista rilasciata il 4 agosto 2023 a *Dimt.it*, dal titolo *Intervista al Prof. Avv. Tommaso della Massara. Il caso ChatGPT: tra vecchi e nuovi paradigmi della responsabilità*.

<sup>10</sup> Tale principio non è nuovo al diritto europeo. Lo stesso informa altre discipline. Si pensi al *GDPR* (Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016), *Digital Service Act* (Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE, *Data Governance Act* (Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724).

<sup>11</sup> v. nota 9.

<sup>12</sup> G. COMANDÉ, *Multilayered (Accountable) Liability for Artificial Intelligence*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, 4<sup>ed.</sup>, Baden-Baden, 2019, p. 165-183. L'a. suggerisce che l'IA richiede un approccio graduale e stratificato alla responsabilità basato sui principi di responsabilità (già incorporati nei sistemi giuridici dell'UE). L'IA richiede l'uso della tecnologia stessa per dispiegare un sistema di responsabilità multilivello e risolvere il «problema di dividere il conto» (p. 177).

<sup>13</sup> *Ibidem*.

ragion d'essere nell'intendimento di orientare i comportamenti degli operatori del mercato verso determinati obiettivi, tra cui l'affidabilità e l'eticità dell'IA.

## Bibliografia

- N. ABRIANI e G. SCHNEIDER, *Diritto delle imprese e intelligenza artificiale. Dalla Fintech alla Corpotech*, Bologna, 2021, p. 21 ss.
- E. AL MUREDEN, *La sicurezza dei prodotti e la responsabilità del produttore*, Torino, 2017.
- A. ALBANESE, *La sicurezza generale dei prodotti e la responsabilità del produttore nel diritto italiano ed europeo*, in *Egr. dir. priv.*, 2005, p. 977 ss.
- A. ALBANESE, *La responsabilità civile per i danni da circolazione di veicoli ad elevata automazione*, in *Contr. impr.*, 2020, p. 995 ss.
- G. ALPA, U. CARNEVALI, F. DI GIOVANNI, G. GHIDINI, U. RUFFOLO e C.M. VERARDI, *La responsabilità per danno da prodotti difettosi*, Milano, 1990.
- G. ALPA e M. BESSONE, *La responsabilità civile*, Milano, 2001.
- G. ALPA (a cura di), *La responsabilità del produttore*, Milano, 2019.
- G. ALPA, *Quale modello normativo europeo per l'intelligenza artificiale?*, in V. CUOCCI, F.P. LOPS e C. MOTTI (a cura di), *La responsabilità civile nell'era digitale*, Bari, 2022, p. 3 ss.
- E. ALPAYDIN, *Machine learning. The new AI*, Cambridge, MA, 2016.
- S. AMATO, *Il soggetto e il soggetto di diritto*, Torino, 1990.
- A. AMIDEI, *Intelligenza artificiale e diritto – Intelligenza artificiale e product liability: sviluppi del diritto dell'Unione Europea*,” *Giur. It.*, 2019, VII, p. 1657 ss.
- A. AMIDEI, *Le responsabilità da intelligenza artificiale tra product liability e sicurezza del prodotto*, in U. RUFFOLO (a cura di), *XXVI lezioni di Diritto dell'Intelligenza Artificiale*, Torino, 2021, p. 149 ss.
- A. AMIDEI, *La proposta di Regolamento UE*, in *Tecn. Dir.*, 2022, I, p. 26,
- A. AMIDEI, *Il (futuro) nuovo volto della responsabilità da prodotto difettoso*, in *Tecn. Dir.*, 2023, I, p. 27- 53.
- M. ANDERSON, S.L. ANDERSON, *Machine ethics*, Cambridge, 2011.
- J. ARMOUR e H. EIDENMÜLLER, *Self-Driving Corporations?*, in *Harvard Business L. Rev.*, 2020, X, p. 87 ss.
- P.M. ASARO, *Robots and responsibility from a legal perspective*, in *Proceedings of the IEEE*, 2007, p. 20- 24.
- G. AVANZINI, *Decisioni amministrative e algoritmi informatici. Predeterminazione analisi predittiva e nuove forme di intellegibilità*, Napoli, 2020, p. 3-13.
- M. BARCELLONA, *Trattato della responsabilità civile*, Torino, 2011.
- M. BASILE, «Persona giuridica (diritto privato)», in *Enc. dir.*, XXXIII, Milano, 1983.
- A. BERTOLINI, *Robots as products the case for a realistic analysis of robotic applications and liability rules*, in *Law Innov. Technol.*, 2013, vol. 5, 2, p. 214-247.
- A. BERTOLINI, *Robot companions: a legal and ethical analysis*, in *The Information Society*, 2018, vol. 34, 3, p. 130-14.
- A. BERTOLINI, *Artificial Intelligence and Civil Liability*, in *Publications Office of the European Union*, 2020.

- A. BERTOLINI, *Artificial intelligence does not exist! Defying the technology-neutrality narrative in the regulation of civil liability for advanced technologies*, in *Europa dir. priv.*, 2022, II, p. 370 ss.
- L. BERTUZZI, *Intelligenza Artificiale intelligenza artificiale definizione, governance*, in *euractiv.com*, 2022.
- C.M. BIANCA, *Diritto civile*, V, *La responsabilità*, Milano, 2021.
- N. BOBBIO, *Sulla funzione promozionale del diritto*, in *Riv. trim.*, 1969, p. 1313 ss.
- J.J. BRYSON, M.E. DIAMANTIS, T.D. GRAND, *Of, for, and by the people: the legal lacuna of synthetic persons*, in *Artif. Intell. Law*, 2017, vol. 25, 3, p. 273-291.
- F.D. BUSNELLI, *voce Illecito civile*, in *Enc. giur. Treccani*, Roma, 1989, p. 33 ss.
- M. BUSSANI, *L'illecito civile*, in *Tratt. dir. civ. CNN*, diretto da P. Perlingieri, Napoli, 2020.
- G. CALABRESI, *Costo degli incidenti e responsabilità civile, Analisi economico-giudiziarica* (tr.it), Milano, 1975.
- D.J. CALVERLEY, *Android science and animal rights: does an analogy exist?*, in *Connect. Sci.*, 2006, vol. 18, 4, p. 403-417.
- C. CAMARDI, *Prime osservazioni sulla Direttiva (UE) 2019/770 sui contratti per la fornitura di contenuti e servizi digitali. Operazioni di consumo e circolazione di dati personali*, in *Giust. civ.*, 2019, p. 499 ss.
- U. CARNEVALI, *Produttore e responsabilità per danno da prodotto difettoso nel codice del consumo*, in *Resp. civ. prev.*, 2009, 10, p. 1938 ss.
- U. CARNEVALI (a cura di), *Commentario del Codice civile. Dei fatti illeciti, I, Art. 2043 - art. 96 c.p.c.*, Torino, 2011.
- F. CARINGELLA e L. BUFFONI, *Manuale di diritto civile*, VI<sup>ed.</sup>, Roma, 2016.
- G. CARULLO, *Decisione amministrativa e intelligenza artificiale*, in *Dir. Inf. Inform.*, 2021, III, p. 431-461.
- D. CARUSO e R. PARDOLESI, *Per una storia della Direttiva 1985/374/CEE*, in *Danno resp.*, 2012, p. 9-17.
- T. CASADEI e S. PIETROPAOLI, *Intelligenza artificiale: fine o confine del diritto?*, in T. CASADEI e S. PIETROPAOLI (a cura di), *Diritto e tecnologie informatiche. Questioni di informatica giuridica, prospettive istituzionali e sfide sociali*, Trento, 2021, p. 219-232.
- C. CASTRONOVO, *La nuova responsabilità civile*, 3<sup>ed.</sup>, Milano, 2006.
- C. CASTRONOVO, *La responsabilità civile*, Milano, 2018.
- P. CAVALIERI, *La questione animale: perché gli animali non umani meritano i diritti umani*, Oxford, 2001.
- C. CHAVALIER, *Fictio e natura: due metodi alternativi per la realizzazione dell'aequitas*, in *Teoria e storia del dir. priv.*, XIV, 2021, p. 1- 60.
- G. CHINÈ e A. ZOPPINI, *Manuale di diritto civile*, IX<sup>ed.</sup>, Molfetta, 2020.
- A. CHITI, *Explainable AI: come andare oltre la black box degli algoritmi*, in *agendadigitale.eu*, 2022.
- S. CHOPRA e L.F. WHITE, *A Legal Theory for Autonomous Artificial Agents*, Ann Arbor, 2011.

- G. CIAN, *Antigiuridicità e colpevolezza. Saggio per una teoria dell'illecito civile*, Padova, 1966.
- G. COMANDÉ, *Multilayered (Accountable) Liability for Artificial Intelligence*, in S. LOHSSE, R. SCHULZE e D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, 4<sup>ed.</sup>, Baden-Baden, 2019, p. 165-183.
- G. COMANDÉ, *Il "dividendo tecnologico" della responsabilità civile, prospettive sul contenzioso*, in *Danno Resp. civ.*, 2021, III, p. 269-282.
- M. COMPORZI, *Esposizione al pericolo e responsabilità civile*, Napoli, 1965.
- M. COMPORZI, *Fatti illeciti: le responsabilità oggettive*, in *Cod. civ. Comm.* Schlesinger Busnelli, Milano, 2009
- A. CONDELLO, *Metodo giuridico e algoritmo*, in *Riv. crit. dir. priv.*, 2020, 3, p. 339-368.
- L. COPPINI, *Robotica ed intelligenza artificiale: questioni di responsabilità civile*, in *Il Mulino*, 2018, p. 728;
- L. CORSARO, *Responsabilità civile I, Diritto civile*, in *Enc. giur. Treccani*, XXVI, 1991.
- L. CORSARO, *Responsabilità da cose*, in *Dig. disc. priv., Sez. civ.*, XVII, 1998, Torino, p. 103 ss.
- M. COSTANZA, *L'intelligenza Artificiale e gli stilemi della responsabilità civile*, in *Giur. it.*, 2019, VII, p. 1688.
- S. COTTA, *Soggetto di diritto*, in *Enc. dir.*, XLII, Milano, 1990.
- G. D'ACQUISTO, *Intelligenza artificiale*, Torino, 2021.
- G. D'ALFONSO, *Il regime di responsabilità da cose in custodia tra questioni tradizionali e "responsabilità da algoritmo"*, in *EJPLT*, 2022, I, p. 1 ss.
- K. DARLING, *Extending legal protection to social robots: the effects of anthropomorphism, empathy, and violent behavior towards robotic objects*, in R. CALO, M.A. FROOMKIN e I. KERR (a cura di), *Robot law*, Northampton, 2016, p. 213-231.
- J. DE BRUYNE, E. VAN GOOL e T. GILS, *Tort law and Damage Caused by AI Systems*, J. DE BRUYNE, C. VANLEENHOVE (a cura di) in *Artificial Intelligence and the law*, Cambridge, 2023, p. 395- 448.
- A. DE CUPIS, *Il danno. Teoria generale della responsabilità civile*, Milano, 1954.
- A. DE FRANCESCHI, *La vendita di beni con elementi digitali*, Napoli, 2019.
- G. DE ROSA, *Quali regole per i sistemi automatizzati "intelligenti"?*, in *Riv. dir. civ.*, 2021, p. 842 ss.
- R. DE RUGGIERO, *voce Finzione legale*, in *Diz. prat.*, 3, Milano, 1923, p. 184 ss.
- O. DHEU, J. DE BRUYNE e C. DUCUING, *The European Commission's approach to extra-contractual liability and AI—An evaluation of the AI liability directive and the revised product liability directive*, in *Comp. L. Sec. Rev.*, 2023, I, p. 51.
- O. DHEU e J. DE BRUYNE, *Artificial Intelligence and Tort Law: A 'Multi-faceted' Reality*, in *Eu. Rev. Priv. Law*, 2023, II e III, p. 261-298.
- O. DHEU e J. DE BRUYNE, *Liability for Damage Caused by Artificial Intelligence—Some Food for Thought and Current Proposals*, in *Tort Liability and Autonomous Systems Accidents*, Cheltenham, 2023, p. 27-62.
- F. DI CIOMMO, *Evoluzione tecnologica e regole di responsabilità civile*, Napoli, 2003.

- L. DI DONNA, *Intelligenza artificiale e rimedi risarcitori*, Milano, 2022, p. 21 ss.
- V. DI GREGORIO, *Intelligenza artificiale e responsabilità civile: quale paradigma per le nuove tecnologie?*, in *Danno resp.*, 2022, p. 51 ss.
- F. DI MARZIO, *Codice della responsabilità civile commentato*, Milano, 2017.
- R. ESPOSITO, *Le persone e le cose*, Torino, 2014.
- E. ESPOSITO, *The future of prediction: from statistical uncertainty to algorithmic forecasts*, in *artificialcommunication.mitpress.mit.edu*, 2022, p. 87 ss.
- A. FALZEA, voce *Capacità (teoria generale)*, in *Enc. dir.*, VI, Milano, 1960.
- A. FALZEA, *Il civilista e le sfide del millennio*, in Aa.Vv., *Studi in onore di Pietro Schlesinger*, I, Milano, 2004, p. 15 ss.
- M. FARINA, *Brevi riflessioni sullo status delle "persone elettroniche"*, in *L'Ircocervo*, 2021, 2, p. 106 ss.
- G. FASANO, *L'intelligenza artificiale nella cura dell'interesse generale*, in *Giornale dir. amm.*, 2020, VI, p. 724.
- F. FERRARA JR., *Le persone giuridiche*, Torino, 1956.
- M. FERRARI, *Il vantaggio della responsabilità concorsuale da uso "organizzato" di algoritmi*, in *Il Mulino*, 2021, II, p. 424 ss.
- F. FIDANZA, *Sulla distinzione tra intelligenza artificiale e algoritmi*, in *Riv. dir. impr.*, 2022, II, p. 389- 406.
- G. FINOCCHIARO, *Intelligenza artificiale e responsabilità*, in *Contr. e impresa*, 2020, II, p. 716 ss.
- L. FLORIDI, *L'estensione dell'intelligenza. Guida all'informatica per filosofi*, Roma, 1996.
- E. FRANCESCONI, *The winter, the summer and the summer dream of artificial intelligence in law. Presidential address to the 18th International Conference on Artificial Intelligence and Law*, in *Artificial Intelligence and Law*, 30, 2022, p. 147–161.
- M. FRANZONI, *La responsabilità oggettiva. Il danno da cose e da animali*, Padova, 1988.
- M. FRANZONI, *Dei fatti illeciti. Artt. 2043-2059*, in *Commentario c.c. Scialoja-Branca-Galgano*, 2<sup>a</sup> ed., Bologna, 2020.
- F. FRATTARI, *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in *Contr. impr.*, 2020, I, p. 452 ss.
- V. FROSINI, *Soggetto del diritto*, in *Noviss. dig. it.*, XVII, Torino, 1970, p. 813 ss.
- M. GAL, *Algorithmic Challenges to Autonomous Choice*, in *Michigan Telecom. Tech. Law Rev.*, 2018, 25, p. 59.
- F. GALGANO, *Delle persone giuridiche*, in *Comm. Scialoja e Branca*, Bologna-Roma, 1969.
- F. GALGANO, voce «*Persona giuridica*», in *Dig. disc. priv., Sez. civ.*, XIII, Torino, 1995.
- A. GAMBARO, voce *Finzione giuridica nel diritto positivo*, in *Dig. disc. priv., Sez. civ.*, 8, Torino, 1992.
- M. GEISTFELD, E. KARNER, B.A. KOCH e C. WENDERHOSTE, *Civil Liability for Artificial Intelligence and Software*, in *Tort and Insurance law*, 37, Vienna, 2023.
- E. GIORGINI, *Algorithms and Law*, in *ItaLJ*, 2019, I, p. 131-149.
- A. GORASSINI, *Il valore della cultura giuridica digitale*, in *Tecn. Dir.*, 2020, p. 38 ss.

- J-S. GORDON, *Remarks on a disability-conscious bioethics*, in J-C. PÖDER, H. BURCKHART e J- S. GORDON (a cura di), *Human rights and disability. Interdisciplinary Perspectives*, London, 2017, p. 9- 20.
- J-S. GORDON, *What do we owe to intelligent robots?*, in *AI Soc.*, 2020, vol. 35, 2, p. 209-223.
- J-S. GORDON e F. TAVERA-SALYUTOV, *Remarks on disability rights legislation*, in *Equal. Divers. Incl.*, 2018, vol. 37, 5, p. 506-526.
- J-S. GORDON, O. RAUPRICH e J. VOLLMAN, *Applying the four principles approach*, in *Bioethics*, 2011, vol. 25, 6, p. 293-300.
- D. GUNKEL, *The machine question: critical perspectives on AI, robots, and ethics*, Cambridge Mass., 2012.
- D. GUNKEL, *A vindication of the rights of machines*, in *Philos. Technol.*, 2014, vol. 27, 1, p. 113-132.
- D. GUNKEL, *Robot rights*, Cambridge Mass., 2018.
- D.J. GUNKEL, J. BRYSON, *Introduction to the special issue on machine morality: the machine as moral agent and patient*, in *Philos. Technol.*, 2014, vol. 27, 1, p. 5-142.
- P. HACKER, *The European AI liability directives – Critique of a half-hearted approach and lessons for the Future*, in *Comp. law sec. rev.*, 2023, 51, p. 1 ss.
- P. HACKER, *Sustainable AI Regulation, Working Paper* presentato alla Privacy Law Scholars Conference, 2023, [arxiv.org](https://arxiv.org).
- G. HALLEVY, *The criminal liability of artificial intelligence entities*, in *ssrn.com*, 2010, p. 1- 42.
- G. HALLEVY, *Virtual criminal responsibility*, in *ssrn.com*, 2010, p. 1-22.
- G. HALLEVY, *When robots kill. Artificial intelligence under criminal law*, Boston, 2013.
- E. HAHANA, *The AI Utility Levels Schema – Building an AI Classification*, in [law.stanford.edu](https://law.stanford.edu), 2022.
- M. HILDEBRANDT, *Criminal liability in a smart environment*, in R. DUFF e S.P. GREEN (a cura di) *Philosophical foundations of criminal law*, Oxford, 2011, p. 507-532.
- F.C. HOLLET, *Deep learning with Python*, 2<sup>ed.</sup>, Shelter Island, New York, 2021, p. 7.
- G. HOWELLS, C. TWIGG-FLESNER e C. WILLETT, *Protecting the values of consumer law in the digital economy the case of 3D-printing*, in A. DE FRANCESCHI e R. SCHULZE (a cura di), *Digital revolution – Neu challenges for Law*, Monaco, 2019, p. 224 ss.
- D. KIM, *Artificial Intelligence Should not Become a “Black Hole” for Human Agency in Tort Law*, in *Tort Law Rev.*, 2023, 29, p. 152 ss.
- J. KNETSCH, *Are existing tort theories ready for AR*, in L.A. DI MATTEO, C. PONCIBÒ e M. CANNARSA (a cura di), *The Cambridge Handbook of Artificial Intelligence. Global perspective on law and ethics*, Cambridge, 2022, p. 108 ss.
- B.A. KOCH, *Product liability 2.0 Mere update or new version?*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things*, Baden-Baden, 2019, p. 99 ss.

- B.J. KOOPS, M. HILDEBRANDT e D.O. JAQUET-CHIFFELLE, *Bridging the accountability gap: rights for new entities in the information society?*, in *Minnesota J. Law Sci. Technol.*, 2010, vol. 11, 2, p. 497-561.
- V.A.J. KURKI e T. PIETRZYKOWSKI (a cura di), *Legal personhood: animals, artificial intelligence and the unborn*, Berlin, 2017.
- T.L. JAYNES, *Legal personhood for artificial intelligence: citizenship as the exception to the rule*, in *AI Soc.*, 2020, vol. 35, 2, p. 343 ss.
- B. LATOUR, *Reassembling the Social. An Introduction to Actor-Network-Theory*, Oxford, 2005, p. 21 ss.
- J. LAUX, S. WACHTER e B. MITTELSTADT, *Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk*, in *ssrn.com*, 2022.
- N. LIPARI, *Le categorie del diritto civile*, Milano, 2013.
- Z.C. LIPTON, *Il mito dell'interpretabilità dei modelli: In machine learning, the concept of interpretability is both important and slippery*, in *researchgate.net*.
- G. LO SAPIO, *La black box: l'esplicabilità delle scelte algoritmiche quale garanzia di buona amministrazione*, in *Federalismi.it*, 2021, 16, p. 136-149.
- S. LOHSE, R. SCHULZE e D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, 4<sup>ed.</sup>, Baden-Baden, 2019.
- A. LUPOLI, *La responsabilità per prodotti difettosi*, in F. BOCCHINO (a cura di), *Diritto dei consumatori e nuove tecnologie*, II, *Il mercato*, Torino, 2003, p. 39 ss.
- C. MAIORCA, *Responsabilità (teoria generale)*, in *Enc. dir.*, XXXX, 1988.
- C. MAIORCA, *I fondamenti della responsabilità*, Milano, 1990.
- G. MAIRA, *Intelligenza umana e intelligenza artificiale* in *Federalismi.it*, VII, 2021.
- A.W. JR. MACHEN, *Corporate Personality*, in *Harv. L. Rev.*, 1911, 24, p. 347-350.
- L. MENGONI, *Ancora sul metodo giuridico*, in *Riv. trim.*, 1984, p. 337 ss.
- L. MENGONI, *Diritto e tecnica*, in *Riv. trim.*, 2001, p. 1 ss.
- M.M. MOLLICONE, *Il rischio dell'intelligenza artificiale applicata. Modelli di allocazione a confronto*, in *AJI*, 2023, 18, p. 2108-2133.
- P.G. MONATERI, *La responsabilità civile*, in *Le fonti delle obbligazioni*, III, in *Tratt. dir. civ. Sacco*, Torino, 1998.
- P.G. MONATERI, *Responsabilità civile nel diritto comparato*, in *Dig. disc. priv., Sez. civ.*, XVII, Torino, 1998.
- M.L. MONTAGNANI, *Liability and emerging digital technologies: an EU perspective*, in V. FALCE (a cura di), *Fairness e innovazione nel mercato unico digitale*, Torino, 2020, p. 85 ss.
- R. MONTINARO, *Responsabilità da prodotto difettoso e tecnologie digitali tra soft law e hard law*, in *Pers. merc.*, 2020, p. 370 ss.
- T. MUNAKATA, *Fundamentals of the New Artificial Intelligence*, London, 2008.
- A. MURRAY, *Almost Human: law and human agency in the time of artificial intelligence*, The Hague, 2021.
- F. NADDEO, *Intelligenza artificiale: profili di responsabilità*, in *Comparazione e diritto civile*, 2020, III, p. 1150 ss.

- N. NAFFINE, *Who are law's persons? From Cheshire cats to responsible subjects*, in *Modern Law Review*, 2003, vol. 66, 3, p. 346-367.
- N. NAFFINE, *Law's meaning of life: philosophy, religion. Darwin and the legal person*, Oxford and Portland, 2009.
- N. NILSSON, *The quest for artificial intelligence. A history of ideas and achievements*, Cambridge, 2009.
- P.H. PADOVAN, C.M. MARTINS e C. REED, *Black is the new orange: how to determine AI liability*, in *A.I. and Law*, Berlino, 2023, p. 133-167.
- U. PAGALLO, *The laws of robots: Crimes, contracts, and torts*, Heidelberg, 2013.
- E. PALMERINI, voce "Robotica" (parte giuridica), in *Enciclopedia di bioetica e scienza giuridica*, X, Napoli, 2016, p. 1104.
- E. PALMERINI, *Soggettività e agenti artificiali: una soluzione in cerca di un problema?*, in *Oss. Dir. civ. comm.*, 2020, p. 464 ss.
- P. PARDOLESI, *Riflessioni sulla responsabilità da prodotto difettoso in chiave di analisi economica del diritto*, in *Riv. dir. priv.*, 2017, p. 87 ss.
- F. PASQUALE, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Roma, 2020.
- S. PATTI, *Probabilità e verosimiglianza nella disciplina del danno da prodotto*, in *Riv. dir. civ.*, 1990, p. 705 ss.
- C. PERLINGIERI, *Enti e diritti della persona*, Napoli, 2008.
- C. PERLINGIERI, *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in *Rass. dir. civ.*, 2015, IV, p. 1235 ss.
- C. PERLINGIERI, *Diritto privato delle nuove tecnologie: contenuti e competenze*, in *Tecn. dir.*, 2021, p. 83 ss.
- P. PERLINGIERI, *La personalità umana nell'ordinamento giuridico*, Camerino-Napoli, 1972.
- P. PERLINGIERI (a cura di), *Soggetti e norma, individuo e società*, Napoli, 1987.
- P. PERLINGIERI, *Valori normativi e loro gerarchia. Una precisazione dovuta a Natalino Irti*, in *Rass. dir. civ.*, 1999, p. 802 ss.
- P. PERLINGIERI, *Dialogando con due filosofi, ermenenti del diritto*, in *Rass. dir. civ.*, 2001, p. 669 ss.
- P. PERLINGIERI, *Interpretazione ed evoluzione dell'ordinamento*, in *Riv. Dir. Priv.*, 2011, p. 159 ss.
- P. PERLINGIERI, *Dogmatica giuridica e legalità costituzionale*, in *Annali S.i.s.d.i.c.*, 2019, 3, p. 1 ss.
- P. PERLINGIERI, *Il diritto civile nella legalità costituzionale secondo il sistema italo-europeo delle fonti*, 4<sup>a</sup> ed., Napoli, 2020.
- P. PERLINGIERI, *Presentazione*, in ID., S. GIOVA e I. PRISCO (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*, Atti del 15° Convegno Nazionale Sisdic, Napoli, 2020.
- P. PERLINGIERI, *Sul trattamento algoritmico dei dati*, in *Tecn. dir.*, 2020, p. 181 ss.
- D.L. POOLE e A.K. MACKWORTH, *Artificial Intelligence: Foundations of Computational Agents*, 2<sup>ed.</sup>, Cambridge, 2017, p. 3-7.

- A. PROCIDA MIRABELLI DI LAURO, *Le intelligenze artificiali tra responsabilità civile e sicurezza sociale*, Atti del 15° Convegno Nazionale Sisdic. Napoli 14-15-16 maggio 2020, Napoli, 2020, p. 297.
- S. PUGLIATTI, voce *Finzione* in *Enc. dir.*, XVII, Milano, 1968.
- M. RATTI, *Riflessioni in materia di responsabilità civile e danno cagionato da dispositivo intelligente alla luce dell'attuale scenario normativo*, in *Contr. impr.*, 2020, III, p. 1174-1191.
- P. RESCIGNO, *Capacità giuridica* (dir. civ.), in *Noviss. Dig. it.*, II, Torino, 1988.
- S.K. RIPKEN, *Corporations are people too: a multi-dimensional approach to the 1080 corporate personhood puzzle*, in *Fordham Journal of Corporate & Financial Law*, 2010, 15, p. 97- 177.
- S. RODOTÀ, *Il problema della responsabilità civile*, Milano, 1967.
- S. RODOTÀ, *Dal soggetto alla persona*, Napoli, 2007.
- U. RUFFOLO, *Intelligenza artificiale e responsabilità*, Milano, 2017.
- U. RUFFOLO, *Intelligenza artificiale e responsabilità. Responsabilità da "algoritmo"? A.I. e automobili self-driving, automazione produttiva, robotizzazione medico-farmacologica. A.I. e attività contrattuali. Le tendenze e discipline unionali*. Convegno del 29 novembre 2017, Università degli stranieri di Perugia, Milano, 2017
- U. RUFFOLO, *Intelligenza Artificiale e diritti della persona: le frontiere del 'transumanesimo'*, in *Giur. it.*, 2019, VII, p. 1658- 1670.
- U. RUFFOLO, *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 2019, p. 1689 ss.
- U. RUFFOLO, *Responsabilità da produzione e gestione di A.I. self-Learning*, in S. GIOVA e I. PRISCO (a cura di), *Rapporti civilistici e intelligenze artificiali: attività e responsabilità*, Atti del 15° Convegno Nazionale Sisdic Napoli 2020.
- S. RUSSEL e P. NORVIG, *Intelligenza artificiale in Intelligenza artificiale. Un approccio moderno*, Londra, 2021.
- C. SALVI, *Responsabilità extracontrattuale (diritto vigente)*, in *Enc. dir.*, XXXIX, 1998.
- C. SALVI, *La responsabilità civile*, in *Tratt. dir. priv.* Iudica e Zatti, 3<sup>ed</sup>, Milano, 2019.
- A. SANTOSUOSSO, C. BOSCARATO e F. CAROLEO, *Robot e diritto. Una prima ricognizione*, in *Nuova giur. civ. comm.*, 2012, II, p. 494-516.
- V. SCALISI, *Ermeneutica dei diritti fondamentali e principio «personalista» in Italia e nell'Unione europea*, in *Riv. dir. civ.*, 2010, I, p. 145 ss.
- V. SCALISI, *L'ermeneutica della dignità*, Milano, 2018.
- R. SCOGNAMIGLIO, *Responsabilità civile*, in *Noviss. Dig. it.*, XV, Torino, 1968.
- R. SCOGNAMIGLIO, *Illecito (diritto vigente)*, in *Noviss. Dig. it.*, VIII, 1962.
- B. SCHÜTTE, L. MAJEWSKI e K. HAVU, *Damages Liability for Harm Caused by Artificial Intelligence – EU Law in Flux*, in *Legal Studies Research Paper Series*, 2021, 69, p. 26.
- E. SCHAEERER, R. KELLEY e M. NICOLESCU, *Robots as animals: A framework for liability and responsibility in human-robot interaction. In Robot and human interaction communication. RO-MAN 2009—The 18th IEEE international symposium on robot and human interactive communication*, in *Journal Advanced Robotics*, 2009, vol. 24, 13, p. 1861-1871.

- R. SCOGNAMIGLIO, *Illecito (diritto vigente)*, in *Noviss. Dig. it.*, VIII, 1962.
- R. SCOGNAMIGLIO, *Responsabilità civile*, in *Noviss. Dig. it.*, XV, Torino, 1968.
- P. SINGER, *Speciesism and moral status*, in *Metaphilosophy*, 2009, vol. 40, 3-4, p. 567- 581.
- B. SMITH, *Legal personality*, in *Yale Law Journal*, 1928, vol. 37, 3, p. 283-299.
- S.M. SOLAIMAN, *Legal personality of robots, corporations, idols and chimpanzees: a quest for legitimacy*, in *Artif. Intell. Law*, 2017, vol. 25, 2, p. 155-179.
- L.B. SOLUM, *Legal personhood for artificial intelligences*, in *N. Carolina Law Rev.*, 1992, 70, p. 1231-1287.
- J.F. STAGL, *Da 'qualcosa' a 'qualcuno', da 'qualcuno' a 'qualcosa'. Percorsi esatti ed errati del concetto di persona*, in P. BUONGIORNO e S. LOHSSE (a cura di), *Fontes iuris. Atti del VI Jahrestreffen Junger Romanistinnen und Romanisten (Lecce, 30-31 marzo 2012)*, Napoli, 2013, p. 87 ss.
- J.F. STAGL, *Domani l'impersonale sarà il genere umano?*, in *REHJ*, 2016, XXXVIII, p. 377-391.
- J. TALLINN e R. NGO, *Automating Supervision of AI delegates*, in S. VOENEKY, P. KELLMEYER, O. MUELLER e W. BURGARD, (a cura di), *The Cambridge Handbook of Responsible Artificial intelligence. Interdisciplinary perspectives*, Cambridge, 2022, p. 19-30.
- G. TEUBNER, *Rights of Non-humans? Electronic agents and animals as new actors in politics and law*, Badia Fiesolana, 2007.
- G. TEUBNER, *Ibridi ed attanti. Attori collettivi ed enti non umani nella società e nel diritto*, trad. it. di L. Zampino, Milano, 2015.
- G. TEUBNER, *Soggetti giuridici digitali*, Napoli, 2019.
- P. TRIMARCHI, *Rischio e responsabilità oggettiva*, Milano, 1961.
- P. TRIMARCHI, *Illecito (diritto privato)*, in *Enc. dir.*, XX, Milano, 1970, p. 90.
- P. TRIMARCHI, *La responsabilità civile: atti illeciti, rischio, danno*, 3<sup>ed.</sup>, Milano, 2021
- R. TUR, *The "Person" in Law*, in A. PEACOCKE, G. GILLET (a cura di), *Persons and personality: a contemporary inquiry*, 1987, p. 116-121.
- R. VAN DEN HOVEN VAN DEN GARDEN, *Do We Need New Legal Personhood in the Age of Robots and AI?*, in M. CORRALES, M. FENWICK, N. FORGÓ (a cura di) *Robotics, AI and the Future of Law*, Singapore 2018, p. 15-55.
- G. VISINTINI, *Trattato breve della responsabilità civile: fatti illeciti, inadempimento, danno risarcibile*, XV, Padova, 2005.
- S. VOENEKY ET AL. (a cura di), *The Cambridge Handbook of Responsible Artificial Intelligence. Interdisciplinary Perspectives*, Cambridge, 2022.
- W. WALLACH e C. ALLEN, *Moral machines: Teaching robots right from wrong*, Oxford, 2009.
- E.M. WEITZENBOECK, *Electronic Agents and the Formation of Contracts*, in *International Journal of Law and Information Technology*, 2001, 3, p. 204-234.
- J. WROBLEWSKI, *Dilemmi dell'età tecnologica: il diritto e l'omeostasi dell'esistenza umana*, in F. RICCOBONO (a cura di), *Nuovi diritti dell'età tecnologica*, Milano, 1991, p. 195 ss.

- A. ZACCARIA, *La responsabilità del “produttore” di software*, in *Contr. impr.*, 1993, p. 303 ss.
- P. ZATTI, *Persona giuridica e soggettività*, Padova, 1975.
- H. ZECH, *Liability for AI: Public Policy Considerations*, in *ERA Forum*, 2021, 22, p. 152.
- P. ZELLINI, *La matematica degli dei e gli algoritmi degli uomini*, Milano, 2016.
- G. WAGNER, *Robot liability*, in S. LOHSSE, R. SCHULZE e D. STAUDENMAYER (a cura di), *Liability for Artificial Intelligence and the Internet of Things: Munster Colloquia on Eu Law and the Digital Economy*, 4<sup>ed.</sup>, Baden-Baden, 2019, p. 41 ss.
- E.M. WEITZENBOECK, *Electronic Agents and the Formation of Contracts*, in *International Journal of Law and Information Technology*, 2001, III, p. 204-234.
- C. WENDEHORST, *AI liability in Europe: anticipating the EU AI Liability Directive*, in *adalovelaceinstitute.org*, 2022, p. 9.
- N. WIENER, *The Human Use of Human Beings: Cybernetics and Society*, Boston, 1950.

### Documenti Istituzionali

- ACM TECHNOLOGY POLICY COUNCIL, *Computing and Climate Change*, in *TechBrief*, 2021, I.
- HLEG AI, *Ethics Guidelines for Trustworthy AI*, 2019.
- COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI), *Consolidated working draft of the framework convention on artificial intelligence, human rights, democracy and the rule of law*, Strasbourg 7 luglio 2023, CAI(2023)18.
- EPDS, *Opinion 42/2023 on the Proposals for two Directives on AI liability rules*, in *edps.europa.eu*, 2023.
- GRUPPO DI LAVORO ART. 29, *Linee guida sul processo decisionale individuale automatizzato e sulla profilazione ai fini del regolamento 2016/679 (WP 251)*.
- IPSOS, *Indagine sulle imprese europee sull'uso delle tecnologie basate sull'IA*, 2020.

### Sitografia

- [www.adalovelaceinstitute.org](http://www.adalovelaceinstitute.org)
- [www.agendadigitale.eu](http://www.agendadigitale.eu)
- [www.artificialcommunication.mitpress.mit.edu](http://www.artificialcommunication.mitpress.mit.edu)
- [www.arxiv.org](http://www.arxiv.org)
- [www.congress.gov](http://www.congress.gov)
- [www.doi.org](http://www.doi.org)
- [www.edps.europa.eu](http://www.edps.europa.eu)
- [euractiv.com](http://euractiv.com)
- [www.eur-lex.europa.eu](http://www.eur-lex.europa.eu)
- [www.foroplus.it](http://www.foroplus.it)
- [www.law.stanford.edu](http://www.law.stanford.edu)
- [www.mise.gov.it](http://www.mise.gov.it)
- [www.ssrn.com](http://www.ssrn.com)

