



**CONSULTA ONLINE**

PERIODICO TELEMATICO ISSN 1971-9892



2020 FASC. II

(ESTRATTO)

**PAOLO ZUDDAS**

**PREGIUDIZI DIGITALI E PRINCIPIO DI PRECAUZIONE**

9 LUGLIO 2020

**IDEATORE E DIRETTORE RESPONSABILE: PROF. PASQUALE COSTANZO**

**Paolo Zuddas**  
**Pregiudizi digitali e principio di precauzione\***

**ABSTRACT:** *The investigation examines the possibility of transposing the precautionary principle - in its various forms - from the original fields in which it developed, to the Artificial Intelligence sector. In particular, the principle will be applied to the complex problems posed by using machine learning systems to make decisions, faced with the risk that such tools produce selective decisions.*

*The application of the precautionary principle to machine learning technologies can help answer questions such as: when it is advisable not to adopt artificial intelligence; when it is deemed possible to use it, within what limits and with how much caution; what are the responsibilities to be stressed in this area and on which subjects they fall.*

SOMMARIO: 1. Note introduttive. – 2. I richiami al principio di precauzione nei principali atti normativi che disciplinano l’impiego dell’intelligenza artificiale. – 3. I molteplici significati del principio di precauzione e l’influenza che ciascuna declinazione del principio può dispiegare sull’utilizzo dell’intelligenza artificiale. – 4. (Segue) Le due “varianti” essenziali del principio di precauzione e il loro potenziale impatto sui sistemi di intelligenza artificiale, alla luce dei criteri applicativi individuati dalla Commissione europea – 5. Note conclusive: l’intelligenza artificiale tra precauzione, prevenzione e innovazione.

### 1. Note introduttive

Uno dei rischi più rilevanti, legati al ricorso crescente all’intelligenza artificiale per assumere decisioni in grado di influenzare in modo significativo la vita individuale e collettiva, è rappresentato dalla possibilità che i *software* impiegati – segnatamente nell’ipotesi di utilizzo di sistemi di apprendimento automatico – formulino previsioni errate che penalizzino sistematicamente specifiche categorie di soggetti, generando decisioni discriminatorie. In particolare, le *AI-driven discriminations*, cioè le discriminazioni indotte dall’intelligenza artificiale, possono derivare dalla presenza di “pregiudizi”, in grado di influenzare la decisione formulata dal *software*, che possono originarsi in tre diversi momenti in cui si sviluppa il processo decisionale algoritmico. La decisione algoritmica, infatti, può risentire: in primo luogo, dei pregiudizi del programmatore (e dell’organizzazione in cui opera), il quale può “ordinare” al sistema di *machine learning* di dare rilievo, all’interno del set di dati di riferimento, a caratteri che identificano o rinviano ad una categoria protetta, a rischio di discriminazione; in secondo luogo, delle distorsioni presenti nel *data set* utilizzato per alimentare il sistema di apprendimento automatico, che possono riflettere, ad esempio, i pregiudizi di chi ha formulato le decisioni in base alle quali l’intelligenza artificiale “impara”, elaborando i conseguenti schemi decisionali; in terzo luogo, dei margini di autonomia tipici dei sistemi di *machine learning*. Tali sistemi, infatti, potrebbero individuare, all’interno del set di dati, caratteri ricorrenti che rinviano indirettamente a categorie protette, associando ai loro detentori un trattamento sfavorevole: in quest’ultimo caso i “pregiudizi” sarebbero frutto di un’elaborazione (parzialmente) autonoma del sistema informatico<sup>1</sup>.

Tra i rimedi di norma suggeriti per far fronte ai rischi evocati, non è infrequente il richiamo – all’interno dei principali atti di *soft law* che rappresentano ad oggi il punto di riferimento fondamentale per la regolazione dell’uso di sistemi di AI nelle decisioni riguardanti persone – alla necessità di adottare adeguate misure precauzionali: si tratta, tuttavia, per lo più – come si illustrerà

---

\* Contributo sottoposto a referaggio.

<sup>1</sup> Su questi aspetti sia consentito di rinviare a P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, in questa [Rivista](#), *Liber amicorum per Pasquale Costanzo*, 16 marzo 2020.

nel dettaglio più avanti – di richiami generici, raramente accompagnati da specifiche indicazioni operative.

Muovendo quindi dall'esigenza di individuare con maggiore precisione l'insieme delle misure precauzionali potenzialmente attuabili, scopo dell'indagine che si condurrà nelle pagine seguenti sarà quello di trasporre il principio di precauzione, nelle diverse declinazioni in cui è stato formulato e applicato, dagli originari ambiti materiali nei quali ha trovato sviluppo – ambiente e salute, anzitutto, estendendosi poi a settori ulteriori, quali la sicurezza nel lavoro, la tutela dei consumatori, o le applicazioni biotecnologiche – alla tematica dell'intelligenza artificiale (e in particolare ai complessi problemi posti dal ricorso ai sistemi di *machine learning* per assumere decisioni), con specifico riguardo alle potenzialità discriminatorie proprie di tali strumenti. Si cercherà pertanto di valutare in quali termini il principio in esame può trovare applicazione nel settore delle tecnologie di apprendimento automatico e con quali effetti, verificando in particolare se esso sia in grado di offrire spunti utili a rispondere ad interrogativi quali: quando è consigliabile non ricorrere all'intelligenza artificiale; quando invece si ritiene possibile ricorrervi, ma entro quali limiti e con quali cautele; quali sono le responsabilità che vengono in rilievo in questo ambito e su quali soggetti ricadono.

## 2. I richiami al principio di precauzione nei principali atti normativi che disciplinano l'impiego dell'intelligenza artificiale

Il principio di precauzione, come è noto, nasce nel diritto internazionale, all'interno della disciplina in materia ambientale<sup>2</sup>, ma assume progressivamente una portata più generale<sup>3</sup>; connotazione – che ha indotto taluni osservatori a postularne l'inclusione nel diritto internazionale

---

<sup>2</sup> In particolare, il principio viene delineato per la prima volta nella Carta mondiale della natura approvata dall'ONU nel 1982, in cui si afferma che «se i potenziali effetti avversi, non possono essere pienamente valutati, le attività [che comportano rischi significativi per la natura] non potranno essere realizzate» (par. 11b): una formulazione, dunque, che lo configura come rigido «principio-moratoria» (così D. AMIRANTE, *Il principio precauzionale fra scienza e diritto. Profili introduttivi*, in *Dir. gest. amb.*, 2001, n. 2, 22). Viene poi riproposto nella Seconda conferenza internazionale per la protezione del mare del Nord del 1987 e poi nella Terza del 1990. Viene quindi riconfermato nella Convenzione quadro sui cambiamenti climatici – c.d. Dichiarazione di Rio – del 1992 (principio n. 15) e nella Convenzione sulla biodiversità dello stesso anno (frutto anch'essa della Conferenza sull'Ambiente e sullo Sviluppo delle Nazioni Unite, tenutasi a Rio de Janeiro nel 1992), consolidandosi come principio qualificante il diritto internazionale dell'ambiente e trovando conferma nei principali atti negoziali successivamente adottati in materia.

La medesima “matrice ambientale” del principio si riscontra, peraltro, sia – come si dirà meglio *infra* – nel diritto comunitario e nell'ordinamento italiano, sia anche in altri ordinamenti europei, come quello tedesco e francese: in particolare, in Germania l'origine del principio è stata individuata nell'art. 20a della Legge fondamentale, laddove si assegna allo Stato la responsabilità di proteggere le basi naturali della vita e gli animali, nell'interesse delle generazioni future (cfr. sul punto G. MANFREDI, *Note sull'attuazione del principio di precauzione nel diritto pubblico*, in *Dir. pubbl.*, 2004, n. 3, 1077), mentre in Francia il principio precauzionale compare nel codice ambientale, dove è confluita la c.d. legge Barnier, n. 95-101 del 2 febbraio 1995 (in materia di rafforzamento della salvaguardia dell'ambiente), che l'ha introdotto per prima nell'ordinamento (cfr. *ivi*, 1079).

Segnatamente, nella Legge fondamentale tedesca emerge un nesso evidente con il principio di prevenzione, del quale – come si illustrerà meglio più avanti – il principio in esame rappresenta, per molti versi, un'evoluzione (l'introduzione del c.d. *Vorsorgeprinzip* – che postula la tutela dell'interesse delle generazioni future – nasce infatti dall'esigenza di elevare il livello di tutela offerta dal principio di prevenzione: cfr. in argomento S. GRASSI, A. GRAGNANI, *Il principio di precauzione nella giurisprudenza costituzionale*, in L. CHIEFFI (cur.), *Biotecnologie e tutela del valore ambientale*, Torino, 2003, 151). Parimenti, nelle Costituzioni di Spagna, Grecia e Portogallo si richiama espressamente il solo principio di prevenzione, deducendosi da questo l'istanza precauzionale, intesa come rafforzamento della protezione (cfr. sul punto T. MAROCCO, *Il principio di precauzione e la sua applicazione in Italia e in altri Stati membri della Comunità europea*, in *Riv. it. dir. pubbl. com.*, 2003, n. 5, 1244; si v. in particolare le considerazioni in merito all'art. 45 della Costituzione spagnola, in cui il principio di prevenzione viene declinato con riferimento alla tutela dell'ambiente).

<sup>3</sup> Cfr. in proposito S. GRASSI, *Prime osservazioni sul 'principio di precauzione' come norma di diritto positivo*, in *Dir. gest. amb.*, 2001, n. 2, 50.

consuetudinario<sup>4</sup> – confermata dalla disciplina europea<sup>5</sup>, all'interno della quale il principio precauzionale esibisce un carattere prevalentemente programmatico, che gli consente di operare sia in funzione ispiratrice della disciplina di dettaglio, sia in funzione interpretativa dei Trattati in vigore<sup>6</sup>: notevole, peraltro, è il rilievo detenuto all'interno dell'ordinamento dell'UE, nel quale il principio – introdotto dal Trattato di Maastricht attraverso l'inclusione nel Trattato istitutivo del nuovo art. 130 R<sup>7</sup>, in materia ambientale – trova poi attuazione in ambiti ulteriori, in quanto principio dell'ordinamento comunitario applicabile in via generale<sup>8</sup>. Anche nell'ordinamento italiano il principio di precauzione, pur essendo evocato espressamente soprattutto nella normativa in materia di tutela ambientale<sup>9</sup>, ha visto progressivamente estendersi il proprio ambito di applicazione: un percorso favorito anche dalla giurisprudenza sviluppata in argomento dalla Corte costituzionale, la quale, pur esprimendosi sull'istanza precauzionale in riferimento alla sua applicazione in settori specifici<sup>10</sup>, ha ragionato sulla natura del principio riguardato nei suoi termini più generali, offrendo spunti utili ad una sua estensione anche ad ambiti ulteriori a quelli originari<sup>11</sup>.

<sup>4</sup> Cfr. sul punto G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1077.

<sup>5</sup> Si v. in proposito la *Comunicazione* della Commissione del 2 febbraio 2000, che, in linea con l'orientamento espresso sul punto dalla giurisprudenza comunitaria (v. *infra*, nota 8), ha definito il principio in oggetto (nella formulazione che compare – come si preciserà *infra* – nell'art. 174 del Trattato CE) un principio di diritto internazionale «di portata generale», suscettibile dunque di essere esteso a settori ulteriori rispetto a quello ambientale (cfr. par. 4; in merito all'ampiezza dell'ambito di applicazione del principio, v. anche le considerazioni svolte nel sommario, al punto n. 3).

In argomento cfr. S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 150 (che, tuttavia, segnalano anche la presenza di posizioni contrarie all'interno della letteratura internazionalistica in materia).

<sup>6</sup> Cfr. sul punto T. MAROCCO, *Il principio di precauzione*, cit., 1236.

<sup>7</sup> Poi transitato nell'art. 174 del Trattato CE ed ora nell'art. 191 del Trattato sul funzionamento dell'Unione europea.

<sup>8</sup> Cfr. in tema G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1078, S. GRASSI, *Prime osservazioni*, cit., 50 e S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 150. All'interno della copiosa giurisprudenza sul punto, si v., ad esempio, la sentenza sul caso “Artegodan” (Trib. CE, 26 novembre 2002, cause riunite T-74/00, T-76/00, T-83/00, T-84/00, T-85/00, T-132/00, T-137/00 e T-141/00), nel quale si afferma che «nonostante sia menzionato nel Trattato solamente in relazione alla politica ambientale, il principio di precauzione ha [...] un ambito di applicazione più ampio. Esso è destinato ad applicarsi, al fine di assicurare un livello elevato di protezione della salute, della sicurezza dei consumatori e dell'ambiente, in tutti gli ambiti di azione della Comunità [...]» (punto 183); «ne consegue che il principio di precauzione può essere definito come un principio generale del diritto comunitario che fa obbligo alle autorità competenti di adottare provvedimenti appropriati al fine di prevenire taluni rischi potenziali per la sanità pubblica, per la sicurezza e per l'ambiente, facendo prevalere le esigenze connesse alla protezione di tali interessi sugli interessi economici. Infatti, essendo le istituzioni comunitarie responsabili, in tutti i loro ambiti d'azione, della tutela della salute, della sicurezza e dell'ambiente, il principio di precauzione può essere considerato come un principio autonomo che discende dalle menzionate disposizioni del Trattato» (punto 184). Si v. anche la sentenza sul caso “Solvay” (Trib. CE, 21 ottobre 2003, T-392/02), in cui si ribadisce che «il principio di precauzione costituisce un principio generale del diritto comunitario che fa obbligo alle autorità interessate di adottare [...] provvedimenti appropriati al fine di prevenire taluni rischi potenziali per la sanità pubblica, per la sicurezza e per l'ambiente, facendo prevalere le esigenze connesse alla protezione di tali interessi sugli interessi economici» (punto 121).

<sup>9</sup> In particolare, la precauzione compare tra i principi generali del codice dell'ambiente: cfr. l'art. 3-ter del d.lgs. 3 aprile 2006, n. 152.

<sup>10</sup> Principalmente, oltre all'ambiente, il settore sanitario. In particolare, tra le pronunce che hanno originariamente contribuito a delineare il principio, si v., *ex multis*: la sentenza n. 185 del 1998, sul diritto a farmaci sperimentali, in relazione al c.d. caso Di Bella (cfr. in argomento P. VERONESI, *Le cognizioni scientifiche nella giurisprudenza costituzionale*, in *Quad. cost.*, 2009, n. 3, 598 ss.); la sentenza n. 351 del 1999, sul trattamento dell'encefalopatia animale; la sentenza n. 382 del 1999, relativa alla fissazione, da parte della Regione Veneto, di limiti alla esposizione a campi elettromagnetici più restrittivi rispetto a quelli previsti dalla legislazione statale (sull'elettrosmog, peraltro, si sono espressi ampiamente anche i giudici comuni: cfr. in argomento G.D. COMPORTELLI, *Amministrazioni e giudici sull'onda dell'elettrosmog*, in *Foro amm.*, 2001, 2455 ss.).

<sup>11</sup> V. sul punto – oltre alla succitata sentenza n. 185 del 1998, in cui si propone una lettura “inversa” del principio in oggetto (come sottolinea opportunamente P. VERONESI, *Le cognizioni scientifiche*, cit., 599, «il principio di precauzione viene qui fatto funzionare in modo esattamente rovesciato rispetto al comune sentire: non si vieta bensì s'impone l'estensione mirata di un trattamento della cui efficacia ancora non v'è certezza. In divenire, non si esclude comunque un più penetrante sindacato di legittimità allorché i riscontri scientifici fossero disponibili») – la sentenza n. 282 del

I “dati genetici” sinteticamente richiamati appaiono sufficienti a far emergere due elementi essenziali ai fini dell’applicazione del principio di precauzione all’ambito dell’intelligenza artificiale: la nascita all’interno del diritto internazionale ne marca una vocazione “globale”, utile per la regolamentazione soprattutto dei sistemi di apprendimento automatico che si alimentano dei *Big data* acquisiti dalla rete Internet, mentre la sua qualificazione come principio generale lo rende pienamente applicabile anche al di fuori degli ambiti nei quali è espressamente richiamato.

Una duplice “attitudine” che trova conferma nella frequente evocazione del principio precauzionale all’interno dei più rilevanti atti normativi – in prevalenza, come ricordato, di *soft law* – che disciplinano l’utilizzo dell’intelligenza artificiale. In questo ambito, vanno richiamate anzitutto le dichiarazioni di intenti di studiosi e operatori, relativi a singoli settori delle nuove tecnologie: in particolare, l’applicazione del principio in oggetto è stata invocata nel campo delle nanotecnologie<sup>12</sup>, della telefonia mobile<sup>13</sup> e in relazione all’inserimento di impianti ICT nel corpo umano<sup>14</sup>.

L’unico documento in materia che riveste una valenza più generale è la Dichiarazione di Asilomar del 2017, predisposta da accademici e rappresentanti delle più importanti multinazionali operanti nel campo dell’intelligenza artificiale (quali, ad esempio, *Facebook*, *Google* e *Tesla*)<sup>15</sup>: gli *Asilomar AI Principles*, pur non evocando esplicitamente il principio di precauzione, sollecitano

---

2002, in cui la Corte prende posizione contro l’eccesso di precauzione: in particolare, la sentenza 282/2002, nel mettere in guardia contro gli interventi limitativi del legislatore, delinea una «competenza costituzionalmente fondata dei medici, sulla base delle conoscenze scientifiche acquisite da istituzioni specializzate e nel rispetto delle regole di deontologia professionale», a svolgere attività curative, preventive e riabilitative (così G. GEMMA, *Giurisprudenza costituzionale e scienza medica*, in A. D’ALOIA (cur.), *Bio-tecnologie e valori costituzionali. Il contributo della giustizia costituzionale*, Torino, 2005, 65), che individua uno spazio normativo scientificamente riservato all’autonomia e responsabilità del medico.

Volendo applicare il modello tratteggiato in questa pronuncia al settore informatico, potrebbe postularsi un analogo spazio normativo scientificamente riservato all’ingegnere informatico (e, segnatamente, al programmatore del *software*); senonché, le «regole di deontologia professionale» operanti in questo campo paiono, al momento, ancora troppo poco strutturate per offrire sufficienti garanzie (cfr. il *Codice etico-deontologico degli ingegneri robotici*, allegato alla Risoluzione del Parlamento europeo del 16 febbraio 2017, recante *Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, 2015/2103(INL) e gli *Orientamenti etici per un’intelligenza artificiale affidabile* adottati dal Gruppo di esperti ad alto livello sull’intelligenza artificiale istituito dalla Commissione europea nel giugno 2018, nel testo reso pubblico l’8 aprile 2019, entrambi illustrati *infra*). E comunque, affidarsi alla sola autoregolamentazione, in una fase così delicata dello sviluppo informatico, considerate le tante incognite ancora presenti, sembra del tutto improprio; senza contare che l’evoluzione dei sistemi di *machine learning* prospetta scenari – invero assai poco rassicuranti – di macchine pensanti dotate di ampi margini di autonomia. Pertanto, affidare la delega del bilanciamento in concreto ai “tecnici” e non ai giudici, nel settore informatico, appare ad oggi una scelta che nasconde molte insidie.

Sembra invece pienamente applicabile al settore informatico il principio, anch’esso desumibile dalla pronuncia in commento, per cui il legislatore – sempre basandosi sullo stato delle conoscenze scientifiche – può comunque dettare regole dirette a «prescrivere procedure particolari per l’impiego di mezzi terapeutici “a rischio”» (sentenza n. 282 del 2002, punto 5 considerato in diritto); sul punto cfr. S. PENASA, il quale osserva che tanto maggiore è l’incertezza scientifica, tanto più ampio è lo spazio occupato dal legislatore, «operando l’incertezza medico-scientifica come “valvola di regolazione” del rapporto tra autonomia medico-scientifica ed eteronomia politico-legislativa» (*Il dato scientifico nella giurisprudenza della Corte costituzionale: la ragionevolezza scientifica come sintesi tra dimensione scientifica e dimensione assiologica*, in *Pol. dir.*, 2015, n. 2, 318). Come si preciserà più avanti, infatti, il ricorso a procedure di gestione del rischio – sulla base di una interpretazione “debole” o “procedurale” del principio di precauzione – risulta una delle soluzioni più appropriate di fronte alle incognite poste dall’impiego di sistemi di intelligenza artificiale per assumere decisioni incidenti sulla sfera giuridica dei destinatari.

<sup>12</sup> Si segnala in proposito la proposta, formulata dall’*ETC Group (Action Group on Erosion, Technology and Concentration)*, di moratoria della ricerca nel campo delle nanotecnologie: cfr. *ETC Group*, in cui compaiono numerosi richiami al principio di precauzione (v. sul punto S. CLARKE, *Future technologies, dystopic futures and the precautionary principle*, in *Ethics and Information Technology*, 2005, 121).

<sup>13</sup> Cfr. in argomento lo *Stewart Committee Report* del 2000 (sul quale S. CLARKE, *Future technologies*, cit., 122).

<sup>14</sup> Cfr. la proposta avanzata in proposito dall’*European Group on Ethics* del 2005 (anche sul punto si rinvia a S. CLARKE, *Future technologies*, cit., 122).

<sup>15</sup> Cfr. *Asilomar AI Principles*, adottati a conclusione della *Asilomar Conference on Beneficial AI* tenutasi in California dal 5 all’8 gennaio 2017.

l'introduzione di meccanismi procedurali di gestione del rischio e raccomandano un'evoluzione "controllata" dei sistemi di intelligenza artificiale<sup>16</sup>. In particolare, il documento indica – tra le finalità della ricerca da finanziare – l'aggiornamento della disciplina normativa in tema di gestione del rischio associato all'AI<sup>17</sup> e – tra gli obiettivi a lungo termine – la predisposizione di rigorose misure di controllo e di sicurezza alle quali dovrebbero essere sottoposti i sistemi di AI a sviluppo autonomo ed autoreplicanti<sup>18</sup>.

Il principio di precauzione viene richiamato anche nei principali documenti europei che si occupano di nuove tecnologie. Un primo richiamo compare nella risoluzione del Parlamento europeo del 2017 che detta norme di diritto civile sulla robotica, in cui il principio viene evocato in due occasioni: con riferimento all'immissione sul mercato dei sistemi di AI e dei robot, da effettuarsi «dopo le opportune valutazioni della sicurezza e nel rispetto del principio di precauzione»<sup>19</sup> (invitandosi a tal fine la Commissione e gli Stati membri ad incentivare la ricerca sui possibili rischi)<sup>20</sup>; e con riferimento alla identificazione delle aree in cui testare i robot in condizioni reali (essenziali per operare una corretta valutazione dei rischi)<sup>21</sup>. Segnatamente, la risoluzione invita la Commissione a elaborare criteri uniformi valevoli in tutta l'Unione, che i singoli Stati membri dovrebbero utilizzare per identificare le aree in cui autorizzare gli esperimenti con robot, «nel rispetto del principio di precauzione»<sup>22</sup>. Inoltre, alla risoluzione è allegato un *Codice etico-deontologico degli ingegneri robotici*, che individua il principio di precauzione tra i principi qualificanti la ricerca nel campo della robotica, ribadendo la necessità di adottare le debite precauzioni «proporzionalmente al livello di protezione» (stabilito)<sup>23</sup>; nel codice si ribadisce inoltre che il processo di valutazione dei rischi, su cui deve basarsi il funzionamento di un sistema robotico, «dovrebbe essere sempre improntato ai principi di proporzionalità e di precauzione»<sup>24</sup>.

Alla risoluzione del 2017 ha fatto seguito un analogo atto di indirizzo adottato dal Parlamento europeo nel 2019, in tema di politica industriale sulla robotica e l'AI<sup>25</sup>: anche la risoluzione del 2019 richiama espressamente il principio in esame con riferimento all'attività di ricerca sull'intelligenza artificiale, sottolineando che «la ricerca sull'IA e altre attività correlate dovrebbero

---

<sup>16</sup> Interventi che possono considerarsi espressione del principio in oggetto, nella sua versione "debole", come si preciserà meglio *infra*, par. 4.

<sup>17</sup> Cfr. il principio n. 2.

<sup>18</sup> Cfr. il principio n. 22.

<sup>19</sup> Cfr. la Risoluzione del Parlamento europeo del 16 febbraio 2017, recante *Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, cit., punto 7.

<sup>20</sup> In particolare, la risoluzione invita la Commissione e gli Stati membri a «promuovere i programmi di ricerca, a incentivare la ricerca sui possibili rischi e sulle possibili opportunità a lungo termine dell'intelligenza artificiale e delle tecnologie robotiche e a promuovere quanto prima l'avvio di un dialogo pubblico strutturato sulle conseguenze dello sviluppo di tali tecnologie» (punto 7).

<sup>21</sup> La risoluzione, peraltro, contiene numerose sollecitazioni a considerare attentamente i rischi legati all'uso di sistemi robotici: si v., ad esempio, il punto 10, nel quale si osserva che «le possibilità di realizzazione personale che derivano dall'uso della robotica sono relativizzate da un insieme di tensioni o rischi e dovrebbero essere valutate in modo serio dal punto di vista della sicurezza delle persone e della loro salute, della libertà, la vita privata, l'integrità, la dignità, dell'autodeterminazione e la non discriminazione nonché della protezione dei dati personali». In questa prospettiva, particolare attenzione viene dedicata all'utilizzo di robot all'interno dei luoghi di lavoro (mettendosi in evidenza, oltre alle «grandi potenzialità offerte dalla robotica per migliorare la sicurezza sul posto di lavoro mediante il trasferimento di alcuni compiti pericolosi e dannosi dagli esseri umani ai robot», anche la possibilità che l'impiego di tali tecnologie crei «anche una serie di nuovi rischi, dovuti al numero crescente di interazioni fra esseri umani e robot sul luogo di lavoro»: punto 46) e all'impiego delle tecnologie robotiche per compiere interventi riparativi e migliorativi del corpo umano (sottolineandosi i «rischi correlati alla possibilità di *hacking*, disattivazione o cancellazione della memoria dei CPS integrati nel corpo umano, dato che possono mettere in pericolo la salute o, in casi estremi, anche la vita umana»: punto 39).

<sup>22</sup> Ivi, punto 23.

<sup>23</sup> E comunque «incoraggiando allo stesso tempo il progresso a vantaggio della società e dell'ambiente»: *Codice etico-deontologico degli ingegneri robotici*, par. "Precauzione".

<sup>24</sup> Ivi, par. "Massimizzare i vantaggi e ridurre al minimo il danno".

<sup>25</sup> Risoluzione del Parlamento europeo del 12 febbraio 2019 su *una politica industriale europea globale in materia di robotica e intelligenza artificiale*, 2018/2088(INI).

essere condotte nel rispetto del principio di precauzione e dei diritti fondamentali» e ribadendo inoltre che «tutti i soggetti coinvolti nello sviluppo, nell’attuazione, nella diffusione e nell’utilizzo dell’AI dovrebbero [...] anticipare le potenziali conseguenze sulla sicurezza e prendere le dovute precauzioni, in proporzione al livello di protezione, compresa la tempestiva divulgazione di fattori che potrebbero mettere a rischio i cittadini e l’ambiente»<sup>26</sup>. In particolare, l’analisi dei rischi (sulla base dei quali calibrare lo sviluppo potenziale della tecnologia) dovrebbe essere affidata ad una futura Agenzia europea di regolamentazione per l’AI (di cui il Parlamento europeo sollecita la costituzione), incaricata, tra l’altro, di «elaborare una matrice di valutazione del rischio in grado di classificare i tipi di algoritmo e i campi di applicazione in funzione della loro capacità di produrre potenzialmente un impatto negativo significativo sui cittadini»<sup>27</sup> (indagine verosimilmente operabile anche al fine di valutare possibili rischi ignoti, in coerenza con il modello applicativo del principio di precauzione)<sup>28</sup>. Utili indicazioni vengono altresì offerte in merito alla soglia di rischio accettabile, che viene correlato a quello riscontrato nella vita normale<sup>29</sup>. Infine, compare un richiamo – estremamente significativo ai fini della nostra indagine – al tema dei *bias*, che vengono espressamente inclusi tra i rischi valutabili in sede di applicazione del principio di precauzione: il parlamento, infatti, «invita la Commissione, gli Stati membri e le autorità di protezione dei dati ad adottare tutte le misure per prevenire o ridurre al minimo le *discriminazioni* e le *distorsioni algoritmiche*»<sup>30</sup>.

Sorprende invece l’assenza di una menzione esplicita del principio in oggetto all’interno degli *Orientamenti etici per una IA affidabile*, predisposti dalla Commissione europea nel 2018, che rappresenta per molti aspetti un punto di riferimento essenziale per il futuro sviluppo delle tecnologie di intelligenza artificiale: nel documento, infatti, sono frequentissimi i richiami espressi al principio di prevenzione<sup>31</sup>, operati in relazione a rischi noti e alla loro valutazione, mentre il principio precauzionale viene evocato solo in via indiretta, laddove si fa riferimento alla gestione dei rischi ignoti. Tuttavia, i riferimenti impliciti al principio in oggetto sono diversi e significativi: in particolare, compare un accenno alla necessità di valutare i rischi connessi all’AI in relazione ad effetti negativi «anche *difficili da prevedere, individuare, o misurare*», sottolineando che occorre, «se necessario, adottare provvedimenti adeguati ad attenuare tali rischi, in modo proporzionato alla loro portata»<sup>32</sup>. Si aggiunge inoltre che vanno ridotti al minimo «i danni non intenzionali e *imprevisti*»<sup>33</sup>; si ipotizza, infine, l’emergere, nel lungo termine, di «gravi preoccupazioni», le quali, secondo «l’approccio basato sul rischio», dovrebbero «essere tenute in considerazione alla luce di eventuali *incognite* sconosciute dei cosiddetti “cigni neri”». Di fronte a scenari così allarmanti, tuttavia, la risposta appare sostanzialmente debole: pur riconoscendosi infatti «la natura dirompente

<sup>26</sup> Così il par. 19. Si noti in particolare come il principio di precauzione si presenti, in questo caso, nella forma della “anticipazione preventiva” (cfr. lo schema interpretativo proposto *infra*, par. 3), richiamandosi, peraltro, due criteri applicativi del principio stabiliti dall’UE: il livello di protezione (adeguato) e gli obblighi informativi.

<sup>27</sup> Così il par. 187.

<sup>28</sup> Anche sul punto si rinvia ad *infra*, par. 3.

<sup>29</sup> Cfr. in proposito il par. 24, nel quale si precisa che, se i rischi sono inevitabili e fanno parte integrante della ricerca sull’IA, occorre sviluppare solidi protocolli di valutazione e gestione dei rischi, tenendo conto che il rischio di danno non dovrebbe essere superiore a quello riscontrato nella vita normale.

<sup>30</sup> Così il par. 157; corsivi aggiunti.

<sup>31</sup> Il principio di «prevenzione dei danni» viene individuato come uno dei principi etici fondamentali (v. pag. 9 e punto 48), soprattutto in relazione ai potenziali danni a carico dei soggetti deboli, a rischio di discriminazione (punto 51) e viene evocato espressamente con riferimento, in particolare: al principio di robustezza dell’IA (consistente nella capacità di prevenire effetti indesiderati e che non causino danni involontari) (punto 27), alla riservatezza dei dati (punto 71), alla sostenibilità ecologica (punto 83) e alla riproducibilità dell’AI (punto 70).

<sup>32</sup> Così il punto 15; corsivi aggiunti. Si noti il richiamo al principio di proporzionalità, che – come sarà precisato meglio *infra* – rappresenta uno dei criteri applicativi del principio di precauzione nella versione “debole” accolta all’interno della disciplina europea. In proposito si v. anche i punti 10 e 43, in cui si ribadisce che i rischi devono essere «gestiti in modo appropriato e vanno previsti e minimizzati».

<sup>33</sup> Così il punto 66 (corsivo aggiunto), parimenti in linea con la versione “debole” del principio in esame.

di tali preoccupazioni, assieme all'attuale *incertezza* sui relativi sviluppi», si ritiene infatti che essa imponga (solo) una «valutazione regolare di questi temi»<sup>34</sup>.

Eguale, va rilevata l'assenza di riferimenti espliciti al principio di precauzione all'interno dei principali documenti in materia di AI adottati da istituzioni pubbliche italiane: nessun richiamo compare nel *Libro Bianco sull'AI al servizio del cittadino* predisposto dall'AgID nel marzo del 2018<sup>35</sup> (che non menziona espressamente neanche il principio, in certa misura contiguo, di prevenzione); nessun richiamo, parimenti, è presente nel documento che illustra la *Strategia nazionale per l'IA*, approvato dal Ministero per lo Sviluppo Economico nel luglio del 2019<sup>36</sup>, a parte un genericissimo riferimento allo «sviluppo sostenibile» (nozione ampia, che può includere, nella sua interpretazione più estesa, la necessità di considerare e gestire, nella progettazione dello sviluppo economico, anche i rischi ignoti) come prospettiva dalla quale valutare l'uso dei sistemi di IA<sup>37</sup>.

Tuttavia, al di là dei rinvii indiretti, anche quando – negli atti richiamati – il principio di precauzione viene espressamente evocato, esso appare sempre configurato come principio di diritto<sup>38</sup> dotato di un ampio tasso di “vaghezza”<sup>39</sup>; o come un valore, che indica essenzialmente un fine, senza una compiuta individuazione dei mezzi per conseguirlo. Una tendenziale genericità che riflette peraltro il tratto genetico del principio precauzionale, che nasce vago e si mantiene «perplesso»<sup>40</sup> e sostanzialmente privo di un'univoca definizione<sup>41</sup>.

### *3. I molteplici significati del principio di precauzione e l'influenza che ciascuna declinazione del principio può dispiegare sull'utilizzo dell'intelligenza artificiale*

---

<sup>34</sup> Così il punto 135; corsivi aggiunti.

<sup>35</sup> Cfr. *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino*, a cura della Task force sull'Intelligenza Artificiale dell'Agenzia per l'Italia Digitale, Versione 1.0 marzo 2018.

<sup>36</sup> Cfr. *Strategia Nazionale per l'Intelligenza Artificiale*, bozza per la consultazione pubblicata dal Ministero per lo Sviluppo Economico in data 31 luglio 2019.

<sup>37</sup> Cfr. p. 4 (par. “*Visioni e obiettivi*”).

Nel documento predisposto dal MISE compare invece un richiamo al principio di prevenzione: vi si ribadisce infatti la «ferma volontà del Governo di prevenire ogni forma di IA che possa inasprire le ineguaglianze sociali» e, a tal fine, si manifesta l'intento di promuovere «forme di certificazione che consentano di verificare il pieno allineamento dei sistemi di IA con i principi di affidabilità sanciti dalle linee guida europee per un'IA etica» (p. 6, par. 6: un richiamo, quest'ultimo, che consentirebbe di far valere le istanze, latamente collegate al principio di precauzione, contenute nel documento europeo).

<sup>38</sup> Così G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1097. Una connotazione, peraltro, che rappresenta il carattere tipico delle “algo-regole”: come osserva in proposito A. CELOTTO, «dal punto di vista contenutistico non abbiamo norme puntuali e precettive, ma piuttosto principi che si richiamano ai valori di fondo. Sono formulati principi generali di equità, dignità, giustizia, al fine di tutelare anche rispetto alle nuove tecnologie il rispetto delle libertà e dei principali diritti individuali e collettivi. Più nello specifico, tali principi tendono a indirizzare anche in chiave etica alcuni problemi sollevati dal funzionamento dell'Intelligenza Artificiale, vale a dire quelli della qualità e della neutralità dei dati, della responsabilità di chi utilizza gli algoritmi, della trasparenza e dell'apertura di questi ultimi, nonché della tutela della privacy. L'obiettivo è mostrare come un uso scorretto delle tecnologie a nostra disposizione possa contribuire allo sviluppo di una società più ingiusta che alimenta le disuguaglianze, mentre la consapevolezza dei rischi che si corrono nell'affidarsi alle macchine «intelligenti», può aiutarci a minimizzare questi ultimi e a progettare un mondo migliore.

In buona sostanza, queste Carte sui fenomeni degli algoritmi non pongono vere e proprie norme, ma piuttosto principi, anche con un buon grado di vaghezza, cioè di elasticità, che sta a mostrare non tanto l'esigenza di porre regole aperte, quanto piuttosto il grado ancora acerbo di regolazione giuridica di questi fenomeni, anche perché del tutto nuovi per il giurista e per gran parte ancora sconosciuti» (*Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in *An. giur. ec.*, 2019, n. 1, 57).

<sup>39</sup> Per riprendere l'espressione suggerita da CELOTTO (*supra*, nota precedente).

<sup>40</sup> Così G. MANFREDI, il quale osserva che «a tutt'oggi si ritiene che il principio di precauzione sia connotato da una grave perplessità» (*Note sull'attuazione del principio di precauzione*, cit., 1088).

<sup>41</sup> In tema v. D. BODANSKY, *Scientific Uncertainty and the Precautionary Principle*, in *Environment*, 1991, n. 7, 4 ss.



Come dianzi accennato, il principio di precauzione difetta di una precisa definizione, che non viene offerta né dalle norme di diritto internazionale che lo richiamano, né da quelle del diritto dell'UE<sup>42</sup>: emerge certo un tratto caratterizzante, che rinvia ad un generale principio di cautela<sup>43</sup> ed alla connessa esigenza di regolazione del rischio<sup>44</sup>. Ma la sua concreta fisionomia disegna un principio multidimensionale<sup>45</sup>, declinabile in una molteplicità di significati, ognuno dei quali riflette un diverso atteggiamento verso l'innovazione scientifica e tecnologica<sup>46</sup>. Tra questi diversi significati, frutto anche dell'applicazione del principio ad ambiti ulteriori a quelli in cui ha trovato originaria applicazione, possono sinteticamente individuarsi<sup>47</sup>: la promozione dei diritti naturali intrinseci, la salvaguardia degli ecosistemi, l'anticipazione preventiva, la proporzionalità della risposta (o efficacia rispetto ai costi dei margini di errore), l'inversione dell'onere della prova, l'obbligo di pagare il debito (ecologico) causato nel passato.

Le prime due declinazioni del principio risentono di un più intenso legame con le tematiche ambientali che ne hanno rappresentato la matrice originaria e paiono dunque difficilmente applicabili all'impiego di sistemi di intelligenza artificiale per trattare informazioni costituite in prevalenza da dati personali. In particolare, la promozione dei diritti «naturali» intrinseci fa riferimento specificamente ai «diritti della natura» e si presenta direttamente correlata alla nozione legale di danno ambientale<sup>48</sup>. Parimenti, la finalità della massima «salvaguardia degli ecosistemi» rappresenta una declinazione utile per definire il possibile livello di tutela e, in certo senso, la possibile «intensità» che può assumere il principio di precauzione: essendo infatti fondata sulla «consapevolezza che in alcuni casi non possono essere accettati [...] margini di tolleranza»<sup>49</sup>, la declinazione in esame tende a dispiegare un effetto di totale preclusione delle attività a rischio; anche in questo caso, tuttavia, il valore tutelato è rappresentato principalmente dall'ambiente e dalla salute.

Una declinazione, invece, estremamente utile ai nostri fini è quella che configura il principio di precauzione come «anticipazione preventiva». La formula – che non a caso fa emergere la marcata contiguità con il principio di prevenzione – si riferisce ad un'applicazione del principio precauzionale che impone di anticipare il momento in cui opera la preclusione dell'attività a rischio «rispetto alla prova scientifica dell'evidenza del bisogno di intervenire»<sup>50</sup>: si anticipano, in sostanza, le misure cautelative «al momento in cui le conoscenze scientifiche non sono ancora certe»<sup>51</sup>. Un chiaro esempio della declinazione proposta si rinviene nella Dichiarazione di Rio del 1992, che, al principio n. 15, afferma che «in presenza del rischio di danni gravi o irreversibili, l'assenza di certezza scientifica assoluta non deve costituire il pretesto per rimandare l'adozione di misure

<sup>42</sup> Cfr. in argomento T. MAROCCO, *Il principio di precauzione*, cit., 1237. Qualche elemento, come pure si è evidenziato *supra*, emerge dagli atti di *soft law* in materia di intelligenza artificiale che lo richiamano, ma si tratta di indicazioni non univoche.

<sup>43</sup> Cfr. G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1075.

<sup>44</sup> Esigenza, peraltro – come rilevato da M. GRAZIADEI, *La regolazione del rischio e il principio di precauzione: Stati Uniti ed Europa a confronto*, in *Sist. intell.*, 2017, n. 2, 499 – che intercetta una delle funzioni fondamentali del diritto.

<sup>45</sup> Sulla multidimensionalità del principio v. meglio *infra*, nota 76.

<sup>46</sup> Cfr. sul tema G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1076 e 1090.

<sup>47</sup> Eucleati in origine S.B. CHRISTIANSENS, *The precautionary principle in Germany: enabling government*, in T. O'RIORDAN, J. CAMERON (curr.), *Interpreting the precautionary principle*, London, 1994 e poi discussi da altri Studiosi, tra cui S. GRASSI, *Prime osservazioni*, cit., 45 ss. e G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1076 ss.).

<sup>48</sup> Così S. GRASSI (*Prime osservazioni*, cit., 46), il quale osserva in proposito che «la promozione di diritti naturali intrinseci [allarga] la nozione legale di danno ambientale, per includere l'esigenza di consentire ai processi naturali di svolgere le loro funzioni, in modo tale da mantenere il supporto essenziale per la vita della terra (secondo una concezione che dà prevalenza alla scelta etica che si riconduce al riconoscimento dei c.d. «diritti della natura»)».

<sup>49</sup> Così S. GRASSI, *Prime osservazioni*, cit., 45.

<sup>50</sup> *Ibidem*.

<sup>51</sup> Così T. MAROCCO, *Il principio di precauzione*, cit., 1233.

adeguate ed efficaci, anche in rapporto ai costi, volte a prevenire il degrado ambientale»: la formulazione richiamata viene considerata da diversi osservatori<sup>52</sup> rivelatrice dell'essenza del principio di precauzione, almeno colto nella sua connotazione originaria come espressione del modello "anticipatorio" nella gestione delle azioni in materia ambientale<sup>53</sup>. In questa prospettiva, il principio di precauzione rappresenterebbe, in certo senso, un rafforzamento del principio di prevenzione, imponendo l'attivazione di cautele anche in situazioni di incertezza in cui è solo ipotizzabile un rischio di lesione<sup>54</sup>.

L'applicazione del principio di precauzione così configurato all'uso di sistemi di intelligenza artificiale che presentano il rischio di produrre effetti discriminatori – rispetto al cui impiego è, per l'appunto, solo ipotizzabile il rischio di *bias* – deve tuttavia tener conto che la finalità che anima l'anticipazione preventiva è quella di anticipare gli interventi cautelativi (e dunque, in certa misura, preclusivi o limitativi) di fronte ad un possibile danno *irreparabile*, come quelli sovente causati all'ambiente o alla salute<sup>55</sup>: da qui il riferimento, valido soprattutto per il danno irreversibile all'ambiente, alle generazioni future<sup>56</sup>. Nell'ipotesi in cui il bene da proteggere sia il principio di non discriminazione, varrà allora la pena di chiedersi se possa individuarsi un danno totalmente irreparabile e se assumano rilievo, in questo caso, i diritti delle generazioni future: in proposito, vanno sottolineati i pericoli derivanti dall'impiego di sistemi di *machine learning* in grado di dar vita autonomamente a nuove categorie di soggetti discriminati, sottoponendoli ad un trattamento deteriore che rischia – nell'ipotesi in cui tale discriminazione non venga rilevata tempestivamente, mettendo rapidamente in atto le necessarie misure di contrasto e di compensazione – di consolidarsi nel tempo, provocando pertanto danni potenzialmente irreversibili. Un'irreparabilità, dunque, che in questo frangente si presenta legata alla possibilità che l'effetto dannoso – di per sé astrattamente riparabile – non venga rilevato, o venga rilevato in un momento in cui gli interventi non sono più in grado di rimuovere tutte le conseguenze sfavorevoli.

Un'ulteriore declinazione del principio di precauzione estremamente utile ai fini della presente indagine è quella che *limita* l'effetto delle misure impeditive di attività che comportano rischi imprevedibili, richiedendo una risposta proporzionata e imponendo, di conseguenza, di contemperare l'efficacia della misura preclusiva con i costi dei margini di errore: questa versione del principio di precauzione, che prescrive in sostanza che la restrizione non sia eccessivamente

---

<sup>52</sup> V. per tutti N. DE SADELEER, *Les principes du pollueur-payeur, de prévention et de précaution. Essai sur la genèse et la portée juridique de quelques principes juridiques du droit de l'environnement*, Bruxelles, 1999, 39 ss.

<sup>53</sup> V. in proposito anche S. GRASSI, *Prime osservazioni*, cit., 43, il quale fa notare che la denominazione del principio in oggetto, nell'ordinamento tedesco, come *Vorsorgeprinzip* (composto da *vor*, che significa "prima di", e da *sorge*, che significa "timore") reca in sé l'idea di una «anticipazione».

<sup>54</sup> Cfr. in proposito S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 152.

Vi è tuttavia chi lo interpreta – propendendo, al contrario, per una lettura del principio, secondo quanto verrà meglio precisato *infra*, par. 4, in senso "debole" (e intendendo in senso "forte" solo formule come, ad esempio, la moratoria indiscriminata prevista dalla succitata Carta dell'ONU sulla natura del 1982) – come norma procedurale di gestione del rischio incerto: così ad esempio, D. AMIRANTE, *Il principio precauzionale*, cit., 23, il quale deduce dalla formula della Dichiarazione di Rio la necessità di individuare un «percorso, anche procedurale, che devono seguire i pubblici poteri nelle situazioni di incertezza; innanzi tutto la rilevazione del rischio [...], poi una sua analisi, ed infine la decisione sulle misure da adottare, che potrà anche prescindere dal raggiungimento di una certezza scientifica in merito alla quantificazione dei danni potenziali. Si avvia in tal modo la prefigurazione di regole generali per la "gestione" di rischi potenziali o incerti».

<sup>55</sup> In proposito si v. ad esempio la sentenza della Corte di Giustizia del 5 maggio 1998, causa C-180/96 (sulle misure di emergenza per fronteggiare l'encefalopatia spongiforme), che individua tra i criteri guida della decisione – oltre alla verifica dell'incertezza scientifica sul rischio, dell'adeguatezza e della proporzionalità della misura preclusiva – la valutazione della irreparabilità del danno ipotizzato (cfr. sul punto S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 168).

<sup>56</sup> Sul tema cfr., *ex multis*, S. GRASSI, *Prime osservazioni*, cit., 52 e S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 153 ss.

costosa<sup>57</sup>, non comporta un divieto assoluto di svolgere attività rischiose, ma richiede il compimento di una valutazione di costi e benefici dell'intervento limitativo<sup>58</sup>.

In particolare, nell'ipotesi di applicazione dello schema dianzi tracciato all'AI potenzialmente discriminatoria, il rapporto costi-benefici vede contrapposti i vantaggi derivanti dall'adozione delle misure preclusive dell'impiego della tecnologia algoritmica – in grado di mettere al riparo dai rischi di *AI-driven discriminations* – ai costi conseguenti alla rinuncia ad utilizzare gli algoritmi nelle procedure decisionali, consistenti principalmente nella riduzione dell'efficienza e della rapidità della decisione<sup>59</sup>. Tra l'altro, in ordine ai costi dei margini di errore, vale precisare che il rischio che l'algoritmo produca decisioni discriminatorie va comparato con il rischio che la discriminazione sia operata dall'essere umano: tale valutazione consente di collegare la soglia di rischio al c.d. «danno accettabile», che è tale se è pari ai rischi che si corrono nella «vita normale»<sup>60</sup>. Da questo punto di vista, non può affermarsi a priori che le macchine siano più – o meno – “inclinati a discriminare” rispetto a quanto lo siano gli uomini<sup>61</sup>: in questa prospettiva, potrebbe al limite sostenersi che, laddove i sistemi di apprendimento automatico creassero autonomamente – e in maniera sostanzialmente inconsapevole – nuove categorie di soggetti discriminati, probabilmente produrrebbero danni maggiori rispetto a quelli causati dall'essere umano, se non altro perché, in questo caso, potrebbe risultare estremamente complicato (e talvolta impossibile) rilevare le cause – e, in certo senso, le “ragioni” – del trattamento differenziato.

All'interno della valutazione costi-benefici, un rilevante fattore di riduzione del rischio legato all'impegno di sistemi di intelligenza artificiale è rappresentato dalle limitazioni imposte alla “significatività” del ruolo svolto dall'algoritmo nella decisione, limitazioni che si sostanziano di norma in interventi mirati a ridurre l'influenza dispiegata sulla decisione finale dall'elaborazione dei dati realizzata dal *software*, o dalle soluzioni eventualmente suggerite dal sistema informatico<sup>62</sup>: la presenza di un controllo umano sulla decisione finale, infatti – come peraltro espressamente stabilito dall'art. 22 del Regolamento UE 2016/679<sup>63</sup> e ribadito da alcuni tra i più rilevanti

---

<sup>57</sup> Così S. GRASSI, *Prime osservazioni*, cit., 45, osservando, con riferimento ai costi dei possibili margini di errore non prevedibili, che si inserisce «la considerazione dell'ignoranza nell'analisi costi/benefici» (*ibidem*).

<sup>58</sup> Un esempio concreto di declinazione del principio di precauzione che include un'analisi costi-benefici è contenuto nella Convenzione quadro sul cambiamento climatico del 1992, il cui art. 3, par. 3, riprende sostanzialmente il contenuto della Dichiarazione di Rio (stabilendo che «le Parti devono adottare misure precauzionali per rilevare in anticipo, prevenire o ridurre al minimo le cause dei cambiamenti climatici e per mitigarne gli effetti negativi» e che «qualora esistano rischi di danni gravi o irreversibili, la mancanza di un'assoluta certezza scientifica non deve essere addotta come pretesto per rinviare l'adozione di tali misure»), ma precisando che ciò deve avvenire «tenendo presente che le politiche e i provvedimenti necessari per far fronte ai cambiamenti climatici devono essere il più possibili efficaci in rapporto ai costi, in modo da garantire vantaggi mondiali al più basso costo possibile».

Un esempio analogo è presente all'interno del Codice ambientale francese del 2000, che consente il ricorso al principio di precauzione a condizione che comporti «costi economicamente accettabili» (cfr. sul tema D. AMIRANTE, *Il principio precauzionale*, cit., 29).

<sup>59</sup> I costi della rinuncia, peraltro, potrebbero rivelarsi assai più elevati laddove si considerasse – estendendo lo sguardo al di là della singola procedura decisionale – che «l'innovazione tecnologica, se attentamente monitorata, può apportare immensi benefici sociali, creando le premesse per una società più aperta e inclusiva» (così G. RESTA, *Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza*, in *Pol. dir.*, 2019, n. 2, 234).

<sup>60</sup> Cfr. in argomento la Risoluzione del Parlamento europeo del 2019 sulla robotica, cit., punto 24, e il Codice etico-deontologico degli ingegneri robotici, allegato alla precedente risoluzione del 2017, cit., par. “*Massimizzare i rischi e ridurre al minimo il danno*”.

<sup>61</sup> Cfr. sul punto F.Z. BORGESIU, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making*, Council of Europe, Strasbourg, 2018, 18.

<sup>62</sup> Anche sul punto cfr., volendo, P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, cit., 15 ss.

<sup>63</sup> L'art. 22, primo par., del [Regolamento \(UE\) 2016/679](#) del Parlamento europeo e del Consiglio del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati - GDPR), nell'ipotesi in cui si effettui un trattamento automatizzato di dati personali, vieta decisioni totalmente affidate al *software*. In particolare, è stabilito che «l'interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o

documenti di *soft law* in materia di intelligenza artificiale<sup>64</sup> – dovrebbe consentire di verificare la neutralità del risultato e correggere le eventuali distorsioni, operando a pieno titolo come misura precauzionale nel senso qui indicato.

Permangono tuttavia margini di incertezza sull'effettiva possibilità che la supervisione umana sia sempre in grado di far emergere la presenza di effetti distorsivi, che risultano difficilmente riconoscibili nell'ipotesi in cui un sistema di apprendimento automatico dia origine a *proxy discriminations*<sup>65</sup>.

Spunti egualmente utili a tracciare un quadro delle possibili applicazioni del principio di precauzione ai sistemi di AI potenzialmente discriminatori derivano dalla configurazione del principio in esame come «dovere di cautela» o «inversione dell'onere della prova» a carico di coloro che producono i cambiamenti (o, nel nostro caso, che propongono l'introduzione della tecnologia informatica)<sup>66</sup>: l'applicazione del principio di precauzione così declinato individuerrebbe una precisa responsabilità in capo al soggetto fornitore del *software* (in ipotesi il proprietario o il programmatore), chiamato a dimostrare che l'algoritmo non produrrà effetti distorsivi o che comunque il margine di errore si manterrà entro una soglia di tollerabilità. Una responsabilità certamente gravosa<sup>67</sup>, ma in grado di superare in parte il “diaframma” rappresentato dalla normativa sulle opere di ingegno e il segreto industriale, dietro al quale di celano le società private, che impedisce sovente di assicurare la piena trasparenza in ordine al funzionamento degli algoritmi<sup>68</sup>. Inoltre – considerato che, nel caso in cui si ricorra a sistemi di apprendimento automatico, una garanzia sull'assenza di *bias* può essere offerta solo se si è in grado di prevedere come il *machine learning* si evolverà – l'applicazione del principio di precauzione così configurato potrebbe stimolare un maggiore controllo sull'evoluzione autonoma dell'algoritmo, incrementando il livello di sicurezza delle nuove tecnologie informatiche.

Un'ultima possibile declinazione del principio di precauzione, in grado di offrire qualche spunto ulteriore ai fini di un'applicazione all'ambito informatico, è rappresentata dall'obbligo di pagare il debito ecologico causato nel passato. Tale obbligo, di norma riassunto nella formula “chi inquina paga”, prefigura un intervento riparatorio *ex post* sui danni ambientali, che rende in realtà controversa la sua ascrivibilità al novero delle misure precauzionali<sup>69</sup>. Ciò non di meno, il principio potrebbe rivelarsi utilmente applicabile all'ipotesi di utilizzo di sistemi di AI potenzialmente discriminatori se declinato come obbligo di pagare il debito *sociale* conseguente alle discriminazioni operate dall'algoritmo: una procedura adottabile, tuttavia, a condizione che sia

---

che incida in modo analogo significativamente sulla sua persona» (cfr. anche il successivo par. 2, che individua i casi in cui il trattamento automatizzato integrale dei dati personali è consentito).

Tuttavia, va segnalato che la profilazione può essere effettuata anche attraverso il ricorso ad un trattamento dei dati personali solo *parzialmente* automatizzato (cfr. sul punto G. RESTA, *Governare l'innovazione tecnologica*, cit., 206).

<sup>64</sup> Così, ad esempio, la Dichiarazione di Asilomar (*Asilomar AI Principles*, cit.), stabilisce, al principio n. 16 (*controllo umano*), che «gli uomini devono scegliere *come e se* delegare decisioni ai sistemi di intelligenza artificiale»; e il documento UE per un'AI affidabile (*Orientamenti etici per un'Ai affidabile*, cit.) individua, tra i requisiti necessari per garantire l'affidabilità dei sistemi di intelligenza artificiale, «l'intervento e la sorveglianza umana» (p. 2).

<sup>65</sup> Le *proxy discriminations* rappresentano una forma di discriminazione “indiretta”, basata sulla individuazione di caratteri distintivi (in base ai quali operare un trattamento differenziato) che, pur formalmente “neutri”, di norma ricorrono con particolare frequenza nelle categorie protette (in argomento, con specifico riferimento alle *proxy discriminations* originate dai sistemi informatici, cfr., *ex multis*, A. PRINCE, D.B. SCHWARCZ, *Proxy Discrimination in the Age of Artificial Intelligence and Big Data*, in *Iowa Law Review*, 5 agosto 2019).

<sup>66</sup> Cfr. in argomento S. GRASSI, *Prime osservazioni*, cit., 45.

<sup>67</sup> Dimostrare in anticipo che l'algoritmo non genererà effetti discriminatori, stanti le difficoltà nel prevedere compiutamente le conseguenze applicative dei sistemi di apprendimento automatico, costituisce per molti versi una *probatio diabolica* (sulle difficoltà connesse all'assolvimento degli obblighi informativi imposti dall'inversione dell'onere della prova v. G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1093).

<sup>68</sup> Cfr. sul tema A. VENANZONI, *La valle del perturbante: il costituzionalismo alla prova delle intelligenze artificiali e della robotica*, in *Pol. dir.*, 2019, n. 2, 238.

<sup>69</sup> In proposito si v., ad esempio, N. DE SADELEER (*Les principes du pollueur-payeur*, cit.), che riconduce il principio “chi inquina paga” al modello curativo di azione ambientale, distinto da quello preventivo e anticipatorio, in cui si collocherebbe il principio di precauzione.

possibile verificare *ex post* la correttezza delle previsioni elaborate dal *software*<sup>70</sup>. Qualora sia possibile operare una valutazione del genere, il principio consentirebbe di applicare il modello civilistico del risarcimento del danno ingiusto. In proposito, indicazioni utili ad affrontare la complessa problematica legata alla responsabilità civile per l'uso delle nuove tecnologie sono offerte dalla succitata Risoluzione del Parlamento europeo del 16 febbraio 2017, recante Raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica<sup>71</sup>: il documento – nel delineare gli scenari possibili di un futuro “diritto robotico” (che prefigurano, tra l'altro, l'obbligo di registrazione dei robot, l'assicurazione obbligatoria e la costituzione di un fondo di garanzia per i robot non assicurati) – introduce la nozione di “soggettività elettronica”, riferita ai robot in grado di compiere scelte autonome e ne postula la connessa responsabilità<sup>72</sup>. Un concetto, peraltro assai controverso<sup>73</sup>, che presenta anche declinazioni più attenuate, come ad esempio quella che configura le intelligenze artificiali e i robot come “semi-soggetti” ai quali andrebbero riconosciuti pretese, diritti, doveri e responsabilità, nella prospettiva di una sempre più intensa integrazione tra uomo e macchina<sup>74</sup>.

#### 4. (Segue) *Le due “varianti” essenziali del principio di precauzione e il loro potenziale impatto sui sistemi di intelligenza artificiale, alla luce dei criteri applicativi individuati dalla Commissione europea*

<sup>70</sup> Si consideri ad esempio il “caso COMPAS”, uno dei casi più noti di applicazione di algoritmi per assumere decisioni risultate, ad una successiva analisi, nettamente discriminatorie: il nome deriva dalla sigla identificativa di un *software* prodotto da una società privata americana, che consentiva di calcolare le probabilità di recidiva di un imputato – in base ai precedenti giudiziari, a questionari appositamente somministrati ed a caratteri ulteriori selezionati dall'algoritmo non rivelati dalla società fornitrice – e che produsse decisioni fortemente *biased*, sovrastimando le probabilità di recidiva degli imputati di colore e sottostimando quella degli imputati bianchi (cfr. in argomento, *ex multis*, A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma nell'intelligenza artificiale e il suo impatto sul diritto costituzionale*, in *Riv. fil. dir.*, 2019, n. 1, 94 ss.). In quella circostanza l'effetto discriminatorio era stato calcolato operando un confronto tra il tasso di recidiva previsto dall'algoritmo e quello effettivamente realizzatosi per ciascun imputato ed aggregando infine i dati per sesso e per gruppo etnico.

In tema, si v. le considerazioni di A. VENANZONI, *La valle del perturbante*, cit., 247, in merito al trattamento automatizzato dei dati giudiziari a fini predittivi e a fini decisori.

<sup>71</sup> In argomento v., da ultimo, A. CELOTTO, *Come regolare gli algoritmi*, cit., 50 ss.

<sup>72</sup> Cfr. il punto 59, lettera b). In tema v. A. AMIDEI, *Robotica intelligente e responsabilità: profili e prospettive evolutive del quadro normativo europeo*, in U. RUFFOLO (cur.), *Intelligenza artificiale e responsabilità*, Milano, 2017, 63 ss.

<sup>73</sup> Si v., ad esempio, il parere negativo espresso sul tema nel 2017 dal Comitato economico e sociale dell'UE, che si è dichiarato «contrario a introdurre una forma di personalità giuridica per i robot o per l'IA. In tal modo verrebbe infatti compromesso l'effetto di correzione preventiva della nozione di responsabilità giuridica, con possibilità di azzardo morale, sia nello sviluppo che nell'impiego dell'IA, e conseguentemente di abusi»: *Parere del Comitato economico e sociale europeo su «L'intelligenza artificiale – Le ricadute dell'intelligenza artificiale sul mercato unico (digitale), sulla produzione, sul consumo, sull'occupazione e sulla società»* (parere d'iniziativa) (2017/C 288/01), par. 1.12.

<sup>74</sup> Cfr. in argomento A. VENANZONI, *La valle del perturbante*, cit., 273, il quale osserva a riguardo che «porre le basi di un costituzionalismo che inglobi nella sua narrazione anche IA e robot, come semi-soggetti cui riconoscere pretese, diritti, doveri e responsabilità, non significa voler tracciare una linea oscura puntata verso un orizzonte distopico; significa, molto più semplicemente, prendere atto della complessità stordente della socialità ibrida, in cui silicio e carne, anima e processori, sono incistati in maniera ormai indistinguibile, e arrivare alla conclusione per cui qualunque regolazione costituzionale che inglobi anche la macchina intelligente è presidio innanzitutto dell'umano. Ciò che va costituzionalizzato è quindi il sistema complesso, a network, che assomma progettazione, realizzazione, *output* finale, elementi intersezionali e integrativi tra uomo e macchina, e ipostatizzi l'ambiente complesso risultante da questi incistamenti».

Il tema appare destinato ad assumere un rilievo crescente negli anni a venire, alla luce del fatto che i «costrutti matematici o computazionali di per sé non dispiegano un impatto negativo sui diritti umani, ma tale impatto è determinato dalla loro implementazione ed applicazione alle interazioni umane» (B. WAGNER, *Algorithms and human rights. Study on the human rights dimensions of automated data processing techniques and possible regulatory implications*, prepared by the Committee of Experts on internet intermediaries (MSI-NET) for the Council of Europe, 2018, 8).

Emerge, dalla sintetica rassegna che si è proposta nelle pagine che precedono, l'estrema eterogeneità dei significati che il principio di precauzione ha assunto nel tempo<sup>75</sup>, molti dei quali presentano a loro volta rilevanti margini di indeterminatezza; la complessità risulta ulteriormente accentuata dal carattere multidimensionale del principio, che tende ad assumere significati diversi nell'ambito di ciascuna dimensione (scientifica, economica, politica, etica, giuridica) in cui viene in rilievo<sup>76</sup>.

Tuttavia, la varietà dei significati che si sono individuati possono tendenzialmente ricondursi<sup>77</sup> a due fondamentali "varianti" del principio, che sottendono due diversi modi di affrontare i rischi ignoti. Il principio di precauzione, cioè, può essere inteso essenzialmente in due sensi. In senso "forte" (o sostanziale), come "regola per decidere": in questo caso postula un obbligo di astensione dalle attività di cui siano ignoti i potenziali effetti negativi. Ovvero in senso "debole" (o formale), come "regola per procedere": in questo caso esso postula un obbligo di "presa in considerazione" dell'incertezza scientifica nell'ambito delle analisi di rischio.

In particolare, il principio di precauzione nel suo significato "forte" implica la preclusione totale dell'attività che comporti rischi ignoti: a questo esito conduce la declinazione del principio nei termini di inversione dell'onere della prova<sup>78</sup> e di anticipazione preventiva<sup>79</sup>.

La versione "debole" del principio – che rappresenta peraltro il modello prevalente in Italia e in Europa<sup>80</sup> – comporta invece la preclusione parziale dell'attività potenzialmente dannosa, che, in una situazione di incertezza scientifica sugli effetti potenziali, va ridotta al minimo ma non vietata<sup>81</sup>: a questo esito conduce la declinazione del principio in termini di proporzionalità della risposta (o dell'efficacia) rispetto ai costi dei margini di errore (o dei rischi). Una declinazione che – riferita all'impiego di algoritmi potenzialmente discriminatori – pone al centro il problema del *quantum* e impone sempre una valutazione preliminare della significatività dell'algoritmo nella decisione<sup>82</sup>; valutazione certo non sempre facile, sia perché coincidente talvolta con una *probatio diabolica*<sup>83</sup>, sia in quanto avente ad oggetto decisioni spesso sottoposte a "cattura" da parte del

---

<sup>75</sup> Le ragioni di tale eterogeneità vanno verosimilmente ricercate – come prima ricordato – nella vaghezza della formulazione originaria del principio all'interno del diritto internazionale, così configurato allo scopo di scongiurare eccessive compressioni della sovranità degli Stati aderenti alle convenzioni internazionali in materia ambientale: cfr. in argomento G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1089, il quale sottolinea che la vaghezza deriva anche dalla caratterizzazione della precauzione come principio di politica del diritto (ivi, 1098), «talché, in definitiva, pure la mancata determinazione a priori delle misure precauzionali, e la necessaria graduazione delle stesse in relazione alle specificità del caso concreto, paiono costituire un'implicazione logica del principio in parola. V'è però da dire che da ciò discende un dato – questo sì – decisamente peculiare, ovvero sia l'attribuzione di una sfera di discrezionalità amplissima, e, soprattutto, del tutto inusitata, alle autorità deputate all'attuazione del principio che qui interessa»; per cui «esso indica *quando* occorre intervenire, e, tutt'al più, *come* bisogna intervenire, ovvero sia i parametri secondo cui vanno commisurati gli interventi cautelativi: ma non individua con precisione *cosa* bisogna fare nel contesto di detti interventi, ossia le misure che di volta in volta vanno adottate in concreto» (*ibidem*).

L'eterogeneità risulta peraltro confermata dalla varietà dei criteri applicativi, come si dirà meglio *infra*.

<sup>76</sup> Cfr. in proposito S. GRASSI, *Prime osservazioni*, cit., 37-38.

<sup>77</sup> Cfr. sul tema F. FONDERICO, *Tutela dall'inquinamento elettromagnetico e amministrazione 'precauzionale'*, in *Giorn. dir. amm.*, 2004, 339. e J.B. WIENER, *Precaution*, in D. BODANSKY, J. BRUNNÉE, E. HEY (curr.), *The Oxford handbook of international environmental law*, Oxford University Press, Oxford, 2008, 597 ss.

<sup>78</sup> Così ad esempio S. GRASSI, *Prime osservazioni*, cit., 46, il quale sottolinea in proposito il rischio che la responsabilità per danni imprevedibili – e dunque senza colpa – soffochi l'innovazione.

<sup>79</sup> Ma sul punto si v. le perplessità espresse da D. AMIRANTE, *supra*, nota 54.

<sup>80</sup> Pur in presenza di significative variazioni nazionali in relazione all'apprezzamento di determinati rischi, come ricordato da M. GRAZIADEI, *La regolazione del rischio*, cit., 505.

<sup>81</sup> Cfr. in tema S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 154.

<sup>82</sup> Sul tema cfr. par. precedente.

<sup>83</sup> Così A. SIMONCINI, S. SUWEIS, *Il cambio di paradigma*, cit., 100.

sistema informatico e dunque influenzate *di fatto* dall'algoritmo in misura ampiamente superiore rispetto a quanto formalmente previsto<sup>84</sup>.

Ciascuna delle varianti del principio di precauzione sconta peraltro delle specifiche limitazioni: le versioni "forti" riescono ad imporre adeguate restrizioni alle tecnologie future (a garanzia degli interessi che intendono proteggere), ma conducono sovente a direttive operative contraddittorie<sup>85</sup>. Le versioni "deboli", invece, talvolta si rivelano inefficaci, non riuscendo ad opporre alcuna restrizione allo sviluppo di tecnologie potenzialmente dannose<sup>86</sup>.

Muovendo dalle coordinate tracciate, spunti utili a chiarire il significato e a identificare il valore del principio in esame si ricavano dalla *Comunicazione della Commissione UE sul principio di precauzione* del 2 febbraio 2000<sup>87</sup>.

In primo luogo, il documento offre una chiara indicazione dei presupposti che giustificano il ricorso al principio, affermando che esso trova applicazione «in condizioni in cui le informazioni scientifiche sono insufficienti, non conclusive o incerte» e vi sono indicazioni che i possibili effetti dell'attività che si intende sottoporre a misure precauzionali «possono essere potenzialmente pericolosi»<sup>88</sup>.

In secondo luogo, individua due diversi momenti nei quali il principio può trovare applicazione: anzitutto, la fase preliminare della decisione politica, nella quale occorre stabilire se adottare o meno misure precauzionali. In questa fase sono essenzialmente due i fattori che impongono di agire: l'esistenza di un rischio potenziale derivante da un fenomeno, un prodotto o un procedimento e l'incertezza scientifica sulla portata del rischio<sup>89</sup>. Il secondo momento è quello della predisposizione delle misure precauzionali, che andrebbero costruite sulla base dei criteri indicati nel documento: proporzionalità, non discriminazione, coerenza con le misure adottate in situazioni analoghe, valutazione del rapporto costi-benefici, attenzione all'evoluzione scientifica.

L'analisi dei criteri richiamati può rivelarsi estremamente proficua, non solo perché essi forniscono indicazioni utili a guidare la concreta applicazione del principio di precauzione, ma anche perché operano come altrettanti limiti ad un ricorso "drastico" alla sola preclusione delle tecnologie innovative, favorendo ulteriormente la lettura "debole" del principio in esame.

Il primo criterio applicativo del principio di precauzione individuato dalla Comunicazione della Commissione è quello di *proporzionalità* (della misura restrittiva) rispetto al livello di protezione ricercato o «*standard* precauzionale» di tutela<sup>90</sup>. L'individuazione di un livello di protezione "voluta" o "adeguata" comporta l'implicito riconoscimento che non sempre si deve ricercare il livello di protezione massimo e, per contro, che non sempre si deve conseguire un livello di rischio zero: il livello di rischio zero, infatti, può essere raggiunto assai di rado (e, segnatamente, nel caso di possibili discriminazioni di origine algoritmica, occorre tener conto che neanche l'essere umano è, per così dire, "a rischio zero", essendo i pregiudizi espressione di una diffusa "inclinazione" dell'animo umano). Tale criterio può rappresentare un fondamentale elemento di temperamento

---

<sup>84</sup> Tra le cause principali del fenomeno della "cattura" sono state individuate l'esonero della responsabilità, la riduzione del carico di lavoro e l'«aura di razionalità e infallibilità» che accompagna l'algoritmo (così F.Z. BORGESIU, *Discrimination, Artificial Intelligence*, cit., 8).

<sup>85</sup> V. in proposito l'analisi di S. CLARKE, *Future technologies*, cit., il quale sottolinea, tra l'altro, che interrompere le sperimentazioni in settori scientifici considerati pericolosi potrebbe impedire l'individuazione di soluzioni adeguate a prevenire rischi potenziali.

<sup>86</sup> *Ibidem*.

<sup>87</sup> Sulla quale si v., *ex multis*, le analisi di T. MAROCCO, *Il principio di precauzione*, cit., 1237 ss. e di G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit. 1092 ss.

<sup>88</sup> Così la *Comunicazione della Commissione del 2 febbraio 2000*, cit., par 1 ("Introduzione").

<sup>89</sup> Ivi, punto 5.1.3. Il rischio, inoltre, deve risultare «socialmente accettabile» in relazione ai benefici (cfr. punto 5): un criterio che appare particolarmente rilevante ai fini di una possibile applicazione del principio di precauzione all'uso dell'intelligenza artificiale, considerato il peso detenuto dai benefici sociali legati allo sviluppo tecnologico.

<sup>90</sup> Riprendendo l'espressione impiegata da S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 149, con riguardo al livello di protezione cui fa riferimento l'art. 174, par. 2, trattato C.E.

contro l'eccesso di precauzione (o contro un impiego "assolutistico" del principio)<sup>91</sup> che condurrebbe a vietare indiscriminatamente il ricorso all'intelligenza artificiale: talvolta, infatti, un divieto totale può non essere la risposta proporzionale ad un rischio potenziale<sup>92</sup>.

Il secondo criterio indicato dalla Comunicazione in esame è quello di *non discriminazione* nell'applicazione del principio precauzionale, allo scopo di garantire il rispetto delle regole di concorrenza nel mercato interno europeo<sup>93</sup>: criterio utile a scongiurare un ricorso improprio al principio di precauzione per limitare lo sviluppo delle nuove tecnologie di intelligenza artificiale a fini anticoncorrenziali.

Il terzo criterio richiede che le misure precauzionali siano *coerenti* con le misure adottate in situazioni analoghe – caratterizzate parimenti dalla mancanza di dati scientifici in grado di definire il rischio – o che esse riflettano comunque il medesimo approccio al problema: in questa prospettiva, qualunque analisi volta ad individuare tali analogie può rappresentare uno spunto non solo utile, ma in certa misura doveroso nella prospettiva di una futura applicazione del principio di precauzione all'intelligenza artificiale.

Il quarto criterio prescrive l'*analisi* preliminare dei *vantaggi e degli oneri* derivanti dall'azione o dall'inazione, a conferma che la ricerca di un equilibrio tra le ragioni della prudenza e quelle dell'innovazione – la quale implica anche, «se appropriata e fattibile, un'analisi economica di costi e benefici»<sup>94</sup> – è intrinseca allo stesso principio di precauzione (almeno nella versione "debole" che caratterizza la disciplina europea e italiana)<sup>95</sup>.

Il quinto e ultimo criterio applicativo impone di *rimodulare* le misure precauzionali *in relazione all'evoluzione scientifica*<sup>96</sup>: ciò comporta che tali misure vanno mantenute solo finché i dati scientifici rimangono insufficienti o non concludenti<sup>97</sup> e finché il rischio sia ritenuto sufficientemente elevato per non accettare di farlo sostenere alla società. Anche in questo caso si ammette che non va necessariamente perseguito il rischio zero, ma occorre invece ridurre i rischi ad un livello accettabile, temperandoli con i benefici attesi in applicazione del principio di proporzionalità<sup>98</sup>. In particolare, i responsabili delle decisioni sono chiamati a stabilire il livello accettabile del rischio per la società<sup>99</sup>: una valutazione eminentemente politica<sup>100</sup>, intesa a realizzare un bilanciamento tra una lesione *attuale* prodotta dall'adozione delle misure precauzionali (alla libertà d'impresa, alle prospettive di sviluppo economico, al progresso tecnologico) e una lesione *virtuale* derivante dalla mancata adozione delle stesse (con riferimento ai valori messi a rischio, che includono a pieno titolo il principio di non discriminazione). E proprio la necessità di tale

<sup>91</sup> Si v. in argomento le convincenti considerazioni di G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1106.

<sup>92</sup> Così T. MAROCCO, *Il principio di precauzione*, cit., 1239.

<sup>93</sup> Cfr. sul punto S. GRASSI, *Prime osservazioni*, cit., 59.

<sup>94</sup> Così in proposito S. GRASSI, *Prime osservazioni*, cit., 60.

<sup>95</sup> Come osserva opportunamente a riguardo S. GRASSI, *Prime osservazioni*, cit., 60, «si tratta di applicare, con riferimento al principio di precauzione, i criteri del bilanciamento tra i valori costituzionali e comunitari che entrano in gioco».

<sup>96</sup> Implicando di conseguenza il loro necessario riesame alla luce dell'evoluzione scientifica: cfr. sul punto anche M. GRAZIADEI, *La regolazione del rischio*, cit., 504.

<sup>97</sup> Il che significa che le misure limitative sono tendenzialmente temporanee, anche perché la "scientificità" delle ragioni che le giustificano risente dell'andamento fallibilista e graduale della scienza: così G.D. COMPORI, *Amministrazioni e giudici*, cit., 2466.

<sup>98</sup> Vanno nel senso della (sola) riduzione del rischio – sottintendendo che la tecnologia va, per quanto possibile, utilizzata – anche le concrete misure di intervento suggerite dalla Comunicazione in esame (cfr. sul punto anche le osservazioni di G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1093). Si tratta peraltro di misure assai blande, come «finanziare un programma di ricerca» (al fine di ridurre l'incertezza scientifica) e «informare l'opinione pubblica sui possibili effetti negativi di un prodotto o procedimento» (evitando, in questo caso, l'eccesso di cautela, qualora sia adottata al solo scopo di rassicurare l'opinione pubblica: cfr. sul punto G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1104).

<sup>99</sup> Cfr. sul punto S. GRASSI, *Prime osservazioni*, cit., 58.

<sup>100</sup> Spettante – come precisa proposito S. GRASSI, *Prime osservazioni*, cit., 41 – all'opinione pubblica, ai cittadini e agli organi rappresentativi.



bilanciamento – nel quale risulta «decisiva la considerazione della proporzionalità della misura di cautela»<sup>101</sup> – fa propendere, ancora una volta, per l'accoglimento di una interpretazione “debole” del principio di precauzione.

La Comunicazione in esame stabilisce infine che le misure precauzionali devono allocare correttamente la *responsabilità* per la produzione delle *prove scientifiche* necessarie ad una più completa valutazione del rischio<sup>102</sup>: un principio che rinvia al profilo, dianzi evidenziato, della «inversione dell'onere della prova», intesa come «forma di attribuzione della responsabilità di fornire evidenze scientifiche [sull'assenza di rischi] a chi intenda agire [nel nostro caso, introducendo sistemi di intelligenza artificiale nelle procedure decisionali]. Sotto questo profilo, il principio di precauzione si collega con il principio civilistico dell'art. 2050 del codice civile e con i criteri classici della prevenzione cui sono sottoposte le attività industriali particolarmente pericolose»<sup>103</sup>.

##### 5. Note conclusive: l'intelligenza artificiale tra precauzione, prevenzione e innovazione

Il principio di precauzione – come più volte rilevato nelle pagine precedenti – presenta una sostanziale affinità col principio di prevenzione, derivante dalla comune attitudine ad operare come strumenti di gestione del rischio. Tuttavia, la natura del rischio considerato marca la distanza tra i due principi: il principio di precauzione si applica in presenza di rischi ignoti, che assumono il carattere di minaccia, in quanto non è possibile determinare né l'entità né il momento in cui l'evento dannoso potrebbe verificarsi<sup>104</sup>. Esso rappresenta di conseguenza una risposta agli inconvenienti derivanti dalla incertezza scientifica<sup>105</sup>. Il principio di prevenzione, invece, si applica in presenza di rischi noti, che assumono il carattere di pericolo, riferibili ad eventi dannosi attestati da evidenze scientifiche<sup>106</sup>.

La distanza tra i due principi, in particolare, si manifesta sia sul piano, per così dire, “qualitativo”, riflettendo due diverse percezioni dei pericoli derivanti dal progresso (in relazione al carattere certo o incerto del rischio)<sup>107</sup>, sia sul piano, in certo senso, “quantitativo”: il principio di precauzione, infatti – soprattutto nella versione originaria, che lo vede configurato in termini di anticipazione preventiva – anticipa la soglia al di là della quale si attivano le misure preventive<sup>108</sup> e determina di conseguenza un innalzamento del livello di tutela<sup>109</sup>.

Venendo allora al ruolo che possono svolgere i due principi nella gestione dei rischi di discriminazione legati all'impiego di sistemi di intelligenza artificiale, può osservarsi che il principio di prevenzione può trovare applicazione sia nella fase di elaborazione dell'algoritmo – che va impostato in modo tale da impedire la produzione di effetti discriminatori diretti e indiretti<sup>110</sup> –,

<sup>101</sup> Così S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 158.

<sup>102</sup> Cfr. in argomento M. GRAZIADEI, *La regolazione del rischio*, cit., 504.

<sup>103</sup> Così S. GRASSI, *Prime osservazioni*, cit., 60.

<sup>104</sup> Così D. AMIRANTE, *Il principio precauzionale*, cit., 21.

<sup>105</sup> Così G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1085. Cfr. in tema S. GRASSI, *Prime osservazioni*, cit., 39, che riconduce il principio in esame al «diritto della scienza incerta».

<sup>106</sup> Anche se va rammentato che tra «certezza» e «incertezza» scientifica permangono sempre delle «zone grigie» (così S. GRASSI, A. GRAGNANI, *Il principio di precauzione*, cit., 154).

<sup>107</sup> Cfr. sul punto N. DE SADELEER, *Gli effetti del tempo, la posta in gioco e il diritto ambientale*, in *Riv. giur. amb.*, 2001, n. 5, 598.

<sup>108</sup> Cfr. in proposito M. CECCHETTI, *Principi costituzionali per la tutela dell'ambiente*, Milano, 2000, 189 ss.

<sup>109</sup> Cfr. sul punto S. GRASSI, *Prime osservazioni*, cit., 50, con specifico riferimento al «livello di protezione elevato» prescritto dalla normativa UE in materia di salute, ambiente e tutela dei consumatori. Si v. in proposito le osservazioni formulate *supra* in ordine al *Vorsorgeprinzip*.

<sup>110</sup> In particolare, è possibile intervenire nella fase di programmazione, “ordinando” al sistema di apprendimento automatico di non prendere in considerazione specifici caratteri, presenti nel *data set* di riferimento: o che identificano in via immediata categorie protette (quali ad esempio la razza o l'orientamento sessuale), scongiurando in tal modo il prodursi di discriminazioni algoritmiche “dirette”; o che si presentano formalmente come caratteri “neutri”, ma che

sia nella fase di selezione dei dati da immettere nel sistema informatico – che va operata anche attraverso adeguati interventi di “depurazione” dalle informazioni sulle quali potrebbero fondarsi trattamenti discriminatori<sup>111</sup>. Il principio di precauzione trova invece una sede ideale di applicazione nell'imposizione di limiti all'evoluzione autonoma (e, per certi versi, incontrollata) dei sistemi di *machine learning*. Il processo di selezione dei caratteri e degli elementi rilevanti per la decisione, infatti, può essere in parte indirizzato in sede di programmazione, ma il percorso che conduce ad individuare correlazioni significative all'interno dei dati, specialmente in presenza di un *data set* aperto, sfugge al pieno controllo dei programmatori: è questa peraltro la fase della decisione algoritmica nella quale risulta più elevato il rischio di *bias*, legato alla possibilità che il sistema generi nuove classi di individui discriminati in base a caratteri ricorrenti individuati autonomamente dal sistema informatico<sup>112</sup>.

Vale la pena, in conclusione, di soffermarsi sulla possibilità di applicare al settore informatico il principale istituto qualificante il principio di prevenzione nella sua declinazione “classica”, sviluppatasi originariamente all'interno del diritto dell'ambiente: la valutazione d'impatto ambientale, la quale mira – come è noto – a determinare in modo (tendenzialmente) certo gli effetti sull'ambiente di opere e interventi<sup>113</sup>.

Il tema appare particolarmente rilevante, in quanto una specifica “valutazione d'impatto” riferita all'uso delle tecnologie informatiche è prevista espressamente dal diritto europeo: il richiamo va all'art. 35 del Regolamento UE 2016/679, che prescrive, tra l'altro, una valutazione d'impatto del trattamento automatizzato di dati personali anche con riferimento a potenziali effetti discriminatori. In particolare, la disposizione evocata (rubricata *Valutazione d'impatto sulla protezione dei dati*), stabilisce, al par. 1, che «quando un tipo di trattamento, allorché prevede in particolare l'uso di nuove tecnologie, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il titolare del trattamento effettua, prima di procedere al trattamento, una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali [...]». Tra i trattamenti che prevedono «l'uso di nuove tecnologie» rientrano a pieno titolo i trattamenti automatizzati: lo stabilisce espressamente il successivo par. 3, che richiede la valutazione d'impatto «in particolare» nel caso di «una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche» (e segnatamente nell'ipotesi di trattamento, su larga scala, di dati sensibili)<sup>114</sup>; mentre i rischi «elevat[i] per i diritti e le libertà delle persone fisiche» includono senz'altro il rischio di discriminazioni<sup>115</sup>.

---

ricorrono di norma con particolare frequenza nelle categorie protette (quali ad esempio il codice postale, utilizzato come dato indicativo dell'appartenenza ad una specifica comunità discriminata o per ragioni razziali o per ragioni socio-economiche), scongiurando in tal modo il prodursi di discriminazioni “indirette” (o *proxy discriminations*): sul punto si consenta di rinviare a P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, cit., 5 ss.

V. in proposito le prescrizioni contenute nel *Codice etico-deontologico degli ingegneri robotici*, cit. e le considerazioni sviluppate in tema da A. CELOTTO, *Come regolare gli algoritmi*, cit., 55 ss. V. inoltre i suggerimenti formulati da A. SIMONCINI, S. SUWEIS (*Il cambio di paradigma*, cit., 104) in ordine alla necessaria educazione dei programmatori ai valori etici (all'interno dei quali deve considerarsi ricompreso a pieno titolo il principio di eguaglianza formale).

<sup>111</sup> Maggiori difficoltà sorgono tuttavia nell'ipotesi di acquisizione di *Big data* dalla rete Internet: in questo caso i rischi a cui viene esposto il sistema di apprendimento automatico, progettato per individuare autonomamente caratteri ricorrenti all'interno di un *set* di dati “aperto”, non sono pienamente prevedibili (o, meglio, è altamente probabile – e dunque ampiamente prevedibile – che sistema di *machine learning* “assorbirà” i pregiudizi diffusi in Rete, anche nelle versioni più esasperate).

<sup>112</sup> Anche sul punto cfr., volendo, P. ZUDDAS, *Intelligenza artificiale e discriminazioni*, cit., 11.

<sup>113</sup> Cfr. in argomento S. GRASSI, *Prime osservazioni*, cit., 40.

<sup>114</sup> Cfr. lett. a) e lett. b).

<sup>115</sup> Cfr. in argomento L. EDWARDS, M. VEALE, *Slave to the algorithm: Why a right to an explanation is probably not the remedy you are looking for*, in *Duke Law & Technology Review*, 2017.

Ma non solo: l'invito ad introdurre e sviluppare metodologie di valutazione dell'impatto algoritmico compare in alcuni tra i più importanti atti di *soft law* dedicati all'uso di sistemi di intelligenza artificiale. Così, ad esempio, nei succitati *Orientamenti etici per un'AI affidabile*, si invita ad effettuare, nelle fasi preliminari allo sviluppo del sistema informatico, una valutazione d'impatto sui diritti fondamentali<sup>116</sup> e si sollecita la predisposizione di una valutazione di impatto algoritmica al fine di ridurre al minimo l'impatto negativo in misura proporzionale ai rischi<sup>117</sup>; parimenti, nella succitata Risoluzione del Parlamento europeo del 2019 si propone l'introduzione di una Analisi di Impatto Algoritmica, da svolgersi preliminarmente all'impiego degli algoritmi nei sistemi decisionali quando questi hanno un «impatto significativo sulla vita dei cittadini»<sup>118</sup>. Analoghe finalità sono presenti nei documenti in materia adottati in Italia: in particolare, nel Libro Bianco dell'AgID si indica, tra le «sfide» da raccogliere per il futuro sviluppo dell'IA, la «misurazione dell'impatto nell'uso di soluzioni di intelligenza artificiale nella PA»<sup>119</sup>, mentre nella Strategia Nazionale predisposta dal MISE si propone di «istituire una cabina di regia interministeriale e multi-stakeholder» al fine di assicurare, tra l'altro, la «valutazione [dell']impatto economico e sociale» dei sistemi di intelligenza artificiale<sup>120</sup>.

Un accenno, infine, al principio di innovazione, che intrattiene col principio precauzionale un rapporto assai controverso: se infatti, da un lato, si è affermato – abbracciando *in toto* la tesi della contrapposizione – che l'equilibrio tra i due principi riflette l'oscillazione tra pensiero scienziato e anti-scienziato<sup>121</sup>, dall'altro lato, può sostenersi che tale equilibrio è in realtà già implicato dal principio di precauzione (segnatamente nella sua versione “debole”), che propugnerebbe una sorta di “equidistanza” tra accettazione acritica e rifiuto altrettanto acritico dell'innovazione<sup>122</sup>. E comunque, anche volendo collocare i due principi su piani contrapposti, non si tratterebbe di una contrapposizione netta: il principio di precauzione, infatti, riflette in realtà un atteggiamento di coraggio, non di paura rispetto al rischio<sup>123</sup> in quanto, pur nascendo dall'incertezza per il futuro, esige tuttavia l'attivazione di forme razionali e procedurali che superino la paura derivante dalla mancanza di dati<sup>124</sup>. In questa prospettiva si colloca l'idea, d'anzì affermata, di un'applicazione del principio di precauzione che consenta di controllare e rendere “sicura” l'evoluzione autonoma dei sistemi di *machine learning*, espressione della concezione “debole” del principio legata alla declinazione che mira a contemperare efficacia e costi; e che disegna un principio «volto non tanto a rallentare il progresso e la tecnologia, ma ad individuare strumenti di azione e di gestione nei confronti dei rischi prodotti da una tecnologia che non riesce più a controllare sé stessa»<sup>125</sup>.

<sup>116</sup> Cfr. *Orientamenti etici per un'AI affidabile*, cit., punto 63.

<sup>117</sup> Ivi, punto 89.

<sup>118</sup> Risoluzione del Parlamento europeo del 12 febbraio 2019 su *una politica industriale europea globale in materia di robotica e intelligenza artificiale*, cit., punto 154.

<sup>119</sup> *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino*, cit., pp. 66-68.

<sup>120</sup> *Strategia Nazionale per l'Intelligenza Artificiale*, cit., p. 7. In particolare, va sottolineato come nei documenti che delineano la strategia nazionale in materia di IA si concentri l'attenzione soprattutto sull'uso delle tecnologie in esame nell'ambito della pubblica amministrazione e sul loro impatto sociale.

<sup>121</sup> Cfr. sul punto l'efficace sintesi delle posizioni in campo proposta da G. MANFREDI, *Note sull'attuazione del principio di precauzione*, cit., 1082-84.

<sup>122</sup> Cfr. in argomento M.V. LUMETTI, *Il principio di precauzione nella legislazione, nella giurisprudenza e nelle recenti sentenze della Corte costituzionale*, in *Giust. amm.*, 2004, n. 1, 43.

<sup>123</sup> Su questi temi v. N. LUHMANN, *Sociologia del rischio*, Milano, 1996.

<sup>124</sup> Cfr. S. GRASSI, *Prime osservazioni*, cit., 39. Un chiaro esempio di tale impostazione è offerto dalla Dichiarazione di Rio del 1992, che inserisce il principio di precauzione tra i principi idonei a garantire uno sviluppo sostenibile, che dunque non va arrestato bensì accompagnato con le cautele necessarie affinché proceda senza causare danni irreparabili.

<sup>125</sup> Così D. AMIRANTE, *Il principio precauzionale*, cit., 20.