

ROBERTO REDAELLI

## IL PROFILO NORMATIVO DELL'INTELLIGENZA ARTIFICIALE. UN'INDAGINE SULLA RELAZIONE TRA SISTEMI INTELLIGENTI E NORMATIVITÀ MORALE

1. Normatività morale e agenti artificiali
2. La questione normativa: uno sguardo d'insieme sulle diverse teorie della moralità
3. Forme di normatività digitale
4. L'intelligenza artificiale da una prospettiva postfenomenologica e la nozione di quasi-normatività
5. La normatività morale tra uomo e intelligenza artificiale

**ABSTRACT: THE NORMATIVE PROFILE OF ARTIFICIAL INTELLIGENCE: A STUDY ON THE RELATIONSHIP BETWEEN INTELLIGENT SYSTEMS AND MORAL NORMATIVITY**

The paper aims to investigate so-called digital normativity understood as the binding force exerted on our lives by the predictions and standards produced by intelligent systems. In order to investigate such normative force in the moral field, we make use of the postphenomenological reflections of D. Ihde and P.P. Verbeek. In particular, starting from the definition of artificial intelligence in the terms of quasi-other proposed by Ihde and the function of technological mediation highlighted by Verbeek, we propose to define the normative force exerted by intelligent systems in the terms of quasi-normativity. By this expression we intend to highlight the peculiar status of the normative demands raised by such systems. In fact, according to the proposed thesis, such instances, although not morally neutral, cannot be considered normative in the strong sense, since this kind of strong normativity requires human approval. In this direction of inquiry, we suggest rethinking as the (hybrid) origin of moral normativity the human-technology interaction.



By this expression we intend to highlight the peculiar status of the normative demands raised by such systems. In fact, according to the proposed thesis, such instances, although not morally neutral, cannot be considered normative in the strong sense, since this kind of strong normativity requires human approval. In this direction of inquiry, we suggest rethinking as the (hybrid) origin of moral normativity the human-technology interaction.

### 1. Normatività morale e agenti artificiali

La normatività informa le nostre vite: le nostre azioni e i nostri pensieri seguono regole. Le norme del diritto e dell'etica, così come quelle della logica e del linguaggio, strutturano i vari ambiti della nostra esistenza mondana tanto sociale quanto privata, dando forma alla nostra quotidianità<sup>1</sup>. Così quando facciamo la spesa o guidiamo l'auto ci atteniamo a norme ben

<sup>1</sup> In questo senso si potrebbe definire goethianamente il dovere (*Pflicht*) nei termini di «ciò che il giorno esige» (J.W. Goethe, *Massime e riflessioni*, tr. it. Theoria, Roma-Napoli 1990, p. 99).

precise e ci aspettiamo che gli altri membri della comunità agiscano allo stesso modo<sup>2</sup>.

Un esempio ancora più prossimo a noi della pervasività propria della sfera normativa nella nostra esistenza è rappresentato dal testo che avete sotto mano: esso è il risultato di una peculiare pratica di scrittura che si fonda su regole, a cui chi scrive deve attenersi se intende dar forma a uno scritto, sia esso un saggio, un articolo, un libro o una semplice email. Specularmente, chi legge non può, al fine di comprendere il messaggio dell'autore, che conformarsi alle rispettive norme imposte dalla lingua impiegata dallo scrivente: in questo caso specifico, egli leggerà da sinistra a destra, avrà una certa dimestichezza con la lingua usata, wittgensteinianamente conoscerà le regole del gioco linguistico, e seguirà le norme proprie della coordinazione senso-motoria al fine di girare le pagine e focalizzare lo sguardo sulle frasi. Nelle complesse operazioni di leggere e scrivere, così come in quelle rappresentate dal guidare l'auto o fare la spesa, è, dunque, attivo un complesso intreccio di pratiche<sup>3</sup>, guidate da diversi set di regole necessarie all'esercizio di tali operazioni.

---

<sup>2</sup> Fare la spesa e guidare l'auto rappresentano degli esempi emblematici della complessità normativa che informa la nostra esistenza, tanto che entrambi sono spesso chiamati in causa, quali esemplificazioni di istituzioni complesse, in campi di ricerca tra loro disparati. Celebre, a questo proposito, è la definizione di istituzione offerta da Douglass North per cui «institutions are the rules of the game in a society or more formally, the humanly devised constraints that shape human interactions. [...] They are a guide to human interaction so that when we wish to greet friends on the street, drive an automobile, buy oranges, borrow money, form a business, bury our dead, or whatever, we know (or can learn easily) how to perform these tasks» (D. North, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge 1990, pp. 3-4). Un altro celebre esempio è offerto, nel campo dell'antropologia evolutiva, da M. Tomasello, che, nella scia delle riflessioni di Searle sulla costruzione della realtà sociale, distingue *foraging versus shopping*, riconoscendo nel fare la spesa una realtà istituzionale (*institutional reality*) in cui emerge «a uniquely human sense of 'we,' a sense of shared intentionality» (M. Tomasello, *Why we cooperate*, The MIT Press, Cambridge 2009, p. 57).

<sup>3</sup> Si intende qui la nozione di pratica, di certo legata a quella wittgensteiniana di gioco linguistico, in senso siniano: «Every life practice is a [form of] 'wisdom' *sui generis*. At the least, it is knowing how to do this and that (to stand, walk, grasp, and so on); then, it is knowing how to say; and finally, it is knowing how to write, in all the sense of this expression» (C. Sini, *Ethics of Writing*, SUNY, New York 2009, p. 104).

Queste regole di varia natura portano in primo piano un peculiare carattere dell'esistenza umana che coinvolge tanto le cosiddette funzioni superiori, quali la memoria, il linguaggio, il ragionamento quanto la stessa pratica sensoriale<sup>4</sup>, facendo dell'uomo, oltretutto un animale sociale, un essere eminentemente normativo.

Alla natura normativa delle nostre forme di vita molteplici discipline rivolgono oggi giorno particolare attenzione a partire da differenti prospettive. Dalla linguistica alla giurisprudenza, dall'antropologia alla filosofia, dall'economia alle neuroscienze, il tema della normatività costituisce un nodo gordiano dell'epoca presente, verso cui tendono gli sforzi di comprensione delle scienze e della filosofia. Emblematico del crescente interesse per tale tema è il fiorire di numerose teorie della normatività, che si richiamano, proprio in campo filosofico, a pensatori del presente e del passato. Accanto a teorie normative che si appellano alla fenomenologia<sup>5</sup> o all'antropologia filosofica<sup>6</sup>, si stagliano, infatti, nell'attuale orizzonte filosofico, alcune proposte teoretiche che si rifanno, solo per fare qualche esempio, al pensiero di Aristotele<sup>7</sup>, Hegel e Kant. Per tale motivo ancora oggi riecheggia, in vari settori della filosofia, questa volta in riferimento alla sfera normativa, la *vexata questio*: Aristotele o Platone? Kant o Hegel?<sup>8</sup>

---

<sup>4</sup> Sulla natura normativa della pratica percettiva si veda, ad esempio, l'indagine di S.G. Crowell, *The Normative in Perception*, in R. Baiasu, G. Bird, A.W. Moore (eds), *Contemporary Kantian Metaphysics. New Essays on Space and Time*, Palgrave Macmillan, London 2012, pp. 81-106.

<sup>5</sup> Su questo punto si veda S.G. Crowell, *Normativity and Phenomenology in Husserl and Heidegger*, Cambridge University Press, Cambridge 2013; S. Loidolt, *Experience and Normativity: The Phenomenological Approach*, in A. Cimino and C. Leijenhorst (eds), *Phenomenology and Experience: New Perspectives*, Brill, Leiden/Boston 2019.

<sup>6</sup> Si confronti M. Schloßberger, *Phänomenologie der Normativität. Entwurf einer materialen Anthropologie im Anschluss an Max Scheler und Helmuth Plessner*, Schwabe Verlag, Basel/Berlin 2019.

<sup>7</sup> Un esempio è offerto da M. LeBar, *Aristotelian Constructivism*, in «Social Philosophy and Policy», 25, 2008, pp. 182-213.

<sup>8</sup> La domanda è posta in Ch. Krijnen (ed.), *Concepts of Normativity: Kant or Hegel?*, Brill, Leiden 2019.

A proposito di quest'ultima domanda, almeno nel campo dell'etica di stampo analitico, sembra avere una certa preminenza l'opera kantiana. Difatti, tale opera, seppur accusata dalle correnti avversarie di formalismo e individualismo, riveste, in questi ultimi anni, un posto di primissimo piano, facendo del filosofo di Königsberg un interlocutore privilegiato per tutti coloro che si accostano ai problemi dell'etica e della metaetica. Come si sa, a tale vera e propria Kant-Renaissance ha dato l'abbrivio il saggio di John Rawls *Kantian Constructivism in Moral Theory*<sup>9</sup>, con cui venne inaugurata la posizione teorica rubricata sotto il titolo, per l'appunto, di costruttivismo. Secondo tale posizione, «nella misura in cui ci sono verità normative, esse non sono determinate da fatti normativi che sono indipendenti da ciò intorno a cui gli agenti razionali concorderebbero in alcune specifiche condizioni di scelta»<sup>10</sup>.

A questa posizione, che si è declinata in numerose forme kantiane e non kantiane, ha aderito una delle più promettenti allieve di Rawls, Christine Korsgaard. A Korsgaard va il merito di aver avviato, in occasione delle Tanner Lectures on Human Values del 1992, un acceso dibattito su quali siano le fonti della normatività morale. La filosofa ha, infatti, posto al centro delle sue lezioni ciò che ha rubricato sotto il titolo di domanda normativa, la quale assume di volta in volta forme diverse: perché dovrei agire moralmente?<sup>11</sup> Che cosa *giustifica* le pretese che la moralità avanza nei nostri confronti?<sup>12</sup>

Intorno alla risposta offerta da Korsgaard a tali quesiti – risposta che si appella alla nozione kantiana di autonomia e al concetto marcatamente esistenzialistico di identità pratica – hanno dibattuto alcuni dei maggiori rappresentanti del variegato

---

<sup>9</sup> J. Rawls, *Kantian Constructivism in Moral Theory: The Dewey Lectures 1980*, in «Journal of Philosophy», 77, 1980, pp. 515–572.

<sup>10</sup> C. Bagnoli, *Constructivism in Metaethics*, in E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition).

<sup>11</sup> C. Korsgaard, *Le origini della normatività*, tr. it. ETS, Pisa 2014, p. 29.

<sup>12</sup> Cfr. *ibid.*, p. 30.

orizzonte filosofico attuale. Accanto alle celebri osservazioni di G.A. Cohen, R. Geuss, T. Nagel e B. Williams, che hanno costituito un commentario alle Tanner Lectures di Korsgaard, se ne sono affiancate, infatti, alcune più recenti di stampo fenomenologico ed ermeneutico, mediante cui si è cercato di comprendere l'origine della forza cogente delle ragioni morali e quindi della loro capacità di guidare le nostre azioni. Ma, oltre alle diverse risposte offerte dall'indagine filosofica ve ne sono altre che, seppur non direttamente rivolte al testo di Korsgaard, hanno oltrepassato i confini del dibattito squisitamente filosofico e investito campi d'indagine tra loro disparati. Alcune di esse hanno rivitalizzato l'ambizioso progetto nicciano di tracciare una genealogia della morale. Tale genealogia ha assunto ai giorni nostri, talvolta, un punto di vista filogenetico, che mette capo a una storia naturale della morale umana<sup>13</sup>, talvolta una prospettiva ontogenetica, sia di stampo sociologico<sup>14</sup> sia fondata sui risultati della psicologia infantile<sup>15</sup>. In tali casi, psicologi, sociologi e antropologi si sono impegnati a identificare le diverse fonti della normatività, mettendo in luce il ruolo svolto nel processo d'insorgenza della sfera normativa da fenomeni quali la cooperazione, l'educazione e l'aggressività<sup>16</sup>.

Ad arricchire l'orizzonte tracciato da tali indagini volte direttamente (o indirettamente) alla questione normativa si affianca ai giorni nostri il problema dello statuto morale degli agenti artificiali<sup>17</sup> e del valore morale delle loro azioni. In

---

<sup>13</sup> M. Tomasello, *A Natural History of Human Morality*, Harvard University Press, Cambridge/London 2016.

<sup>14</sup> E. Fittipaldi, *Everyday Legal Ontology*, LED, Milano 2012.

<sup>15</sup> M. Tomasello, *Becoming Human: A Theory of Ontogeny*, Harvard University Press, Cambridge 2019.

<sup>16</sup> R. Wrangham, *The Goodness Paradox: the Strange Relationship Between Virtue and Violence in Human Evolution*, Pantheon, New York 2019.

<sup>17</sup> Per una panoramica su questo dibattito si vedano J. Llorca Albareda, P. García, F. Lara, *The Moral Status of AI Entities*, in F. Lara, J. Deckers (eds), *Ethics of Artificial Intelligence*, The International Library of Ethics, Law and Technology, vol. 41, Springer, Cham 2023, pp.59-83; R. Redaelli, *Different Approaches to the Moral Status of AI: a Comparative Analysis of Paradigmatic*

modo particolare i sistemi dotati di intelligenza artificiale (IA) sollevano istanze etiche sempre più urgenti, per le quali è necessaria una riflessione che consideri lo statuto morale non solo del vivente (umano e non umano), ma anche del non vivente, e in questo caso dell'oggetto tecnologico con la sua (eventuale) capacità di veicolare istanze normative. Proprio a quest'ultima capacità è rivolto il presente scritto che mira a comprendere l'origine e la natura della forza prescrittiva veicolata dall'IA. A tal fine è necessario dapprima presentare la questione normativa posta da Korsgaard e offrire uno sguardo d'insieme sulle principali teorie relative alle fonti della normatività morale. Con tale panoramica desideriamo tracciare preventivamente le coordinate entro cui si muoverà la nostra indagine sulla capacità morale dell'IA. Entro tali coordinate intendiamo poi focalizzare l'attenzione sui pregi e i limiti della recente trattazione della normatività digitale proposta da E. Fournieret e B. Yvert, per infine apportare alle loro riflessioni un decisivo correttivo mediante l'introduzione della nozione di quasi-normatività. Allo scopo di introdurre e sviluppare tale nozione nell'ambito d'indagine volto alla sfera digitale ci avvarremo delle riflessioni postfenomenologiche di D. Ihde e P.P. Verbeek sulla natura dell'agente artificiale inteso come quasi-altro e sul ruolo di mediazione delle tecnologie. Grazie all'ausilio di tali riflessioni e all'introduzione della nozione di quasi-normatività riteniamo sia possibile gettare nuova luce sulla forza normativa dell'IA.

## ***2. La questione normativa: uno sguardo d'insieme sulle diverse teorie della moralità***

Il problema della normatività morale sollevato da Korsgaard riguarda la fonte dell'autorità dei nostri standard morali, della

loro capacità di guidare il nostro agire. Con tale questione ciò che vi è in gioco è, più precisamente, l'origine della forza che essi esercitano su di noi nell'intricata trama ordita dalla nostra vita quotidiana. Per Korsgaard, difatti, le ragioni e i concetti morali non hanno natura descrittiva, bensì prescrittiva, che spinge l'agente, posto di fronte alla loro autorità, a sollevare la questione inerente alla loro giustificazione. Per cui innanzi alle pressanti pretese che la moralità avanza nei nostri confronti sorge la seguente domanda: perché dovrei agire moralmente?<sup>18</sup>

Con tale questione si dischiude una costellazione di problemi che riguarda il diritto della morale di imporci leggi; un diritto che, occorre precisare fin da subito, ha ben poco a che fare con la questione se e perché ci convenga essere morali<sup>19</sup>. Difatti, l'autorità della morale di cui Korsgaard va alla ricerca è incondizionata, ossia indipendente dal fatto che le sue ingiunzioni siano o meno corroborate dai nostri interessi e dalle nostre inclinazioni. Così posta, la questione normativa da un lato prende le distanze da qualsivoglia forma di utilitarismo, mentre dall'altro lato fa propria quella ricerca kantiana dell'incondizionato, con cui la filosofa intende porre fine allo scetticismo morale<sup>20</sup>, che reitera all'infinito la domanda «perché devo farlo?». All'individuazione di tale incondizionato, Korsgaard perviene, come vedremo, non mediante l'appello a qualche entità intrinsecamente normativa, bensì attraverso il ricorso a quello che lei definisce nei termini di approvazione riflessiva. La domanda normativa fa, dunque, questione dell'autorità della morale, evitando di chiamare in causa tanto una qualche forma di realismo sostanziale<sup>21</sup> quanto una qualsiasi versione

---

<sup>18</sup> C. Korsgaard, op. cit., p. 29.

<sup>19</sup> *Ibid.*

<sup>20</sup> *Ibid.*, p. 58.

<sup>21</sup> A differenza del realismo sostanziale, per cui vi sono valori là fuori che possono essere intuiti, Korsgaard definisce la propria posizione filosofica nei termini di *procedural realism*: «la forma di realismo che sto adottando qui è il realismo procedurale piuttosto che il realismo sostanziale: i valori vengono



dell'utilitarismo: essa, come vedremo ora, trova risposta in un kantismo, per così dire, esistenziale<sup>22</sup>.

Proprio per la peculiare piega impressa alla questione normativa e per la forza teoretica del suo costruttivismo, la proposta filosofica di Korsgaard ha avuto e ha tuttora una forte eco entro il frastagliato orizzonte filosofico contemporaneo, in seno al quale è possibile ravvisare tre differenti approcci mediante cui si è inteso far fronte a tale questione. Di questi approcci W.H. Smith ha offerto una preziosa panoramica, che riprendiamo qui, seppur schematicamente, al fine di definire preventivamente le coordinate entro cui si muoverà la nostra indagine sulla capacità morale dell'IA.

Il primo approccio, che è quello impiegato per l'appunto da Korsgaard, è definito dalla stessa autrice alla prima persona. In tale approccio, viene assunta una prospettiva internalista - caratterizzata grammaticalmente dall'impiego del pronome io. A partire da tale prospettiva si deve chiarire perché una certa ragione sia vincolante alla prima persona, vale a dire dalla prospettiva soggettiva dell'agente. Al fine di offrire una risposta adeguata a tale questione, Korsgaard si appella a una caratteristica propria dell'umano, vale a dire all'autoriflessione<sup>23</sup>. A differenza degli animali non-umani, che agiscono in conformità ai loro istinti senza che tali istinti siano oggetto di attenzione e deliberazione, l'uomo non può accontentarsi di tali incentivi, bensì ha bisogno di ragioni per

---

costruiti secondo una procedura, la procedura che consiste nel produrre leggi per noi stessi» (*ibid*, p. 154).

<sup>22</sup> Rispondendo alle osservazioni di Nagel, Korsgaard confessa ne *Le origini della normatività* che Nagel definisce il «[mio] appello all'identità 'piuttosto esistenzialistico' (a mio avviso correttamente) e, quindi, anche non kantiano (in modo a mio avviso non corretto)» (*ibid*, p. 309).

<sup>23</sup> La nozione di autoriflessione chiamata in causa da Korsgaard è stata oggetto di pressanti critiche, tra le quali spiccano quelli di Mark Okrent, riprese in parte da S. Crowell. Secondo Okrent la nozione di autoriflessione così come presentata da Korsgaard dà luogo a tre differenti interpretazioni, lasciando tale nozione per lo più indeterminata (M. Okrent, *Heidegger and Korsgaard on Human Reflection*, in «Philosophical Topics», 27, 2, pp. 47-76; C. Crowell, *Normativity and Phenomenology in Husserl and Heidegger*, cit., pp. 246 ss.).



credere e per agire<sup>24</sup>. Tali ragioni, per Korsgaard, non sono altro che un tipo di successo riflessivo (*reflective success*)<sup>25</sup>, vale dire le nostre percezioni e i nostri desideri offrono ragioni qualora superino il vaglio di un attento esame riflessivo<sup>26</sup>: è, difatti, «nel mondo interno creato dalla coscienza di sé che sorge una ragione»<sup>27</sup>. La riflessione sembra così affrancare il soggetto umano dall'autorità degli istinti che domina il regno animale e aprire, allo stesso tempo, uno spazio normativo in cui egli si lega ai propri standard morali. In questo senso, Korsgaard chiama in causa nella sua argomentazione la nozione kantiana di autonomia, pur riconoscendo nell'autonomia solo una condizione formale per l'agire morale. All'autonomia, difatti, l'autrice affianca ciò che definisce nei termini di identità pratiche intese come «immagini di sé fondate nelle forme di vita concrete»<sup>28</sup>. Tali identità pratiche (di padre, madre, marito, manager ecc.) forniscono «constitutive standards» rispetto a cui possiamo fallire o avere successo, o meglio standard alla luce dei quali le nostre ragioni e obbligazioni sorgono. Chiaramente questa prospettiva alla prima persona, che unisce autonomia e identità pratica, presenta il limite, sottolineato da diversi critici (tra cui ricordiamo S.G. Crowell e lo stesso Smith), di non rendere ragione della forza normativa dell'altro: l'altro, il tu è

---

<sup>24</sup> Sulla differenza tra uomo e animale si veda C. Korsgaard, *Self-Constitution*, laddove l'autrice distingue due tipi di autonomia, quella animale «governated by the principles of your own causality» e quella umana caratterizzata dalla scelta dei principi che definiscono il nostro volere (C. Korsgaard, *Self-Constitution. Agency, Identity, and Integrity*, Oxford University press, Oxford 2009, p. 108).

<sup>25</sup> A partire da tale prospettiva, osserva Korsgaard, «lo scetticismo verso il buono e il moralmente giusto non è scetticismo verso l'esistenza di entità intrinsecamente normative. È la posizione per la quale i problemi che la riflessione ci pone sono irrisolvibili e le domande che pone non hanno risposta» (C. Korsgaard, op. cit., p. 131).

<sup>26</sup> Più precisamente Korsgaard afferma che «ogni impulso, così come si presenta alla volontà, deve superare un tipo di test della normatività prima che lo si possa adottare come ragione per agire» (*ibid.*, p. 128).

<sup>27</sup> C. Korsgaard, *Self-Constitution. Agency, Identity, and Integrity*, cit., p. 116.

<sup>28</sup> A. Staiti, *Praktische Identität aus phänomenologischer Sicht: Korsgaard und Husserl*, in «Phänomenologische Forschungen», 1, 2015, p. 172.

ridotto, come ben osservato da Smith, a «una fonte di ragioni proto-morali - analoghe agli incentivi o alle inclinazioni corporee nella deliberazione pratica»<sup>29</sup>.

Differentemente dall'impostazione di Korsgaard, la prospettiva alla seconda persona, di cui anche Smith si fa portavoce, è contraddistinta dal punto di vista del tu, laddove preminente non è più il rivolgersi dell'io a sé stesso, bensì l'incontro con l'altro. In questa prospettiva, la forza cogente dell'ingiunzione morale non sorge dall'autoriflessione, bensì originariamente dall'istanza, dalla domanda che l'altro rivolge all'io. Si tratta qui, in termini levinasiani, di un incontro faccia a faccia con l'altro e di una conseguente fondazione interpersonale della morale, che ha assunto e assume tutt'oggi forme tra loro disparate, talvolta radicalmente differenti l'una dall'altra, come quelle proposte da S. Darwall<sup>30</sup> e dallo stesso Smith. Riferendoci in modo particolare a Smith, si può osservare che egli, seguendo tale linea di pensiero, tenta di elaborare una teoria che raccolga in modo armonico un momento normativo alla prima persona e una prospettiva alla seconda persona. In tal modo Smith intende offrire un fondamento ontologico alla riflessione levinasiana del vis-à-vis attraverso il ricorso alle nozioni heideggeriane di *Mitsein* e risolutezza. In virtù di tale commistione, all'altro è assegnato un ruolo centrale nell'etica<sup>31</sup>, senza per tale motivo venir meno un momento alla prima persona.

La prospettiva alla terza persona è, invece, quella incarnata dalla figura dello spettatore disinteressato, che osserva per così dire dall'esterno, *from nowhere*, ciò che avviene nel mondo. Questa prospettiva è assunta, in modo paradigmatico, nel campo delle scienze, laddove si mira a una visione oggettiva del mondo.

---

<sup>29</sup> W.H. Smith, *The Phenomenology of Moral Normativity*, Routledge, London/New York 2012, p. 37.

<sup>30</sup> S. Darwall, *The Second-Person Standpoint: Morality, Respect, and Accountability*, Harvard University Press, Cambridge 2006.

<sup>31</sup> Sulla centralità dell'altro in ambito etico e non solo si veda C. Di Martino, *Figure della relazione. Saggi su Ricoeur, Patočka e Derrida*, Edizioni di pagina, Bari 2018.

Nell'ambito dell'etica, il massimo rappresentante di tale approccio è, com'è noto, Thomas Nagel, la cui riflessione mette capo a una forma di realismo, secondo cui vi sono *agent-neutral reasons*: dall'oggettività di tali ragioni deriva la loro forza normativa. In questa e altre forme di realismo, la domanda normativa sembra, tuttavia, perdere di interesse poiché la risposta a tale quesito, come osserva correttamente Wedgwood, «corrisponde a una verità o a un fatto normativo appropriato»<sup>32</sup>. Una volta tracciata a grandi linee la cornice entro cui intendiamo collocare le nostre riflessioni sull'IA è opportuno richiamare le recenti riflessioni di Fournernet e Yvert<sup>33</sup> che hanno il merito di lumeggiare la questione normativa in relazione ai sistemi intelligenti. Tali riflessioni costituiranno il punto di partenza per la nostra riflessione che mira ad arricchire e integrare la posizione dei due autori. A tale scopo faremo ricorso dapprima ad alcune istanze presenti nel dibattito sulle fonti della normatività morale qui richiamato per accenni, per poi introdurre, a partire da una prospettiva postfenomenologica, la nozione di quasi-normatività.

### **3. Forme di normatività digitale**

La presenza pervasiva di sistemi intelligenti nella nostra vita sta dando luogo a quelle che sono state definite da Fournernet e Yvert forme di normatività digitale. Con tale espressione gli autori intendono «la capacità degli algoritmi di definire standard che gli esseri umani incorporano come ciò che dovrebbe essere ritenuto normale nella loro vita e che guidano le loro azioni»<sup>34</sup>. Questi standard originati dall'IA presentano quindi un potere normativo, vale a dire la capacità di guidare, di indirizzare le

---

<sup>32</sup> R. Wedgwood, *The Nature of Normativity*, Oxford University Press, Oxford 2007, p. 7.

<sup>33</sup> E. Fournernet, B. Yvert, *Digital Normativity: A Challenge for Human Subjectivation*, in «Frontiers in Artificial Intelligence», 3 (<https://doi.org/10.3389/frai.2020.00027>)

<sup>34</sup> *Ibid.*, p. 1.

nostre azioni e credenze. Essi non sono una mera descrizione di ciò che già facciamo e crediamo, bensì mirano a farci agire secondo determinate direzioni. Gli algoritmi sono, infatti, capaci di restituire i *trends* che riconoscono nei dati e creare così «una visione normalizzata» dei problemi che si trovano ad affrontare, mettendo capo a una prima forma di normatività digitale. Nella terminologia del dibattito appena ripercorso, potremmo dire che tali algoritmi ci offrono ragioni per agire e decidere in determinate situazioni<sup>35</sup>.

Per Fourneret e Yvert, alla prima forma di normatività digitale se ne lega strettamente una seconda, che coinvolge in modo particolare gli algoritmi predittivi. Questi algoritmi suggeriscono determinati comportamenti sulla base dei nostri comportamenti precedenti e su quelli degli altri utenti, senza alcuna considerazione delle ragioni dei nostri comportamenti. In tal modo l'individuo è «oggettivato (normalizzato) dall'algoritmo»<sup>36</sup>. In linea con la prima forma di normatività digitale, che crea una visione normalizzata dei problemi, la seconda forma conduce a una visione normalizzata di noi stessi. In questo senso, l'algoritmo riduce l'individuo e il suo comportamento ai suoi dati (e a quelli degli altri individui), senza alcuna considerazione delle motivazioni dei suoi comportamenti e della portata etica delle sue azioni. Tale seconda forma di normatività digitale ha un carattere ricorsivo e dinamico che, a nostro avviso, è presente anche nella prima, per cui «le raccomandazioni, avanzate dagli algoritmi, basate sulle azioni umane precedenti influenzano a loro volta le azioni successive»<sup>37</sup>.

---

<sup>35</sup> Occorre qui sottolineare che l'azione degli artefatti tecnologici non è passibile di *reasons explanation* e per questo, come osserva Johnson, essi non sono agenti morali (D.G. Johnson, *Computer systems: Moral entities but not moral agents*, in «Ethics and Information Technology», 8, 2006, pp. 195-204). Tuttavia, tale limite non impedisce a tali agenti artificiali di offrire all'uomo ragioni per agire e per credere.

<sup>36</sup> E. Fourneret, B. Yvert, *Digital Normativity: A Challenge for Human Subjectivation*, cit., p. 2.

<sup>37</sup> *Ibid.*

Una terza forma di normatività digitale deriva dall'efficienza (*efficiency*) predittiva degli algoritmi. Tale capacità supera, ormai in molti campi, quella umana, imponendo norme a cui l'uomo si adegua, data per l'appunto l'efficienza dell'IA. Questa subordinazione all'efficienza degli algoritmi avviene anche qualora non sia chiaro il procedimento attraverso cui il sistema sia pervenuto al risultato. Con ciò emerge il problema della responsabilità morale e legale per gli effetti prodotti dall'IA. Proprio per far fronte al cosiddetto problema del *black box* si tende oggi a indicare come valore indispensabile per lo sviluppo di una *trustworthy* IA la trasparenza. In questa direzione si è proposto di introdurre *checkpoint* e/o la ripetizione di *stress-test* nei sistemi intelligenti<sup>38</sup>.

Ora, ciò che è importante sottolineare delle riflessioni svolte da Fournernet e Yvert è il ruolo che, secondo gli autori, possono svolgere questi tipi di normatività digitale entro il complesso processo di soggettivazione umana, laddove con tale termine si intende il «processo di costruzione che porta qualcuno a diventare e ad essere consapevole di essere un soggetto, ossia di essere libero e responsabile delle proprie azioni, delle proprie rappresentazioni e dei propri giudizi»<sup>39</sup>.

Per gli autori, infatti, i sistemi intelligenti sono in grado d'esercitare una decisiva influenza su tale processo secondo una peculiare dinamica: da un lato tali sistemi possono condurre l'uomo a non svolgere più alcuni compiti gravosi, ma dall'altro lato possono anche indurci a non prendere più decisioni in alcuni campi particolarmente significativi della nostra vita data l'efficienza previsionale degli algoritmi. Un esempio ormai divenuto celebre è quello del giudice che in alcuni sistemi legali si avvale dell'ausilio di algoritmi per stabilire il rischio di

---

<sup>38</sup> Su questo punto si veda N. Cristianini, *The Shortcut: Why Intelligent Machines Do Not Think Like Us*, CRC Press, Boca Raton/London/New York 2023, pp. 155-159.

<sup>39</sup> E. Fournernet, B. Yvert, *Digital Normativity: A Challenge for Human Subjectivation*, cit., p. 2.

recidiva. In questo caso, si potrebbe creare una conflittualità tra la decisione del giudice e quella dell'algoritmo, oppure, come nel caso Compas, vi potrebbero essere *biases* che influenzano negativamente la decisione del giudice supportata dai sistemi intelligenti.

Al di là di tali limiti, le straordinarie possibilità offerte dagli algoritmi predittivi, secondo la diagnosi degli autori, potrebbero attivare nelle future generazioni un'«AI governmentality»<sup>40</sup> con un conseguente processo di deresponsabilizzazione, per cui l'uomo si affiderebbe ciecamente ai risultati ottenuti dai sistemi intelligenti. Infatti, benché rimanga sempre aperta la possibilità di resistere alla forza normativa che sorge dall'IA, l'efficacia di tale tecnologia potrebbe comportare, in ultima istanza, un'imposizione, seppur non violenta, delle scelte dei sistemi intelligenti. Tale imposizione avverrebbe, e avviene già adesso, attraverso un radicale cambiamento delle nostre pratiche di vita e di sapere. Di fronte a tale rischio, gli autori, in conclusione al loro studio, suggeriscono di accompagnare lo sviluppo dell'IA con la riflessione etica promossa da Verbeek. Tale etica del design ha, infatti, il merito di identificare il sistema di valori di cui la tecnologia è portatrice e di riflettere sui principi da seguire in fase di progettazione per proteggere il processo di soggettivazione dell'umano.

Al fine di mettere in luce pregi e difetti delle riflessioni di Fournieret e Yvert, è ora necessario leggere la proposta avanzata dagli autori alla luce di alcuni elementi delle teorie alla prima e alla seconda persona a cui ci siamo riferiti nel paragrafo precedente. Innanzitutto occorre osservare che l'intelligenza artificiale, in virtù della sua efficienza, presenta istanze dotate di un carattere ingiuntivo a cui l'uomo è chiamato a rispondere. In questo senso, l'IA è riconosciuta come l'altro che

---

<sup>40</sup> *Ibid.*

si presenta all'io, per cui nell'incontro con il volto dell'altro, come sostiene Lévinas, vi è un'interdizione etica o la messa in discussione della mia libertà. A tale messa in questione della libertà umana, gli autori rispondono appellandosi a un'etica del design che svolge chiaramente la funzione di autoriflessione proposta dalla teoria di Korsgaard. Secondo tale teoria, le istanze dell'altro necessitano di superare il vaglio riflessivo al fine di divenire per l'io ragioni per agire e credere. Allo stesso modo, mediante un'etica del design, identifichiamo quali valori siano per noi vincolanti e decidiamo quali valori tali tecnologie debbano promuovere nelle nostre società.

Seguendo questa direzione d'indagine, si può dunque riconoscere nelle riflessioni di Fournieret e Yvert un momento alla seconda persona, per cui la normatività morale sorge nel faccia a faccia con l'altro, e un momento alla prima persona, per cui l'ingiunzione dell'altro non ha forza normativa finché non ottiene l'approvazione dell'io. Dall'intreccio tra questi due momenti si stagliano chiaramente due questioni capitali, la cui risoluzione può contribuire a chiarire lo statuto della normatività digitale. La prima questione riguarda la riconduzione dell'IA all'altro: si può equiparare l'IA all'altro (vivente umano e non umano)? Le ingiunzioni ad agire dell'IA hanno la stessa forza normativa di quelle umane? Questa ultima domanda si lega a una seconda questione relativa alla natura della normatività digitale: si può basare la forza normativa delle istanze sollevate dall'IA sulla mera efficienza, quindi secondo una prospettiva a suo modo utilitaristica, che la posizione kantiana di Korsgaard sembra estromettere dal discorso morale?

Al fine di rispondere alla prima questione, vale a dire al problema dell'identificazione dell'IA con l'altro, ci avvarremo delle analisi postfenomenologiche di Ihde relative al ruolo di mediazione svolto dalle tecnologie nella nostra vita, mentre per la risoluzione della seconda questione impiegheremo la nozione di



protonormatività o quasi-normatività, che abbiamo già elaborato altrove<sup>41</sup> e che ci sembra utile richiamare qui allo scopo di chiarire ciò che lega (legittimamente o illegittimamente) la forza normativa dell'IA all'efficienza.

#### **4. L'intelligenza artificiale da una prospettiva postfenomenologica e La nozione di quasi-normatività**

L'approccio postfenomenologico inaugurato da Ihde ha l'indiscusso merito di unire analisi filosofica e investigazione empirica nell'indagine sulle relazioni sussistenti tra esseri umani e artefatti tecnologici. In tale tipo d'indagine Ihde offre un decisivo contributo al campo dei *Science and Technology Studies* (STS) mettendo in luce diverse forme di mediazione tecnologica. Per l'autore, difatti, le tecnologie mediano, in differenti modi, la nostra relazione al mondo e indirizzano il nostro agire. In virtù di tale ruolo, la relazione uomo-mondo assume la forma uomo-tecnologia-mondo.

La riflessione di Ihde è ormai ampiamente conosciuta, per cui un'esposizione sinottica delle relazioni uomo-tecnologia riconosciute dal filosofo è sufficiente per i nostri fini. Come si sa, Ihde individua quattro forme di mediazione tecnologica: *embodiment relations*, *hermeneutic relations*, *background relations*, e *alterity relations*. Tale elenco non intende essere esaustivo e una forma di mediazione non esclude l'altra, perciò una stessa tecnologia può realizzare diverse forme di mediazione.

Il primo tipo di relazione uomo-tecnologia è caratterizzato dal fatto che, come afferma il filosofo, «io accolgo le tecnologie *nella* mia esperienza in un modo particolare, vale a dire percependo *attraverso* tali tecnologie» e trasformando in tal modo

---

<sup>41</sup> Sulla nozione di quasi-normatività mi permetto di rimandare a R. Redaelli, *A Relational Account of Moral Normativity: The Neo-Kantian Notion of We-Subject*, in «Journal of Transcendental Philosophy», 2-3, 2021, pp. 303-320.

il mio senso percettivo corporeo<sup>42</sup>. In questa relazione, dunque, la tecnologia diventa un'estensione del corpo umano, diventa un quasi-me. Esempi di tale tipo di relazione sono offerti da tecnologie ottiche come il telescopio o gli occhiali. Entrambe svolgono un ruolo di mediazione tra colui che vede e ciò che è visto, per cui, in entrambi i casi, vediamo attraverso lo strumento ottico<sup>43</sup>. Lo stesso si può affermare per gli apparecchi acustici che diventano un'estensione del proprio corpo. Tuttavia, al di là del diverso tipo di tecnologia a cui ci riferiamo, ciò che caratterizza, secondo Ihde, la relazione di incorporazione è il fatto che la tecnologia coinvolta in tale relazione deve in qualche modo retrocedere dalla nostra attenzione al fine di essere incorporata. Perciò più la tecnologia è vicina alla trasparenza e più essa consente l'estensione del nostro senso corporeo<sup>44</sup>.

Il secondo tipo di relazione uomo-tecnologia è definito da Ihde nei termini di relazione ermeneutica. In questo caso, le tecnologie offrono una rappresentazione del mondo che richiede un'interpretazione. Per tale ragione, Ihde parla per l'appunto di relazione *ermeneutica* riferendosi esplicitamente allo studio dell'interpretazione. Un esempio è fornito dal termometro che mediante dei valori ci permette di "leggere" la temperatura<sup>45</sup>.

Un altro tipo di relazione è individuata dal filosofo nella *background relation*. In questo tipo di relazione la tecnologia, benché sia per così dire non-tematizzata, forma il nostro ambiente. Il termostato con cui regoliamo la temperatura nella stanza e la luce accesa che ci consente di vedere gli oggetti intorno a noi sono entrambi parte del nostro ambiente, benché non intratteniamo con tali tecnologie una relazione diretta. Queste tecnologie rimangono difatti sullo sfondo.

---

<sup>42</sup> D. Ihde, *Technology and the Lifeworld: From Garden to Earth*, Indiana University Press, Bloomington/Indianapolis 1990, p. 72.

<sup>43</sup> *Ibid.*, p. 73.

<sup>44</sup> *Ibid.*, p. 74.

<sup>45</sup> *Ibid.*, p. 85.

L'ultimo tipo di relazione è quello su cui desideriamo focalizzare la nostra attenzione perché chiama in causa esplicitamente l'intelligenza artificiale: si tratta della *relazione di alterità* (*alterity relation*). In questo tipo di relazione, la tecnologia non media il nostro rapporto al mondo, bensì noi instauriamo una relazione con la tecnologia<sup>46</sup>. Perciò, in questo caso, la tecnologia è il *terminus* della relazione; gli artefatti tecnici si stagliano dallo sfondo mondano e sono in primo piano.

Riprendendo da Lévinas il termine alterità per indicare la differenza radicale che separa ogni uomo dall'altro uomo, Ihde definisce, entro l'*alterity relation*, la tecnologia come quasi-altro, riferendosi alla stessa IA in tali termini. Pur mettendo in luce i rischi dell'antropomorfizzazione della tecnologia nel riconoscere in essa un quasi-altro, Ihde sottolinea come l'*automaton* esemplificato dal robot sia un quasi-altro a cui possiamo prestare la nostra attenzione<sup>47</sup>, perché, benché il robot abbia una diversa "esperienza" del mondo rispetto a noi<sup>48</sup>, esso si presenta come polo di una possibile relazione sociale.

Al fine di sostenere la propria argomentazione, Ihde cita l'esempio del videogioco che coinvolge non solo una relazione di *embodiment* (il joystick come estensione corporea) ed ermeneutica (il contesto del gioco, ad es. la simulazione di uno sport), ma anche il senso di «*interagire con qualcosa di diverso da me*», vale a dire con ciò che Ihde definisce «il *concorrente tecnologico*»<sup>49</sup>. Nella competizione che prende forma nel videogioco vi è, infatti, uno scambio o un dialogo con il quasi-altro.

A proposito di dialogo, al fine di comprendere meglio la relazione di alterità possiamo pensare, oltre ai robot, anche ai chatbot con cui ormai la maggior parte di noi interagisce ogni giorno. Tali

---

<sup>46</sup> *Ibid.*, p. 97.

<sup>47</sup> *Ibid.*, p. 104.

<sup>48</sup> Ihde sottolinea la differenza tra uomini e robot, mettendo in luce come l'«esperienza» che il robot ha del mondo sia significativamente diversa da quella umana già a livello sensoriale.

<sup>49</sup> *Ibid.*, p. 100.

sistemi, esemplificati dagli assistenti digitali, simulano diversi tipi di conversazione, rispondendo alle nostre richieste e personalizzando sempre più la loro attività in base alle nostre preferenze. Essi impiegano l'intelligenza predittiva e realizzano un'analisi di enormi quantità di dati mediante cui sono in grado di offrirci suggerimenti sempre più accurati, anticipando e indirizzando le nostre esigenze. Proprio per tali caratteristiche potremmo dire che noi avvertiamo tali chatbot come quasi-altri allo stesso modo dei robot chiamati in causa da Ihde. Difatti, i ro(bot)<sup>50</sup>, come appena accennato, benché non siano umani, non ci appaiono come meri strumenti: essi sono qualcosa di diverso da noi, un altro da noi con cui possiamo intrattenere una relazione sociale. I robot e chatbot co-formano le nostre pratiche di vita e di sapere attraverso un certo grado di autonomia e interazione che altri strumenti non hanno. Basti pensare ai casi estremi offerti dai care robot o sex robot con cui gli individui intrattengono relazioni che coinvolgono la sfera affettiva o alcuni tipi di chatbot così evoluti per cui è difficile comprendere se stiamo interagendo con un agente umano o artificiale.

A partire da tale caratterizzazione dei robot e chatbot, possiamo ora approfondire un punto cruciale sviluppato da Fournernet e Yverte che si lega alla questione della normatività digitale. Sintetizzando la riflessione degli autori, potremmo dire che più elevata è l'efficienza dei sistemi intelligenti nel relazionarsi a noi e più le risposte e le istanze da essi sollevate hanno forza normativa, vale a dire la forza di indirizzare le nostre azioni. E, allo stesso modo, più performante è la capacità predittiva dell'IA più noi ci fidiamo dei risultati da essa offerti fino a giungere ad accettare ciecamente i risultati ottenuti dall'IA. In questo senso, di certo Fournernet e Yverte hanno ragione nel legare la capacità normativa dell'IA alla sua efficienza con un

---

<sup>50</sup> Ci avvaliamo qui della grafia usata da Wallach e Allen per indicare sia i robot che i chatbot (W. Wallach, C. Allen, *Moral Machines: Teaching Robots Right from Wrong*, OUP, Oxford 2009).

conseguente rischio per il processo di soggettivazione. Eppure se da un lato, secondo gli autori, l'efficienza dei sistemi intelligenti conferisce una forza normativa ai risultati da loro ottenuti, dall'altro lato rimane pur sempre aperta la possibilità di prendere posizione di fronte a tali risultati. Tale possibilità, richiamata dagli autori solo di passaggio e implicita nell'appello all'etica, mostra in realtà come il legame efficienza-normatività digitale non risolva del tutto la questione normativa nei sistemi intelligenti. Difatti, se si intende con normatività la forza cogente, l'autorità che esercitano su di noi gli standard (che possono essere morali o no) creati dai sistemi predittivi, offrendoci (o meno) ragioni per agire in determinate situazioni, si può osservare che le istanze promosse dall'IA non sollevano una vera propria forma di normatività, perché essa si ottiene solo superando il test della riflessione. Tali sistemi danno luogo invece a quelle che proponiamo di chiamare istanze protonormative o quasi-normative, laddove il "quasi" sta qui a indicare che tali istanze non sono normative in senso pieno, ma neppure sono moralmente neutrali. Tali istanze sono, infatti, caratterizzate dal fatto che, benché richiedano una nostra approvazione al fine di essere normative in senso forte, possiedono già una forza ingiuntiva, legata all'efficienza predittiva degli algoritmi, che ridefinisce in un certo grado il nostro spazio di libertà e indirizza la nostra azione.

Un esempio a riprova della bontà di tale tesi, per cui l'IA e, più in generale, le tecnologie avanzate sono portatrici di un'istanza quasi-normativa e non normativa *in senso forte*, ci è offerto in campo postfenomenologico proprio da Verbeek che, in apertura al suo *Moralizing Technology*, introduce la sua proposta - le tecnologie hanno un significato morale in quanto mediatrici della nostra esperienza del mondo - a partire dagli effetti dell'utilizzo di alcuni dispositivi medici. Un buon esempio è

offerto dall'ecografia ostetrica<sup>51</sup>. Per il filosofo, non si può attribuire a questa tecnologia un ruolo meramente funzionale (quello di rendere visibile un bambino non ancora nato nel grembo materno), piuttosto essa ridefinisce sia lo statuto ontologico del nascituro sia la nostra esperienza del feto. Infatti, tale tecnologia, capace di prevedere se il bambino sarà affetto o meno da determinate malattie, trasforma il nascituro in un possibile paziente ed i suoi genitori in soggetti che decidono della sua vita. In questo modo, i dispositivi medici ridefiniscono il rapporto tra figlio e genitori in termini finora inaspettati, per cui la gravidanza diventa un processo di scelta<sup>52</sup>. Infatti, la possibilità di «fare ecografie, e quindi di individuare difetti congeniti prima della nascita, cambia irreversibilmente il carattere di ciò che un tempo si chiamava ‘aspettare un bambino’»<sup>53</sup>. Nei nostri termini potremmo dire che tale tecnologia è portatrice di istanze quasi-normative, vale a dire dotate di una certa capacità di indirizzare l'azione; istanze che, come già detto, possono diventare normative, cioè cogenti, solo con l'approvazione umana. Difatti, se da un lato è necessario considerare che dall'introduzione di queste tecnologie anche la stessa “scelta di *non* sottoporsi a un'ecografia è una scelta”, dall'altro lato anche in presenza di gravi patologie diagnosticate al nascituro, questi sistemi predittivi non esercitano una forza coercitiva sui genitori tale da indurli ad abortire. In questo

---

<sup>51</sup> Si veda P.P. Verbeek, *Obstetric ultrasound and the technological mediation of morality: a postphenomenological analysis*, in «Human Studies», 31, 2008, pp. 11-26. È importante osservare che il caso dell'ecografia ostetrica richiamato da Verbeek si riferisce ad una tecnologia non dotata di IA eppure provvista di un chiaro profilo normativo. Questo caso mostra come le tecnologie *in diversa misura* ridefiniscano il piano normativo, e l'intelligenza artificiale, con i suoi caratteri di interattività, autonomia e adattabilità, aumenti la forza normativa di tali tecnologie. A questo proposito si osservi come, proprio nel campo della diagnostica per immagini, l'ecografia assistita dall'IA aumenti l'efficienza predittiva dell'ecografia e così la forza normativa delle sue previsioni (su tale tecnologia si confronti R. Horgan, L. Nehme, A. Abuhamad, *Artificial Intelligence in Obstetric Ultrasound: A Scoping Review*, in «Prenatal Diagnosis», 43, 2023, pp. 1176-1219).

<sup>52</sup> P.P. Verbeek, *Moralizing Technology: Understanding and Designing the Morality of Things*, University of Chicago Press, Chicago/London 2011, p. 25.

<sup>53</sup> *Ibid.*

senso, se è chiaro come la ridefinizione della realtà da parte delle tecnologie comporti l'emergere di una quasi-normatività (in questo caso, un dovere scegliere per motivi medici), dovrebbe essere ancora più chiaro come qualsiasi previsione che emerga dall'uso di tali tecnologie non abbia l'autorità propria delle ragioni normative che hanno superato il vaglio riflessivo.

### ***5. La normatività morale tra uomo e intelligenza artificiale***

A partire da tali considerazioni si può osservare che l'emergere della normatività, come forza vincolante che prescrive al soggetto un certo comportamento, non è attribuibile alla sola tecnologia, bensì è l'effetto dell'interazione tra umani e non umani. In questo senso, la normatività è indissolubilmente legata alla sfera protonormativa dell'apparato tecnologico, senza tuttavia imputare a quest'ultimo un ruolo diverso da quello di mediazione assegnatogli da Verbeek e prima ancora da Ihde. In altri termini ancora, possiamo dire che la forza normativa delle ragioni ha un carattere relazionale: essa sorge nella relazione tra l'io e l'altro o, nel caso dell'IA, il quasi-altro. Ed è proprio questa relazione - possiamo aggiungere come postilla al nostro discorso - che dà luogo a quel noi-società, di cui fanno parte oggi giorno non solo i viventi umani e non-umani, ma anche i sistemi dotati di intelligenza artificiale. Questi ultimi, infatti, fanno parte della nostra società, sono quasi-altri, perché le relazioni che intratteniamo con loro condividono con gli altri tipi di relazione sociale il fatto che esse, usando le parole di Margalit e Raz, «dipendono per la loro esistenza dalla condivisione di modelli di aspettative, da tradizioni che conservano conoscenze implicite su come fare cosa, da convenzioni tacite su ciò che fa parte di questa o quella impresa e ciò che non ne fa parte, su ciò che è appropriato e ciò che non lo è, su ciò che ha valore e ciò che non



lo ha»<sup>54</sup>. In questo preciso senso, i sistemi intelligenti sono parte della nostra società di cui tra l'altro si fanno portavoce e da cui deriva la necessità di quel richiamo all'etica del design che ci permette di decidere di quali istanze tali robot *debbano* essere portatori; una decisione che affranca, seppur in parte, il nostro processo di soggettivazione dai rischi di degenerazione messi ben in luce da Fournieret e Yvert. Difatti, finché la forza dirompente dell'IA può essere riconosciuta come una forma di quasi-normatività, l'uomo rimane responsabile degli standard che intende adottare e la stessa efficienza dell'IA non può sostituire l'intelligenza umana, semmai può potenziarla, ampliando il campo umano della libertà e della responsabilità. In questo senso, parafrasando le celebri parole di G. Anders, le possibilità dischiuse dalla tecnologia, e nel nostro caso dischiuse dall'IA, non *devono* (sempre e comunque) essere realizzate, bensì *sono possibilità aperte di fronte a cui occorre prendere posizione* nel continuo processo di autoformazione dell'umano.

ROBERTO REDAELLI è Ricercatore di Tipo B in Filosofia morale presso l'Università degli studi di Milano

[roberto.redaelli@unimi.it](mailto:roberto.redaelli@unimi.it)

---

<sup>54</sup> A. Margalit, J. Raz, *National Self-Determination*, in «Journal of Philosophy», 87, 1990, pp. 439-461.