

*vai alla
scheda
del libro*

Daniele Verdesca

l'autore

Etica delle intelligenze artificiali per leader aziendali e startupper

Artificial Intelligence (AI) Ethics Model Canvas

 **EPC**
EDITORE

Daniele Verdesca

Etica delle intelligenze artificiali per leader aziendali e startupper

Artificial Intelligence (AI) Ethics Model Canvas

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER
ISBN: 978-88-9288-001-6

Copyright © 2020 EPC S.r.l. Socio Unico

EPC S.r.l. Socio Unico – Via Clauzetto, 12 – 00188 Roma

www.epc.it

Servizio clienti: 06 33245277 - Fax 06 33245248

Redazione: Tel. 06 33245264/205

Proprietà letteraria e tutti i diritti riservati alla EPC S.r.l. Socio Unico. La struttura e il contenuto del presente volume non possono essere riprodotti, neppure parzialmente, salvo espressa autorizzazione della Casa Editrice. Non ne è altresì consentita la memorizzazione su qualsiasi supporto (magnetico, magneto-ottico, ottico, fotocopie ecc.).

La Casa Editrice, pur garantendo la massima cura nella preparazione del volume, declina ogni responsabilità per possibili errori od omissioni, nonché per eventuali danni risultanti dall'uso dell'informazione ivi contenuta.

Il codice QR che si trova sul retro della copertina, consente attraverso uno smartphone di accedere direttamente alle informazioni e agli eventuali aggiornamenti di questo volume.

SOMMARIO

PREMESSA.....	7
---------------	---

capitolo 1

PERCHÉ È NECESSARIA UN'ETICA DELLE IA	11
1.1. Duplice vantaggio di un approccio etico alla IA	19
1.2. IA eticamente affidabile come fattore di competitività	27
1.3. Sfide etiche per la governance delle IA nelle Smart Cities	33
1.4. Sfide etiche per la governance delle IA nelle aziende.....	39
1.4.1. <i>Rischi delle IA per aziende e startup</i>	41
1.4.2. <i>Governance aziendale delle IA</i>	47

capitolo 2

I PRINCIPI ETICI ALLA BASE DEL CANVAS	49
2.1. Beneficenza (fare del bene)	51
2.2. Non maleficenza (non fare del male)	52
2.3. Autonomia (il potere di decidere).....	53
2.4. Giustizia	55
2.5. Esplicabilità	57
2.5.1. <i>La Dichiarazione sull'Etica e la protezione dei dati nelle IA</i>	58

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

2.5.2. <i>La Carta Etica Europea per l'uso delle IA nei sistemi giudiziari</i>	63
2.5.3. <i>Raccomandazioni dell'OCSE sulla IA</i>	67

capitolo 3

GLI ORIENTAMENTI ETICI DELL'UNIONE EUROPEA	73
3.1. Introduzione e sguardo d'insieme	73
3.2. I fondamenti UE per una IA "affidabile"	78
3.3. I principi per la realizzazione di una IA "affidabile"	81
3.3.1. <i>Metodi tecnici e metodi non tecnici</i>	84
3.3.1.1 <i>Metodi tecnici</i>	85
3.3.1.2 <i>Metodi non tecnici</i>	89

capitolo 4

IL CANVAS - ARTIFICIAL INTELLIGENCE ETHICS MODEL	93
4.1. I 10 Moduli costituenti il Canvas	94
Modello 1	
Autonomia Umana – AU	97
Check List del Modulo 1	100
Modello 2	
Governo dei dati – GD	103
Check List del Modulo 2	110
Modello 3	
GDPR e Privacy – GP	113
Check List del Modulo 3	117
Modello 4	
Responsabilità – RES	119
Check List del Modulo 4	130

Modello 5	
Impatto Socio Ambientale – ISA	133
Check List del Modulo 5	145
Modello 6	
Impatto Umano – IU	147
Check List del Modulo 6	15
Modello 7	
Benefici e Benessere – BB	153
Check List del Modulo 7	157
Modello 8	
Equità e Uguaglianza – EU	159
Check List del Modulo 8	168
Modello 9	
Robustezza e Sicurezza – RS	171
Check List del Modulo 9	173
Modello 10	
Trasparenza e Tracciabilità – TT	177
Check List del Modulo 10	180
capitolo 5	
CONCLUSIONI PER LEADERS E STARTUPPER	183
5.1. Mettere a fuoco ed esaminare i valori aziendali	185
5.2. Oltre i valori: cinque aree che richiedono la leadership dall’alto	188
capitolo 6	
GLOSSARIO DEI TERMINI	195

PREMESSA

L'idea del libro nasce dalla necessità di risolvere un'esigenza concreta nell'era della rivoluzione digitale: trovare una sintesi ottimale per leader aziendali e startupper tra i "voli pindarici" dell'approccio filosofico all'etica e l'"accanimento terapeutico" della statistica algoritmica nel cercare di prevenire i potenziali rischi di errore o di risultati non programmati. Per ogni azienda o startup che volesse prendere seriamente in considerazione i dilemmi etici che pone l'uso commerciale delle c.d. "intelligenze artificiali" (ancora troppo "acerbe" e lungi dall'essere regolamentate) si troverebbe di fronte alla diarchia polarizzante dei filosofi e degli statistici-matematici. In ogni convegno, seminario, documento o articolo elaborato sull'argomento, infatti, se prodotto da un filosofo, vi sono incardinate delle affascinanti narrazioni etico-morali, ma con principi e regole difficilmente poi traducibili in procedure aziendali o di governance dell'innovazione; di contro, se prodotti da statistici e/o matematici e/o programmatori sono difficilmente comprensibili (anche da un soggetto normalmente buon conoscitore del lessico matematico), e comunque per loro stessa natura (operando per sintassi e non per semantica) impossibilitate a risolvere formalmente i dilemmi della moralità umana.

Il libro è un tentativo di provare a tracciare un nuovo percorso, una sintesi delle tante "mappe" prodotte dai punti di vista del legislatore, del programmatore, del filosofo, così come dallo specialista di marketing e comunicazione o dal responsabile della cybersecurity. Da qui le due scelte di:

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

1. cercare di tradurre i tantissimi documenti (istituzionali, della ricerca e di enti indipendenti) in procedure aziendali o comunque di ricondurre sempre il tutto (con linguaggio non specialistico) ai sistemi di governance aziendale o di startup.
2. utilizzare un modello già ben noto ai leader aziendali e in generale a tutti coloro che vogliono definire il proprio “piano d’impresa” (il *Business Model Canvas*) e di rileggerlo (con identica impronta grafica) sulla scorta dei diversi Moduli necessari a governare eticamente l’uso commerciale delle intelligenze artificiali (l’*Ethics Model Canvas*).

Per ottenere un risultato concreto, ossia spendibile nelle procedure aziendali, è fondamentale comprendere come il Canvas (e i Moduli di cui è composto) non siano finalizzati a trovare una “algoretica”⁽¹⁾ (come detto, elemento solo per “iniziati”), ma a **valutare i rischi etici dell’uso commerciale delle intelligenze artificiali**, ossia il loro utilizzo all’interno del mercato mondiale per soddisfare specifiche esigenze industriali, sanitarie, editoriali, della logistica, ecc... E quindi la loro conformità normativa alle istanze del GDPR, alle responsabilità penali e assicurative, al benessere dei suoi utilizzatori, alla sicurezza dei sistemi e alla loro accessibilità. Nessuna formula magica, quindi, tantomeno dilemmi morali: solo un sano pragmatismo manageriale, in un equilibrio procedurale tra il livello digitale e quello etico.

Perché questo pragmatismo manageriale? Non solo per forma mentis dello scrivente, ma soprattutto per convenienza aziendale. È ben chiaro a tutti, soprattutto ai grandi leader delle imprese multinazionali (del calibro di Google, Facebook, Microsoft, Amazon, Apple), che la reale possibilità di affermazione dei loro servizi basati sulle “intelligenze artificiali” sia determinata non solo dalla raffinatezza statistico-matematica degli algoritmi, ma soprattutto dalla **FIDUCIA** che i potenziali utenti/cittadini porranno in essi, ossia dal non riceverne danno, diretto o indiretto che sia. Se dovesse prevalere, infatti, il pensiero “luddista” di “paura” nei confronti delle “macchine digitali”; i servizi basati sulle intelligenze artificiali verrebbero sostanzialmente “rigettati” perché

1. Termine coniato dal teologo francescano Padre Paolo Benanti, della Pontificia Università Gregoriana.

non affidabili ⁽²⁾; ancor più se potenzialmente pericolose per lo stesso genere umano come lascia prefigurare la narrazione cinematografica hollywoodiana (Terminator o Matrix su tutti), o come già fanno intuire i casi concreti degli scandali sulla privacy come “Cambridge Analytica” o l’uso fraudolento delle nostre identità digitali per generare “Fake News”. Ecco quindi che, al pari di quanto accaduto per il GDPR, l’Europa si prefigura come area commerciale “più sicura”, più “affidabile”, perché meno piegata sulla logica commerciale e meglio orientata a proteggere i propri cittadini (ma la stessa considerazione si può fare per l’aspetto ambientale e sociale). La fiducia, quindi, come “valore aggiunto” rispetto a prodotti americani o cinesi o russi, di principio non affidabili perché non progettati eticamente o perché non conformi moralmente al benessere dell’uomo e della sua progenie. Tutto qui: **un’etica pragmatica per una strategia di posizionamento del prodotto**. Come i prodotti biologici o ecofriendly sono sempre di più ben accettati dai consumatori (soprattutto dalle giovani generazioni), anche le intelligenze artificiali conformi ai principi etici avranno maggiori possibilità di essere accettate e di diffondersi all’interno delle società più evolute e benestanti.

Il libro, di conseguenza, avendo come scopo il pragmatismo etico per i leader aziendali e gli startupper, è stato scientemente suddiviso in tre parti. Nella prima, sempre secondo la logica della “convenienza aziendale”, sono analizzati i potenziali rischi etici delle intelligenze artificiali e la convenienza commerciale a utilizzarle secondo logiche etiche (umane, sociali, economiche, ambientali, tecnologiche). Nella seconda parte, invece, vengono presentati tutti i principi etici che sottostanno al *Modello di Canvas Etico*, partendo dagli indirizzi dell’Unione Europea e con una panoramica globale che va da quelli indicati dall’OCSE (e quindi accettati dai governi) sino alle indicazioni di gruppi indipendenti come quelli di ricercatori e imprenditori riuniti sotto la dichiarazione di Asilomar. Nella terza parte, il cuore innovativo del libro, viene presentato nel dettaglio il **Modello di Canvas Etico**, rielaborazione di quello originale di Alex Osterwalder (*Business Model Canvas*), costituito da 10 Moduli base su cui costruire lo scenario etico dell’uso commerciale delle intelligenze artificiali: Autonomia umana; Governo dei dati; GDPR e privacy; Responsabilità; Impatto socio-ambientale; Impatto umano; Benefici e

2. Non è un caso che le Linee Guida UE sull’Etica delle Intelligenze Artificiali sono incentrate proprio sul tema della IA “affidabile”.

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

benessere; Equità ed uguaglianza; Robustezza e sicurezza; Trasparenza e tracciabilità. Per essere coerenti con l'approccio manageriale, ogni singolo Modulo del Canvas è dotato di una sua specifica Check-list (sullo stile degli audit di certificazione), così che gli enunciati generali descritti in ogni singolo componente siano poi traducibili in procedure di governance aziendale e dei processi. È ovvio che le Check-list hanno un approccio generale e, quindi, in alcuni punti non possono essere specifiche per i diversi settori di potenziale uso commerciale delle intelligenze artificiali: sarà mia personale cura sviluppare ulteriormente il Canvas per singoli comparti (e-commerce; logistica, moda e lusso; alimentare; ecc.), così da poter poi dare elementi ancora più concreti di applicabilità nelle procedure di aziende o startup. Conclude il libro un'appendice di approfondimento (sempre con linguaggio pragmatico e non specialistico) dei tanti temi trattati nei vari Moduli, cercando di dare maggiori informazioni su cosa siano le intelligenze artificiali e le metodologie che sottostanno al loro sviluppo.

Un ringraziamento particolare, infine, lo debbo a Oliviero Vittori e tutti i soci dell'editore Piazza Copernico (e della sua azienda controllata Conformity) per aver creduto nel mio progetto e per aver investito in esso, con grande pazienza e pragmatismo. Un grazie particolare anche a Martina Longo per le sue traduzioni dei testi specialistici internazionali, senza la quale non avrei potuto "analizzare" una mole così vasta di documenti. Un ringraziamento particolare, infine, va alla mia famiglia (mia moglie Patrizia e i miei figli Corrado e Liliana) che non solo mi hanno pazientemente sopportato per tutto il tempo di elaborazione del libro (oltre un anno), ma soprattutto per i tanti spunti "quotidiani" che mi hanno dato, proprio dall'ottica di chi "non conosce" il mondo delle intelligenze artificiali ma le usa inconsapevolmente ogni giorno con il loro immancabile telefonino.

Per chi volesse rimanere aggiornato sul tema dell'etica delle intelligenze artificiali e l'aggiornamento del libro consiglio il mio neonato gruppo su Facebook:

www.facebook.com/groups/etica.intelligenze.artificiali

capitolo 1

PERCHÉ È NECESSARIA UN'ETICA DELLE IA

Dal punto di vista prettamente tecnico-scientifico, l'**Intelligenza Artificiale (IA)**, tra le tante definizioni coniate nel tempo, è stata descritta come “*La scienza che si propone di sviluppare macchine intelligenti*”⁽¹⁾. Tale definizione, però, rischia di essere fuorviante, in quanto AlphaGO (l'algoritmo di Google che ha battuto il campione mondiale Lee Sedol al gioco cinese Go) e Deep Blue (l'algoritmo di IBM che ha battuto il campione mondiale di scacchi Gary Kasparov), non avevano alcuna consapevolezza di quello che stessero effettivamente facendo (non avevano un capitale “semantico” di riferimento). Erano degli ottimi algoritmi (intelligenze c.d. “deboli” o “specialistiche”), in grado sicuramente di realizzare performance notevolmente migliori degli umani in ambiti predefiniti di regole del gioco (anche sino al livello di rivoluzionare totalmente le strategie di gioco stesso sino ad ora storicamente sviluppate dai maestri), ma in ogni caso ben lontani dall'essere una intelligenza in grado di pensare consapevolmente come il cervello umano (intelligenze c.d. “forti” od “olistiche”). Le moderne applicazioni delle IA non mostrano, infatti, alcuna forma di *intelligenza* che in qualche modo si avvicini a quella umana, sia per capacità “istintiva” che per consapevolezza. In altre parole, le moderne applicazioni di IA non comportano la creazione di macchine in grado di “leggere dentro” (dal latino *intus legere* o *intelligere*) e, di conseguenza, sviluppare una forma di coscien-

1. Russel S.J., Norvig P. (2003), *Artificial Intelligence: a modern approach*, Prentice Hall.

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

za autonoma, una valutazione consapevole del contesto in cui opera o addirittura, una sua propria “visione del mondo” (Weltanschauung) ⁽²⁾. Ne sono dimostratori inconfutati filosofi del livello di John R. Searle ⁽³⁾ che, molto acutamente, sottolinea come le IA lavorino per “sintassi” (codici di programmazione e statistica) e non per “semantica” (tanto meno ontologica) ⁽⁴⁾ o dell’italiano Luciano Floridi ⁽⁵⁾ che, nelle sue conferenze rimarca il “divorzio” tra tecnologia e intelligenza, poiché gli algoritmi sono indubbiamente migliori dell’uomo nella computazione statistica, ma sicuramente incapaci di cogliere il senso delle connessioni causali tra eventi diversi e il “capitale semantico” di cui le interazioni socio-antropiche sono costituite ⁽⁶⁾. Molto più pragmaticamente, per altri autori, le IA, specialmente nelle sue versioni più comuni, vengono assimilate a “software che sono semplicemente capaci di processare una quantità tale di dati da riuscire a metterli in relazione tra loro, identificando eventuali collegamenti o *pattern*, calcolando statisticamente quale mossa abbia probabilità di avere successo” ⁽⁷⁾. Ciò non vuol dire altro che forme di “apprendimento” basate sulla ripetuta esposizione a forme esperienziali (ad esempio, l’osservazione di enormi quantità di immagini), abbinata alla definizione di algoritmi di apprendimento complessi (come le “reti neurali”). Si “allena” dunque la macchina a riconoscere immagini, interpretare il linguaggio, monitorare rischi, individuare tendenze di cui l’essere umano non è in grado (numericamente) di poter calcolare, aumentando di fatto la nostra capacità di interpretare (o distorcere) la realtà. Tali applicazioni delle IA, nella loro forma attuale,

2. Il termine Weltanschauung appartiene alla lingua tedesca ed esprime un concetto fondamentale nella filosofia ed epistemologia tedesca, spesso applicato in vari altri campi, in primis nella critica letteraria e della storia dell’arte. Non è letteralmente traducibile in lingua italiana perché non esiste nel suo lessico una parola che le corrisponda appieno. Essa esprime un concetto di pura astrazione che può essere restrittivamente tradotto con “visione del mondo”, “immagine del mondo” o “concezione del mondo” e può essere riferito a una persona, a un gruppo umano o a un popolo, come a un indirizzo culturale o filosofico o a un’istituzione ideologica in generale e religiosa in particolare.

3. Professore di filosofia e linguaggi mentali all’Università di Berkeley (California).

4. John R. Searle, “Il Mistero della Realtà”, Raffaello Cortina Editore, (2019).

5. Professore di filosofia ed etica digitale all’università di Oxford (Gran Bretagna).

6. Fondazione “Leonardo” (la civiltà delle macchine), Conferenza su “Status giuridico delle Intelligenze Artificiali”, Roma, 2019.

7. Signorelli A.D. (2017), Rivoluzione artificiale. L’uomo all’epoca delle macchine intelligenti. Informant.

PERCHÉ È NECESSARIA UN'ETICA DELLE IA

sono estremamente dipendenti dalla disponibilità di basi dati con archiviate una mole significativa di informazioni; queste ultime, tra l'altro, anche correttamente annotate per permettere all'algoritmo la piena interpretabilità e utilizzabilità nella predetta "fase di apprendimento". Uno degli esempi più importanti è sicuramente il database "Imagenet" per il riconoscimento delle immagini, che contiene 14 milioni di immagini, tutte debitamente annotate.

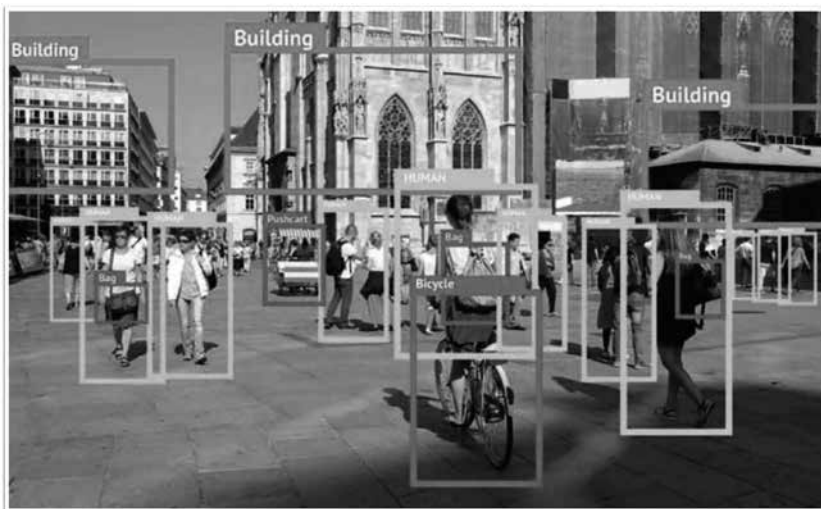


Fig. 1.1 – *Classico esempio di riconoscimento degli oggetti in una immagine con una macchina "allenata" precedentemente con database correttamente annotati* ⁽⁸⁾

In effetti, ad oggi, due sono stati gli approcci principali utilizzati per sviluppare IA:

- Il primo è **up-bottom**, ossia dalle regole generali e funziona solo in domini in cui le regole stesse e le definizioni sono molto chiare (ossia effettivamente processabili da un computer): matematica, scacchi, (video)giochi in generale.
- Il secondo è **bottom-up**, ossia dai dati empirici, apprendendo quindi

8. *Ibidem.*

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

con l'esperienza derivante da prove, tentativi ed errori, che rafforzano o indeboliscono le connessioni tra reti che elaborano le informazioni (l'esempio classico è quello delle c.d. "reti neurali").

Per ottenere i risultati funzionali (alla vincita del gioco o alla guida autonoma dell'auto) le IA fanno riferimento essenzialmente a due metodi:

- 1) il **Machine Learning** (o apprendimento automatico);
- 2) il **Deep Learning** (o apprendimento approfondito).

Il primo approccio (*Machine Learning*) utilizza un set di dati preimpostato (*Big Data*) e apprende da esso, sino a divenire in grado di fare previsioni o trarre conclusioni su qualcosa che però già esiste nel mondo reale. Il secondo invece (*Deep Learning*), pur partendo sempre da un set di dati preimpostato, affida a singole reti specialistiche la ricerca di pattern (modelli) all'interno delle informazioni a disposizione, affidando il risultato ad altre reti specialistiche che a loro volta ricercano altri modelli, sino ad arrivare all'identificazione di un possibile riconoscimento finale di una identità (sia esso un viso o un comportamento), sempreché quest'ultima sia identificabile nel mondo reale (le astrazioni simboliche non sono ad oggi contemplate). I vari livelli lavorano tra loro a cascata: le caratteristiche di livello più alto derivano da quelle di livello più basso, allo scopo di creare una rappresentazione gerarchica che procede per crescenti livelli di raffinazione del dato, compreso l'intervento finale di verifica del risultato o del suo errore (che permette le successive correzioni). Questa tecnica, in pratica, ha consentito di innovare drasticamente il modo tradizionale con cui i software processano i dati e portano a compimento la funzione per cui sono stati sviluppati. A differenza dei programmi classici, che eseguono un codice che spiega loro cosa deve essere fatto, gli attuali algoritmi alla base della IA (debole) apprendono da soli e correggono/perfezionano i loro risultati nel tempo, diventando sempre più affidabili progressivamente ma, ed è questo il primo problema etico, sempre meno "comprensibili" dagli esseri umani su come e perché siano arrivati a determinate conclusioni. Non è un caso, infatti, che questo processo di miglioramento continuo "non trasparente" venga comunemente denominato come "**Black Box**", ossia "Scatola Nera", non essendo facile nemmeno ai loro creatori

comprendere i processi che si sono sviluppati all'interno delle reti e dei loro processi di apprendimento. Ed è in questa "zona grigia" che l'Etica inizia a divenire fondamentale per governare un processo i cui contorni e i suoi effetti non sono più sotto controllo.

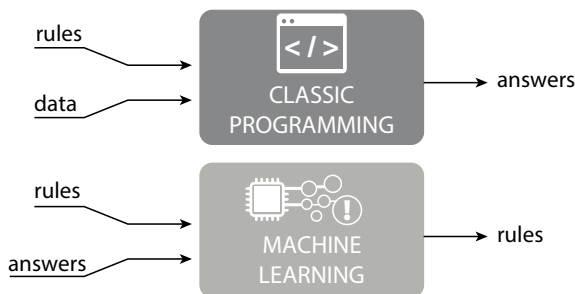


Fig. 1.2 – Schema sintetico che rappresenta graficamente la differenza tra i software tradizionali e l'attuale sistema delle IA basate sul Machine Learning ⁽⁹⁾

Acclarato quindi che la IA sicuramente avranno, e già hanno, un impatto sulla società (sia comportamentale che economico), diviene necessario comprendere se si tratta di un impatto positivo o negativo, e in quale modalità questo può manifestarsi. Per tale motivo diviene altrettanto fondamentale porsi non meno di quattro domande strategiche sulle quali è necessario focalizzarsi, poichè legate alla questione della dignità e della prosperità umana:

- 1) **chi possiamo diventare?** (auto-realizzazione e ricerca della felicità).
- 2) **cosa possiamo fare?** (azione umana potenziata e trascendenza).
- 3) **cosa possiamo ottenere?** (migliori capacità individuali e benefici socioeconomici).
- 4) **come possiamo interagire tra di noi e con il mondo?** (coesione sociale, sviluppo sostenibile, riduzione della polarizzazione delle disuguaglianze, vita extramondo).

9. JRC, Artificial Intelligence: a European perspective, 2008, Luxemburg (cit. "Paradigm shift in ML", Chollet, 2017).

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

In realtà, nonostante le suddette domande presuppongono implicitamente che i sistemi di IA siano destinati al buon uso e in generale a una loro applicazione positiva, è consapevolezza condivisa che di fatto essi possono essere sfruttati in diversi modi:

- possono essere utilizzati per promuovere la natura umana e le sue potenzialità, creando così delle opportunità;
- possono essere sottoutilizzati, non sfruttando le opportunità concesse dalla tecnologia e creando così costi opportunità;
- possono essere utilizzati eccessivamente o in modo improprio, portando di conseguenza alla generazione di rischi o danni al genere umano.

Sono proprio la paura e le preoccupazioni mal riposte che possono indurre la società a sottostimare tali tecnologie e, di conseguenza, a non riuscire a realizzare pienamente quelli che sono gli enormi vantaggi offerti. Come anche l'ignoranza etica sulle conseguenze degli usi "sprejudicati" delle IA è un fattore che potrebbe procurare molti più danni di quanti ne potrebbero arrecare usi in partenza deliberatamente "illeghi" o "malevoli" ⁽¹⁰⁾. Risulta perciò imprescindibile valutare attentamente ciò che può rappresentare la possibilità di progresso sociale rispetto, invece, al rischio che l'IA possa agevolare una manipolazione malevola e causare danni non di poco conto (pensiamo ad esempio a truffe via e-mail o addirittura a guerre cibernetiche su larga scala). Allo stesso modo, di converso, un rischio potrebbe essere anche il fatto che la IA venga sottoutilizzata per paura di farne un uso eccessivo o improprio. Proviamo allora a dare una spiegazione più dettagliata circa la relazione tra opportunità e rischi descritti precedentemente, sviluppandola secondo una logica affine ai leader aziendali e agli startupper.

Chi possiamo diventare: consentire l'autorealizzazione senza svalutare le capacità umane.

Elemento chiave nell'uso di sistemi di IA è la loro capacità di lavorare

10. Sebbene questi rischi derivino in gran parte da conseguenze non intenzionali e riguardino in genere buone intenzioni fallite, è anche opportuno tenere in considerazione casi in cui essi sono frutto di un uso eccessivo o intenzionale di tecnologie di IA, che ad esempio trovano le loro radici in incentivi disallineati, avidità, geopolitica contraddittoria o intenzioni malevole.

PERCHÉ È NECESSARIA UN'ETICA DELLE IA

a sostegno dei bisogni dell'uomo, promuovendone l'autorealizzazione: si tratta dunque di offrire all'uomo la possibilità di realizzarsi e prosperare in termini di caratteristiche, interessi, capacità o abilità potenziali, aspirazioni e progetti di vita (ricerca della felicità).

Infatti, la possibilità di utilizzare "un'automazione intelligente" per essere assistiti nel lavoro (o in altri aspetti della vita) e liberarsi da un lavoro pesante (ma anche da un lavoro domestico), offre sicuramente l'opportunità di avere più tempo a disposizione per dedicarsi ad attività culturali, intellettuali e sociali e, magari, ad un lavoro più interessante e gratificante. Il rischio, in questo caso, non è tanto il degrado di alcune vecchie competenze e l'emergere di nuove, ma piuttosto il ritmo con cui questo accade e le ineguali distribuzioni dei costi e dei benefici che ne derivano. Tale svalutazione e la conseguente interruzione del mercato del lavoro, possono essere analizzate da due punti di vista diversi:

- quello *individuale*, dove vediamo il lavoro come strettamente legato all'identità personale, all'autostima, al ruolo o alla posizione sociale: fattori che possono essere sicuramente influenzati in chiave negativa dalla ridondanza e quindi causare potenziali danni economici.
- quello *sociale*, dove il lavoro in settori più particolari e sensibili, che richiedono competenze specifiche (come possono essere, ad esempio, l'assistenza sanitaria o l'aviazione), potrebbe creare vulnerabilità pericolose in caso di malfunzionamento del sistema di IA o di un attacco malevolo.

È chiaro, dunque, come divenga necessario sviluppare e studiare attentamente un piano d'azione per riuscire a creare una sorta di **equilibrio** tra lo sviluppo della IA a sostegno di nuove abilità e competenze e l'impatto che può avere su quelle vecchie; promuovendo in questo senso idee potenzialmente radicali, al fine poi di sostenere una transazione tra il presente e il futuro nel modo più equo possibile.

Cosa possiamo fare: migliorare la capacità di azione senza rimuovere la responsabilità umana.

Se posta al servizio dell'intelligenza umana, la IA può senza dubbio migliorare la nostra capacità di azione: l'individuo potrebbe essere in grado di fare di più, meglio e più velocemente, con la possibilità di gode-

ETICA DELLE INTELLIGENZE ARTIFICIALI PER LEADER AZIENDALI E STARTUPPER

re di opportunità e benefici illimitati, migliorando allo stesso tempo la società tutta. In questo senso è essenziale il concetto di responsabilità, in relazione a come sviluppiamo, usiamo e condividiamo i benefici di questa tecnologia. Allo stesso tempo però, tale responsabilità si traduce anche in un rischio, nel momento in cui, ad esempio, interagiamo con una IA attraverso una mentalità da *black box*, considerando cioè il processo decisionale del sistema al di là della nostra comprensione e quindi anche del nostro controllo. Queste preoccupazioni si possono riscontrare non solo in casi di alto profilo, come i decessi causati da veicoli autonomi, ma anche in usi più banali ma comunque significativi, come nelle decisioni automatizzate sulla libertà condizionale o sull'affidabilità creditizia. Se correttamente sviluppata la IA può offrire grandi opportunità di perfezionare e ampliare l'azione umana: è necessario però integrare determinati "frameworks di facilitazione", progettati per migliorare la probabilità di risultati moralmente positivi, nell'insieme di funzioni che deleghiamo ai sistemi di IA.

Cosa possiamo ottenere: aumentare le capacità della società senza ridurre il controllo umano.

Un uso maggiore dei sistemi di IA a supporto delle capacità degli individui e della società, in generale, comporta anche il raggiungimento di obiettivi più ambiziosi. Che si tratti di prevenire e curare malattie o di ottimizzare i trasporti e la logistica, l'uso delle tecnologie di IA presenta innumerevoli possibilità di reinventare la società stessa, attraverso il miglioramento radicale di ciò di cui gli esseri umani sono capaci. Ma proprio perché tecnologie come queste hanno potenziali così grandi, significa che possono anche introdurre rischi ad essi connessi. Il desiderio sempre crescente di delegare i propri compiti al sistema di IA, porta l'uomo a non dover essere sempre "in/on the loop" (cioè come parte del processo o del suo controllo), riducendo ovviamente la sua capacità di monitorare le prestazioni del sistema e, di conseguenza, non di poter più prevenire o correggere gli errori che si presentano ("post loop"). Allo stesso modo, potrebbe accadere che tali errori si accumulino tra di loro e diventino poi difficili da eliminare. Per questo motivo, anche qui diviene indispensabile riuscire a trovare un equilibrio tra perseguire le numerose opportunità offerte da una IA e assicurare che l'uomo resti in controllo di questi sviluppi e dei relativi risultati.

Pagine omesse dall'anteprima del volume