# Diritto dell'era digitale

### Lezione 13 Intelligenza artificiale e robotica

Prof. Giovanni Pascuzzi



# Una definizione di intelligenza artificiale Comunicazione della Commissione al parlamento europeo, al consiglio, al comitato

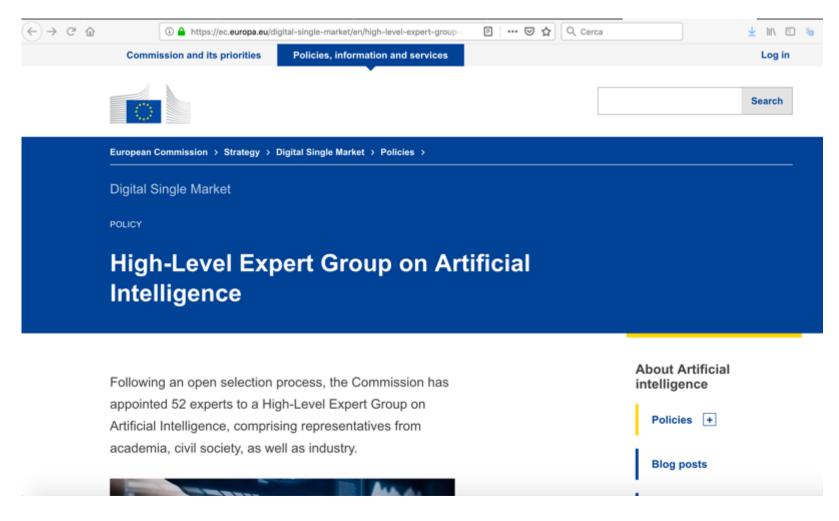
economico e sociale europeo e al comitato delle regioni l'intelligenza artificiale per l'europa (swd(2018) 137 final)

- Cos'è l'intelligenza artificiale?
- "Intelligenza artificiale" (IA) indica sistemi che mostrano un comportamento intelligente analizzando il proprio ambiente e compiendo azioni, con un certo grado di autonomia, per raggiungere specifici obiettivi.
- I sistemi basati sull'IA possono consistere solo in software che agiscono nel mondo virtuale (per esempio assistenti vocali, software per l'analisi delle immagini, motori di ricerca, sistemi di riconoscimento vocale e facciale); oppure incorporare l'IA in dispositivi hardware (per esempio in robot avanzati, auto a guida autonoma, droni o applicazioni dell'Internet delle cose).
- Utilizziamo l'IA quotidianamente, per esempio per tradurre le lingue, generare sottotitoli nei video o bloccare lo spam delle email.
- Molte tecnologie di IA richiedono dati per migliorare le loro prestazioni. Raggiunto un buon livello di prestazioni, esse possono contribuire a migliorare e automatizzare il processo decisionale nello stesso campo. Per esempio, un sistema di IA verrà addestrato e in seguito utilizzato per rilevare gli attacchi informatici sulla base dei dati provenienti dal sistema o dalla rete interessati.



#### La Commissione ha istituito....

https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence





Fonte: The European Commission's HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE A DEFINITION OF AI: MAIN CAPABILITIES AND SCIENTIFIC DISCIPLINES

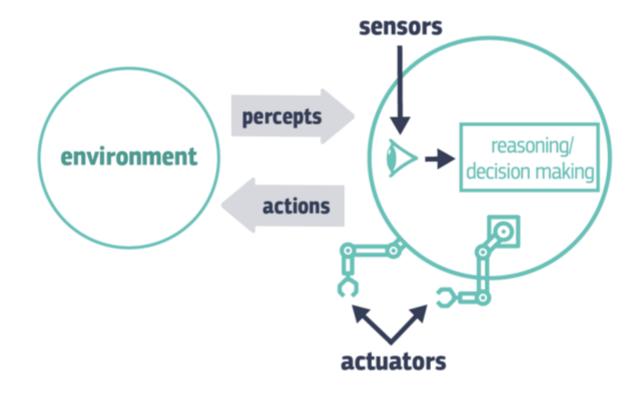


Figure 1: Schematic depiction of an Al system.



Fonte: The European Commission's HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE A DEFINITION OF AI: MAIN CAPABILITIES AND SCIENTIFIC DISCIPLINES

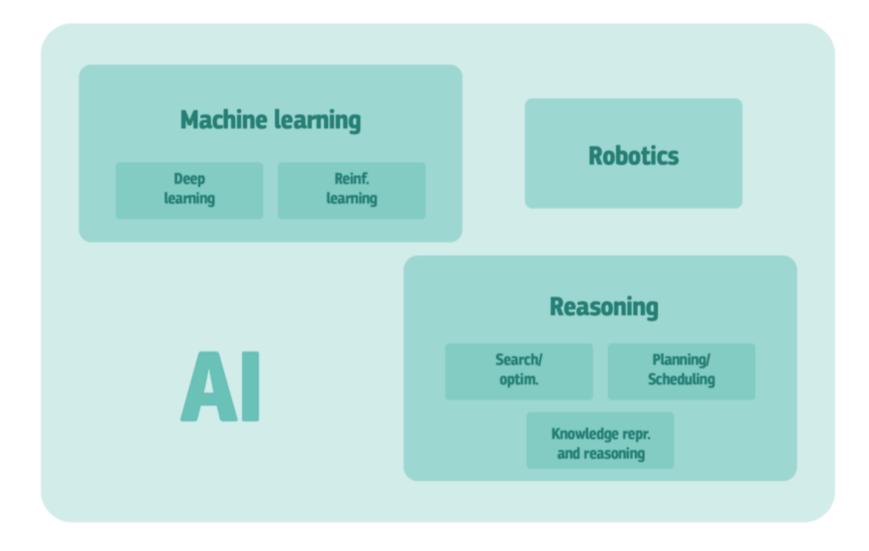


Figure 2: Al's sub-disciplines and their relationship.



#### Fonte: The European Commission's HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE A DEFINITION OF AI: MAIN CAPABILITIES AND SCIENTIFIC DISCIPLINES

- Definizione aggiornata di AI (versione 2018)
- "L'intelligenza artificiale (AI) si riferisce a sistemi progettati dall'uomo che, dato un obiettivo complesso, agiscono nel mondo fisico o digitale percependo il loro ambiente, interpretando i dati strutturati o non strutturati raccolti, ragionando sulle conoscenze derivate da questi dati e decidendo le migliori azioni da intraprendere (in base a parametri predefiniti) per raggiungere l'obiettivo indicato.
- I sistemi di intelligenza artificiale possono anche essere progettati per imparare ad adattare il loro comportamento analizzando come l'ambiente è influenzato dalle loro azioni precedenti.
- Come disciplina scientifica, l'intelligenza artificiale include diversi approcci e tecniche, come l'apprendimento automatico (di cui l'apprendimento profondo e l'apprendimento per rinforzo sono esempi specifici), il ragionamento automatico (che comprende pianificazione, programmazione, rappresentazione e ragionamento delle conoscenze, ricerca e ottimizzazione) e robotica (che comprende controllo, percezione, sensori e attuatori, nonché l'integrazione di tutte le altre tecniche nei sistemi cyberfisici). "

- Definizione aggiornata di Al (versione 2019)
- "I sistemi di intelligenza artificiale (AI) sono sistemi software (e possibilmente anche hardware) progettati da esseri umani che, dato un obiettivo complesso, agiscono nella dimensione fisica o digitale percependo il loro ambiente attraverso l'acquisizione dei dati, interpretando i dati strutturati o non strutturati raccolti, ragionando sulla conoscenza o sull'elaborazione delle informazioni derivate da questi dati e decidendo delle migliori azioni da intraprendere per raggiungere l'obiettivo prefissato.
- I sistemi di intelligenza artificiale possono <u>utilizzare regole</u> <u>simboliche o apprendere un modello numerico</u> e possono anche adattare il loro comportamento analizzando il modo in cui l'ambiente è influenzato dalle loro azioni precedenti.
- Come disciplina scientifica, l'intelligenza artificiale include diversi approcci e tecniche, come l'apprendimento automatico (di cui l'apprendimento profondo e l'apprendimento per rinforzo sono esempi specifici), il ragionamento automatico (che comprende pianificazione, programmazione, rappresentazione e ragionamento delle conoscenze, ricerca e ottimizzazione) e robotica (che comprende controllo, percezione, sensori e attuatori, nonché l'integrazione di tutte le altre tecniche nei sistemi cyberfisici).

#### Definizione di deep learning (apprendimento profondo)

Comunicazione della commissione l'intelligenza artificiale per l'europa {swd(2018) 137 final}

- L'apprendimento profondo è stato un elemento rivoluzionario per l'IA, che ha comportato un incredibile miglioramento delle prestazioni per compiti specifici quali il riconoscimento di immagini o vocale o la traduzione automatica.
- Addestrare un algoritmo di apprendimento profondo a classificare gli oggetti significa fornirgli una grande quantità di esempi etichettati (per esempio immagini) che sono correttamente categorizzati (per esempio immagini di aeroplani).
- Una volta addestrati, gli algoritmi possono classificare correttamente oggetti che non hanno mai visto, in alcuni casi con una precisione che supera quella umana.
- Progressi significativi in queste tecnologie sono stati ottenuti mediante l'impiego di grandi set di dati e una potenza di elaborazione senza precedenti.



#### Dal dato alla conoscenza

- Dati personali e dati non personali
- I big data
- Dato informazione conoscenza



#### Data analytics

• Concetto composito costituito da acquisizione, raccolta, convalida, elaborazione, quantificazione, visualizzazione e interpretazione dei dati (l'analisi dei dati viene utilizzata per comprendere gli oggetti rappresentati dai dati, per fare previsioni per una determinata situazione e consigliare sui passi da fare per raggiungere gli obiettivi. Le intuizioni ottenute dalla data analytics sono utilizzate per vari scopi come il processo decisionale, la ricerca, lo sviluppo sostenibile, la progettazione, la pianificazione, ecc.).

- 1) Il fenomeno.
  - Con l'espressione «data mining» si fa riferimento a quel processo computazionale che crea modelli analizzando i dati quantitativi da diverse prospettive e dimensioni, classificandoli ed enucleando potenziali relazioni e impatti.
  - «Data mining» significa estrarre conoscenza dai dati: è il terreno proprio della «data science».
  - L'aggettivo «educational» ci dice che tale attività viene applicata al mondo della formazione, in particolare a quello della didattica online e dei cosiddetti MOOC (acronimo che sta per «Massive Open Online Courses», ovvero «Corsi aperti online su larga scala»).
  - Gli studiosi di «Educational data mining» sostengono che è possibile dedurre conoscenze accurate dalla mole enorme di dati generati da ogni studente nelle proprie attività di apprendimento a distanza.
  - In particolare essi ritengono che si possano prevedere le prestazioni e gli apprendimenti degli allievi in base alla valutazione dei loro risultati, della partecipazione, dell'impegno, dei voti e delle conoscenze disciplinari in un lavoro di apprendimento.



# Esempio: Educational data mining ovvero «Learning analytics»

H. Aldowah et al. Telematics and Informatics 37 (2019) 13-49

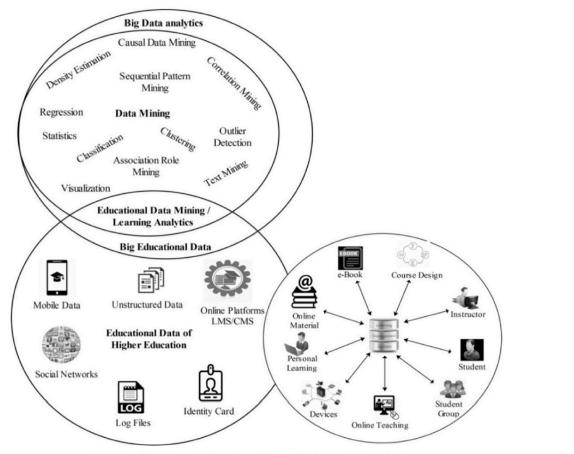


Fig. 1. An illustration of data mining (EDM and LA) use in higher education.



- 2) Le modalità.
  - Ogni studente impegnato in attività di apprendimento a distanza «genera» tanti dati.
    - il tempo trascorso sulle piattaforme che erogano corsi online,
    - il tipo di interazioni che si sono poste in essere,
    - l'attività svolta sulla singola pagina consultata,
    - le modalità di utilizzo del mouse,
    - la velocità e il tipo di movimento degli occhi sullo schermo.
  - Tali dati vengono elaborati attraverso le tecniche tipiche del «data mining».
    - «classificazione» attraverso la quale si possono mappare i dati in diverse classi predefinite così da conoscere i rendimenti, gli abbandoni, i comportamenti problematici, il livello di motivazione e così via;
    - «clustering» che consente di trovare gruppi di studenti con caratteristiche di apprendimento simili in base al contenuto delle pagine visitate e dei loro modelli di navigazione online;
    - «l'association rule» usata per scoprire regole di apprendimento in base alle caratteristiche e alle competenze degli studenti.



- 3) I vantaggi.
  - Gli studenti possono personalizzare l'apprendimento e migliorare il loro rendimento aumentando la consapevolezza di sé;
  - I professori, grazie ai continui feedback sul lavoro svolto, possono fare interventi mirati per migliorare l'efficacia dell'insegnamento e fronteggiare per tempo la possibile emersione di difficoltà (a tacere della possibilità di sperimentare strategie didattiche innovative come il «collaborative learning»);
  - Le agenzie educative sono in grado di monitorare l'intero processo formativo per migliorarlo anche attraverso la riprogettazione dei piani di studio, la riduzione del tasso di abbandono, l'individuazione degli «studenti a rischio»;
  - I genitori possono ottenere informazioni più specifiche circa le attività di studio dei propri figli.



- 4) Le criticità. Ci sono problemi:
  - a) tecnici: davvero possiamo ritenere attendibili dei dati raccolti in questo modo?
     Quale qualità, quale accuratezza, quale significato statistico possono avere dati accumulati in maniera massiva ma grezza?
  - b) giuridici: c'è ovviamente il tema della tutela dei dati personali: chi può accedere ai dati degli studenti? per quanto tempo? che garanzie ci sono che essi non vengano utilizzati per profilare i ragazzi per le più svariate finalità più o meno commendevoli?
  - c) etici: siamo sicuri che sia moralmente corretto affidare la valutazione degli apprendimenti a procedimenti totalmente automatizzati? può l'essere umano essere giudicato da una macchina?
  - d) pedagogici: la valutazione degli apprendimenti è legata a filo doppio alle diverse teorie dell'apprendimento; se non c'è una chiara definizione delle finalità dell'insegnamento non si può avviare nessuna procedura di valutazione meno che mai una procedura basata sulla mera raccolta massiva di dati disomogenei.



#### Machine learning (o apprendimento automatico)

- Processo mediante il quale un'unità funzionale migliora le sue prestazioni acquisendo nuove conoscenze o abilità o riorganizzando le conoscenze o le abilità esistenti.
- Le macchine apprendono elaborando i dati per estrarne dei modelli ovvero applicando regole conosciute per:
  - a) categorizzare o catalogare persone o cose;
  - b) predire dei risultati o delle azioni basandosi su schemi identificati;
  - c) identificare modelli e relazioni attualmente sconosciuti;
  - d) rilevare comportamenti anomali o inattesi.



- – supervised learning (apprendimento supervisionato): la macchina apprende tramite esempi. (Vengono forniti esempi degli input e degli output desiderati).
  - Essi vengono usati per determinare le correlazioni e la logica che è possibile utilizzare per prevedere la risposta. Una logica che potrà essere adoperata in ogni caso simile.
  - Tra le tecniche di questo tipo di approccio (che può essere usato, ad esempio, nella valutazione del rischio, nella scoperta delle frodi, nel riconoscimento di testi, discorsi e immagini) sono ricomprese: la statistica bayesiana; gli alberi decisionali; le reti neurali; la regression analysis;
- – semi-supervised learning (apprendimento semisupervisionato): viene utilizzato per affrontare problemi simili a quelli appena visti ma con una differenza.
  - Alla macchina vengono forniti alcuni dati con la risposta definita (nota anche come «risposta etichettata») insieme a dati aggiuntivi che non sono etichettati con la risposta: alcuni dei dati di input sono contrassegnati con l'output desiderato, mentre il resto no.
    - Si usa quando ci siano troppi dati o sottili variazioni nei dati così da rendere gravoso fornire un set completo di esempi.
    - Gli input e gli output forniti definiscono il modello generale che la macchina può estrapolare e applicare ai dati rimanenti;



- unsupervised learning (apprendimento non supervisionato): la macchina studia i dati per estrapolare i modelli senza chiavi preconfezionate di comprensione.
  - Determina le relazioni e le correlazioni analizzando i dati disponibili.
    - L'approccio è quello che noi usiamo per catalogare il mondo: cogliamo relazioni tra le cose e le raggruppiamo in categorie. Più cresce l'esperienza meglio siamo in grado di classificare.
  - Tra le tecniche di questo tipo di approccio (che può essere usato, ad esempio, per riconoscere gli oggetti simili, ovvero per studiare le abitudini di acquisto dei consumatori, o per rilevare i comportamenti inattesi) ci sono:
    - l'analisi di affinità;
    - il clustering o analisi dei gruppi;
    - la nearest neighbor mapping che consente di determinare le relazioni tra due set di dati non correlati quantificandone la somiglianza (vicinanza);



- reinforcement learning (apprendimento rinforzato): alla macchina sono fornite una serie di regole, di azioni consentite e di potenziali risultati finali.
  - Applicando le regole, esplorando diverse azioni e osservando le reazioni risultanti, la macchina impara a sfruttare le regole per ottenere il risultato desiderato.
    - È come insegnare a qualcuno le regole del gioco: infatti è la tecnica usata nella robotica ovvero per permettere ai computer di giocare a scacchi.
  - Tra le tecniche proprie di questo tipo di machine learning ci sono
    - il learning automata (consente di selezionare l'azione da compiere sulla base delle esperienze passate); il processo decisionale di Markov;
    - il Q-learning (permettere al sistema di adattarsi all'ambiente che lo circonda migliorando la scelta delle azioni da eseguire cercando di massimizzare il valore del passo successivo);



- deep learning (apprendimento profondo): questo metodo fa un uso avanzato delle reti neurali usando un ammontare molto più ampio di dati.
  - Comunicazione della Commissione europea sull'intelligenza artificiale per l'Europa
    - «L'apprendimento profondo è stato un elemento rivoluzionario per l'IA, che ha
      comportato un incredibile miglioramento delle prestazioni per compiti specifici quali il
      riconoscimento di immagini o vocale o la traduzione automatica. Addestrare un
      algoritmo di apprendimento profondo a classificare gli oggetti significa fornirgli una
      grande quantità di esempi etichettati (ad esempio, immagini) che sono correttamente
      categorizzati (ad esempio, immagini di aeroplani). Una volta addestrati, gli algoritmi
      possono classificare correttamente oggetti che non hanno mai visto, in alcuni casi con
      una precisione che supera quella umana. Progressi significativi in queste tecnologie sono
      stati ottenuti mediante l'impiego di grandi set di dati e una potenza di elaborazione
      senza precedenti»



- cognitive computing: ci sono poi sistemi che cercano di comprendere ed emulare il comportamento umano utilizzando anche il linguaggio naturale
  - (si pensi agli assistenti vocali come Alexa di Amazon, Siri di Apple o l'assistente vocale di Google).
- La computazione cognitiva usa una serie di tecniche compreso il processamento del linguaggio naturale, gli algoritmi avanzati di machine learning, il deep learning e la generazione di linguaggio naturale.



# Max Tegmark è professore di fisica al MIT e presiede il "<u>Future of Life Institute</u>" una istituzione che si propone di "catalizzare e sostenere la ricerca e le iniziative per salvaguardare la vita e sviluppare visioni ottimistiche del futuro, compresi modi positivi per l'umanità di orientare il proprio percorso considerando nuove tecnologie e sfide" (futureoflife.org).

- "vita": ogni processo capace di mantenere la sua complessità e di replicarsi.
- Essa può svilupparsi in tre stadi.
- stadio biologico (1.0):
  - hardware e software sono soggetti all'evoluzione.
- stadio culturale (2.0):
  - la vita può progettare il proprio software (essenzialmente attraverso l'apprendimento).
- stadio tecnologico (3.0):
  - la vita può progettare anche il proprio hardware "diventando padrona del proprio destino"



#### Max Tegmark

- Tegmark stabilisce una equivalenza tra l'intelligenza e la capacità di realizzare fini complessi.
- Attualmente l'intelligenza artificiale disponibile è molto ristretta.
  - Ci sono, ad esempio, computer imbattibili nel gioco degli scacchi: ma sanno fare solo quello. L'uomo, invece, ha uno spettro molto ampio di abilità rispetto a tutti i fini.
- l'intelligenza artificiale ci stia facendo fare progressi nel campo
  - dell'esplorazione spaziale,
  - della produzione industriale,
  - dei trasporti,
  - dell'energia ("reti intelligenti"),
  - della sanità (con i robot che praticano operazioni chirurgiche anche a distanza),
  - delle comunicazioni ("internet delle cose").
- Mette in guardia su situazioni già oggi problematiche (la riduzione dei posti di lavoro con la necessità di trovare il "senso" in attività diverse da quelle lavorative) o addirittura inquietanti: armi, come i droni, in grado di scegliere da sole l'obiettivo umano da distruggere.



#### Max Tegmark

- L'intelligenza artificiale può davvero assumere il controllo del mondo o consentire agli umani di farlo (p. 179)?
- se riusciremo a realizzare una IA di livello umano è probabile che si inneschi una "esplosione di intelligenza".
  - possibili scenari: una superintelligenza che conviva con gli umani (perché "buona" o perché "costretta"); ovvero che li sostituisca del tutto.
- Nocciolo dei problemi: quello dei fini.
- Prima che la situazione ci sfugga di mano è indispensabile fare in modo che l'intelligenza artificiale: a) comprenda i nostri fini; b) adotti i nostri fini; c) conservi i nostri fini.
- Il rischio vero non è quello di ritrovarsi una IA "cattiva", bensì una IA estremamente brava nel realizzare propri fini non coincidenti con i nostri. È facile intuire che il problema di fondo assume una dimensione etica: quali sono i nostri "fini ultimi"? In che modo dobbiamo cercare di plasmare il nostro universo?



# I fenomeni I campi di operatività dell'intelligenza artificiale

# a) L'analisi e la predisposizione automatica di atti e documenti.

- Luminance (https://www.luminance.com).
  - Fondata da matematici dell'Università di Cambridge,
  - Hanno sviluppato il Legal Inference Transformation Engine (LITE), un'applicazione dell'apprendimento automatico al mondo giuridico, combinando la tecnologia di riconoscimento dei modelli con l'apprendimento automatico supervisionato e non supervisionato per leggere e comprendere il linguaggio umano, a velocità e in volumi ben oltre la capacità umana.
  - Luminance consente agli avvocati di analizzare rapidamente set di documenti e contratti.



#### b) L'analisi predittiva

- Capacità di elaborare previsioni mediante un calcolo probabilistico effettuato da algoritmi operanti su base semplicemente statistica o su base logica.
  - Prevenzione della criminalità
    - inserendo una serie di dati (ad esempio, relativi a rapine o furti verificatisi nelle stesse zone e con modalità analoghe), il sistema è in grado di prevedere luoghi e orari in cui verosimilmente potranno essere commessi altri reati della stessa specie.
  - Previsione dell'esito di un giudizio
    - Nel 2016 è stato svolto uno studio che, grazie ai progressi nell'elaborazione del linguaggio naturale e nell'apprendimento automatico, si proponeva di costruire modelli predittivi utili a svelare gli schemi che guidano le decisioni giudiziarie. Il lavoro ha previsto l'esito dei casi analizzati dalla Corte europea dei diritti umani basandosi sul loro contenuto testuale. La previsione è riuscita nel 79% dei casi
  - Previsione dei comportamenti di tutti gli attori del sistema giuridico.
    - Lex Machina, ad esempio, combina dati e software per creare set di dati su giudici, avvocati, parti e soggetti di cause legali, su milioni di pagine di informazioni sulle controversie. Con questi dati gli avvocati possono prevedere i comportamenti e gli esiti che produrranno le diverse possibili strategie legal.



#### Altre applicazioni

- Nella <u>Risoluzione del Parlamento europeo del 16 febbraio 2017</u> recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL)) si citano:
- I veicoli autonomi
- I droni
- I robot impiegati per l'assistenza
- I robot medici
- Gli interventi riparativi del corpo umano



#### Altre applicazioni

Esempi di Al nella <u>Comunicazione della Commissione L'intelligenza artificiale per l'Europa Bruxelles,</u> 25.4.2018 COM(2018) 237 final

- trattamento delle malattie croniche
- riduzione dei tassi di incidenti stradali mortali
- lotta contro il cambiamento climatico
- prevenzione delle minacce alla sicurezza informatica.
- In Danimarca, l'IA aiuta a salvare vite umane permettendo ai servizi di emergenza di diagnosticare gli arresti cardiaci o altre patologie in base al suono della voce di chi chiama.
- In Austria, aiuta i radiologi a individuare i tumori in modo più accurato confrontando istantaneamente le radiografie con un'elevata quantità di altri dati medici.
- Molte aziende agricole in tutta Europa impiegano già l'IA per controllare il movimento, la temperatura e il consumo di mangime degli animali. Un sistema di IA può adattare automaticamente il macchinario che provvede a riscaldare e ad erogare il mangime, permettendo agli allevatori di dedicarsi ad altro pur monitorando il benessere dei loro animali.
- L'IA aiuta anche i produttori europei a diventare più efficienti e contribuisce a riportare gli stabilimenti in Europa.



# lrobot



# Asimov: le tre (poi quattro) leggi della robotica





- Legge zero: Un robot non può recare danno all'umanità, né può permettere che, a causa del proprio mancato intervento, l'umanità riceva danno.
- 1. Un robot non può recar danno a un essere umano né può permettere che, a causa del proprio mancato intervento, un essere umano riceva danno. Purché questo non contrasti con la Legge Zero
- 2. Un robot deve obbedire agli ordini impartiti dagli esseri umani, purché tali ordini non contravvengano alla Legge Zero e alla Prima Legge.
- 3. Un robot deve proteggere la propria esistenza, purché questa autodifesa non contrasti con la Legge Zero, la Prima Legge e la Seconda Legge.



#### Il parlamento europeo cita le leggi di Asimov

- Norme di diritto civile sulla robotica
- Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL))
- Principi generali
- T. considerando che le **leggi di Asimov** devono essere considerate come rivolte ai progettisti, ai fabbricanti e agli utilizzatori di robot, compresi i robot con capacità di autonomia e di autoapprendimento integrate, dal momento che tali leggi non possono essere convertite in codice macchina;



- Definizione e classificazione dei "robot intelligenti"
- È opportuno stabilire una definizione comune europea di robot autonomo intelligente, comprese eventualmente le definizioni delle sue sottocategorie, tenendo conto delle seguenti caratteristiche:
  - la capacità di acquisire autonomia grazie a sensori e/o mediante lo scambio di dati con il proprio ambiente (interconnettività) e l'analisi di tali dati;
  - – la capacità di apprendimento attraverso l'esperienza e l'interazione;
  - la forma del supporto fisico del robot;
  - – la capacità di adeguare il suo comportamento e le sue azioni all'ambiente.



- Registrazione dei robot intelligenti
- Ai fini della tracciabilità e onde agevolare l'applicazione di ulteriori raccomandazioni, è opportuno prevedere un sistema di registrazione dei robot avanzati, sulla base dei criteri fissati per la classificazione dei robot. Il sistema di registrazione e il registro dovrebbero essere istituiti a livello di Unione, coprendo il mercato interno, e potrebbero essere gestiti da un'agenzia designata dell'UE per la robotica e l'intelligenza artificiale, qualora tale agenzia sia istituita.



- Responsabilità civile
- Qualsiasi soluzione giuridica si scelga da applicare alla responsabilità per i robot e l'intelligenza artificiale in casi diversi da quelli di danni alle cose non dovrebbe in alcun modo limitare il tipo o l'entità dei danni che possono essere risarciti, né dovrebbe limitare le forme di risarcimento che possono essere offerte alla parte lesa per il semplice fatto che il danno è provocato da un soggetto non umano.
- Il futuro strumento legislativo dovrebbe essere fondato su una valutazione approfondita della Commissione che stabilisca se applicare l'approccio della responsabilità oggettiva o della gestione dei rischi.
- Dovrebbe inoltre essere istituito un regime assicurativo obbligatorio, che potrebbe basarsi sull'obbligo del produttore di stipulare una copertura assicurativa per i robot autonomi che produce.
- Il regime assicurativo dovrebbe essere integrato da un fondo per garantire il risarcimento dei danni in caso di assenza della copertura assicurativa.
- Qualunque decisione politica relativa alle norme sulla responsabilità civile applicabili ai robot e all'intelligenza artificiale dovrebbe essere presa di concerto con un progetto di ricerca e sviluppo di portata europea dedicato alla robotica e alla neuroscienza, con scienziati ed esperti in grado di valutarne tutti i rischi correlati e le possibili conseguenze.



- Interoperabilità, accesso al codice e diritti di proprietà intellettuale
- È opportuno garantire l'interoperabilità dei robot autonomi collegati in rete che interagiscono tra di loro. Dovrebbe essere possibile accedere al codice sorgente, ai dati di input e ai dettagli costruttivi in caso di necessità, per poter indagare su eventuali incidenti e danni causati dai robot intelligenti e per garantirne il funzionamento ininterrotto, la disponibilità, l'affidabilità e la sicurezza.

#### Norme di diritto civile sulla robotica Risoluzione del Parlamento europeo del 16 febbraio 2017

#### PROFILI ETICI

- I ricercatori del settore della robotica dovrebbero impegnarsi a tenere un comportamento etico e deontologico quanto più rigoroso possibile e a rispettare i seguenti principi:
  - beneficenza: i robot devono agire nell'interesse degli esseri umani;
  - non-malvagità: la dottrina del "primum, non nocere", in virtù della quale i robot non devono fare del male a un essere umano;
  - autonomia: la capacità di adottare una decisione informata e non imposta sulle condizioni di interazione con i robot;
  - giustizia: un'equa ripartizione dei benefici associati alla robotica e l'accessibilità economica dei robot addetti all'assistenza a domicilio e, in particolare, a quelli addetti alle cure sanitarie.



#### Norme di diritto civile sulla robotica Risoluzione del Parlamento europeo del 16 febbraio 2017

- Cosa i ricercatori devono garantire
- Diritti fondamentali
- Precauzione
- Inclusione
- Rendicontabilità
- Sicurezza
- Reversibilità
- Vita privata
- Massimizzare i vantaggi e ridurre al minimo il danno



#### Sul piano etico esistono già molti decaloghi di regole

- La citata <u>Risoluzione del Parlamento europeo</u>
- Future of life institute <a href="https://futureoflife.org/ai-principles/">https://futureoflife.org/ai-principles/</a>
- Università di Montreal <a href="https://www.montrealdeclaration-responsibleai.com/the-declaration">https://www.montrealdeclaration-responsibleai.com/the-declaration</a>
- The future of world <u>http://www.thefutureworldofwork.org/media/35420/uni\_ethical\_ai.</u>
   <u>pdf</u>
- Analizzando questi documenti, si scopre che diversi sono i "valori" cui dovrebbero ispirarsi la robotica e l'intelligenza artificiale.



| lstituzioni<br>Valori   | Unione<br>Europea<br>(Parlamento) | Università<br>di<br>Montreal | Future of<br>life<br>institute | Uni Global<br>Union |
|---|-----------------------------------|------------------------------|--------------------------------|---------------------|
| Allineamento con i valori umani (rispettare le leggi vigenti)   | 332                               |                              | SI                             | SI                  |
| Assicurare una equa transizione   |                                   |                              | 7                              | SI                  |
| Autonomia   | SI                                | SI                           |                                |                     |
| Beneficenza (agire nell'interesse dell'essere umano)  | SI                                |                              |                                | SI                  |
| Benessere (promuovere il benessere delle creature senzienti)  |                                   | SI                           | SI                             |                     |
| Conoscenza (promuovere il pensiero critico ed evitare la manipolazione)   |                                   | SI                           |                                |                     |
| Costituire meccanismi di governo globale  |                                   | -                            |                                | SI                  |
| Democrazia  |                                   | SI                           |                                |                     |
| Essere al servizio delle persone e del pianeta  |                                   |                              |                                | SI                  |
| Evitare discriminazioni (basate su sesso, razza, etc).  |                                   |                              | *                              | SI                  |
| Evitare la corsa agli armamenti   |                                   |                              | SI                             | SI                  |
| Giustizia   | SI                                | SI                           | - 2                            |                     |
| Inclusione  | SI                                |                              |                                |                     |
| Massimizzare i vantaggi e ridurre al minimo il danno  | SI                                |                              |                                |                     |
| Non attribuire la responsabilità ai robot   |                                   |                              |                                | SI                  |
| Non malvagità   | SI                                |                              |                                |                     |
| Precauzione   | SI                                |                              |                                |                     |
| Privacy   | SI                                | SI                           | SI                             |                     |
| Promuovere la prosperità  |                                   |                              | SI                             |                     |
| Rendicontabilità/Responsabilità (dei progettisti)   | SI                                | SI                           | SI                             |                     |
| Reversibilità   | SI                                |                              |                                |                     |
| Rispetto diritti fondamentali   | SI                                |                              |                                | SI                  |
| Rispetto ordine costituito (il potere conferito dal controllo di sistemi IA<br>altamente avanzati dovrebbe rispettare e migliorare, piuttosto che sovvertire, i processi sociali e<br>civici da cui dipende la salute della società). |                                   |                              | SI                             |                     |
| Sicurezza   | SI                                |                              | SI                             |                     |
| Sistemi Al equipaggiati con una "Scatola nera etica"  |                                   |                              |                                | SI                  |
| Trasparenza   |                                   |                              | SI                             | SI                  |

Fonti



i. Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL))

ii.Future of life institute https://futureoflife.org/ai-principles/

iii.Università di Montreal https://www.montrealdeclaration-responsibleai.com/the-declaration

iv.Uni global unione http://www.thefutureworldofwork.org/media/35420/uni\_ethical\_ai.pdf

#### Su due "valori" vorrei soffermare l'attenzione.

- Il primo è la progettazione consapevole.
  - Software e robot devono essere progettati in modo che siano rispettate determinate regole, ad esempio il rispetto della privacy. Si chiama "privacy by design", ed è previsto dall'articolo 25 del <u>nuovo regolamento sulla privacy.</u>
- Il secondo è la trasparenza (apertura versus chiusura).
  - Nella citata <u>Risoluzione del Parlamento europeo del 16 febbraio 2017 recante</u> raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica (2015/2103(INL)), si legge testualmente:
    - (2015/2103(INL)), si legge testualmente: "pone l'accento sul principio della trasparenza, nello specifico sul fatto che dovrebbe sempre essere possibile indicare la logica alla base di ogni decisione presa con l'ausilio dell'intelligenza artificiale che possa avere un impatto rilevante sulla vita di una o più persone; ritiene che debba sempre essere possibile ricondurre i calcoli di un sistema di intelligenza artificiale a una forma comprensibile per l'uomo; ritiene che i robot avanzati dovrebbero essere dotati di una "scatola nera" che registri i dati su ogni operazione effettuata dalla macchina, compresi i passaggi logici che hanno contribuito alle sue decisioni".

