

# Pensare senza testa

Sintesi dello studio «Wenn Algorithmen für uns entscheiden:  
Chancen und Risiken der künstlichen Intelligenz» di TA-SWISS



TA-SWISS, Fondazione per la valutazione delle scelte tecnologiche e centro di competenza delle Accademie svizzere delle scienze, intende riflettere sulle ripercussioni – opportunità e rischi – dell' uso di nuove tecnologie.

Questa sintesi si basa su uno studio scientifico effettuato per conto di TA-SWISS da un gruppo di progetto interdisciplinare guidato da Markus Christen (Digital Society Initiative, Università di Zurigo), Clemens Mader (Empa – Laboratori federali di prova dei materiali e di ricerca) e Johann Čas (ÖAW – Accademia austriaca delle scienze). Essa presenta i principali risultati e le raccomandazioni dello studio in forma condensata e si rivolge a un pubblico non specializzato.

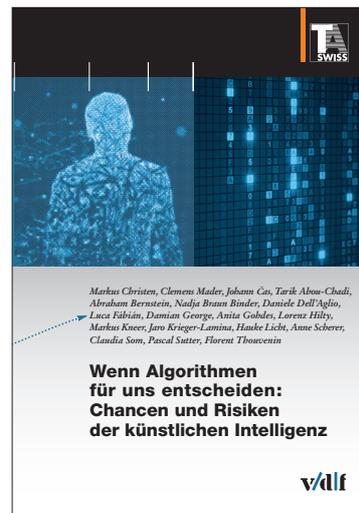
## Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz

Markus Christen, Clemens Mader, Johann Čas, Tarik Abou-Chadi, Abraham Bernstein, Nadja Braun Binder, Daniele Dell'Aglio, Luca Fábrián, Damian George, Anita Gohdes, Lorenz Hilty, Markus Kneer, Jaro Krieger-Lamina, Hauke Licht, Anne Scherer, Claudia Som, Pascal Sutter, Florent Thouvenin

TA-SWISS, Fondazione per la valutazione delle scelte tecnologiche (a cura di).  
vdf Hochschulverlag an der ETH Zürich, 2020.  
ISBN 978-3-7281-4001-2

Lo studio può essere scaricato gratuitamente:  
[www.vdf.ch](http://www.vdf.ch)

E' disponibile in rete anche questa sintesi:  
[www.ta-swiss.ch](http://www.ta-swiss.ch)



<b>L'intelligenza artificiale in estrema sintesi</b>	4
Le opportunità ...	4
... i rischi...	4
... e alcune raccomandazioni	5
<b>L'intelligenza artificiale è figlia della digitalizzazione</b>	5
Dal problema matematico al modello complesso	5
Superare l'umano	6
Distorsioni, problemi di fairness e mancanza di fiducia	7
Svariati campi di applicazione	8
<b>IA e consumo: lo shopping personalizzato</b>	8
Preferenze personali come punti di riferimento	8
Uomo o macchina?	9
Trappole digitali	9
Tutelare i dati tutela anche le persone	9
Difficile equilibrio tra personalizzazione e protezione dei dati	10
<b>IA e comunicazione pubblica: manipolazione delle informazioni nella cassa di risonanza mediatica</b>	10
Sempre più nella bolla	10
Teorie complottistiche fabbricate dai troll	11
Democrazia in pericolo	11
Sguardo scettico sui social network	11
<b>L'IA nel mondo del lavoro: quando le macchine svolgono lavori di concetto</b>	12
Carriera sotto l'egida dell'IA	12
Lavoro a ritmo di macchina	13
Provvedimenti contro la polarizzazione del mercato del lavoro	13
<b>L'IA nell'istruzione e nella ricerca: un sostegno all'intelligenza umana</b>	14
Un super insegnante per ogni alunno	14
Sulle tracce delle eccellenze del futuro	14
Crescente influenza delle aziende private sull'istruzione	15
Intelligenza artificiale: oggetto e motore della ricerca	15
Sviluppare competenze chiave e garantire trasparenza	15
<b>L'IA nell'amministrazione: poteri sovrani all'intelligenza artificiale?</b>	16
In azione contro la frode e la criminalità	16
Diritti fondamentali vs arbitrarietà dello Stato	16
La Svizzera non si sbilancia	17
Nessuna delega del potere decisionale alla macchina	17
Macchine «intelligenti» – una sfida per la giurisprudenza	18
Mettere la museruola ai sistemi di IA?	18
Creatività di prima e seconda mano	18
«Fairness», concetto difficile da quantificare	19
Linee guida etiche a livello europeo	19
Distorsioni e mancanza di trasparenza pesano	19
<b>Raccomandazioni: strutturare ad ampio raggio</b>	20
Esaminare le applicazioni nello specifico	20
Aspettative elevate nei confronti delle autorità	20
Trasparenza e informazioni «digeribili»	20
Sigillo di qualità anziché generiche autorizzazioni di immissione in commercio	20
Non concentrarsi esclusivamente sulle questioni tecniche	20

# L'intelligenza artificiale in estrema sintesi

**A differenza dei programmi informatici convenzionali che risolvono compiti ben precisi attenendosi a passaggi predefiniti, l'intelligenza artificiale (IA) agisce in modo creativo. Grazie all'addestramento con grandi volumi di dati, molti sistemi di IA sono in grado di individuare regolarità occulte e di suggerire soluzioni a cui gli utenti da soli non arriverebbero. I sistemi di IA avanzati combinano diverse competenze, come l'elaborazione del linguaggio, l'organizzazione delle informazioni e l'apprendimento automatico, giungendo persino a svolgere attività impegnative attualmente delegate a professionisti qualificati.**

## Le opportunità ...

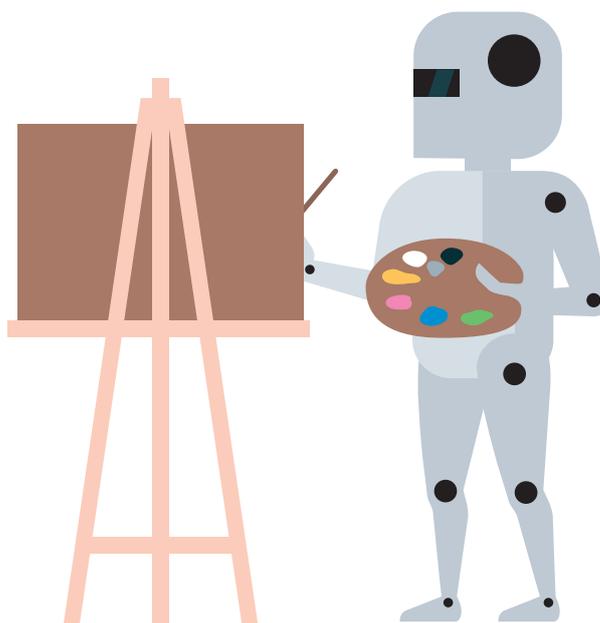
I sistemi di IA trovano impiego nei campi applicativi più vari. Sono in grado di svolgere molte operazioni più velocemente degli esseri umani, e spesso con maggiore precisione; nell'ambito di compiti complessi potrebbero aiutarli a sbrigare numerose attività con un livello di efficienza molto superiore rispetto a ora.

Spesso i sistemi di IA consentono un migliore adattamento dei servizi e delle offerte alle esigenze e alle capacità dei singoli. Uno dei principali punti di forza dell'IA consiste proprio nella personalizzazione delle offerte.

## ... i rischi...

Molti sistemi di IA devono essere addestrati con enormi volumi di dati per riuscire a ottenere le capacità richieste: quest'immensa esigenza di informazioni rischia di minare i concetti di privacy e protezione dei dati.

Se le serie di dati con cui opera l'IA contengono errori, anche gli esiti risultano compromessi. A ciò si somma il problema dei dati non equilibrati, che generano risultati matematicamente corretti, ma viziati sul piano del contenuto: in determinate circostanze possono penalizzare sistematicamente determinati gruppi di persone.



I sistemi di IA autoapprendenti sono soggetti a un'evoluzione costante e in alcuni casi arrivano a risultati incomprensibili sia a chi li ha progettati, sia soprattutto agli interessati.

### ... e alcune raccomandazioni

L'utilizzo di sistemi di IA dovrebbe essere reso facilmente e chiaramente riconoscibile, in modo che gli interessati sappiano appunto di interagire con un sistema di IA e non con un essere umano.

Importanti decisioni che riguardano le persone non vanno affidate a un sistema di IA senza che ne vengano valutati approfonditamente pregi e difetti. Se si tratta di questioni importanti che coinvolgono degli individui, di regola il risultato del sistema deve essere verificato e motivato da un essere umano.

L'uso dell'IA ha conseguenze che vanno ben oltre le implicazioni tecniche. Chiunque sviluppi o utilizzi sistemi di IA o lavori con i suoi risultati dovrebbe quindi disporre anche di nozioni etiche e giuridiche nonché essere disposto a impegnarsi, con le relative competenze, nella cooperazione interdisciplinare con rappresentanti di altri settori scientifici.

Sul fronte metodologico lo studio ha previsto una prima fase basata sull'analisi ad ampio spettro della letteratura specialistica, che ha fornito il materiale per la seconda fase di sondaggio in cui vari esperti si sono espressi a più riprese sui campi problematici e le tesi individuate. I risultati dell'analisi della letteratura e dei sondaggi sono confluiti in raccomandazioni rivolte a vari destinatari.

## L'intelligenza artificiale è figlia della digitalizzazione

**L'intelligenza artificiale (IA) ha cessato da tempo di essere fantascienza per inserirsi gradualmente nella nostra vita quotidiana – ragione sufficiente non solo per apprezzarne i vantaggi, come l'aumento dell'efficienza, ma anche per metterne in luce i lati oscuri, e in particolare la sua «voracità» di dati.**

Il termine «intelligenza» designa, stando all'etimologia, la libertà mentale degli esseri viventi di non seguire passivamente un percorso predeterminato, ma di elaborare alternative e scegliere quale seguire. Che cosa significa però esattamente IA, acronimo dell'espressione «intelligenza artificiale»?

L'IA cominciò ad attirare l'attenzione del mondo scientifico nel 1950, quando l'informatico britannico Alan Turing dichiarò in un articolo apparso sulla rivista «Mind» che, se una macchina era in grado di imitare una conversazione con un essere umano a un tale livello di perfezione da fargli credere di interagire con una persona reale, la si poteva definire «intelligente». Il cosiddetto test di Turing – che viene eseguito a video con una tastiera e senza contatto visivo o uditivo con l'interlocutore – fu promosso a

cartina al tornasole dell'intelligenza dei sistemi fin dalla prima conferenza scientifica sull'IA tenutasi a Dartmouth (USA) nel 1956, nel corso della quale fu coniato anche il termine «intelligenza artificiale».

### Dal problema matematico al modello complesso

Da allora l'IA si è evoluta in un campo di ricerca informatica molto più complesso. Mentre inizialmente ci si era concentrati sugli algoritmi per la risoluzione di problemi matematici e geometrici, a partire dagli anni '90 la crescente potenza di calcolo dei computer ha reso possibile affrontare processi più complicati. Per poter estrapolare modelli significativi da segnali acustici, immagini o altri dati, infatti, molti sistemi di IA richiedono di essere addestrati con un'enorme quantità di dati. La diffusione della digitalizzazione e la crescente disponibilità di grandi volumi di informazioni costituiscono una base importante per l'IA.

Oggi l'IA interessa vari campi di indagine scientifica che provano a utilizzare artefatti – ossia programmi informatici – per riprodurre l'azione razionale o

umana. Un sottogruppo di attività si prefigge per esempio di riprodurre il pensiero e l'apprendimento degli esseri umani. Oltre alle procedure che richiedono grandi quantità di dati per individuare regolarità e tracciare modelli con l'aiuto di algoritmi, vengono utilizzate procedure basate su conoscenze specialistiche e sequenze logiche. I sistemi di IA avanzati combinano varie competenze, tra cui l'elaborazione del linguaggio, l'organizzazione delle informazioni e l'apprendimento automatico.

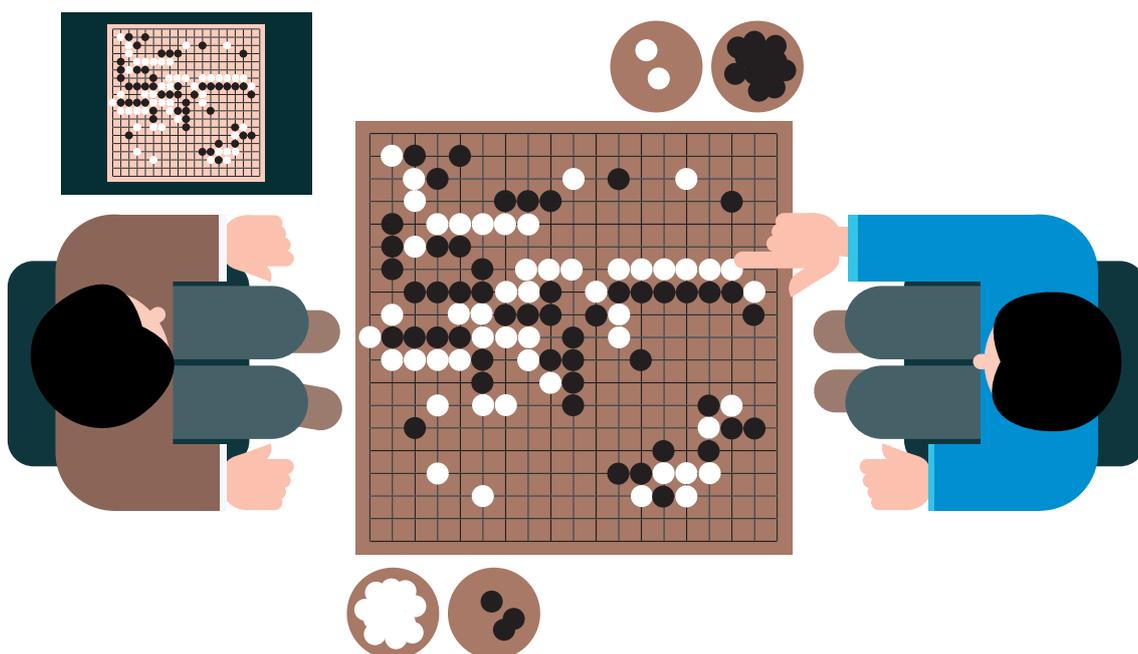
Negli ultimi anni l'IA ha molto contribuito ad accrescere il comfort e la sicurezza della nostra vita quotidiana: si basano su sistemi di IA, per esempio, le formule di mobilità che collegano il trasporto pubblico con le auto a noleggio e le vetture private; le compagnie ferroviarie accrescono la sicurezza attraverso il riconoscimento intelligente dei modelli di danno e l'industria utilizza l'IA per processi più efficienti. La ricerca fondamentale sulla natura del cosmo, così come la conduce il CERN, sarebbe impossibile senza l'analisi di enormi quantità di dati supportata dall'IA, e altrettanto irrealizzabili sarebbero le simulazioni del cambiamento climatico. In genere se i sistemi di IA utilizzano simili dati oggettivi non emergono problemi e, anzi, la società ne ricava spesso notevoli vantaggi. La situazione si complica però quando

l'IA lavora con i dati personali; dunque lo studio di TA-SWISS si concentra particolarmente su questo genere di applicazioni.

## Superare l'umano

I sistemi di IA basati sull'apprendimento automatico e sulle reti neurali sono in grado di risolvere particolarmente bene problemi in cui i programmi informatici convenzionali falliscono. In compiti complessi e sfaccettati come la traduzione, i giochi, il riconoscimento e l'analisi delle immagini oggi l'IA è in grado non solo di tenere testa agli esseri umani, ma a volte anche di superarli.

Una pietra miliare in questo senso è stata posta nel marzo 2016, quando un'IA ha sconfitto il campione del mondo in carica nel complesso gioco di strategia asiatico «Go». Ma non è tutto: in seguito la software house DeepMind, consociata di Google, ha continuato a lavorare sul programma e ne ha sviluppato una versione in grado di imparare il gioco da sola, basandosi esclusivamente sulle regole e non più sull'analisi di un'enorme quantità di mosse ricavate da partite tra campioni umani, come invece accadeva nelle versioni precedenti. Cominciando con partite giocate in modo



casuale al livello base, il software denominato «Alpha Go Zero» ha raggiunto il livello professionale in tre giorni e ha poi superato la versione precedente che aveva battuto il numero uno al mondo.

Lo sviluppo rapidissimo sta rendendo i sistemi di IA sempre più capaci di svolgere compiti un tempo riservati all'uomo: è naturale dunque che diventino la tecnologia di base di innumerevoli applicazioni. L'impiego di nuove forme di IA modifica il ruolo che l'uomo svolge nella definizione e nella risoluzione dei problemi: finora spettava a lui individuare criticità e priorità, redigere istruzioni specifiche di intervento per risolvere i problemi (e apportare le necessarie correzioni dopo i fallimenti iniziali) e infine implementare le istruzioni in campi applicativi relativamente circoscritti; se la valutazione dimostrava che le misure adottate non avevano conseguito un esito positivo nella risoluzione del problema, era sempre l'essere umano a disporre le necessarie rettifiche. L'apprendimento automatico, invece, può sostanzialmente fare a meno dell'intervento umano: grandi volumi di dati forniscono una struttura preliminare al problema, mentre le regole di intervento per la sua soluzione – il cosiddetto algoritmo – vengono sviluppate sulla base dei modelli estratti dai dati e il parallelo apprendimento di rinforzo della macchina. Quest'ultima esegue in modo indipendente anche la retroregolazione, adattando e raffinando in autonomia l'algoritmo – se occorre – sulla base degli ulteriori dati acquisiti; l'essere umano è escluso da gran parte del processo, sebbene spetti a lui respingere eventualmente il risultato del sistema di IA.

Il tutto, per quanto possa sembrare molto astratto, produce conseguenze tangibili nella vita di tutti i giorni, per esempio nella concessione dei prestiti da parte delle banche: la procedura tradizionale prevede che un professionista qualificato esamini una richiesta di finanziamento sulla base della propria esperienza e di svariati criteri; il software intelligente invece esamina i dati degli altri clienti della banca, analizza i loro dati personali e la loro situazione finanziaria e sa chi di loro rimborsa i prestiti puntualmente. Da questi dati crea poi un modello previsionale della probabilità che un nuovo cliente rimborsi il prestito.

I sistemi di apprendimento automatico che migliorano costantemente attraverso l'esperienza sollevano un problema di fondo: a chi spetta la responsabilità di una decisione, se la concatenazione dei processi che l'hanno generata non è chiara perché nessuno è più in grado di ricostruirla? Anche se in

definitiva è una persona a valutare i risultati del software e a prendere la decisione finale, se lo strumento che utilizza è sempre meno comprensibile la situazione si complica.

## **Distorsioni, problemi di fairness e mancanza di fiducia**

Quando l'IA elabora strategie per risolvere problemi, emergono una serie di difficoltà di fondo.

Se per alimentare l'IA si utilizzano grandi volumi di dati che nascondono errori o imperfezioni, il lavoro di elaborazione è di per sé viziato all'origine. Eventuali elementi di parzialità possono influenzare il comportamento dell'algoritmo, se non persino fuorviarlo di proposito nel caso in cui i dati siano stati manipolati. Il possibile squilibrio di grandi masse di dati – raccolti per esempio in base alle query nei motori di ricerca – non è necessariamente deliberato. In determinate circostanze potrebbe essere l'esito di uno sviluppo storico, riflettendo giudizi di valore e abitudini dati acriticamente per acquisiti. In questi casi, se utilizzata con la dovuta accuratezza, l'IA può anche contribuire a individuare le distorsioni.

Un altro problema è costituito dalla crescente indipendenza dell'IA. Se è vero che il software parte da uno stato iniziale noto al programmatore, è altrettanto vero che – soprattutto nelle cosiddette reti neurali artificiali – nel corso dei numerosi cicli di allenamento l'intensità e la ponderazione delle singole connessioni cambiano al punto che ci si ritrova con soluzioni dai fondamenti logici e fisici imperscrutabili. All'atto pratico questi impenetrabili algoritmi di IA, che agiscono come una vera e propria scatola nera, sono di scarsa utilità se si desidera capire in che modo un sistema giunge a determinate conclusioni.

Non solo i dati, ma anche gli stessi algoritmi possono presentare distorsioni, dal momento che i modelli vengono definiti da sviluppatori che in certi casi privilegiano determinati interessi o ideologie a scapito di altri. Algoritmi «ingiusti» hanno già prodotto risultati non corretti, ad esempio stabilendo che persone provenienti da aree residenziali più povere avessero una tendenza più spiccata a commettere reati.

Studi empirici hanno dimostrato che molte persone ripongono più fiducia nelle decisioni umane, anche se sanno che in un determinato campo è comprovato che quelle dell'IA sono più obiettive e fondate. Esistono però anche studi che rivelano il contrario,

ossia che a volte la gente si fida all'eccesso degli esiti di processi decisionali automatizzati. Risulta dunque piuttosto complesso valutare correttamente il livello di fiducia che merita l'IA.

Sta inoltre emergendo l'idea che l'IA possa contribuire a consolidare la supremazia economica. Dato che molte nuove forme di IA si basano su enormi volumi di dati, le aziende che godono dell'accesso a queste informazioni ottengono un vantaggio sulla concorrenza. Ecco perché le maggiori aziende tecnologiche statunitensi e cinesi con accesso ai dati dei clienti collaborano alla ricerca nello sviluppo dell'IA. Rischia dunque di rafforzarsi ulteriormente il noto fenomeno dell'economia online, secondo cui aziende già ben posizionate vanno rafforzandosi rapidamente creando oligopoli.

## Svariati campi di applicazione

L'IA viene già utilizzata in settori disparati. Uno dei punti di forza dei sistemi di IA risiede nella possibilità di personalizzare i servizi e le offerte per adattarli al singolo consumatore.

L'IA viene utilizzata con frequenza sempre maggiore per lo shopping, dai media (nuovi e tradizionali), nel mondo del lavoro, nell'istruzione e nella ricerca, nella pubblica amministrazione. Lo studio TA-SWISS intende approfondire l'analisi di questi cinque campi applicativi. L'ordine adottato nella presente sintesi si discosta da quello scelto per il rapporto completo. Riflette l'esperienza quotidiana della maggior parte dell'utenza che effettua acquisti online e legge le notizie a video più spesso di quanto non entri in contatto con le autorità pubbliche. Tuttavia la scelta non riflette in alcun modo l'importanza dei relativi ambiti d'indagine.

# IA e consumo: lo shopping personalizzato

**Chi ordina prodotti su Internet entra inevitabilmente in contatto con l'IA. Ormai l'esperienza di acquisto personalizzata è diventata la regola, i processi informatici lavorano con discrezione in background e non vengono quasi mai messi in discussione.**

«Ti potrebbe interessare anche», «ordinali insieme!» e «completa l'outfit» sono proposte che gli acquirenti sono abituati a ricevere dai negozi online: segnalano ulteriori prodotti che i clienti potrebbero trovare utili o interessanti. Questi sistemi di raccomandazione utilizzano le «tracce digitali» lasciate dai consumatori: traggono conclusioni sulle preferenze di determinati target basandosi sull'analisi supportata dall'IA degli acquisti effettuati sulla piattaforma in questione e dei dati personali, oltre che di comportamenti come il tempo trascorso sulla pagina o i movimenti del mouse.

Anche la pubblicità è sempre più personalizzata. Operano in questo campo Google, marketplace di prima grandezza come Alibaba e Amazon e social network del calibro di Facebook. Pubblicizzano in modo mirato una vasta gamma di prodotti e servizi, spesso per conto di altre aziende. La tecnologia che

lavora in background è simile a quella dei sistemi di raccomandazione.

## Preferenze personali come punti di riferimento

I sistemi di raccomandazione supportati dall'IA semplificano ai consumatori l'orientamento nell'universo del consumo. Invece di ricevere pubblicità di ogni genere, gli utenti si vedono apparire sullo schermo solo annunci di prodotti appositamente scelti in base ai loro interessi.

Il meccanismo che da un lato aiuta la clientela a orientarsi nel vasto mondo dei prodotti commerciali, permette al contempo ai venditori di ridurre le spese. Grazie all'uso dell'IA è possibile contenere anche del 20% i costi amministrativi; inoltre, se concentrata sul giusto gruppo di destinatari, la pubblicità subisce una minore dispersione. Con l'approccio sempre più personalizzato anche l'assistenza clienti sta raggiungendo una nuova dimensione. A tutto ciò va aggiunta la possibilità di massimizzare i profitti, per esempio attraverso il pricing dinamico: l'IA consente infatti ai fornitori di quantificare il DAP (dispo-

nibilità a pagare) della clientela e di utilizzare il dato per calcolare in automatico il prezzo della merce.

## Uomo o macchina?

L'IA è utilizzata sempre più spesso per acquistare servizi, oltre che prodotti. Per esempio nel maggio 2018 Google ha presentato l'assistente personale digitale Duplex, che fissa appuntamenti dal parrucchiere o prenota un tavolo al ristorante per conto del/della suo/a assistito/a con una voce femminile dal suono estremamente realistico. Dato che l'assistente ha suscitato reazioni inquiete persino negli Stati Uniti, tradizionalmente molto aperti alle nuove tecnologie, Google ha impostato Duplex in modo che si presenti all'inizio di ogni conversazione. Nel suo blog il gruppo espone altre possibili applicazioni della tecnologia, per esempio nei servizi di informazione alla clientela da parte di aziende e uffici pubblici: per ovviare alle lunghe code in attesa di un operatore disponibile, il compito di rispondere alle domande più frequenti potrebbe venire affidato sempre più spesso ad assistenti digitali.

Duplex evidenzia un problema di fondo dell'IA: è sempre più raro che chi ordina prodotti e servizi in Internet o al telefono sappia se a rispondergli è una persona in carne e ossa o se sta interagendo con un algoritmo. Inoltre la clientela trova sempre più difficoltà a capire non solo se e quando vengono utilizzati sistemi di IA, ma anche in che modo e a quale scopo. Persino se gli utenti vengono informati della registrazione delle conversazioni e dei dati personali, non sono in grado di sapere quali informazioni potranno ricavarne gli algoritmi e con quanta sicurezza di giudizio.

## Trappole digitali

La praticità di memorizzare in un account i dati personali e tutte le informazioni sugli ordini effettuati è incontestabile: il portale personalizzato con le immagini degli ultimi acquisti eseguiti trasmette un senso di familiarità e, dato che l'indirizzo e il numero della carta di credito sono già stati archiviati, non c'è neppure il fastidio di doverli inserire di nuovo per ordinare ancora. Con l'andar del tempo, ossia di transazione in transazione, il sistema impara a conoscere sempre meglio la clientela per adeguare con più precisione le raccomandazioni ai gusti dei singoli. Ma c'è anche il rovescio della medaglia: la comodità seduce, spingendo a rimanere fedeli a un venditore anche se magari le offerte di un altro sono

più vantaggiose. Negli ambienti professionali questo effetto è noto come stickiness della piattaforma (letteralmente: «appiccicosità»).

A ciò si aggiunge che già le pure dimensioni di una determinata rete esercitano un certo potere sugli individui, perché è utile far parte di una realtà il più ampia possibile, in cui raggiungere molte persone. Un altro indubbio vantaggio è l'opportunità di svolgere più attività da un'unica piattaforma, in modo da evitare login separati per ogni transazione o interazione. La conseguenza è che alla fine i dati si concentrano nelle mani di pochi grandi gruppi con un notevole potere di mercato.

## Tutelare i dati tutela anche le persone

Le nostre tracce su Internet rivelano molto di noi. Il social network Facebook ha pubblicato uno studio condotto su più di 8000 volontari in cui, basandosi sui «mi piace» ricevuti, un sistema di IA ha valutato le caratteristiche personali degli utenti con una precisione superiore a quella dei loro stessi amici. Alla macchina sono bastati dieci «mi piace» per fornire risultati migliori rispetto ai colleghi umani; con l'analisi di settanta o più «mi piace» la valutazione automatica ha superato quella della cerchia degli amici. Altri studi hanno dimostrato che i sistemi di IA sono in grado di dedurre con precisione sorprendente le caratteristiche individuali di una persona dall'immagine del suo profilo, orientamento sessuale compreso.

L'IA dispiega appieno le sue potenzialità soprattutto se combina dati personali provenienti da fonti diverse. Un vantaggio che aziende come Facebook e Google hanno imparato a sfruttare, essendo arrivate a possedere molti servizi estremamente popolari: Facebook è proprietario del social network di fotografia Instagram e del messenger WhatsApp, mentre Google è proprietario – insieme a molte altre aziende – del portale video YouTube e della società di fitness tracker Fitbit. Anche altri giganti della tecnologia come Amazon e Alibaba hanno acquistato una moltitudine di piattaforme web e set di dati per poter correlare tra loro informazioni di varia origine ed espandere così il proprio vantaggio competitivo rispetto ad aziende più piccole. Non tutti gli esperti, però, ravvisano un pericolo nella concentrazione dei set di dati poiché, dal momento che gli utenti possono metterli a disposizione di più fornitori diversi, il semplice accesso a determinate informazioni non garantirebbe di per sé un vantaggio sui concorrenti.

Rendere la clientela unica responsabile della protezione dei propri dati non è comunque una soluzione accettabile. Sebbene le Condizioni generali di contratto (CGC) debbano specificare per legge quali informazioni vengono raccolte e archiviate, il semplice accettare le CGC non implica necessariamente che la persona stia acconsentendo all'accesso ai propri dati, dal momento che potrebbe non essere consapevole dell'estensione delle informazioni raccolte e dell'utilizzo che si intende farne. L'utente medio impiegherebbe diverse settimane all'anno solo a leggere per intero i termini e le condizioni dei servizi più diffusi.

### Difficile equilibrio tra personalizzazione e protezione dei dati

Secondo il gruppo di esperti intervistati nello studio di TA, è cruciale che l'utenza acquisisca una maggiore consapevolezza del valore dei propri dati. Per le aziende è comunque difficile riuscire a conciliare le esigenze di personalizzazione con quelle di prote-

zione dei dati. Quanto meno questa è la valutazione del panel di esperti sottoposti al sondaggio. Chi opera sul web si trova inoltre ad affrontare problematiche legate al rischio di perdita della fiducia. Gli esperti intervistati ritengono che i dati aperti (open data) con possibilità di controllo da parte della clientela, la comunicazione trasparente dell'uso dell'IA e il diritto di cancellazione dei dati siano misure promettenti per promuovere la fiducia del pubblico.

Tra gli altri rischi significativi potenzialmente associati all'uso dei sistemi di IA, gli intervistati citano la formazione di oligopoli in cui il mercato viene dominato da pochi giganti. Questo perché le strutture di questo tipo possono determinare in ultima analisi una contrazione dell'offerta e una limitazione delle possibilità di cambiare da parte della clientela. È quindi importante garantire che i singoli possano portare con sé i propri dati sensibili quando cambiano fornitore. Infine è fondamentale osservare da vicino le derive del mercato e disporre di opzioni di controllo e intervento che garantiscano il principio di concorrenza.

## IA e comunicazione pubblica: manipolazione delle informazioni nella cassa di risonanza mediatica

**Stampa, radio e televisione hanno perso terreno negli ultimi dieci anni: sempre più persone si informano su social media come Facebook e Twitter. Le notizie che derivano da un lavoro di ricerca professionale e si inseriscono in un contesto più ampio devono riuscire a farsi spazio tra articoli di blog e post sui social network.**

Nel 2017 l'83 per cento della popolazione svizzera ha letto le notizie online. Il 45 per cento ha utilizzato i social media: oltre a Facebook e Twitter, sempre più anche piattaforme come Google News e Apple News che raccolgono contenuti provenienti da altre fonti. La conseguenza è che sempre più utenti condividono articoli e si informano direttamente tra loro, spesso passandosi anche notizie di testate tradizionali. In questo processo l'origine delle notizie che si diffondono a macchia d'olio perde progressivamente trasparenza.

### Sempre più nella bolla

Dai media tradizionali ci si aspetta che forniscano informazioni equilibrate su un ampio ventaglio di argomenti, offrendo una panoramica delle opinioni più diffuse. Chi si informa attraverso la stampa, la radio e la televisione viene inevitabilmente a conoscenza anche di questioni che gli interessano poco.

Le cose vanno diversamente se le notizie vengono condivise sui social network. Quando sul profilo di una persona arrivano molti feed di notizie, è l'algoritmo a decidere quali contenuti mostrare per primi, precisamente in base alle informazioni raccolte sul proprietario del profilo (comunicate volontariamente o in modo indiretto con il comportamento online). I sistemi di IA tendono a suggerire agli utenti informazioni che rispecchiano le loro preferenze di contenuto e dunque gli utenti si orientano tra le

notizie all'interno della loro bolla di filtraggio. Se le persone leggono solo gli articoli suggeriti dalla loro cerchia di amici, aumenta il pericolo di ritrovarsi in una cassa di risonanza che contiene esclusivamente argomenti in linea con i propri interessi e la propria opinione, mentre i punti di vista difforni e le informazioni in controtendenza nemmeno compaiono.

Sebbene le bolle di filtraggio e le camere di risonanza digitali stiano attirando una notevole attenzione pubblica e se ne parli come di una minaccia alla democrazia, non esistono ancora prove scientifiche concrete che la diversità di opinione venga limitata in misura maggiore online rispetto a quanto avviene nei circoli d'opinione tradizionali formati da persone affini. Gli sviluppi appena descritti sono ancora in fase embrionale, ma potrebbero senz'altro diventare un problema per la formazione dell'opinione pubblica nelle società democratiche.

## Teorie complottistiche fabbricate dai troll

Le fake news non sono un fenomeno nuovo: compaiono nell'immaginario collettivo sin dall'inizio dell'Ottocento e sono rappresentate da animali-simbolo nelle varie lingue (in italiano: la bufala). In genere le fake news si diffondono involontariamente e danneggiano la reputazione del giornale che le pubblica. La vera novità è piuttosto la deliberata disinformazione di massa perpetrata attraverso i social media.

Oggi le fake news vengono diffuse in modo perlopiù automatico ed efficientissimo con i social bot, semplici programmi appositamente compilati per diramare notizie in modo mirato sui social media. Nella battaglia elettorale tra Hillary Clinton e Donald Trump negli Stati Uniti (2016) circa un quinto dei tweet inviati proveniva da social bot; la grande maggioranza (cioè fino a tre quarti) degli account di Twitter automatizzati ha votato per Trump.

In futuro l'IA potrebbe essere in grado di imitare i cosiddetti troll, ossia utenti che postano su blog e forum commenti mirati a far leva sulla sfera emotiva, spesso infarciti di informazioni false, per cercare di provocare risposte violente. Con questa strategia i troll, non di rado impiegati in vere e proprie «fabbriche», puntano a incrementare la notorietà di determinati gruppi o addirittura a diffondere teorie cospirative. Generalmente sia i social bot che i troll si appoggiano a siti web specificamente progettati per diffondere la disinformazione e quindi per aumentare la credibilità dei loro «complici» digitali.

Già oggi i sistemi di IA aumentano l'efficienza con cui i troll diffondono i loro messaggi.

## Democrazia in pericolo

Tra le strategie adottate dai propagatori di fake news c'è l'infiltrazione nei canali di informazione consolidati. Negli Stati Uniti è emerso che le fake news vengono diffuse inizialmente sui media locali, in quanto meno capaci di verificare i fatti. Una volta recepite dai notiziari locali, passano in fretta ai media nazionali. Ma l'IA non si limita a diffondere informazioni false, contribuisce anche a renderle sempre più sofisticate. Per esempio si può utilizzare per manipolare immagini, file audio e video, oppure per creare filmati completamente falsi al fine di screditare o ricattare delle persone. Gli esperti chiamano i video falsi creati con espedienti tecnici «deep fakes».

Sotto regimi autoritari che soffocano le notizie indesiderate e controllano i mezzi di comunicazione, i social media rappresentano a volte l'extrema ratio per l'informazione indipendente. Nelle democrazie funzionanti, invece, l'influenza delle fake news sull'opinione pubblica nei social media è considerata una seria minaccia al sistema politico. Da un lato, il pubblico sta perdendo fiducia nei canali d'informazione consolidati: secondo gli ultimi sondaggi, in quasi tutti i paesi il numero delle persone intervistate che si fida ancora della capacità dei media tradizionali di distinguere tra realtà e invenzione si è più che dimezzato. Dall'altro, varie analisi di rilievo hanno dimostrato che le fake news si diffondono addirittura più facilmente delle informazioni corrette. Per quanto riguarda la Germania, ad esempio, si è riusciti a dimostrare che sette delle dieci notizie più diffuse su Angela Merkel sono da considerare fake news.

## Sguardo scettico sui social network

I sondaggi condotti da TA-SWISS nell'ambito di questo studio hanno rivelato un atteggiamento scettico nei confronti dei social media: la maggioranza degli intervistati teme che danneggino la qualità dei contenuti giornalistici. La semplice diffusione online delle notizie è valutata in modo un po' meno negativo. La personalizzazione dei contenuti trova sostenitori anche tra gli esperti; la discriminante sembra essere la possibilità che gli utenti decidano se, come e in che misura il flusso di notizie debba essergli rivolto in forma personalizzata.

Viceversa gli intervistati considerano le fake news un grave pericolo. Solo una minoranza ritiene che il progresso tecnologico possa portare a smascherare più facilmente i fenomeni di disinformazione. Al contrario, la maggioranza teme che tecnologie più sofisticate produrranno fake news sempre più raffinate. I risultati dei sondaggi hanno però evidenziato anche che, nella lotta contro la disinformazione,

occorre non trascurare il diritto alla libertà di espressione. Istituire organi deputati alla verifica dei contenuti è valutato in modo piuttosto critico. Per contro gli intervistati accolgono con favore il fatto che i gestori delle piattaforme cancellino le fake news e ne blocchino gli autori. La verifica automatizzata dei fatti – ossia supportata dall'IA – è considerata più efficiente delle procedure manuali.

## L'IA nel mondo del lavoro: quando le macchine svolgono lavori di concetto

**Per molto tempo si è ritenuto che la razionalizzazione del personale potesse colpire solo attività semplici e ripetitive. Da quando l'IA è penetrata in campi professionali complessi che presuppongono una formazione di alto livello, il timore di perdere il posto si è diffuso anche ai piani alti.**

All'alba del nuovo millennio diverse colonne giornalistiche divertivano i lettori con la traduzione automatica di brevi testi. I programmi fallivano sulle espressioni più semplici, fornendo versioni arbitrarie e spesso stravaganti. Fino a poco tempo fa la traduzione era considerata un compito impossibile da gestire con una macchina. Eppure oggi software come DeepL, che lavora con reti neurali artificiali, forniscono traduzioni assolutamente accettabili, mentre altri programmi sono persino in grado di generare da soli semplici testi giornalistici.

Da molto tempo ormai l'automazione non riguarda più solo banali attività di routine: l'IA viene utilizzata in campo medico, dove nella valutazione di traccati di ultrasuoni, risonanze magnetiche e immagini a raggi X fornisce risultati più accurati rispetto alla maggioranza dei medici professionisti, oppure in campo legale, per aiutare gli avvocati a individuare clausole specifiche all'interno di migliaia di pagine di contratti.

Una serie di studi fornisce cifre relative al rischio di perdere il posto di lavoro o, a seconda del punto di vista, alle possibilità di taglio occupazionale dovuti all'IA. La forbice delle stime è ampia: la potenziale perdita di posti di lavoro oscilla tra il 6 e il 57% a seconda dei presupposti di partenza e dei metodi applicati. La comunità degli esperti concorda sul

fatto che, a differenza dei precedenti sconvolgimenti economici dovuti all'innovazione, non sarà facile compensare con nuovi prodotti e servizi il calo di posti di lavoro causato dall'IA.

### Carriera sotto l'egida dell'IA

L'impatto dell'IA non si avverte solo sulla quantità, ma anche sulla qualità del lavoro. I responsabili delle risorse umane, per esempio, ricorrono sempre più spesso a sistemi di IA per preselezionare i dossier di candidatura durante il reclutamento del personale. A tal fine candidate e candidati devono per prima cosa compilare i questionari standardizzati online. In seguito i responsabili delle risorse umane proseguono il lavoro sulla base della selezione operata dall'IA. Questo non va necessariamente a detrimento dei candidati: se l'esame della documentazione venisse affidato a una persona fin dall'inizio, questa potrebbe attribuire peso eccessivo a criteri decisionali non oggettivi, come l'aspetto fisico ricavato dalla fotografia. Il ricorso a sistemi di IA consente invece di individuare pratiche discriminatorie e di intervenire per modificarle.

Le grandi aziende utilizzano sempre più spesso l'IA per elaborare i profili di competenza dei dipendenti e proporre loro corsi di aggiornamento su misura. Il lato spinoso della questione è che ciò limita l'autonomia dell'individuo, in quanto la richiesta di avvalersi di ulteriori strumenti di qualificazione non è spontanea, bensì suggerita da un sistema con l'intento di ottimizzare le risorse umane (e finanziarie) disponibili. Anche se la scelta finale non è affidata al sistema, le sue proposte possono innescare deter-

minate dinamiche interne all'azienda che rendono difficile ai dipendenti giustificare decisioni difformi. In generale l'uso dell'IA nel mondo del lavoro pone problemi sostanziali di protezione dei dati.

## Lavoro a ritmo di macchina

In futuro l'IA non porterà necessariamente al taglio di interi posti di lavoro, ma piuttosto di singole attività. Le lavoratrici e i lavoratori dovranno imparare a collaborare con macchine «intelligenti», gestendo l'aumento del ritmo lavorativo e la compattazione delle attività. Esperti e istituzioni come l'Organizzazione Internazionale del Lavoro (OIL) sottolineano la necessità di mettere le persone al centro dell'organizzazione del lavoro, in modo che l'utilizzo di sistemi di IA non peggiori la loro condizione. Inoltre, soprattutto nel caso di comportamenti inspiegabili dell'IA, è importante che i dipendenti siano in grado di valutare la plausibilità dei risultati per prevenire errori grossolani.

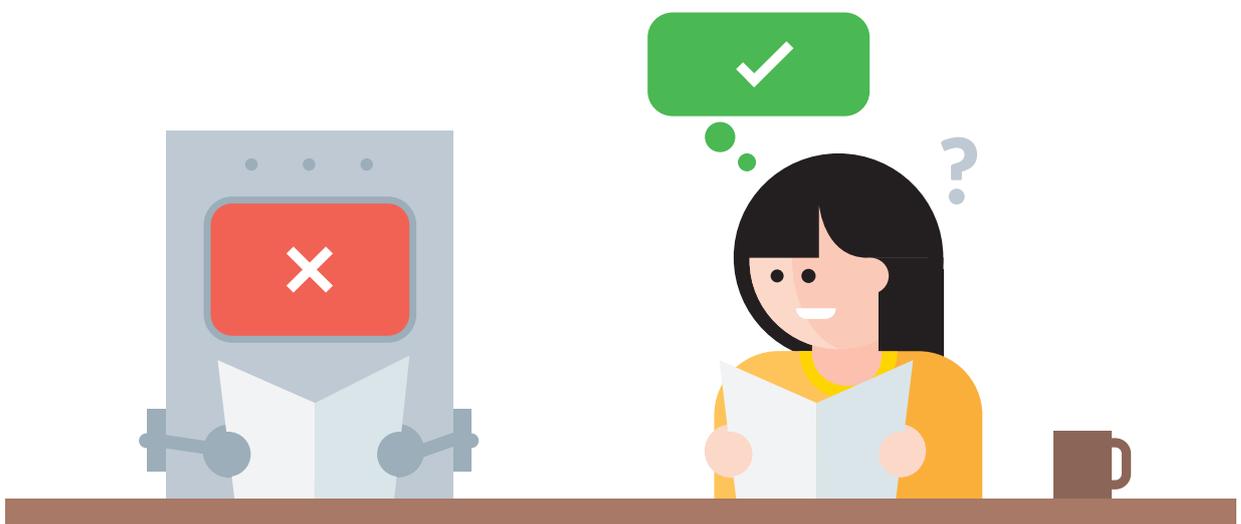
È difficile separare l'influenza dell'IA dagli effetti generali della digitalizzazione sul lavoro. Come ha evidenziato lo studio di TA-SWISS sulla flessibilità del lavoro, la digitalizzazione può contribuire ad aumentare la libertà personale e l'autonomia temporale. Mette però in discussione l'attuale sistema di previdenza sociale. In Internet esistono numerose piattaforme che spezzettano grandi incarichi per

affidarli così suddivisi ai cosiddetti «clickworker». Se ciò da un lato permette agli studenti di guadagnare denaro extra con lavoretti serali, dall'altro mette sotto pressione soprattutto i freelance del settore creativo, che utilizzano queste piattaforme per acquisire commesse sul mercato internazionale e vedono messe a rischio le tutele esistenziali.

## Provvedimenti contro la polarizzazione del mercato del lavoro

Gli esperti intervistati nello studio di TA-SWISS ritengono probabile che l'IA promuoverà la polarizzazione del mercato del lavoro; si aspettano un ulteriore allargamento del divario tra i lavoratori con qualifica elevata e salari consoni e quelli con formazione di basso livello e reddito modesto. Gli esperti prevedono anche un maggiore controllo dei dipendenti e una crescente instabilità dei rapporti di lavoro. Ritengono invece improbabile che l'IA contribuisca a ridurre l'orario e il carico di lavoro. Nel complesso la maggioranza degli esperti ritiene che saranno le grandi aziende a trarre benefici particolarmente elevati da questi cambiamenti del mondo del lavoro.

Gli esperti intervistati hanno individuato nella formazione continua una misura estremamente efficace contro gli effetti della polarizzazione. Quali provvedimenti siano adatti a mitigare i crescenti contrasti nel



mondo lavorativo è una questione tanto controversa quanto quella degli effetti dell'IA. Solo una maggioranza relativamente modesta degli intervistati, infatti, è favorevole ad attenuare le disuguaglianze sociali sviluppando di conseguenza il sistema previdenziale; una maggioranza altrettanto risicata respinge l'idea di ridurre l'orario lavorativo per creare più posti di lavoro da ripartire su un maggior numero di persone, così come l'introduzione di un reddito di base incondizionato è sostenuta da una maggioranza relativa. Gli autori stessi dello studio, invece, sono dell'opinione che la soluzione più sensata nel breve e medio periodo sia distribuire meglio e con più flessibilità il lavoro e il reddito. Occorre inoltre promuovere nuove opportunità occupazionali come le start-up creative e sviluppare forme innovative

di tassazione per continuare a finanziare il bilancio pubblico e il sistema di previdenza sociale.

Gli intervistati ritengono che sia importante anche tutelare i dipendenti da un eccessivo controllo da parte dei datori di lavoro. Le misure contro la sorveglianza sul posto di lavoro e in generale una migliore protezione dei dati sono considerate auspicabili ed efficaci; i dipendenti vanno tutelati anche nei rapporti di lavoro che travalicano i confini nazionali. Vari esperti sottolineano anche il potenziale positivo dell'IA, a condizione che venga utilizzata per automatizzare i lavori monotoni, per aumentare la produttività e la competitività. L'intera comunità potrebbe trarre giovamento dall'equa distribuzione dei vantaggi.

## L'IA nell'istruzione e nella ricerca: un sostegno all'intelligenza umana

**L'IA è un prodotto della ricerca e contribuisce a sua volta al progresso della scienza. A scuola i sistemi di IA possono essere utilizzati per fornire un sostegno individuale a studenti e studentesse.**

Il sostegno individuale dei discenti è già stato inserito nell'agenda della politica dell'istruzione. L'IA può fornire un adeguato supporto tecnico al progetto.

### Un super insegnante per ogni alunno

Numerose start-up hanno già sviluppato strumenti per aiutare singoli studenti a raggiungere i loro obiettivi formativi. Alcune di queste per esempio prevedono il ricorso a chatbot, ossia programmi che rispondono automaticamente alle domande e alle comunicazioni dei discenti, fornendo risposte predefinite sulla base di regole e parole chiave o generate individualmente dal sistema di IA.

Altri programmi guidano passo passo i giovani per mezzo di tutorial che impartiscono loro le nozioni desiderate in materie come matematica, scienze o fisica. Questi sistemi di IA agiscono in base del profilo di competenze e interessi definito inizialmente, utilizzato anche per identificare punti di forza, debolezze e conoscenze pregresse; in seguito analizzano

il percorso didattico svolto e suggeriscono la scelta educativa più indicata per il futuro.

Un altro approccio consiste nel far monitorare al sistema di IA il processo di scrittura dei discenti, fornendo un feedback continuo. Esistono anche sistemi di IA che organizzano lo studio in gruppi riunendo gli studenti con profili compatibili o mettendo in contatto i giovani che non hanno buoni risultati in matematica con tutor di matematica di altre classi, di altre scuole o addirittura di altri paesi.

L'IA è utile anche agli insegnanti, sia per valutare il rendimento dei loro studenti, sia anche per sostenerli individualmente. Grazie all'uso dell'IA, già solo nell'ambito delle mansioni amministrative gli insegnanti possono arrivare a risparmiarsi diverse ore di lavoro alla settimana.

### Sulle tracce delle eccellenze del futuro

Se gli studenti svolgono sempre più spesso i compiti online interfacciandosi con un sistema di IA, sui server del fornitore si riversa una marea di dati personali che rivelano non solo molte informazioni sulle strategie di apprendimento e sulle competenze dei discenti, ma anche sulle loro preferenze e sulle

attività che svolgono nel tempo libero. Queste informazioni risultano interessanti perché permettono di identificare sul nascere potenziali eccellenze. Anche i dati dei ragazzi che non raggiungono il livello di istruzione richiesto sono utili, in quanto potrebbero consentire di identificare i giusti destinatari di corsi di recupero offerti dalla stessa azienda.

Se una scuola svizzera utilizza il software di IA di un fornitore privato, si trova di fronte a una serie di problematiche particolarmente spinose in tema di privacy: quali dati personali possono essere memorizzati, detenuti e consultati? Da chi e per quanto tempo? Come garantire che possano essere cancellati, se necessario? È possibile portare con sé i propri dati quando si cambia scuola? E le nuove scuole sono autorizzate a consultare i dati precedenti, quando ammettono nuovi studenti? Come assicurarsi che in futuro i dati degli esami e dell'apprendimento non influenzino pesantemente le chance di aggiudicarsi un posto di lavoro, nei casi in cui alcuni candidati dispongano di risultati positivi, mentre altri non dispongano di dati oppure ne abbiano solo di negativi? Per annullare i principali rischi legati alla privacy, la Conferenza svizzera dei direttori cantonali della pubblica educazione CDPE ha lanciato la piattaforma Edulog.ch, che consente agli alunni svizzeri di accedere ai programmi didattici in forma anonima. I Cantoni collaborano volontariamente all'iniziativa; resta ora da vedere se all'atto pratico l'offerta verrà utilizzata.

## **Crescente influenza delle aziende private sull'istruzione**

La maggior parte degli esperti intervistati da TA-SWISS nel corso dello studio accoglie con favore la possibilità di trasmettere contenuti didattici su misura per il singolo studente, perché la personalizzazione permette di affrontare i problemi specifici di ciascuno. Gli esperti intervistati, tuttavia, temono che le aziende private esercitino un'influenza sempre maggiore sulle scuole, che di conseguenza potrebbero finire per dipendere da un unico fornitore.

Se da un lato il divieto di utilizzare sistemi di IA privati non è considerato un'opzione percorribile, dall'altro è opportuno negoziare con i fornitori accordi precisi e vincolanti, che includano anche una regolamentazione chiara della privacy. Gli esperti non concordano sul fatto che lo sviluppo dei sistemi di IA scolastici debba essere affidato alle autorità pubbliche, ma in ogni caso ritengono decisamente auspicabile una cooperazione tra pubblica ammini-

strazione e fornitori privati. La maggioranza raccomanda anche una verifica ufficiale delle applicazioni di IA utilizzate dalle scuole.

## **Intelligenza artificiale: oggetto e motore della ricerca**

In cifre assolute, la maggior parte delle pubblicazioni sull'IA appare in Cina e negli Stati Uniti. Ma sono in corso attività di ricerca non trascurabili anche in Svizzera, dove in rapporto al numero di abitanti vengono pubblicati più articoli scientifici sull'IA che in Germania o in Francia. Considerando l'impatto delle citazioni, la Svizzera si trova addirittura in cima alla classifica mondiale.

L'IA viene utilizzata nelle discipline scientifiche più disparate, principalmente se è necessario analizzare enormi quantità di dati; si ricorre per esempio all'IA nella scienza dei materiali, ma anche in biochimica o in astrofisica. È utile anche in ambito accademico, per esempio se si tratta di scoprire imbrogli nei lavori scientifici. Gli strumenti supportati dall'IA rilevano sia plagii (tramite passaggi di testo molto simili ad altri) sia statistiche falsate. L'IA è in grado di rilevare anche cambiamenti nello stile di scrittura, segnali di un potenziale ricorso al ghost writing.

## **Sviluppare competenze chiave e garantire trasparenza**

L'IA sta diventando sempre più importante per la ricerca; spesso è addirittura indispensabile a progressi di una certa importanza. La maggior parte degli esperti intervistati ritiene opportuno che le istituzioni scientifiche allestiscano centri per la formazione continua e l'utilizzo dell'IA nella ricerca, che potrebbero fungere anche da punti di riferimento per il dialogo interdisciplinare su queste tecnologie.

Per quanto riguarda i dati generati dall'IA, la maggioranza degli intervistati concorda sul fatto che debbano rimanere di proprietà del rispettivo istituto di ricerca o che debbano essere messi a disposizione di tutti in formato «open access». Infine i risultati scientifici generati dai sistemi di IA devono essere quanto più trasparenti e riproducibili possibile. Per soddisfare questo requisito, occorrono adeguate linee guida sulla trasparenza.

# L'IA nell'amministrazione: poteri sovrani all'intelligenza artificiale?

**Di regola per svolgere il proprio lavoro le istituzioni si affidano a moduli e procedure standardizzate – presupposti ideali anche per l'uso dell'IA. A fronte di un incremento dell'efficienza amministrativa, emerge il rischio di limitare i diritti fondamentali dei cittadini.**

Nel settembre 2018 un consigliere nazionale svizzero del Partito socialdemocratico ha presentato al Consiglio federale una mozione per l'istituzione di un centro di competenza amministrativo. L'obiettivo era quello di promuovere l'uso dell'IA e dell'apprendimento automatico per incrementare l'efficienza degli organi di amministrazione. Il Consiglio federale ha consigliato di respingere la mozione.

## In azione contro la frode e la criminalità

L'IA si è fatta avanti anche negli uffici governativi. Le autorità fiscali australiane, ad esempio, ricorrono all'IA per riscuotere in automatico i crediti statali. Il sistema tuttavia ha attirato critiche per aver tratto conclusioni parzialmente errate e lasciato ai contribuenti l'onere di difendersi da una tassazione eccessiva. In Germania gli uffici delle imposte utilizzano l'IA per individuare imbrogli nelle dichiarazioni dei redditi.

Sebbene in Svizzera non si discuta (ancora) dell'uso dell'IA in campo fiscale, la digitalizzazione sta facendo progressi anche da noi: i dazi doganali vengono registrati ed elaborati sempre più spesso con mezzi digitali e in futuro anche diverse imposte di consumo – ad esempio sull'alcol, sull'olio minerale e sul tabacco – potrebbero venire rimosse con metodi automatizzati. I fondamenti giuridici in materia vengono forniti nell'allegato al progetto di revisione della Legge sulla protezione dei dati (LPD) varato nel settembre 2015.

La polizia ricorre all'IA per stimare in quale quartiere è più probabile che si verifichi il prossimo crimine. Grazie a queste previsioni, le forze dell'ordine dovrebbero poter pattugliare con maggiore frequenza i luoghi soggetti a maggiore probabilità di

effrazione o rapina. Il lavoro della polizia predittiva viene utilizzato soprattutto in alcune grandi città degli Stati Uniti. Esiste poi un altro approccio che non si concentra tanto sugli spazi, quanto piuttosto sulle persone e sulla loro rete di relazioni: il sistema di IA utilizza i dati dei contatti sociali all'interno di determinate cerchie per stimare la probabilità che una persona faccia parte di una banda criminale.

L'IA aiuta a valutare anche il rischio di recidiva dei delinquenti. Negli USA viene utilizzato un sistema di IA che sfrutta un set di 137 caratteristiche per calcolare la probabilità che un trasgressore commetta un nuovo reato. Il sistema è stato criticato perché ha sistematicamente attribuito rischi di recidiva maggiore ai delinquenti di colore rispetto ai bianchi. È senz'altro vero che anche in Svizzera si ricorre a software per la valutazione criminologica, in particolare per decidere se sia possibile mitigare la pena di un determinato soggetto. Tuttavia questa verifica non si basa su tecnologie di IA avanzate.

## Diritti fondamentali vs arbitrarietà dello Stato

Nell'adempiere le proprie funzioni, le istituzioni pubbliche trattano una grande quantità di dati personali. In quest'ambito ricorrere a sistemi di IA ripromette un notevole aumento dell'efficienza: in futuro documenti e atti potrebbero essere creati automaticamente, mentre a rispondere alle richieste di routine potrebbero essere assistenti digitali con funzione vocale.

D'altra parte però le autorità devono essere particolarmente accurate in tutte le proprie azioni, senza subordinarle soltanto a criteri di efficienza. Per assicurare al singolo i diritti fondamentali di fronte allo Stato, la legge prevede che l'individuo debba essere tutelato da qualsiasi atto lesivo, appunto, di questi diritti.

La protezione dei dati personali è un diritto fondamentale. I dati devono essere conservati in modo sicuro e tutelati dagli abusi. Inoltre tutti gli individui

devono poter fare affidamento sull'accuratezza, la completezza e l'equilibrio dei dati personali raccolti ed elaborati. La verità, tuttavia, è che molte persone non sono del tutto consapevoli di quali dati siano stati effettivamente memorizzati e vengano utilizzati nell'ambito del trattamento supportato dall'IA; di conseguenza non sono concretamente in grado di rintracciare e correggere eventuali errori nelle informazioni che li riguardano.

Quanto alle distorsioni e agli squilibri che possono generarsi nel trattamento dei dati supportato dall'IA, occorre prestare particolare attenzione alle procedure automatizzate, perché potrebbero discriminare determinati gruppi di persone e individui. La legge sulla protezione dei dati revisionata stabilisce che, se le decisioni vengono prese in modo completamente automatico, le persone interessate devono esserne informate.

In definitiva affidare le decisioni a un sistema di IA completamente automatizzato si trova in contraddizione con l'obbligo, per uno Stato di diritto, di motivare provvedimenti amministrativi e sentenze. Qualsiasi atto ufficiale deve essere esposto in modo trasparente per dare modo agli interessati di difendersi, laddove lo ritengano necessario. Se però una macchina autoapprendente basata sull'IA prende una decisione in automatico oppure svolge operazioni preliminari sostanziali in tal senso, può risultare impossibile ricostruire a posteriori il filo logico delle conclusioni del sistema di IA. Ciò rende a sua volta impossibile motivare la decisione in modo da soddisfare i requisiti dello Stato di diritto.

## La Svizzera non si sbilancia

Nel sondaggio condotto nell'ambito dello studio TA-SWISS, la maggioranza degli esperti ritiene che in Svizzera il ricorso a sistemi di IA nella pubblica amministrazione sia meno probabile che in altri paesi e che la completa automatizzazione delle procedure ufficiali complesse sia praticamente da escludere. Nel nostro paese i sistemi di IA potrebbero essere utilizzati tutt'al più per questioni semplici e in processi parzialmente automatizzati.

Per contro la maggioranza degli intervistati ritiene non trascurabile il rischio che le autorità svizzere possano finire per dipendere da fornitori di tecnologia stranieri. Secondo gli esperti si corre anche il pericolo di una certa subordinazione alle macchine, nel senso che il personale amministrativo potrebbe essere in difficoltà a giustificare posizioni diverse da quelle proposte dall'IA. Infine il gruppo degli esperti intervistati ha messo in guardia da un indebolimento della protezione dei dati e dall'impossibilità di garantire un controllo adeguato dei processi automatizzati.

## Nessuna delega del potere decisionale alla macchina

La netta maggioranza degli esperti intervistati concorda su fatto che, quando si tratta di analisi previsionali o di definire un quadro della personalità («profiling»), i sistemi di IA possono solo fornire delle raccomandazioni, ma la decisione deve essere



lasciata alle persone. Una maggioranza risicata degli intervistati ritiene inoltre che l'IA dovrebbe essere utilizzata solo per predisporre decisioni che riguardano luoghi o cose, ma non persone.

La maggioranza accoglie con favore l'obbligo per lo Stato di dichiarare in quali procedure utilizza l'IA. Inoltre le autorità sono tenute a spiegare in modo comprensibile gli elementi fondamentali dei processi che impiegano l'IA. Infine lo Stato deve motivare singolarmente ogni decisione supportata dall'IA.

## Macchine «intelligenti» – una sfida per la giurisprudenza

Se un ingegnere sbaglia a calcolare la portata di un muro o una traduttrice rende in modo errato un testo in un'altra lingua, in sostanza il responsabile dell'errore è chiaro. Ma come ci si comporta se gli errori sono stati favoriti dall'utilizzo di un sistema di IA? L'IA solleva nuovi dubbi di carattere giuridico ed etico, che qui di seguito possiamo solo limitarci a tratteggiare.

Fino ad ora l'ordinamento giuridico è sempre partito dal presupposto che all'origine di un'azione ci sia la decisione di un soggetto, cui si ricollega tutto ciò che accade dopo, comprese le conseguenze legali. L'IA mette in discussione proprio l'idea secondo cui è possibile attribuire un danno al suo autore. Il problema si aggrava mano a mano che i sistemi di IA prendono decisioni in modo sempre più indipendente e imprevedibile.

## Mettere la museruola ai sistemi di IA?

Sebbene l'idea li terrorizzi, tutti i proprietari di cani devono scendere a patti col fatto di doversi assumere la responsabilità di danni che non hanno causato in prima persona: non importa se sia stato il loro cane a scavare buche nel giardino del vicino o a mordere la mano della postina, chi deve pagare i danni e subire le conseguenze è il padrone. Il proprietario può essere esonerato dalla responsabilità solo se è in grado di dimostrare di avere sorvegliato il proprio animale con la dovuta cura.

Vari giuristi ritengono che la responsabilità dei proprietari di animali domestici possa fungere da modello per la responsabilità specifica dei sistemi autonomi di IA, che a volte sono imprevedibili e possono causare danni proprio come gli animali. Chi critica questa analogia sottolinea che l'attività dell'IA

che ha causato il danno è stata delegata alla macchina e quindi all'origine dell'intera catena di azione sta la persona ossia l'utente del sistema.

Adottando una concezione più ampia si potrebbe conferire ai sistemi di IA uno status giuridico proprio. Dal 2017 la Commissione UE sta valutando la creazione di una «persona elettronica» che consentirebbe di attribuire ai sistemi robotizzati avanzati la responsabilità dei danni da loro causati; un approccio analogo sarebbe ipotizzabile anche per i sistemi di IA, ma secondo la valutazione degli esperti è per il momento ancora da respingere tassativamente.

## Creatività di prima e seconda mano

Sulla base della giurisprudenza vigente, i metodi matematici non sono brevettabili. La situazione è meno chiara se si parla di tutela brevettuale di programmi informatici. Secondo la prassi dell'Ufficio europeo dei brevetti, se l'apprendimento automatico fornisce un «contributo tecnico» a un'invenzione – ad esempio se si utilizzano reti neurali per classificare immagini e video digitali – possono sussistere i presupposti per la brevettabilità. La brevettabilità di un sistema di IA deve quindi essere esaminata caso per caso.

Irrisolte sono anche alcune questioni relative al diritto d'autore. Di per sé i programmi informatici sono considerati opere potenzialmente idonee alla tutela del copyright. Tuttavia i sistemi autoapprendenti, evolvendosi di continuo, pongono un problema: questa loro caratteristica retromodifica l'oggetto da tutelare, dato che nel tempo il sistema di IA cambia rispetto a com'era in origine.

Un'ulteriore difficoltà sorge nel caso di beni immateriali generati da sistemi di IA. Non è chiaro se siano tutelabili e, in caso affermativo, chi debba esserne considerato il creatore dal punto di vista della proprietà intellettuale: il programmatore che ha sviluppato il sistema in questione, la persona che lo utilizza o il sistema stesso? Mentre quest'ultima opzione va esclusa a priori per la mancanza di capacità giuridica dei sistemi di IA, per poter attribuire i diritti in questione a una persona fisica occorre compiere una distinzione: teoricamente, in base alla legge sul diritto d'autore, le opere generate in completa autonomia dall'IA non sono tutelabili già solo perché nessun essere umano ha fornito un contributo intellettuale alla loro creazione (uno dei presupposti imprescindibili per la tutelabilità). In pratica, però, non è così facile distinguere opere di

questo genere da prodotti per la cui creazione l'IA è stata utilizzata come semplice strumento ausiliario.

Sul piano del diritto brevettuale, invece, in linea di massima si può presupporre il diritto alla tutela delle invenzioni generate dall'IA. Diventa inventore in senso giuridico la persona fisica che prende atto per prima del risultato e lo interpreta come soluzione a un problema tecnico.

### «Fairness», concetto difficile da quantificare

Oltre a problemi giuridici, l'IA ne solleva di etici, e per di più quasi irrisolvibili. La situazione si può illustrare sulla base del sistema di IA statunitense «COMPAS», che stima la tendenza alla recidiva dei criminali sulla base di 137 caratteristiche. Il sistema è stato criticato quando è emerso che agli afroamericani veniva erroneamente attribuito un rischio di ricaduta quasi doppio rispetto ai bianchi. Viceversa il sistema falliva con frequenza doppia nella previsione di recidive poi effettivamente verificatesi dei bianchi. In tutto ciò, l'algoritmo non aveva ricevuto alcuna informazione sull'appartenenza etnica degli imputati.

L'azienda fornitrice del software si è difesa dall'accusa di razzismo dicendo di aver sviluppato un test per assicurare che il tasso di precisione fosse lo stesso per le persone di colore e per i bianchi e che questo garantiva l'equità del programma. In seguito un gruppo di matematici ha approfondito la questione se un test possa essere contemporaneamente sia giusto che ingiusto, concludendo che ciò è assolutamente possibile – nello specifico quando in una popolazione una caratteristica (nella fattispecie la recidiva) si verifica complessivamente con frequenza diversa a seconda del sottogruppo. Dato che gli afroamericani tendono più spesso a commettere altri crimini rispetto ai bianchi, non è possibile fare previsioni di pari precisione per entrambi i gruppi e allo stesso tempo ottenere tassi di errore della stessa entità per entrambi. Secondo gli esperti è impossibile ottenere giustizia con metodi statistici o matematici. Spetta piuttosto all'etica riflettere su procedure e criteri che garantiscano o per lo meno migliorino la fairness degli algoritmi.

### Linee guida etiche a livello europeo

Nella primavera 2019 un gruppo di esperti di IA ha pubblicato per conto della Commissione Europea le Linee guida etiche per un'IA affidabile. Le linee guida sono destinate sia agli sviluppatori di IA sia alle persone che la usano o lavorano con i suoi prodotti. Intendono garantire il rispetto dei valori fondamentali e delle norme applicabili, facendo sì che l'uso dell'IA rimanga vincolato a obiettivi eticamente sostenibili.

Le linee guida pongono particolare enfasi sulla necessità di garantire che il controllo delle decisioni e delle procedure resti in mano alle persone anche se si utilizza l'IA. Inoltre i sistemi tecnologici devono essere solidi e fornire risultati riproducibili con affidabilità. I dati e la privacy vanno tutelati, i sistemi di IA vanno utilizzati con trasparenza per garantire che i processi decisionali rimangano tracciabili. Infine è importante evitare distorsioni ingiuste nell'uso dei sistemi di IA ed è fondamentale tenere presente il bene della società e dell'ambiente.

### Distorsioni e mancanza di trasparenza pesano

Secondo la maggior parte degli esperti intervistati, serie di dati sbilanciati e mancanza di trasparenza nell'uso dell'IA comportano notevoli rischi. Ciò nonostante gli esperti sono generalmente favorevoli all'uso dell'IA e le riconoscono un'influenza positiva su obiettivi sociali generalmente accettati, come la riduzione della povertà e il miglioramento della salute.

Ciò detto, per gli intervistati resta chiaro che la responsabilità delle decisioni in materia di IA spetta agli esseri umani. Gli utenti devono essere informati quando interagiscono con un'IA e devono anche essere messi nelle condizioni di capire in che modo l'IA è pervenuta alla sua decisione.

# Raccomandazioni: strutturare ad ampio raggio

**I sistemi di IA vengono utilizzati in campi applicativi molto diversi. Le raccomandazioni sull'utilizzo di questa tecnologia devono quindi coprire un raggio altrettanto ampio.**

I sistemi di IA vengono utilizzati per una grandissima varietà di compiti in settori disparati, ognuno dei quali presenta problematiche legali specifiche. Ne consegue che la creazione di una «legge sull'IA» mirante a una regolamentazione uniforme del suo uso non risulta adeguata e va pertanto respinta.

## Esaminare le applicazioni nello specifico

Chi utilizza un determinato sistema di IA? Per che cosa esattamente? Su quali dati si basa e di quali requisiti legali si deve tenere conto? Sono tutte domande a cui bisogna rispondere prima di poter valutare le opportunità e i rischi dell'uso dell'IA. La definizione delle regole di utilizzo dell'IA non può mai prescindere dal rispettivo contesto. Occorre inoltre identificare i problemi concreti e gli sviluppi indesiderati nonché risolverli con provvedimenti adeguati; nei casi in cui la situazione di rischio non è chiara, è opportuno intensificare la ricerca per individuare i potenziali di pericolo.

## Aspettative elevate nei confronti delle autorità

Nell'espletare le proprie funzioni istituzionali, lo Stato agisce da una posizione di superiorità rispetto all'individuo. A maggior ragione è quindi importante che la legalità dell'intervento statale possa essere verificata dagli interessati e che i sistemi di IA vengano utilizzati con la dovuta trasparenza. Se lo Stato utilizza l'IA, è tenuto per principio a soddisfare standard più elevati rispetto agli utenti privati.

## Trasparenza e informazioni «digeribili»

Nei casi in cui l'uso dell'IA coinvolga in modo rilevante delle persone, queste devono essere informate con metodi semplici e di facile comprensione sul fatto di avere a che fare con un sistema di IA. L'eccesso di informazioni, tuttavia, può indurre una certa rassegnazione e quindi a negligenza da parte degli interessati: anche questo va considerato. Occorre garantire che, su richiesta, sia questi ultimi sia le organizzazioni che forniscono assistenza legale alle parti lese ricevano tutte le informazioni necessarie a individuare e valutare eventuali risultati errati.

## Sigillo di qualità anziché generiche autorizzazioni di immissione in commercio

Date le varie possibilità di utilizzo e i diversi livelli di impatto dei sistemi di IA, non sarebbe realistico subordinarne l'impiego a un'autorizzazione generale di immissione in commercio. Laddove – come nel caso dei prodotti medicali – è già comunque prevista un'autorizzazione all'immissione in commercio, la verifica deve comprendere anche l'IA. Se iniziative private sviluppano sigilli di qualità per l'IA, il loro operato va accolto con favore e incoraggiato. Le organizzazioni che tutelano i consumatori dovrebbero essere messe nella condizione di valutare meglio la qualità di tali certificati.

## Non concentrarsi esclusivamente sulle questioni tecniche

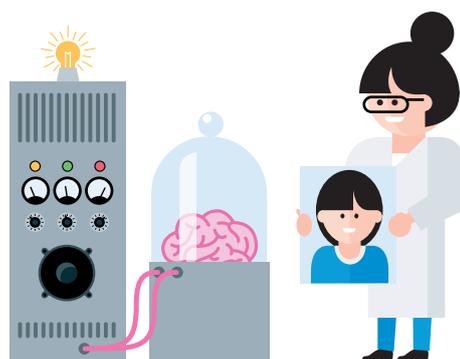
La concezione e lo sviluppo dell'IA sono ben più di semplici questioni tecnologiche. Soprattutto per poterne valutare i rischi correlati, è essenziale che gli sviluppatori siano competenti anche in materia etica e giuridica di base, oltre a essere disposti a collaborare a livello interdisciplinare e disposti a farlo. Competenze adeguate sono indispensabili anche per tutti coloro che utilizzano sistemi di IA e lavorano con i loro risultati.

## Gruppo di accompagnamento

- Prof. Dr. Jean Hennebert, Leitungsausschuss TA-SWISS, Département d'informatique de l'Université de Fribourg, Präsident der Begleitgruppe
- Benjamin Bosshard, Eidgenössische Kommission für Kinder- und Jugendfragen
- Sabine Brenner, Geschäftsstelle Digitale Schweiz, Bundesamt für Kommunikation (BAKOM)
- Dr. Christian Busch, Staatssekretariat für Bildung, Forschung und Innovation (SBFI)
- Dr. Christine Clavien, Institut Ethique Histoire Humanités, Université de Genève
- Daniel Egloff, Staatssekretariat für Bildung, Forschung und Innovation SBFI
- Andy Fitze, SwissCognitive – The Global AI Hub.
- Matthias Holenstein, Stiftung Risiko-Dialog
- Dr. Marjory Hunt, Fonds national suisse de la recherche scientifique (FNS)
- Manuel Kugler, Schweizerische Akademie der Technischen Wissenschaften (SATW)
- Thomas Müller, TA-SWISS Leitungsausschuss, Redaktor Schweizer Radio SRF
- Katharina Prelicz-Huber, TA-SWISS Leitungsausschuss (bis 2019), Präsidentin Gewerkschaft VPOD/SSP, Nationalrätin
- Prof. Ursula Sury, Rechtsanwältin und Professorin, Hochschule Luzern (HSLU)
- Dr. Stefan Vannoni, TA-SWISS Leitungsausschuss, cemsuisse

## Gestione del progetto presso TA-SWISS

- Dr. rer. soc. Elisabeth Ehrensperger, Geschäftsführerin
- Dr. Catherine Pugin, Projektleiterin



## **Impressum**

Pensare senza testa

Sintesi dello studio «Wenn Algorithmen für uns entscheiden:  
Chancen und Risiken der künstlichen Intelligenz» di TA-SWISS  
TA-SWISS, Berna 2020  
TA 72A/2020

Autrice: Lucienne Rey, TA-SWISS, Berna  
Traduzione: Roberta Gado, Lipsia  
Produzione: Fabian Schluep, TA-SWISS, Berna  
Grafica: Hannes Saxer, Berna  
Stampa: Jordi AG – Das Medienhaus, Belp

## **TA-SWISS – Fondazione per la valutazione delle scelte tecnologiche**

Spesso le nuove tecnologie portano netti miglioramenti per la qualità di vita. Talvolta nascondono però anche nuovi rischi, le cui conseguenze non sono sempre prevedibili in anticipo. La fondazione per la valutazione delle scelte tecnologiche TA-SWISS esamina le opportunità e i rischi dei nuovi sviluppi tecnologici in materia di «biotecnologia e medicina», «società dell'informazione» e «mobilità / energia / clima». I suoi studi si rivolgono sia ai decisori nella politica e nell'economia che all'opinione pubblica. TA-SWISS promuove inoltre lo scambio di informazioni e opinioni tra specialisti della scienza, dell'economia, della politica e la popolazione attraverso metodi di partecipazione. Siccome devono fornire informazioni il più possibile obiettive, indipendenti e solide sulle opportunità e sui rischi delle nuove tecnologie, i progetti di TA-SWISS sono elaborati d'intesa con gruppi di esperti composti in modo specifico a seconda del tema. Grazie alla competenza dei loro membri, questi gruppi d'accompagnamento coprono un ampio ventaglio di aspetti della tematica esaminata.

La fondazione TA-SWISS è un centro di competenza delle Accademie svizzere delle scienze.



TA-SWISS  
Fondazione per la valutazione  
delle scelte tecnologiche  
Brunngasse 36  
CH-3011 Berna  
info@ta-swiss.ch  
www.ta-swiss.ch

membro delle  
 **accademie svizzere  
delle scienze**