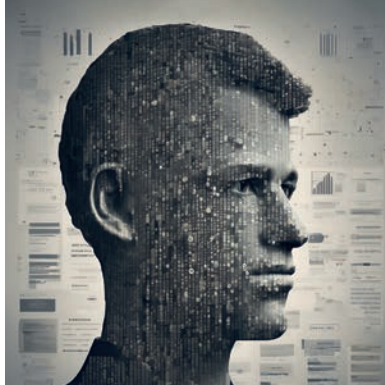


Fabio
Mercorio



Competenze digitali e interazioni

Big Data e IA: come cambierà il mondo del lavoro

Non è semplice stimare l'impatto dell'Intelligenza Artificiale nelle professioni. E dunque rispondere con assoluta certezza a domande del tipo: "Quanti lavori saranno rimpiazzati dall'IA? Il saldo tra lavori dismessi e nuovi lavori sarà positivo? Quali le professioni più a rischio?". Eppure, la questione è centrale. Perché c'è condivisione nel ritenere inevitabile l'impatto di questa tecnologia nelle dinamiche lavorative; allora, forse, è più ragionevole porsi domande su cosa sta accadendo e su quali azioni intraprendere, piuttosto che sperare nell'oracolo che ci riveli il futuro. Scenari ed evidenze. Un percorso.

Da alcuni anni, i termini Big Data e Intelligenza Artificiale sono divenuti di uso comune in molti ambiti, anche al di fuori dell'informatica. Ma cosa vuol dire Big Data? Cosa IA? Perché Big Data e IA sono legati? Come "ragiona" l'IA? Come Big Data e IA stanno cambiando il mondo del lavoro? Quali le sfide in gioco? In questo articolo affronto alcune di queste domande, con l'intento di chiarire alcuni elementi tecnici per i non addetti ai lavori, discutendo gli elementi critici e i punti di attenzione futuri.

Cosa sono i Big Data

Il termine Big Data si riferisce a un insieme di caratteristiche tipiche dei dati e della loro gestione che ne definiscono la complessità. L'aggettivo "big", infatti, non si riferisce alle dimensioni dei dati, ma alle complessità che la loro gestione comporta, in termini di *Volume* (dimensione dei dati), *Velocity* (la velocità di generazione dei dati, informalmente si può pensare al tempo trascorso tra la generazione di un dato e l'altro); la *Variety* ovvero la diversità di formati o di struttura tra i dati. È una caratteristica prettamente tecnica, ma che si può spiegare con un esempio: immaginiamo di dover riporre dei libri di dimensioni diverse in scaffali, ciascuno potenzialmente con ripiani di dimensioni variabili e ignote a priori. Inoltre, ipotizziamo che esistano libri che trattino il medesimo argomento, ma in lingue o da punti di vista diversi. Organizzare una tale "diversità" richiede all'utente di avere presente uno "schema" (criterio) logico comune, che permetta di organizzare i contenuti per autori, per argomenti, sistemando gli scaffali in funzione delle diverse esigenze di accesso alla biblioteca, ecc. Questa diversità è una complessità che definisce i Big Data, spesso sottovalutata, che sovente si presenta quando i dati giungono a noi da fonti diverse ed eterogenee – tipicamente dati del Web, sensori, o dati di aziende diverse – ciascuno con il proprio formato, standard e modello dati. C'è poi la *Veracity* che si riferisce a quanto accurato (veritiero, credibile) sia il dato che si intende processare e successivamente analizzare.

Risulta evidente, quindi, che misurare la veridicità – o la qualità – di un dato oltre la sua natura quantitativa è un'attività tutt'altro che lineare. Soprattutto, una scarsa qualità del dato iniziale può avere effetti imprevedibili sulle analisi che si faranno, il motto "garbage in, garbage out"

intende esattamente sottolineare la crucialità dell'analisi della qualità del dato per rendere le analisi credibili e attendibili. In ultimo, la V di *Value*. È la dimensione la cui metrica può essere "stimata" inizialmente durante il processing, ma misurata esclusivamente al termine del processo di trattamento dei Big Data, ovvero durante l'analytics. In altre parole, il Value intende misurare se e in che misura i Big Data abbiano fornito conoscenza – e quindi valore – per supportare il processo decisionale.

Ma cosa è questa Intelligenza Artificiale

Sebbene sia difficile fornire una definizione olistica, la Commissione Europea ha di recente fornito una definizione che coglie l'attuale stato dell'arte, ovvero: "Sistemi che mostrano un comportamento intelligente analizzando l'ambiente e prendendo decisioni – con diversi gradi di autonomia – per raggiungere obiettivi specifici"¹. È evidente, quindi, che l'IA che si sta diffondendo e di cui parliamo è (per ora) lontana dai robot umanoidi capaci di svolgere le più svariate e generali attività umane. Parliamo piuttosto di sistemi (e quindi software e algoritmi nascosti nei vari dispositivi e servizi web che quotidianamente utilizziamo) che fungono da surrogato dell'intelligenza umana (sembrano intelligenti) grazie alla capacità di analizzare l'ambiente (processare milioni di dati) per raggiungere un obiettivo prestabilito (raccomandare un contenuto da leggere o ascoltare, scrivere un testo, predire un trend, riconoscere un oggetto o un volto, ecc.).

In questo contesto, il machine learning è un particolare tipo di Intelligenza Artificiale che sfrutta i dati per apprendere caratteristiche, ne consegue che la generazione di Big Data in svariati campi ne abilita (e per certi versi nobilita) l'uso.

Si pensi a uno studente a cui bisogna insegnare come funziona l'operatore addizione: si procederà spiegando allo studente un metodo per svolgere le operazioni e successivamente si procederà nell'esercitazione; maggiore il numero di esercizi svolti, maggiore diventerà la confidenza dello studente nella risoluzione corretta. Ovviamente, questa esperienza che lo studente acquisisce non lo rende immune dall'errore. Questo è un modo di ragionare deduttivo: dal generale al particolare.

Immaginiamo ora di insegnare allo studente la somma tra numeri mostrandogli migliaia, milioni di operazioni con l'esito atteso: è possibile che lo studente riesca a derivare autonomamente una o più regole, un metodo appunto, comune a tutte le operazioni osservate. In tal caso, lo studente avrà usato un metodo induttivo (dal particolare inferisce una regola generale). Sappiamo bene che tale metodo è deprecabile nell'insegnamento, perché pone l'attenzione sul fare (risolvere bene le somme) e non sul capire (accogliere i criteri con cui si svolge la somma). Il machine learning è un metodo di apprendimento induttivo che, come lo studente sopra citato, apprende dai dati (gli esercizi di matematica) a svolgere un compito (l'addizione), senza però comprendere il reale funzionamento di ciò che sta facendo. La differenza tra lo studente e la macchina che apprende dagli esempi è una: l'esperienza non garantisce allo studente lo svolgimento corretto; la macchina, invece, sarà sempre più brava quanti più esercizi svolgerà, anche se potrà fallire nel caso si trovi a svolgere un esercizio con caratteristiche mai viste prima. Ma comprendere il funzionamento di una IA è un'altra storia.

Cosa c'entrano i Big Data con l'IA

I dispositivi digitali connessi che possediamo (si pensi ad esempio agli smartphone e le app che quotidianamente utilizziamo) producono continuamente Big Data riguardo le nostre preferenze sociali, musicali, amicali, ecc., raccomandando dei contenuti.

La generazione e la crescente disponibilità di Big Data diventa quindi un enorme patrimonio informativo che – se opportunamente processato – può essere utilizzato per profilare e raccomandare servizi o prodotti. Chiaramente, la pervasività dei dispositivi digitali funge da catalizza-

¹ EU Commission, COM (2018) 795.

tore all'uso dei Big Data, con applicazioni in ambito medico, industriale, sociale, dei servizi alla persona, etc. È cruciale, però, sottolineare che i Big Data stanno svolgendo il ruolo di acceleratore all'uso dell'IA. Potremmo dire che i Big Data stanno al machine learning come il carburante sta all'automobile.

In questo scenario, l'IA generativa è arrivata dirompente nell'autunno del 2022 (il *casus belli* è ChatGPT) consacrando l'IA come "persona dell'anno" 2023. Nonostante la tecnologia alla base di ChatGPT sia la più evoluta, il procedimento dell'apprendimento automatico rimane invariato. Big Data + IA + capacità di calcolo = IA generativa, ovvero una IA che è capace di generare contenuti multimodali (testo, musica, immagini, ecc.) a partire da una richiesta in linguaggio naturale (prompt) perché ha acquisito una competenza del mondo processando l'intero scibile umano (il Web).

La disponibilità in modalità open source di molti di questi modelli pre-addestrati ne ha ulteriormente accelerato la diffusione, soprattutto in ambito business. Un nota bene: sebbene sia argomento di dibattito nella comunità scientifica², al momento non ci sono evidenze che questi modelli generativi "comprendano" quel che producono. Siamo ancora di fronte allo studente che risolve brillantemente una gran quantità di esercizi, anche complessi, a volte meglio del docente, ma c'è ancora il dubbio che "abbia imparato tutto a memoria". Una straordinaria memoria, ma pur sempre solo memoria.

A oggi, infatti, la scienza non è in grado di insegnare concetti alle macchine: è possibile solo mostrare loro dati, tanti dati, sperando che si formi una comprensione dei fenomeni e del mondo. Il tentativo in atto è di combinare i due modi di apprendere per insegnare alle macchine i concetti (l'astrazione, su tutte): è la cosiddetta Neuro-symbolic IA, che nei prossimi dieci anni ci riserverà certamente delle sbalorditive sorprese.

L'IA nel mercato del lavoro: panoramica e sfide in atto

Comprendere che i Big Data sono gli abilitatori della diffusione di massa dell'IA ci permette di capire una caratteristica fondamentale legata al mondo del lavoro: l'impatto dell'IA non è legata ai soli "addetti ai lavori", ma si diffonde pervasivamente in molte professioni, perché l'impatto maggiore dell'IA risiede nella creazione di un ecosistema IA, che include la generazione dei dati, il processamento, il trattamento, la sintesi di valore per il processo decisionale, la rappresentazione della conoscenza, ecc. C'è quindi un impatto dell'ecosistema IA (i) nell'industria, perché anche l'operatore blue collar dovrà interagire con sistemi di IA che controllano macchinari e processi; (ii) nell'agricoltura, perché ci saranno (ci sono già) progetti di micro irrigazione robotizzata con droni per ridurre il consumo d'acqua; (iii) nel commercio, perché saranno richieste capacità ai white collar di interpretare dati e prendere decisioni sulla base degli analytics prodotti; (iv) nella sanità, perché gli algoritmi potranno svolgere screening di massa e raccomandare ai medici azioni da intraprendere; (v) nei servizi bancari, perché gli algoritmi disintermedieranno l'accesso a molti dei servizi (credito), con tutto il tema legato alla discriminazione algoritmica; (vi) legale, perché vengono meno per i giuristi e i legislatori gli elementi di dolo, colpa e intenzione, imputabili questa volta a un algoritmo (o al suo sviluppatore? O a chi ha collezionato i dati?).

Ci sono moltissimi studi che tentano di stimare l'impatto dell'IA sul mondo del lavoro in termini di crescita economica: McKinsey³ ha di recente dichiarato che il valore aggiunto che l'ecosistema "generative IA" produrrà a livello globale si aggira tra i 2.2 e 4.4 trilioni di dollari annui.

Più difficile è invece la stima sull'impatto dell'IA nelle professioni: "Quanti lavori saranno rimpiazzati dall'IA? Il saldo tra lavori dismessi e nuovi lavori sarà positivo? Quali le professioni più a rischio?"

Sebbene le domande siano tutte legittime, personalmente credo che sia più ragionevole porsi

2. Hinton e LeCun, insieme con Bengio, sono tra gli inventori delle reti neurali, costruiti alla base dell'attuale Intelligenza Artificiale e vincitori del Turin Award (il nobel dell'Informatica). Tuttavia, c'è dissenso tra loro, che spesso si esprime sui social e in conferenze di settore: il primo convinto che le macchine comprendano la visione del mondo che generano dai dati, il secondo sostenitore del contrario.

3. McKinsey, *The economic potential of generative AI: The next productivity frontier*, Report, 14 giugno 2023



domande su cosa sta accadendo e su quali azioni intraprendere, piuttosto che sperare nell'oracolo che ci riveli il futuro, incerto per definizione. Ci sono infatti alcune evidenze che possono far luce sullo scenario che ci si prospetta dinanzi:

a. L'IA è qui per restare. Molti studi, tra cui Gartner⁴, hanno definitivamente censito che l'hype tecnologico primordiale legato all'IA è sorpassato: siamo quindi di fronte a un ecosistema AI che si sta affermando e che quindi chiede a imprenditori, aziende e professionisti di ripensare il proprio business, i processi produttivi e decisionali, integrando i "dati" come fattore abilitante per ottenerne un vantaggio competitivo. L'IA si sta infatti affermando come commodity: la difficoltà per le aziende non risiede nello sviluppare algoritmi (che anche i tecnici prendono "dallo scaffale") ma nell'integrare soluzioni di IA nei processi di business affinché contribuiscano a generare valore, arricchiscano il processo decisionale e sostituiscano i lavoratori nelle operazioni ripetitive, valorizzando il capitale umano: in questo senso l'AI è un ecosistema che sta generando una filiera professionale a esso legata.

b. No Data, no IA. È diffusa la convinzione di poter usare l'IA (deduttiva o generativa) senza i dati o, meglio, senza porre attenzione ai dati. Come già detto, invece, l'IA di cui parliamo apprende dai dati, da cui genera un modello di conoscenza. Diventa quindi fondamentale comprendere che la fase di collezionamento, raccolta, selezione, messa in qualità dei dati e competenza di dominio sono fattori fondamentali per permettere all'IA di funzionare. In questo senso, le professionalità che generano e processano i Big Data arricchiscono e contribuiscono l'ecosistema IA. La bontà dei dati da cui l'IA apprende definisce la bontà delle decisioni che ci raccomanderà e sulla base delle quali si prenderanno decisioni, anche strategiche. Non porre attenzione e cura alla filiera di gestione e raccolta dati è come pretendere che un cattivo insegnante produca studenti brillanti.

c. La Generative IA spaventa i lavoratori della conoscenza. Se da un lato la "robotizzazione" ha effetti maggiori sui cosiddetti blue collar, perché automatizza alcuni processi industriali ripetitivi (decisioni operative), sottraendo posti di lavoro, dall'altro l'arrivo di ChatGPT, capace di leggere, scrivere, comporre, disegnare, ecc. sta spaventando i lavoratori della conoscenza, coloro che

4. 2023 Gartner Hype Cycle™ for Artificial Intelligence (AI).

facevano della capacità di generare contenuto un asset per la propria professionalità. Difficile sintetizzare quindi un tema così complesso in poche righe, ma è evidente che una strenua resistenza non sarà una strategia dominante, la storia l'ha confermato in tutte le rivoluzioni o – le evoluzioni – industriali, come in questo caso. Non sappiamo quale sarà l'equilibrio di forze tra un copy writer e ChatGPT, ma certamente il professionista che saprà utilizzare l'AI generativa a suo vantaggio, cioè per rafforzare e meglio esprimere la propria conoscenza del mondo che ha – e che l'IA non ha – ruberà il lavoro a quel professionista che non la saprà utilizzare.

d. Prima erano strumenti, ora sono entità. È fondamentale sottolineare una differenza con “la rivoluzione digitale” che abbiamo vissuto e che non è ancora terminata: se prima l'oggetto che si poneva all'orizzonte era un oggetto con cui interagire (strumenti come word, excel, la mail, il motore di ricerca, ecc.), strumenti quindi al nostro servizio, che rispondono esattamente al nostro input, oggi l'oggetto, ovvero l'IA e in particolare l'IA generativa è una “entità”. Ci rivolgiamo, quindi, ai chatbot come un'entità con cui interloquire e a cui porre una domanda: se la domanda è posta male, la risposta non sarà soddisfacente. Non a caso, una “nuova” professionalità emergente negli annunci di lavoro è il prompt engineer, ovvero colui che è capace di porre domande ai chatbot.

Questi fattori, o evidenze, sono già oggi le basi azotate che compongono il nostro DNA lavorativo del futuro: lavori ibridi in cui l'interazione uomo-macchina sarà chiave; dati da collezionare, selezionare e processare per far sì che l'IA abbia una “visione del business” quanto più pertinente alla realtà; competenze digitali e di interazione IA a tutti i livelli aziendali, dalle decisioni operative a quelle tattiche e strategiche; e, da ultimo, necessità di formazione continua, per fare upskilling e reskilling e mantenere le competenze allineate all'evoluzione tecnologica in cui la capacità di calcolo sarà sempre più demandata alla macchina, mentre la creatività, l'intuizione, l'interazione interpersonale e, non ultimo, la decisione saranno sempre – e ancor di più – una prerogativa prettamente umana.



Fabio Mercorio è professore associato di Computer science presso l'università di Milano-Bicocca.