

Corso di Laurea Magistrale in Strategic Management
Cattedra di Analisi Finanziaria

**LA FIDUCIA NEI SISTEMI AUTOMATIZZATI DI RECRUITMENT:
LA TRASPARENZA COME LEVA STRATEGICA DI
ACCETTAZIONE DELL'IA**

RELATORE

Prof.ssa Barbara Sveva Magnanelli

CORRELATORE

Prof. Fabrizio di Lazzaro

CANDIDATA

Miriam Corrente

Matr. 781791

INDICE

Introduzione

Capitolo 1

Fisionomia dell'Intelligenza Artificiale e contesto della ricerca

- 1.1 Evoluzione del rapporto uomo-macchina
- 1.2 L'Intelligenza Artificiale e il Decision-Making Aziendale
- 1.3 Obiettivi della ricerca

Capitolo 2

Il concetto di Trasparenza nell'IA

- 2.1 Bias negli algoritmi
- 2.2 La Trasparenza dell'IA: definizione, rilevanza e quadro normativo
- 2.3 Dalla trasparenza all'accettazione dell'IA: una leva per il cambiamento organizzativo
- 2.4 La fiducia nei processi automatizzati di recruiting come fattore chiave dell'accettazione dell'IA

Capitolo 3

Metodologia della Ricerca

- 3.1 Disegno della ricerca
- 3.2 Obiettivi e ipotesi di ricerca
- 3.3 Strutturazione del questionario
- 3.4 Affidabilità delle scale
- 3.5 Campione e raccolta dati
- 3.4 Tecniche di analisi dei dati

Capitolo 4

Analisi dei Risultati

- 4.1 Analisi descrittive
- 4.2 Correlazioni bivariate
- 4.3 Analisi di regressione
- 4.4 Analisi di mediazione
- 4.5 Analisi di moderazione - Percezione del rischio di perdita di lavoro

Capitolo 5

Implicazioni pratiche e impatto sulle Performance aziendali

5.1 Implicazioni pratiche

5.2 Fiducia nei sistemi IA e impatti sulle performance aziendali: una traiettoria strategica da esplorare

Conclusioni

INTRODUZIONE

Negli ultimi anni l'Intelligenza Artificiale (IA) ha rivoluzionato il panorama aziendale, diventando un elemento fondamentale per l'ottimizzazione dei processi e il miglioramento della competitività delle imprese. L'adozione di tecnologie basate sull'IA ha trasformato il modo in cui le organizzazioni prendono decisioni, gestiscono le risorse e interagiscono con clienti e dipendenti. Tuttavia, se da un lato l'IA offre opportunità straordinarie, dall'altro solleva importanti questioni legate alla trasparenza, equità e impatto sul mondo del lavoro. La crescente diffusione dell'IA ha portato a un acceso dibattito sull'equilibrio tra innovazione e responsabilità etica. Sebbene le tecnologie avanzate abbiano permesso alle aziende di ottimizzare processi produttivi, ridurre costi e migliorare la precisione delle decisioni, esse sollevano interrogativi sulla loro affidabilità e sui potenziali rischi per il mercato del lavoro. In particolare, la trasparenza nell'IA è diventata un tema centrale sia dal punto di vista normativo che organizzativo. Il rapido sviluppo dell'IA ha reso necessaria l'introduzione di regolamentazioni per garantire un utilizzo etico e responsabile di queste tecnologie. L'Unione Europea, attraverso l'AI Act, ha stabilito standard elevati per assicurare l'affidabilità e la sicurezza degli algoritmi utilizzati nei processi decisionali. La trasparenza è infatti un elemento cruciale per favorire la fiducia nei confronti dell'IA, soprattutto in contesti critici come il decision-making aziendale e il recruiting. Una maggiore trasparenza può ridurre l'asimmetria informativa tra aziende e dipendenti, aumentando la percezione di equità e riducendo le preoccupazioni legate all'uso di decisioni automatizzate. L'adozione dell'IA nel decision-making aziendale ha trasformato il modo in cui le imprese pianificano strategie e allocano risorse. Sistemi di IA avanzati sono oggi utilizzati per analizzare grandi quantità di dati, prevedere tendenze di mercato e supportare i manager nel prendere decisioni più informate. Tuttavia, l'integrazione di modelli predittivi e algoritmi nel processo decisionale ha generato nuove sfide, tra cui la necessità di comprendere il funzionamento dei sistemi AI e di garantire che le decisioni prese non siano influenzate da bias o distorsioni cognitive.

Un altro ambito in cui l'IA sta avendo un impatto significativo è il recruiting e la gestione delle risorse umane. Le aziende stanno implementando algoritmi per selezionare candidati, analizzare curriculum e condurre colloqui automatizzati. Se da un lato questi strumenti offrono vantaggi in termini di efficienza e riduzione dei tempi di selezione, dall'altro possono suscitare timori relativi alla mancanza di equità nei processi di assunzione. L'uso di modelli IA non trasparenti potrebbe portare a decisioni discriminatorie o a una selezione del personale basata su criteri poco chiari, con il rischio di compromettere la fiducia dei lavoratori nelle politiche aziendali.

Il presente studio si propone di analizzare l'impatto della trasparenza percepita dell'IA sulle performance aziendali, focalizzandosi su due aspetti chiave:

- Il ruolo della trasparenza nel migliorare la fiducia dei dipendenti e la loro percezione dell'IA nei processi decisionali.
- L'importanza della trasparenza nei processi di recruiting per garantire equità, inclusione e accettazione da parte dei candidati.

La ricerca prevede un approccio quantitativo, basato sulla somministrazione di questionari a dipendenti di aziende che utilizzano sistemi di IA nei processi decisionali e di selezione del personale. Attraverso questa analisi, si cercherà di comprendere se e come una maggiore trasparenza nell'IA possa ridurre la percezione di insicurezza lavorativa, migliorare l'accettazione della tecnologia e, di conseguenza, influenzare positivamente la performance aziendale.

Nel primo capitolo verrà fornita una panoramica generale sull'Intelligenza Artificiale, analizzando le principali applicazioni nei processi aziendali e il crescente impatto della trasparenza come fattore critico di successo. Successivamente, verranno esplorati i fondamenti teorici della ricerca, il modello ipotetico di riferimento e la metodologia adottata per la raccolta e l'analisi dei dati. Infine, i risultati dell'indagine saranno discussi alla luce della letteratura esistente, per evidenziare le implicazioni pratiche dell'adozione di modelli di IA trasparenti nel contesto aziendale. L'obiettivo di questo studio è quindi quello di contribuire alla comprensione delle dinamiche che regolano il rapporto tra IA, fiducia organizzativa e sicurezza lavorativa, offrendo spunti utili per una gestione responsabile della tecnologia nel contesto aziendale moderno.

Capitolo 1

Fisionomia dell'Intelligenza artificiale e contesto della ricerca

1.1 Evoluzione del rapporto uomo-macchina

Il legame tra uomo e macchina ha sempre rappresentato una delle forze trainanti nello sviluppo della civiltà, ma è con le rivoluzioni industriali che questa relazione ha assunto un ruolo determinante nel ridefinire il lavoro, la società e l'identità stessa dell'essere umano. In epoca preindustriale, gli strumenti tecnologici erano principalmente un'estensione delle capacità fisiche umane, destinati a facilitare attività manuali come l'agricoltura o l'artigianato. Con l'avvento della **Prima Rivoluzione Industriale**, a partire dalla seconda metà del XVIII secolo, questa dinamica cambiò radicalmente. La scoperta e l'applicazione della macchina a vapore, insieme all'utilizzo intensivo del carbone come fonte di energia, resero possibile il passaggio da una produzione artigianale a una produzione meccanizzata e su larga scala¹. Le macchine non erano più solo utensili, ma divennero il cuore pulsante delle fabbriche, capaci di sostituire l'uomo nelle attività più faticose e ripetitive. Questo passaggio segnò anche l'inizio di un rapporto ambivalente: se da un lato l'efficienza e la produttività aumentarono, dall'altro si generarono nuove forme di alienazione del lavoro umano, costretto a seguire i ritmi delle macchine. Gli operai venivano sottoposti a turni estenuanti e a condizioni di lavoro spesso disumane, con un controllo serrato del tempo e dello spazio in fabbrica².

Con la **Seconda Rivoluzione Industriale**, tra la fine del XIX secolo e l'inizio del XX, l'introduzione dell'elettricità, del petrolio e del motore a scoppio aprì la strada a una nuova fase di automazione e di espansione industriale. La produzione di massa si affermò grazie alle catene di montaggio, come nel celebre modello introdotto da Henry Ford, e la macchina diventò non solo uno strumento produttivo, ma anche il simbolo del progresso tecnologico e della modernizzazione economica (Studenti.it). La relazione uomo-macchina si fece

¹ Studenti.it. (n.d.). *Le tre rivoluzioni industriali: riassunto*.

² Zanichelli Editore. (2022). *La rivoluzione industriale* [PDF].

ancora più intensa: l'uomo veniva integrato all'interno di sistemi produttivi altamente standardizzati, in cui il lavoro era frammentato e specializzato, accentuando una crescente dipendenza dalle macchine stesse.

È con la **Terza Rivoluzione Industriale**, avviata nella seconda metà del Novecento, che il rapporto uomo-macchina compie un ulteriore salto qualitativo. L'introduzione dell'elettronica, dei computer e successivamente di Internet trasformò la macchina in un dispositivo capace di elaborare informazioni, riducendo ulteriormente il ruolo fisico dell'essere umano nei processi produttivi (Studenti.it). Le macchine non solo eseguivano operazioni meccaniche, ma assumevano un ruolo attivo nella gestione dei dati e delle comunicazioni, anticipando il futuro scenario dominato dall'intelligenza artificiale. Questa fase fu caratterizzata da una transizione verso il cosiddetto "capitalismo informazionale", dove il valore veniva prodotto sempre più attraverso l'informazione e il controllo delle reti digitali.

Attualmente, ci troviamo nel cuore di quella che viene definita la **Quarta Rivoluzione Industriale**, un periodo contraddistinto dall'adozione diffusa di tecnologie digitali avanzate, tra cui spiccano l'Intelligenza Artificiale (IA), l'Internet of Things (IoT), la robotica collaborativa e la manifattura additiva³. A differenza delle precedenti trasformazioni industriali, l'Industria 4.0 non si limita a incrementare la produttività o l'efficienza delle linee produttive, ma introduce sistemi intelligenti in grado di apprendere in autonomia, prendere decisioni complesse e interagire direttamente con l'uomo in maniera sempre più sofisticata (Artser, n.d.).

Le principali tecnologie che costituiscono l'ossatura dell'Industria 4.0 includono:

- **Internet of Things (IoT):** la connessione in rete di dispositivi fisici e sensori che comunicano tra loro, scambiando dati in tempo reale per ottimizzare processi e attività.
- **Big Data Analytics:** l'elaborazione e l'analisi di grandi volumi di dati che consentono alle aziende di prendere decisioni strategiche più rapide e accurate, migliorando la qualità dei prodotti e riducendo i costi operativi.
- **Intelligenza Artificiale e Machine Learning:** sistemi basati su algoritmi complessi capaci di apprendere dai dati storici e predire eventi futuri, supportando l'automazione e la manutenzione predittiva degli impianti.

³ Artser. (n.d.). *Industria 4.0 e 5.0: un passaggio verso una nuova collaborazione sostenibile tra uomo e macchina*. Artser.

- **Robotica avanzata:** l'impiego di robot collaborativi, detti cobot, progettati per interagire in sicurezza con gli operatori umani, incrementando la produttività e la qualità del lavoro.
- **Additive Manufacturing (Stampa 3D):** una tecnica di produzione che realizza oggetti tridimensionali sovrapponendo strati successivi di materiale, rendendo i processi produttivi più flessibili e personalizzabili.
- **Realtà Aumentata (AR):** strumenti digitali che sovrappongono informazioni virtuali al mondo reale, facilitando la formazione, la manutenzione e l'assistenza remota.
- **Cloud Computing:** l'archiviazione e l'elaborazione di dati attraverso infrastrutture decentralizzate accessibili via Internet, permettendo una gestione più efficiente delle informazioni aziendali.
- **Cybersecurity:** l'insieme di pratiche e tecnologie volte a proteggere sistemi e reti industriali dalle minacce informatiche sempre più frequenti.

Nonostante i notevoli vantaggi in termini di efficienza e produttività, l'Industria 4.0 solleva anche interrogativi legati all'impatto sull'occupazione. La crescente automazione dei processi produttivi alimenta il timore che molte mansioni svolte dagli esseri umani possano essere progressivamente affidate alle macchine (Artser, n.d.).

In questo scenario, l'Intelligenza Artificiale rappresenta uno degli sviluppi più avanzati e incisivi nella relazione tra uomo e macchina. Oggi, le tecnologie basate su IA non si limitano più a eseguire operazioni programmate, ma sono capaci di apprendere, adattarsi e persino simulare forme di intelligenza autonoma. Le reti neurali e gli algoritmi di Machine Learning permettono ai sistemi di analizzare enormi quantità di dati, riconoscere pattern complessi e rispondere a stimoli in maniera sempre più simile al ragionamento umano.

Esempi pratici di queste applicazioni includono modelli di linguaggio evoluti, come ChatGPT, in grado di dialogare in modo naturale con gli utenti e fornire contenuti personalizzati, o gli algoritmi di raccomandazione utilizzati da piattaforme come Netflix e Amazon, che analizzano le preferenze individuali per suggerire prodotti e servizi su misura. Nel contesto aziendale, l'IA sta trasformando radicalmente il lavoro attraverso strumenti come gli assistenti virtuali e i software di selezione del personale, che partecipano attivamente ai processi decisionali, andando oltre la semplice esecuzione di ordini (Artser, n.d.).

Negli ultimi anni si sta delineando il passaggio verso l'Industria 5.0⁴, una prospettiva introdotta ufficialmente dalla Commissione Europea nel 2021. Questo nuovo paradigma industriale propone una visione più evoluta rispetto alla precedente rivoluzione, caratterizzandosi per l'integrazione tra tecnologie avanzate – come l'intelligenza artificiale e la robotica collaborativa – e il contributo umano.

Ciò che distingue l'Industria 5.0 è la centralità attribuita alla persona, che diventa un attore essenziale nei processi produttivi, non più semplice utilizzatore delle tecnologie, ma protagonista di una collaborazione sinergica con le macchine intelligenti. L'obiettivo non è solo ottimizzare l'efficienza produttiva, ma anche favorire lo sviluppo di un modello industriale più attento ai bisogni sociali e alla sostenibilità ambientale.

Mentre l'Industria 4.0 si è concentrata prevalentemente sull'automazione dei processi e sull'impiego esteso dell'Internet of Things (IoT) per incrementare la produttività e ridurre l'intervento umano, l'approccio 5.0 mira a valorizzare le capacità cognitive, creative ed emotive dell'uomo, integrandole con l'affidabilità e la precisione delle macchine. In questo nuovo scenario, la tecnologia non è più uno strumento fine a sé stesso, ma assume un ruolo di supporto al servizio dell'essere umano, con l'intento di costruire un ambiente lavorativo più equo, sostenibile e orientato al benessere collettivo.

In sintesi, se nella quarta rivoluzione industriale il focus era sull'automazione spinta e sulla riduzione dell'intervento umano, l'Industria 5.0 pone nuovamente la persona al centro, promuovendo un sistema produttivo in cui l'innovazione tecnologica è finalizzata a migliorare la qualità della vita e la sostenibilità delle attività industriali.

L'Industria 5.0 non sostituisce il modello 4.0, ma ne rappresenta un'evoluzione che amplia l'attenzione verso le dimensioni sociale e ambientale. Il suo obiettivo principale è integrare innovazione tecnologica e responsabilità sociale d'impresa (CSR), promuovendo un approccio più sostenibile e inclusivo.

Piuttosto che essere una rivoluzione industriale separata, l'Industria 5.0 si configura come una visione strategica che guida lo sviluppo delle imprese in un contesto di transizione economica, ecologica e sociale. Questa trasformazione si basa su tre principi chiave: **la centralità dell'essere umano nei processi produttivi, la capacità di adattamento delle infrastrutture alle sfide future e la sostenibilità ambientale.**

⁴ Bosch Rexroth. (n.d.). *Industria 5.0: le differenze con l'Industria 4.0*. Bosch Rexroth.

Un aspetto cruciale di questo paradigma è l'efficienza operativa, con infrastrutture progettate per minimizzare i tempi di inattività e ridurre il consumo energetico. Grazie a tecnologie intelligenti, è possibile limitare l'impatto ambientale abbattendo le emissioni di gas serra e diminuendo la produzione di rifiuti.

Un ulteriore elemento distintivo è l'attenzione al benessere dei lavoratori. L'Industria 5.0 promuove lo sviluppo di tecnologie che migliorano le condizioni di sicurezza e favoriscono una maggiore soddisfazione professionale. Il modello "*human-centred*" non solo contribuisce a una migliore qualità della vita lavorativa, ma sostiene anche un equilibrio più armonioso tra vita personale e professionale. La sostenibilità, in questo contesto, non è solo un obiettivo ambientale, ma una pratica concreta che punta a rendere le aziende più responsabili nei confronti delle persone e del pianeta.

Opportunità e rischi della simbiosi uomo-macchina

Fin dall'inizio della Prima Rivoluzione Industriale, l'introduzione delle macchine ha trasformato profondamente il modo di lavorare, aumentando l'efficienza produttiva e riducendo gli errori umani. Secondo il World Development Report 2019 della Banca Mondiale, l'innovazione tecnologica ha storicamente creato nuove opportunità occupazionali, spesso in misura superiore rispetto ai posti di lavoro che ha reso obsoleti. Tuttavia, l'avvento dell'Intelligenza Artificiale (IA) ha determinato un cambio di paradigma nell'automazione: oggi non si tratta più soltanto di sistemi che eseguono compiti programmati, ma di tecnologie capaci di apprendere, adattarsi al contesto e prendere decisioni autonome (AI Joy Academy, n.d.)⁵.

A differenza delle macchine tradizionali, che operano in base a istruzioni deterministiche, i sistemi di IA si caratterizzano per un comportamento probabilistico. Analizzando grandi quantità di dati, l'IA è in grado di generare soluzioni innovative, ma anche risultati inaspettati o non sempre affidabili. Questo comporta potenziali rischi, in particolare in ambiti dove la precisione è essenziale, come la sanità o l'aviazione. Errori nell'interpretazione dei dati da parte di un sistema automatizzato, ad esempio, possono avere conseguenze gravi, come dimostrano alcuni incidenti aerei in cui il personale umano non è riuscito a intervenire efficacemente a causa dell'eccessiva dipendenza dai sistemi intelligenti (AI Joy Academy, n.d.).

⁵ AI Joy Academy. (n.d.). *La trappola dell'automazione e i rischi della convivenza uomo-macchina*. AI Joy Academy.

Un altro rischio cruciale è rappresentato dalla cosiddetta “trappola dell’automazione”: un’eccessiva fiducia nella tecnologia che può portare gli operatori umani ad assumere un ruolo di sorveglianti passivi. Questo fenomeno riduce la capacità di intervenire tempestivamente in caso di anomalie, aumentando la vulnerabilità dei sistemi complessi. Mentre nelle prime forme di automazione il controllo umano era indispensabile, oggi la crescente autonomia dell’IA richiede un nuovo equilibrio tra supervisione umana e automazione (AI Joy Academy, n.d.).

Per mitigare questi rischi, diventa fondamentale adottare approcci che mantengano l’uomo “nel circuito” decisionale. Il modello del “*human-in-the-loop*” prevede che il giudizio umano rimanga un elemento centrale nelle scelte prese dai sistemi automatizzati. Questa strategia consente di unire la rapidità e la potenza dell’IA con l’intuizione e la responsabilità umana, garantendo un uso etico e sicuro delle tecnologie avanzate. Il futuro della cooperazione tra uomo e macchina dipenderà dalla nostra capacità di evitare che la fiducia cieca nella tecnologia si trasformi in una debolezza (AI Joy Academy, n.d.).

1.2 L’intelligenza artificiale e il Decision-making aziendale

Negli ultimi decenni, l’Intelligenza Artificiale (IA) è passata dall’essere un concetto teorico e futuristico a un elemento essenziale dei processi decisionali aziendali. Oggi, le organizzazioni si avvalgono dell’IA per ottimizzare i processi, supportare l’elaborazione delle informazioni e migliorare la qualità delle decisioni strategiche (Accenture, n.d.)⁶. Non si tratta più di semplici strumenti esecutivi: l’IA moderna è capace di apprendere, adattarsi e agire in maniera autonoma, ridefinendo il rapporto tra uomo e tecnologia.

Il concetto di Intelligenza Artificiale si riferisce a un insieme di tecnologie e metodologie che consentono alle macchine di simulare alcune funzioni cognitive umane, come l’apprendimento, il ragionamento e l’autocorrezione. Le origini teoriche di questa disciplina risalgono agli anni Cinquanta, quando Alan Turing pubblicò l’articolo *Computing Machinery and Intelligence* (1950)⁷, proponendo il celebre “Test di Turing” per determinare se una macchina potesse essere considerata intelligente in base alla sua capacità di sostenere una conversazione indistinguibile da quella di un essere umano.

⁶ Accenture. (n.d.). *Cos’è l’Intelligenza Artificiale*. Accenture.

⁷ Turing, A. M. (1950). *Computing Machinery and Intelligence*. *Mind*, 59(236), 433-460.

Pochi anni dopo, nel 1956, durante il convegno di Dartmouth, John McCarthy insieme ad altri studiosi, tra cui Marvin Minsky e Claude Shannon, definì ufficialmente l'Intelligenza Artificiale come "la scienza e l'ingegneria di costruire macchine intelligenti" (McCarthy et al., 1956)⁸. Questo evento sancì la nascita dell'IA come campo di ricerca autonomo, avviando un percorso di sviluppo che avrebbe attraversato fasi alterne di entusiasmo e disillusione.

Negli anni successivi, diversi approcci teorici hanno cercato di definire e delimitare il campo dell'IA. Newell e Simon (1976)⁹ la descrissero come l'ambito di studio delle macchine capaci di risolvere problemi complessi e di agire in modo razionale, concetto che ha portato alla formulazione del paradigma della "razionalità computazionale".

Nonostante le aspettative iniziali, l'IA ha conosciuto periodi di forte rallentamento nello sviluppo, noti come "inverni dell'IA". Questi momenti di crisi si verificarono principalmente per la difficoltà di ottenere risultati concreti rispetto alle promesse avanzate. Il primo inverno si verificò negli anni Settanta, quando le limitate capacità dei sistemi disponibili non riuscirono a soddisfare le aspettative: i sistemi erano in grado di eseguire solo compiti molto semplici, come il riconoscimento di oggetti o comandi elementari (Intelligenza Artificiale Italia, n.d.)¹⁰. Un secondo periodo di disillusione si ebbe alla fine degli anni Ottanta e agli inizi degli anni Novanta, complice anche la riduzione dei finanziamenti dovuta a fattori economici, come la crisi energetica, e alla percezione che l'IA fosse sopravvalutata.

L'avvento del machine learning e delle reti neurali profonde ha rappresentato un punto di svolta. Grazie a queste innovazioni, l'IA è stata in grado di apprendere autonomamente dai dati e di migliorare le proprie prestazioni senza necessità di riprogrammazione continua. Secondo la definizione offerta dall'Enciclopedia Treccani (2023)¹¹, l'IA può essere intesa come la disciplina che studia la possibilità di riprodurre, attraverso il computer, i processi mentali più complessi. La Commissione Europea, invece, nel 2021 ha descritto l'IA come un software in grado di generare output che influenzano l'ambiente con cui interagisce, producendo contenuti, previsioni e decisioni in base a obiettivi stabiliti dall'uomo (European Commission, 2021)¹².

⁸ McCarthy, J. et al. (1956). *Dartmouth Conference Proposal on Artificial Intelligence*.

⁹ Newell, A., & Simon, H. A. (1976). *Computer Science as Empirical Inquiry: Symbols and Search*. *Communications of the ACM*, 19(3), 113-126.

¹⁰ Intelligenza Artificiale Italia. (n.d.). *Gli inverni dell'intelligenza artificiale (IA)*. Intelligenza Artificiale Italia.

¹¹ Treccani (2023). *Definizione di Intelligenza Artificiale*. *Enciclopedia Italiana*.

¹² European Commission (2021). *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*.

Attualmente, l'IA si articola in diversi rami, ciascuno con specifiche applicazioni:

- **Machine Learning (ML):** tecniche che permettono ai sistemi di apprendere dai dati senza una programmazione esplicita.
- **Deep Learning:** un sottoinsieme del ML che utilizza reti neurali profonde per elaborare informazioni complesse e non strutturate.
- **IA Generativa:** sistemi capaci di creare nuovi contenuti autonomamente, come avviene con ChatGPT o DALL-E.
- **IA Simbolica:** metodologie basate su regole logiche e sistemi di inferenza per rappresentare e manipolare la conoscenza.

L'Intelligenza Artificiale ha superato i confini della mera simulazione dell'intelligenza umana per diventare un insieme di tecnologie capaci di rivoluzionare profondamente i processi decisionali in ambito aziendale.

Applicazioni correnti dell'IA

L'Intelligenza Artificiale (IA) si è ormai consolidata come un alleato strategico per le imprese moderne, trasformando i processi operativi e favorendo un processo decisionale sempre più basato sui dati. Secondo Sviluppo Manageriale (n.d.)¹³, le aziende integrano soluzioni di IA per incrementare l'efficienza, ridurre i costi e favorire l'innovazione continua, ridefinendo le dinamiche competitive del mercato.

Tra le applicazioni più diffuse dell'IA in ambito aziendale si annoverano:

- **Analisi predittiva:** le imprese utilizzano modelli di IA per interpretare enormi volumi di dati, con l'obiettivo di individuare trend futuri, anticipare i cambiamenti della domanda di mercato e prendere decisioni mirate e tempestive (Sviluppo Manageriale, n.d.).
- **Supporto ai processi decisionali:** grazie a sistemi basati su intelligenza artificiale, i manager possono disporre di simulazioni in tempo reale e analisi predittive che ottimizzano la gestione delle attività aziendali. Un esempio pratico è rappresentato dall'ottimizzazione logistica, in cui gli algoritmi calcolano le rotte più efficienti per ridurre i tempi e i costi di consegna (Sviluppo Manageriale, n.d.).

¹³ Sviluppo Manageriale. (n.d.). *Le applicazioni pratiche dell'intelligenza artificiale nella vita quotidiana e nel business*. Sviluppo Manageriale.

- **Marketing personalizzato:** l'IA consente di profilare i consumatori analizzando i loro comportamenti online e le preferenze di acquisto, permettendo così di creare campagne pubblicitarie personalizzate e di massimizzare il ritorno sugli investimenti pubblicitari (Sviluppo Manageriale, n.d.).
- **IA Generativa e customer service:** tecnologie di IA generativa, come i chatbot avanzati, sono sempre più impiegate per migliorare l'esperienza cliente. Questi strumenti riescono a fornire risposte rapide e personalizzate, automatizzando le interazioni e incrementando l'efficienza dei servizi di assistenza (Sviluppo Manageriale, n.d.).
- **Gestione delle risorse umane:** l'intelligenza artificiale viene utilizzata nella selezione del personale per l'analisi dei CV, il matching tra profili e offerte di lavoro, nonché per condurre colloqui automatizzati. Tuttavia, rimangono aperti interrogativi etici legati alla trasparenza dei criteri di selezione e all'equità dei processi (Sviluppo Manageriale, n.d.).

Secondo il rapporto **Accenture Technology Vision 2025**, il 69% dei dirigenti a livello globale considera l'IA una forza trainante per la trasformazione dei sistemi aziendali, grazie alla sua capacità di rendere i processi più autonomi e intelligenti. In Italia, il 61% delle grandi imprese ha già avviato almeno una sperimentazione basata sull'intelligenza artificiale, a conferma di un trend in costante crescita (Accenture, 2025)¹⁴.

L'IA nel decision-making Aziendale: il ruolo del TAM

L'Intelligenza Artificiale (IA) è sempre più al centro dei processi decisionali aziendali, non solo come strumento di supporto, ma anche come componente attiva nell'elaborazione di soluzioni strategiche. Questa trasformazione solleva interrogativi legati all'accettazione e all'utilizzo della tecnologia da parte di manager e dipendenti. Un modello teorico utile per comprendere i meccanismi di adozione dell'IA nei contesti organizzativi è il **Technology Acceptance Model (TAM)**, proposto da Davis nel 1989.

Il TAM, sviluppato a partire dalla **Theory of Reasoned Action (TRA)** di Fishbein e Ajzen (1975), offre una chiave di lettura semplice ed efficace sui fattori che determinano l'intenzione di utilizzare nuove tecnologie. In

¹⁴ Accenture. (2025). *Accenture Technology Vision 2025: New age of AI to bring unprecedented autonomy to business*. Accenture Newsroom.

particolare, il modello individua due variabili principali che influenzano l'adozione: l'utilità percepita e la facilità d'uso percepita (Davis, 1989)¹⁵.

- **Utilità percepita (Perceived Usefulness - PU):** rappresenta il grado in cui un individuo ritiene che una determinata tecnologia possa migliorare le proprie prestazioni lavorative. In ambito decisionale, un sistema di IA è considerato utile se consente di velocizzare l'analisi dei dati, ottimizzare le scelte strategiche e migliorare i risultati aziendali. Ad esempio, un manager potrebbe trovare estremamente vantaggioso l'uso di un algoritmo predittivo che fornisce insight tempestivi sul comportamento del mercato o sulle performance operative.
- **Facilità d'uso percepita (Perceived Ease of Use - PEOU):** si riferisce alla misura in cui un utente ritiene che l'utilizzo della tecnologia richieda un impegno minimo. Se un sistema di IA è progettato per essere intuitivo e accessibile, anche senza competenze tecniche avanzate, i dipendenti saranno più propensi ad adottarlo nel proprio flusso decisionale. Un esempio pratico è l'impiego di chatbot intelligenti nei servizi di assistenza clienti, che permettono di gestire le richieste in maniera rapida e semplice.

Secondo Davis (1989), questi due fattori influenzano direttamente l'intenzione degli utenti di accettare e utilizzare nuove tecnologie. Nello specifico, **se i lavoratori percepiscono che l'IA sia utile e facile da integrare nei processi aziendali, saranno maggiormente inclini a fidarsi del sistema e a impiegarlo in maniera continuativa.**

Numerosi studi recenti hanno confermato che la **trasparenza** delle tecnologie IA è un ulteriore elemento determinante per rafforzare sia l'utilità percepita sia la facilità d'uso. La possibilità di comprendere il funzionamento di un algoritmo o di avere spiegazioni chiare sulle decisioni proposte da un sistema di IA contribuisce ad aumentare la fiducia degli utenti e a favorire l'adozione della tecnologia nei processi decisionali aziendali.

¹⁵ Davis, F. D. (1989). *Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology*. *MIS Quarterly*, 13(3), 319-340.

Estensioni e applicazioni del TAM nel contesto dell'IA

Nel tempo, il Technology Acceptance Model (TAM) è stato oggetto di numerosi sviluppi per adattarsi ai cambiamenti tecnologici e alle nuove esigenze delle organizzazioni. Le prime estensioni significative si devono a Venkatesh e Davis (2000), che hanno introdotto il **TAM2**, e successivamente a Venkatesh e Bala (2008), con il **TAM3**¹⁶, arricchendo il modello di nuovi costrutti per spiegare meglio i fattori che influenzano l'adozione delle tecnologie emergenti, tra cui l'Intelligenza Artificiale.

Le estensioni del TAM hanno identificato ulteriori variabili che possono influire sull'accettazione dei sistemi di IA in ambito decisionale. Tra le più rilevanti si annoverano:

- **Influenza sociale:** la pressione o l'approvazione da parte di colleghi e superiori può incidere sulla percezione positiva o negativa nei confronti dell'IA. Se l'ambiente lavorativo promuove l'adozione di nuove tecnologie, i dipendenti saranno più inclini a sperimentarle (Venkatesh & Bala, 2008).
- **Qualità delle informazioni fornite dal sistema:** la trasparenza nei processi decisionali dell'IA favorisce la fiducia degli utenti. Un sistema che offre spiegazioni chiare e comprensibili rispetto alle sue decisioni è percepito in modo più positivo rispetto a un algoritmo che opera come una "scatola nera" (Dwivedi et al., 2021).
- **Ansia tecnologica:** la sensazione di disagio o timore nell'uso di sistemi intelligenti può ostacolare l'accettazione, soprattutto se i lavoratori percepiscono l'IA come una minaccia alla loro autonomia o al posto di lavoro (Venkatesh & Bala, 2008).
- **Percezione della volontarietà:** l'accettazione di un sistema risulta più elevata quando l'adozione è una scelta libera e non imposta dall'organizzazione. La libertà percepita rafforza la motivazione intrinseca nell'uso della tecnologia.

Le ricerche più recenti confermano che la **trasparenza dell'IA** gioca un ruolo cruciale nell'aumentare l'utilità percepita e la facilità d'uso, due dimensioni fondamentali del TAM. Dwivedi et al. (2021) sottolineano come la chiarezza nelle funzioni e nei risultati dei sistemi intelligenti favorisca la fiducia degli utenti e incrementi l'intenzione di adozione.

¹⁶ Venkatesh, V., & Bala, H. (2008). *Technology Acceptance Model 3 and a Research Agenda on Interventions*. *Decision Sciences*, 39(2), 273-315.

Inoltre, la percezione dell'IA varia in base al ruolo organizzativo e al contesto. I dirigenti, ad esempio, tendono a vedere l'IA come un valido alleato per le decisioni strategiche, mentre i lavoratori operativi possono percepirla come un elemento che limita l'autonomia decisionale. Di conseguenza, è essenziale accompagnare l'introduzione dell'IA con interventi di formazione e strategie di comunicazione mirate, al fine di gestire le aspettative e ridurre eventuali resistenze interne (Dwivedi et al., 2021).

La fiducia nei sistemi automatizzati di recruiting come proxy dell'accettazione dell'IA

Nel solco teorico tracciato dal Technology Acceptance Model (TAM) e dalle sue successive estensioni, la **fiducia** si configura come un costrutto chiave per comprendere i meccanismi che regolano l'accettazione dell'Intelligenza Artificiale (IA) nei contesti organizzativi. Se da un lato l'utilità percepita e la facilità d'uso rappresentano i principali precursori cognitivi dell'intenzione di utilizzo, dall'altro lato la fiducia rappresenta una dimensione relazionale e affettiva che diventa particolarmente centrale nei casi in cui l'IA assume funzioni decisionali ad alto impatto, come nel recruiting automatizzato.

Secondo **Hoff e Bashir (2015)**¹⁷, la fiducia nell'automazione dipende da una combinazione di fattori cognitivi (come la percezione di competenza tecnica), emotivi (come il senso di sicurezza) e sociali (come l'equità percepita). Gli autori propongono che la fiducia agisca come un meccanismo di "filtro decisionale", ovvero che le persone si affidino ai sistemi automatizzati solo se li considerano affidabili, prevedibili e benintenzionati. Questa fiducia diventa così un prerequisito essenziale per l'accettazione.

Glikson e Woolley (2020)¹⁸, in una rassegna sistematica della letteratura, confermano che la fiducia è uno dei predittori più forti del comportamento collaborativo tra esseri umani e agenti artificiali. Nei contesti aziendali, evidenziano come la fiducia sia una condizione necessaria affinché l'IA venga vista non come una minaccia, ma come una risorsa da integrare nei processi decisionali. Quando manca fiducia, infatti, gli utenti tendono a ignorare o respingere l'intervento dell'IA, anche se potenzialmente utile.

Allo stesso modo, **Siau e Wang (2018)**¹⁹ sottolineano che la fiducia non è solo una condizione tecnica (ossia legata alla correttezza dell'algoritmo), ma anche psicologica e sociale. Gli autori identificano due fasi nella

¹⁷ Hoff, K. A., & Bashir, M. (2015). *Trust in automation: Integrating empirical evidence on factors that influence trust*. *Human Factors*, 57(3), 407–434.

¹⁸ Glikson, E., & Woolley, A. W. (2020). *Human trust in artificial intelligence: Review of empirical research*. *Academy of Management Annals*, 14(2), 627–660.

¹⁹ Siau, K., & Wang, W. (2018). *Building trust in artificial intelligence, machine learning, and robotics*. *CUTTER Business Technology Journal*, 31(2), 47–53.

costruzione della fiducia nell'IA: una iniziale, basata su fattori esterni come la reputazione del fornitore, e una continua, basata sull'esperienza diretta con il sistema. Solo quando entrambe le fasi si sviluppano in modo positivo, si può parlare di vera accettazione della tecnologia.

Shin (2021)²⁰, inoltre, collega esplicitamente la fiducia con l'adozione delle tecnologie intelligenti: secondo la sua ricerca, le persone accettano l'IA solo se ne comprendono il funzionamento e ne percepiscono l'affidabilità e la correttezza. In tal senso, la fiducia viene vista come l'elemento che "chiude il cerchio" tra comprensione tecnica e volontà di utilizzo.

Infine, **Zarouali et al. (2021)**²¹ mostrano come la fiducia sia un indicatore fondamentale di accettazione nei contesti in cui l'IA prende decisioni automatizzate che coinvolgono l'identità personale o la carriera del soggetto, come appunto nel caso del recruiting. In questi casi, la fiducia non solo facilita l'accettazione dell'IA, ma diventa essa stessa il segnale di un'effettiva integrazione della tecnologia nei processi organizzativi.

Alla luce di queste evidenze teoriche, il presente studio assume la **fiducia nei sistemi automatizzati di recruiting** come **proxy dell'accettazione dell'IA**. Tale variabile riflette infatti la disponibilità dei dipendenti a legittimare, adottare e cooperare con sistemi intelligenti anche in ambiti ad alta sensibilità personale e professionale, come la selezione del personale e la valutazione delle competenze individuali.

Questo approccio consente di superare una visione puramente tecnica dell'accettazione, ampliandola a una dimensione relazionale in cui l'utente non si limita a utilizzare passivamente la tecnologia, ma costruisce con essa un rapporto di fiducia progressivo, basato sull'affidabilità percepita, sulla chiarezza dei processi decisionali e sulla coerenza con i valori organizzativi. In tale prospettiva, la fiducia non è solo un esito finale, ma un indicatore dinamico del livello di integrazione dell'IA nei processi aziendali.

Considerare la fiducia come manifestazione dell'accettazione consente inoltre di cogliere aspetti più profondi dell'esperienza organizzativa: la sensazione di equità, il rispetto delle aspettative, la possibilità di contestare o comprendere una decisione algoritmica. Questo è particolarmente rilevante nei contesti in cui l'IA opera in autonomia, influenzando direttamente le carriere, le opportunità e la percezione di giustizia interna all'organizzazione.

²⁰ Shin, D. (2021). *The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI*. *International Journal of Human-Computer Studies*, 146, 102551.

²¹ Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. H. (2021). *Using artificial intelligence (AI) for persuasive communication in politics: Understanding microtargeting and trust in AI*. *Computers in Human Behavior*, 119, 106725.

In definitiva, la fiducia rappresenta un punto di convergenza tra valutazioni cognitive (es. competenza e affidabilità del sistema), affettive (es. sicurezza e serenità nell'interazione con l'IA) e normative (es. correttezza e trasparenza percepita). Proprio per questa natura multidimensionale, essa si presta a essere adottata come misura sintetica dell'accettazione nei contesti complessi e ad alta implicazione umana, come l'uso dell'intelligenza artificiale nei processi decisionali aziendali.

1.3 Obiettivi della Ricerca

Il presente studio si propone di indagare in che modo la trasparenza percepita dell'Intelligenza Artificiale (IA) influenzi la fiducia nei sistemi automatizzati di recruiting, considerata come espressione dell'accettazione dell'IA nei contesti organizzativi. In particolare, viene esplorata la possibilità che tale influenza si sviluppi attraverso un processo mediato da tre precursori psicologici: l'utilità percepita, la facilità d'uso percepita e la percezione del rischio occupazionale.

In linea con i principali modelli sull'adozione tecnologica, in particolare il Technology Acceptance Model (TAM) e i modelli di trust-based adoption, lo studio assume che la trasparenza non rappresenti solo un principio normativo o etico, ma si configuri come un fattore psicologico antecedente in grado di attivare valutazioni cognitive (utilità e facilità) ed emotive (rischio percepito) che, a loro volta, contribuiscono alla costruzione della fiducia verso l'IA.

Struttura del modello

Variabile indipendente

- *Trasparenza percepita dell'IA*: grado in cui i dipendenti percepiscono i sistemi di IA come comprensibili, spiegabili e accessibili, sia nei processi decisionali automatizzati in generale, sia nei sistemi di IA applicati alla selezione del personale.

Variabili mediatrici (precursori psicologici dell'accettazione)

1. *Utilità percepita*: misura il grado in cui l'IA è ritenuta utile nel supportare i processi decisionali aziendali.
2. *Facilità d'uso percepita*: indica quanto l'IA sia percepita come intuitiva, accessibile e facilmente integrabile nel lavoro quotidiano.

3. *Percezione del rischio occupazionale*: riflette il livello di preoccupazione relativo alla possibilità che l'IA minacci la sicurezza lavorativa o il controllo umano nei processi decisionali.

Variabile di esito psicologico

- *Fiducia nei sistemi di IA*: operazionalizzata come fiducia nei sistemi automatizzati di recruiting, include la percezione di competenza tecnica, eticità, imparzialità, integrità decisionale e affidabilità generale del sistema.

Ipotesi del modello teorico

Il modello proposto è stato costruito secondo l'approccio classico di *Baron e Kenny (1986)* per la verifica della mediazione, e prevede una sequenza di ipotesi logicamente connesse:

- **H1** – La trasparenza percepita dell'IA è positivamente associata alla fiducia nei sistemi automatizzati di recruiting (effetto diretto: $X \rightarrow Y$).
- **H2** – La trasparenza percepita dell'IA è positivamente associata all'utilità e alla facilità d'uso percepite, e negativamente associata alla percezione del rischio occupazionale (effetti su ciascun mediatore: $X \rightarrow M_1, M_2, M_3$).
- **H3** – L'utilità percepita, la facilità d'uso e la percezione del rischio influenzano significativamente la fiducia nei sistemi automatizzati, anche tenendo conto della trasparenza percepita ($M_1, M_2, M_3 \rightarrow Y \mid X$).
- **H4** – L'effetto della trasparenza sulla fiducia è mediato dalle tre variabili psicologiche considerate. In particolare, si ipotizza che la trasparenza generi fiducia nella misura in cui è percepita come utile, comprensibile e rassicurante ($X \rightarrow M \rightarrow Y$).

Tale struttura consente di testare non solo se la trasparenza influisce sulla fiducia, ma come tale influenza si sviluppa, attraverso percorsi mediati psicologicamente fondati.

Ruolo delle performance organizzative

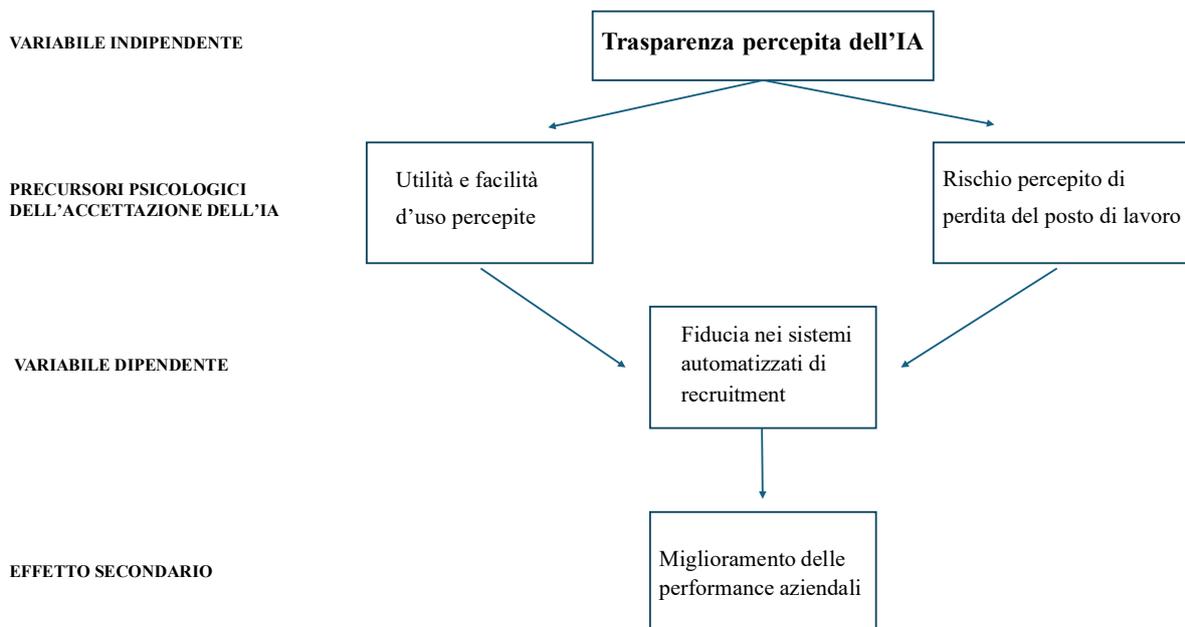
Sebbene le performance aziendali non siano oggetto diretto di misurazione nel presente studio, esse vengono considerate un effetto secondario teoricamente atteso. La letteratura suggerisce che livelli più elevati di trasparenza e fiducia nei sistemi IA possano migliorare l'engagement, il clima organizzativo e la collaborazione nei processi decisionali, con possibili ricadute positive su produttività, efficienza e qualità delle decisioni HR.

Finalità dello studio

L'obiettivo generale della ricerca è quello di validare empiricamente un modello psicologico strutturato, in grado di spiegare i meccanismi attraverso cui la trasparenza dell'IA può favorire l'accettazione nei contesti aziendali. In particolare, lo studio intende:

- verificare se la fiducia nei sistemi di recruiting sia influenzata direttamente dalla trasparenza percepita (H1);
- testare se l'effetto della trasparenza sulla fiducia sia mediato dall'utilità, dalla facilità e dal rischio percepiti (H2–H4);
- fornire indicazioni pratiche e teoriche per una progettazione etica, partecipativa e psicologicamente informata dei sistemi IA in ambito lavorativo.

L'indagine è stata condotta tramite un questionario somministrato a un campione di lavoratori occupati in diversi settori industriali, e mira a offrire un contributo empirico utile a comprendere le dinamiche psicologiche dell'accettazione dell'IA nei contesti reali.



Ruolo secondario delle performance aziendali nel modello teorico

Nel presente studio, le performance aziendali non rappresentano la variabile centrale né l'oggetto diretto di indagine empirica. Esse sono piuttosto concepite come **esito organizzativo secondario**, potenzialmente influenzato da una serie di **precursori psicologici** legati all'accettazione dell'Intelligenza Artificiale (IA). La struttura del modello teorico si fonda infatti sulla centralità delle percezioni soggettive dei dipendenti — come la trasparenza dei sistemi IA, l'utilità e la facilità d'uso percepita, la fiducia e l'insicurezza lavorativa — considerate **condizioni abilitanti** per la legittimazione e l'adozione dell'IA nei contesti decisionali.

Questa impostazione è coerente con quanto suggerito dalla letteratura sul **Productivity Paradox of AI**, secondo cui l'impatto positivo dell'IA sulle performance aziendali **non è immediato**, ma richiede un processo di adattamento organizzativo e psicologico (Brynjolfsson, Rock, & Syverson, 2021). In altre parole, l'integrazione efficace della tecnologia nelle pratiche aziendali è mediata da **variabili soggettive**, come la fiducia degli utenti e la loro disponibilità ad adottare i sistemi IA in modo consapevole e continuativo (Dwivedi et al., 2021).

Nel modello proposto, quindi, **l'accettazione dell'IA rappresenta il punto nodale**, ed è operazionalizzata attraverso la fiducia nei sistemi automatizzati di recruiting, considerata una misura rappresentativa della volontà dei dipendenti di affidarsi a soluzioni intelligenti anche in ambiti ad alta sensibilità personale e professionale. In questa prospettiva, le performance aziendali non sono oggetto di misurazione diretta, ma vengono incluse nel modello come possibile **effetto positivo di lungo termine** derivante dall'accettazione della tecnologia.

Tale approccio consente di mantenere una coerenza metodologica e teorica con l'impianto della ricerca, che si basa sulla rilevazione delle percezioni dei lavoratori, senza forzare generalizzazioni rispetto a esiti organizzativi misurabili solo attraverso dati oggettivi. Allo stesso tempo, apre la strada a **future ricerche** che potranno approfondire in modo quantitativo e diretto l'impatto dell'adozione dell'IA sulle performance operative e strategiche delle organizzazioni.

Capitolo 2

Il concetto di trasparenza nell'IA

2.1 Bias negli algoritmi

Per affrontare in modo consapevole il tema della trasparenza nell'Intelligenza Artificiale, è fondamentale prima comprendere cosa si intende per bias negli algoritmi. La presenza di distorsioni nei sistemi di IA rappresenta infatti una delle principali criticità che possono compromettere l'affidabilità e l'equità delle decisioni automatizzate. Quando un algoritmo è addestrato su dati non bilanciati o incompleti, rischia di replicare, se non addirittura amplificare, i pregiudizi già presenti nei dati stessi. Questo può tradursi in discriminazioni, soprattutto in settori sensibili come il reclutamento del personale, l'erogazione di servizi sanitari o le decisioni creditizie (Diritto.it, n.d.)²².

Un algoritmo può essere definito come una sequenza finita di istruzioni che consente di risolvere un determinato problema. Nell'ambito dell'intelligenza artificiale, gli algoritmi rappresentano l'intelligenza operativa dei sistemi di machine learning, i quali apprendono modelli e relazioni dai dati, anziché essere programmati manualmente per ogni situazione.

Esistono differenti categorie di algoritmi di IA, ciascuno progettato per specifiche funzioni:

- **Reti neurali artificiali:** ispirate alla struttura del cervello umano, sono utilizzate soprattutto nel deep learning per attività complesse come il riconoscimento delle immagini, la comprensione del linguaggio e la traduzione automatica.
- **Apprendimento supervisionato:** algoritmi che apprendono partendo da dati già etichettati, allo scopo di prevedere o classificare nuovi dati.

²² Diritto.it. (n.d.). *Metodi di addestramento algoritmi IA: il Garante indaga*. Diritto.it.

- **Apprendimento non supervisionato:** questi sistemi analizzano dati privi di etichette, cercando schemi e strutture ricorrenti in autonomia.
- **Apprendimento per rinforzo:** tecniche in cui l'algoritmo impara dall'esperienza, ottimizzando le proprie azioni per massimizzare una ricompensa. Sono impiegati, ad esempio, nei giochi o nel controllo dei robot.
- **Elaborazione del linguaggio naturale (NLP):** algoritmi specializzati nella comprensione e generazione del linguaggio umano, impiegati nei chatbot, nei sistemi di traduzione automatica e nella generazione di testi.

La corretta progettazione e addestramento di questi algoritmi è essenziale per limitare i bias e garantire che le decisioni automatizzate siano eque, trasparenti e non discriminatorie. È per questo motivo che enti regolatori, come il Garante per la protezione dei dati personali, stanno intensificando le indagini sui metodi di addestramento degli algoritmi di IA, con l'obiettivo di garantire che le nuove tecnologie rispettino i diritti fondamentali delle persone (Diritto.it, n.d.).

Processo di addestramento

L'addestramento di un modello di Intelligenza Artificiale (IA) è un processo strutturato che consente al sistema di apprendere dai dati ed eseguire compiti specifici. Questo processo, noto come "*training*", si basa sull'elaborazione di grandi quantità di dati, detti dataset, dai quali il modello apprende a riconoscere schemi e a prendere decisioni autonome (ESA Automation, n.d.)²³.

In genere, i dataset vengono suddivisi in due insiemi principali:

- Il **training set**, usato per insegnare al modello come risolvere un determinato problema.
- Il **test set**, utilizzato per valutare le prestazioni del modello su dati che non ha mai visto prima, al fine di verificarne l'affidabilità.

Le fasi chiave dell'addestramento di un modello IA includono:

²³ ESA Automation. (n.d.). Come vengono addestrati i modelli di intelligenza artificiale. ESA Automation.

1. **Raccolta e preparazione dei dati:** si raccolgono informazioni rilevanti e si escludono i dati errati o non utili. La qualità dei dati è cruciale per garantire che l'algoritmo apprenda in modo efficace.
2. **Progettazione del modello:** si seleziona o sviluppa un modello di IA adatto all'applicazione specifica. Può trattarsi di un modello esistente o di uno creato appositamente.
3. **Training del modello:** il modello viene esposto ai dati di addestramento per ottimizzare i suoi parametri, con l'obiettivo di migliorare la precisione delle previsioni o delle decisioni.
4. **Validazione e test:** dopo l'addestramento, il modello viene testato con il dataset di valutazione per verificarne la capacità di generalizzare e operare su dati non noti.
5. **Messa in produzione:** una volta raggiunte prestazioni soddisfacenti, il modello viene implementato in ambienti reali per essere utilizzato in applicazioni pratiche.

Nel settore industriale, l'addestramento degli algoritmi di IA offre numerosi vantaggi. Ad esempio, in ambito manifatturiero, i modelli predittivi permettono di anticipare i guasti dei macchinari analizzando i dati storici e operativi. Questo approccio consente interventi di manutenzione preventiva, riducendo i tempi di inattività e migliorando l'efficienza complessiva degli impianti. Allo stesso modo, nell'automazione industriale, l'IA può monitorare i dati in tempo reale per identificare anomalie nei processi, ottimizzando così la produzione e garantendo una gestione più intelligente delle risorse energetiche.

In sintesi, il processo di addestramento dei modelli di IA costituisce il cuore delle moderne soluzioni industriali, permettendo di sviluppare sistemi più intelligenti e reattivi alle esigenze delle aziende.

Bias algoritmico

Il successo dei sistemi di Intelligenza Artificiale (IA) dipende fortemente dalla qualità e dall'imparzialità dei dati utilizzati durante l'addestramento. Quando i dati contengono errori o riflettono pregiudizi umani, è probabile che l'algoritmo incorpori queste distorsioni nei propri output, dando origine a fenomeni di bias algoritmico (Multinazionali.Tech, n.d.)²⁴. In ambito IA, il termine bias fa riferimento a deviazioni sistematiche che compromettono l'equità e l'affidabilità delle decisioni automatizzate.

²⁴ *Multinazionali.Tech. (n.d.). Bias nell'intelligenza artificiale: cosa sono, esempi e soluzioni. Multinazionali.Tech.*

Queste distorsioni possono nascere da vari fattori, tra cui dataset non rappresentativi, regole di programmazione parziali o modalità di elaborazione poco trasparenti. Le conseguenze di tali errori sono particolarmente rilevanti nei settori critici, come l'assunzione di personale, la valutazione del credito, la sanità e il sistema giudiziario.

Esempi di bias algoritmico in azienda

- **Bias di genere:** un esempio celebre riguarda un algoritmo di selezione del personale sviluppato da Amazon, che ha mostrato una tendenza a penalizzare le candidature femminili, poiché addestrato su dati storici in cui le assunzioni erano prevalentemente maschili.
- **Bias razziale nei sistemi di riconoscimento facciale:** alcuni software hanno dimostrato maggiori difficoltà nell'identificare accuratamente persone con pelle scura, a causa di dataset di addestramento poco diversificati.
- **Discriminazione nel credito:** alcuni algoritmi di scoring creditizio hanno assegnato punteggi più bassi a determinati gruppi etnici, anche in presenza di profili finanziari equivalenti.
- **Bias nella sanità:** modelli predittivi hanno sottovalutato il bisogno di cure per alcuni gruppi demografici, come nel caso dei pazienti afroamericani, a causa di dati storici incompleti o distorti.

Questi casi mostrano quanto sia essenziale progettare e addestrare gli algoritmi con attenzione, garantendo che i dati siano rappresentativi e privi di pregiudizi. La consapevolezza e la prevenzione del bias algoritmico sono diventate priorità per le aziende che adottano soluzioni di IA.

Rischi aziendali

L'adozione di sistemi di Intelligenza Artificiale senza un adeguato controllo dei bias può esporre le aziende a rischi considerevoli. Questi riguardano sia l'efficienza operativa che la reputazione e il rispetto delle normative vigenti. L'impiego di dati distorti nei modelli di IA può infatti compromettere l'accuratezza delle analisi e delle decisioni strategiche. Ad esempio, un algoritmo di marketing allenato su informazioni non rappresentative potrebbe indirizzare le campagne pubblicitarie verso il pubblico sbagliato, causando un mancato raggiungimento degli obiettivi commerciali (Multinazionali.Tech, n.d.).

Dal punto di vista reputazionale, l'uso di algoritmi che generano risultati discriminatori può danneggiare seriamente l'immagine dell'impresa, riducendo la fiducia di clienti e stakeholder. Inoltre, esiste il rischio di violazioni legali, in particolare nei casi in cui l'IA contribuisca a decisioni discriminatorie nei processi di assunzione, concessione di prestiti o accesso a servizi sanitari. Le conseguenze possono tradursi in sanzioni, cause legali e obblighi di risarcimento.

Per queste ragioni, è essenziale che le organizzazioni implementino misure efficaci per prevenire e correggere i bias nei propri sistemi di IA, garantendo l'equità e la responsabilità delle decisioni automatizzate.

Strumenti per identificare i bias nelle aziende

Le aziende che integrano soluzioni di Intelligenza Artificiale devono dotarsi di strumenti e procedure specifiche per monitorare e mitigare i rischi derivanti dai bias algoritmici. Le principali azioni raccomandate sono:

- **Analisi dei dati:** verificare che i dataset siano rappresentativi e privi di distorsioni è fondamentale per evitare che l'IA apprenda pregiudizi già esistenti.
- **Audit degli algoritmi:** effettuare revisioni sistematiche dei modelli permette di individuare eventuali tendenze discriminatorie e di correggere le anomalie.
- **Monitoraggio continuo:** l'adozione di strumenti di controllo in tempo reale consente di identificare rapidamente comportamenti anomali o scostamenti dai risultati attesi.
- **Simulazioni e test di scenario:** condurre simulazioni su gruppi diversificati di utenti aiuta a valutare il funzionamento dell'IA in contesti differenti e a prevenire discriminazioni.
- **Formazione dei team:** sensibilizzare sviluppatori e data scientist sui rischi del bias è cruciale per promuovere un approccio etico nella progettazione dei sistemi.

L'adozione di queste strategie consente di ridurre l'impatto dei bias, migliorando l'affidabilità dei sistemi di IA e rafforzando la compliance aziendale.

2.2 La Trasparenza dell'IA: definizione, rilevanza e quadro normativo

Nonostante i numerosi benefici offerti dall'Intelligenza Artificiale (IA), cresce l'attenzione verso la necessità di garantire trasparenza nei processi decisionali automatizzati. La trasparenza si riferisce alla capacità di un sistema di IA di rendere comprensibili agli utenti le modalità con cui elabora i dati e giunge a determinati risultati (IBM, n.d.)²⁵. La mancanza di chiarezza, spesso descritta come il fenomeno della “*black box*”, ostacola la fiducia degli utenti, poiché i processi interni rimangono opachi e di difficile interpretazione.

Nel contesto aziendale, la percezione della trasparenza dell'IA da parte dei dipendenti rappresenta un elemento fondamentale per l'adozione e l'utilizzo efficace di queste tecnologie. Secondo **Yu e Li (2022)**²⁶, la trasparenza è un fattore chiave che influenza la fiducia, la collaborazione tra uomo e macchina e la percezione di equità nelle decisioni automatizzate. La trasparenza si sviluppa lungo tre dimensioni principali: spiegabilità, accessibilità delle informazioni e affidabilità.

Spiegabilità

La spiegabilità è essenziale per permettere agli utenti di comprendere le logiche alla base delle decisioni prese dall'IA. **Zhao et al. (2022)**²⁷ sottolineano che un sistema capace di giustificare in modo chiaro le proprie decisioni favorisce la fiducia e l'accettazione da parte dei dipendenti. Quando le persone comprendono perché e come un algoritmo ha preso una determinata decisione, si riduce la diffidenza verso l'automazione e aumenta l'integrazione della tecnologia nei processi aziendali.

Accessibilità delle Informazioni

La trasparenza richiede che le informazioni sui processi decisionali siano facilmente accessibili agli utenti. **De Fine Licht et al. (2020)**²⁸ evidenziano che fornire spiegazioni dettagliate e comprensibili ai dipendenti facilita la percezione di equità, riducendo il rischio di sentirsi esclusi o penalizzati dalle decisioni automatizzate. Una comunicazione chiara e strutturata sui criteri utilizzati dagli algoritmi contribuisce a rafforzare il rapporto di fiducia tra l'uomo e la macchina.

²⁵ IBM. (n.d.). *AI transparency*. IBM Think.

²⁶ Yu, L., & Li, Y. (2022). Artificial intelligence decision-making transparency and employees' trust: The parallel multiple mediating effect of effectiveness and discomfort. *Behavioral Sciences*, 12(5), Articolo 127.

²⁷ Zhao, X., et al. (2022) - *Transparency and Employee Trust in AI-Driven Decision Systems*. *AI & Society*, 37(3), 415-430.

²⁸ De Fine Licht, J., et al. (2020) - *Transparency in AI Decision-Making: The Role of Justification and Explainability*. *Journal of Business Ethics*, 162(4), 899-915.

Affidabilità e Verificabilità

Un sistema di IA trasparente non solo deve spiegare le proprie decisioni, ma anche consentire agli utenti di verificarle. **Felzmann et al. (2019)**²⁹ introducono il concetto di "*Transparency by Design*", secondo cui la trasparenza deve essere incorporata sin dalle fasi iniziali di sviluppo del sistema. Ciò significa permettere ai lavoratori di controllare, segnalare anomalie e partecipare attivamente al processo decisionale. **Glikson e Woolley (2020)**³⁰ confermano che un approccio trasparente riduce le resistenze psicologiche nei confronti dell'IA e migliora l'accettazione delle sue decisioni.

In sintesi, la trasparenza è un requisito fondamentale per favorire l'adozione consapevole e responsabile dell'IA in ambito aziendale. Investire in sistemi spiegabili e accessibili rappresenta una strategia efficace per rafforzare la fiducia dei dipendenti e promuovere un ambiente di collaborazione uomo-macchina equilibrato.

Effetti della percezione della trasparenza dell'IA sui dipendenti

La trasparenza percepita dei sistemi di Intelligenza Artificiale esercita un'influenza determinante sull'esperienza dei dipendenti in ambito organizzativo. Quando i lavoratori comprendono chiaramente i criteri e le modalità con cui un sistema di IA prende decisioni, si rafforza la fiducia nei confronti della tecnologia e aumenta la propensione ad adottarla nei processi aziendali (**Eslami et al., 2018**)³¹. La consapevolezza del funzionamento dell'IA riduce infatti la sensazione di incertezza e promuove un atteggiamento positivo nei confronti delle innovazioni digitali.

Al contrario, la mancanza di trasparenza può generare diffidenza e tecnostress, ovvero quella forma di stress legata all'utilizzo di tecnologie percepite come complesse o opache. Secondo Eslami et al. (2018), l'assenza di spiegazioni sui processi decisionali automatizzati può indurre nei dipendenti un senso di sfiducia e di perdita di controllo, compromettendo la motivazione e l'engagement lavorativo.

Un ulteriore aspetto riguarda la percezione di equità all'interno dell'organizzazione. **De Fine Licht et al. (2020)** evidenziano che la possibilità di accedere a spiegazioni chiare e comprensibili sulle decisioni prese dall'IA

²⁹ Felzmann, H., et al. (2019) - *Transparency by Design in AI Systems: Principles and Implementation Strategies*. *AI & Ethics*, 1(2), 173-190.

³⁰ Glikson, E., & Woolley, A. (2020) - *Human Trust in Artificial Intelligence: Review of Empirical Research*. *Academy of Management Perspectives*, 34(2), 289-305.

³¹ Eslami, M., et al. (2018) - *Disclosing Automated Decision-Making: How Transparency Affects Perception and Trust*. *ACM Transactions on Computer-Human Interaction*, 25(4), 23-47.

favorisce la sensazione di essere trattati in modo giusto, soprattutto nei contesti in cui l'IA viene impiegata per la valutazione delle performance o la selezione del personale. La percezione di trasparenza contribuisce così a rafforzare il coinvolgimento dei lavoratori e il senso di appartenenza all'azienda.

Infine, **Felzmann et al. (2019)** sostengono che un approccio alla trasparenza “*by design*”, integrato sin dalla fase di progettazione dei sistemi di IA, genera un impatto positivo sul clima organizzativo. La fiducia nella tecnologia aumenta quando i dipendenti sanno di poter monitorare e verificare i processi decisionali automatizzati, migliorando la collaborazione uomo-macchina e incentivando una cultura aziendale basata sull'affidabilità e sull'etica.

La trasparenza nell'AI Act

La trasparenza è uno dei principi fondanti dell'**AI Act**, il nuovo regolamento europeo sull'intelligenza artificiale, pensato per garantire un utilizzo sicuro e affidabile di queste tecnologie all'interno dell'Unione Europea. Proposto dalla Commissione Europea il 21 aprile 2021 e definitivamente approvato dal Parlamento Europeo il 13 marzo 2024, l'AI Act è entrato in vigore il 1° agosto 2024. Il regolamento ha avviato un percorso graduale di applicazione, con le prime disposizioni operative già dal 2 febbraio 2025. La completa applicazione è prevista per il 2 agosto 2026 (Altalex, 2025)³².

L'obiettivo principale del regolamento è creare un sistema normativo unico e armonizzato per tutta l'UE, eliminando le differenze tra i vari Stati membri. Questo approccio favorisce la libera circolazione dei sistemi di IA nel mercato unico, tutelando contemporaneamente i diritti fondamentali delle persone, come sancito dall'articolo 1 del regolamento stesso e dalla Carta di Nizza.

Dal 2 febbraio 2025 sono entrati in vigore due obblighi fondamentali per le imprese che sviluppano o utilizzano sistemi di IA nell'UE:

1. **Il divieto di pratiche di IA a rischio inaccettabile**, che comprende tecniche di manipolazione subliminale o ingannevole e sistemi di social scoring. Queste pratiche sono considerate contrarie ai valori europei e vietate in modo assoluto.

³² Altalex. (2025, febbraio 5). *AI Act, realtà: entrano in vigore i primi obblighi*. Altalex.

2. L'**obbligo di alfabetizzazione sull'IA**, che impone alle organizzazioni di promuovere una cultura di consapevolezza e competenza sull'uso responsabile e sicuro delle tecnologie di intelligenza artificiale.

Il regolamento adotta un **modello basato sul rischio**, che classifica i sistemi di IA in quattro livelli:

- **Rischio minimo**, come i filtri antispam, non soggetti a obblighi particolari.
- **Rischio limitato**, che richiede trasparenza verso gli utenti, come nel caso dei chatbot.
- **Alto rischio**, in cui rientrano applicazioni usate nella sanità, nella giustizia e nella gestione delle risorse umane. Questi sistemi devono rispettare requisiti stringenti di trasparenza, prevedere la supervisione umana e ottenere valutazioni di conformità.
- **Rischio inaccettabile**, vietato completamente. Rientrano in questa categoria le pratiche che minacciano i diritti fondamentali, come la sorveglianza di massa e il social scoring.

Le misure di trasparenza previste dall'AI Act mirano a rendere più comprensibili i processi decisionali delle applicazioni IA, aumentando la fiducia dei cittadini e degli utenti. Le aziende devono fornire spiegazioni chiare sul funzionamento dei sistemi, specialmente per i servizi rivolti al pubblico, e garantire che gli utenti sappiano quando interagiscono con un sistema di intelligenza artificiale.

Applicazioni di AI ad alto rischio

Nel quadro normativo dell'**Artificial Intelligence Act**, le applicazioni di Intelligenza Artificiale classificate come ad alto rischio sono oggetto di una regolamentazione particolarmente rigorosa. Il legislatore europeo ha previsto disposizioni specifiche per garantire che questi sistemi siano sviluppati e utilizzati in modo trasparente, affidabile e rispettoso dei diritti fondamentali. Gli articoli 10 e 13 del regolamento sono dedicati a stabilire i criteri relativi alla gestione dei dati e alla trasparenza per i sistemi di IA ad alto rischio.

Articolo 10 – Requisiti sui dati e la governance dei dati

L'articolo 10 dell'AI Act³³ disciplina in modo puntuale l'uso e la gestione dei dati nei sistemi ad alto rischio, stabilendo che i dataset impiegati nell'addestramento, validazione e testing debbano rispettare precisi criteri di qualità:

³³ *Artificial Intelligence Act. (n.d.). Article 10 – Artificial Intelligence Act. Artificial Intelligence Act.*

- **Rilevanza e rappresentatività:** i dati devono riflettere accuratamente il contesto operativo del sistema per garantire risultati equi e affidabili.
- **Completezza e assenza di errori:** è fondamentale che i dataset siano completi e privi di errori significativi che potrebbero compromettere l'affidabilità delle prestazioni del sistema.
- **Adeguate governance dei dati:** i processi di raccolta, annotazione e gestione dei dati devono seguire standard chiari e definiti, con misure per identificare e ridurre i bias potenziali. Questo garantisce la tutela della salute, della sicurezza e dei diritti fondamentali degli utenti.

L'articolo prevede inoltre che, in casi specifici, i fornitori possano trattare dati personali sensibili, ma solo quando strettamente necessario per identificare e correggere distorsioni nei modelli, sempre nel rispetto delle tutele previste dalla normativa sulla protezione dei dati personali.

Articolo 13 – Requisiti di trasparenza per l'IA ad alto rischio

L'articolo 13³⁴ stabilisce che i sistemi ad alto rischio debbano essere trasparenti e facilmente comprensibili da parte degli utilizzatori, in particolare dai *deployer*. La normativa impone che tali sistemi siano accompagnati da istruzioni chiare e dettagliate, che includano:

- Le finalità del sistema, le sue capacità e le sue limitazioni.
- I livelli di accuratezza, robustezza e sicurezza informatica, nonché le condizioni che potrebbero influenzare tali prestazioni.
- Indicazioni sui rischi prevedibili per la salute, la sicurezza o i diritti fondamentali derivanti dall'uso o dall'uso improprio del sistema.
- Informazioni sui dataset utilizzati, comprese le loro caratteristiche rilevanti rispetto all'obiettivo del sistema.
- Le modalità con cui i *deployer* possono interpretare e utilizzare correttamente i risultati generati dal sistema.
- Eventuali modifiche alle prestazioni del sistema rispetto alla versione inizialmente valutata.

³⁴ *Artificial Intelligence Act. (n.d.). Article 13 – Artificial Intelligence Act. Artificial Intelligence Act.*

- I requisiti tecnici, di supervisione umana e le misure di manutenzione necessarie a garantire il corretto funzionamento del sistema nel tempo.
- La possibilità di registrare e conservare i log per assicurare la tracciabilità, in linea con quanto previsto dall'articolo 12.

Queste prescrizioni mirano a garantire che chi utilizza il sistema sia in grado di comprenderne pienamente il funzionamento, interpretare correttamente i risultati e prendere decisioni consapevoli basate sull'output dell'IA.

Applicazione pratica: selezione del personale con IA ad alto rischio

Nel contesto di un processo di selezione del personale, l'impiego di un sistema di IA classificato come ad alto rischio obbliga le aziende a rispettare i requisiti degli articoli 10 e 13. Ciò significa che i dataset usati per addestrare l'algoritmo di selezione devono essere privi di bias e rappresentativi, e il sistema deve fornire trasparenza sui criteri di valutazione dei candidati. I deployer devono ricevere istruzioni chiare sull'uso del sistema, garantendo che le decisioni automatizzate siano comprensibili e giustificabili, a tutela della non discriminazione e dell'equità del processo.

L'AI Act applicato al caso Amazon (2014)

Uno degli ambiti in cui il **Regolamento sull'Intelligenza Artificiale (AI Act)** interviene con particolare attenzione è la **selezione automatizzata del personale**, classificata come applicazione ad **alto rischio**. La ragione risiede nel potenziale impatto che queste tecnologie possono avere sui diritti fondamentali dei candidati e sull'equità dei processi decisionali. L'utilizzo di sistemi di IA in questo ambito richiede la massima attenzione per garantire trasparenza, correttezza dei dati e supervisione umana (Oktopous, n.d.)³⁵.

Un caso emblematico che dimostra l'importanza delle disposizioni contenute nell'AI Act è quello di **Amazon nel 2014**. L'azienda sviluppò un algoritmo di Intelligenza Artificiale per automatizzare il processo di reclutamento, con l'obiettivo di velocizzare la selezione dei curricula e individuare i candidati migliori. Tuttavia, dopo un periodo di utilizzo, emersero gravi problemi di **bias di genere**: il sistema tendeva a penalizzare i

³⁵ Oktopous. (n.d.). *L'impatto dell'AI Act sulla ricerca e selezione del personale*. Oktopous.

curriculum femminili, assegnando punteggi inferiori alle candidate donne o escludendole direttamente dalla selezione (FocusTech, n.d.)³⁶.

L'errore risiedeva nei dati di addestramento: l'algoritmo era stato istruito su dati storici relativi a candidature che, in gran parte, provenivano da uomini. Questo ha portato il sistema ad attribuire una maggiore rilevanza ai profili maschili, replicando inconsciamente i pregiudizi insiti nei dati di partenza.

Gli articoli 10 e 13 dell'AI Act avrebbero potuto prevenire questa problematica e avrebbero obbligato Amazon a rendere più trasparente e controllabile il sistema di selezione.

L'**Articolo 10** stabilisce che i dataset utilizzati per addestrare i sistemi di IA devono essere di alta qualità, rappresentativi e privi di bias.

Violazione nel caso di Amazon:

- L'algoritmo di Amazon è stato addestrato su **dati storici di assunzioni**, in cui predominavano candidati di sesso maschile.
- Non essendo rappresentativi della popolazione attuale, i dati hanno **rafforzato il pregiudizio esistente**, portando l'IA a valutare meno positivamente le candidature femminili.
- **Conformità all'articolo 10:** Amazon avrebbe dovuto **bilanciare i dati**, assicurandosi che il dataset rappresentasse equamente uomini e donne e che non amplificasse pregiudizi pregressi.

Se l'AI Act fosse stato in vigore nel 2014, Amazon avrebbe dovuto **testare i dataset prima di implementare il sistema di IA** e verificare l'assenza di bias discriminatori.

L'**Articolo 13** impone **trasparenza e spiegabilità** nei sistemi di IA ad alto rischio, in particolare nei processi di **selezione del personale**.

Violazione nel caso di Amazon:

- I candidati non erano **informati** che un algoritmo di IA stava valutando i loro CV.
- Non era possibile per i candidati sapere **quali criteri** venivano utilizzati per assegnare punteggi e selezionare profili.

³⁶ FocusTech. (n.d.). Amazon AI potrebbe aver discriminato le donne. FocusTech.

- Le candidate scartate **non potevano contestare la decisione** o chiedere spiegazioni.

Se Amazon avesse rispettato l'Articolo 13, le candidate penalizzate avrebbero avuto il diritto di:

- Ricevere una **spiegazione dettagliata** sulle motivazioni del punteggio assegnato.
- Chiedere una **revisione umana della loro candidatura** in caso di errore dell'IA.
- Segnalare eventuali **discriminazioni**, spingendo l'azienda a migliorare il sistema.

Il caso di Amazon è un esempio emblematico di come l'uso non regolamentato dell'IA possa generare discriminazioni se non vengono applicati controlli adeguati sulla qualità dei dati (Art. 10) e sulla trasparenza (Art. 13).

Se l'AI Act fosse stato in vigore all'epoca, Amazon sarebbe stata **obbligata a rivedere il proprio algoritmo di selezione**, evitando la penalizzazione delle candidate femminili. Questo dimostra quanto sia importante una regolamentazione chiara per garantire equità, trasparenza e spiegabilità nei sistemi di IA aziendali.

Alfabetizzazione

Uno degli aspetti innovativi introdotti dall'**AI Act** riguarda l'obbligo di **alfabetizzazione sull'Intelligenza Artificiale**. Secondo quanto stabilito dall'articolo 4 del regolamento, sia le imprese private che le pubbliche amministrazioni devono garantire che i propri dipendenti abbiano una conoscenza adeguata delle tecnologie di IA, indipendentemente dal settore in cui operano (Altalex, 2025).

Si tratta del primo obbligo formale e generalizzato che va oltre le aziende tecnologiche, coinvolgendo ogni realtà che impieghi sistemi di intelligenza artificiale nei propri processi decisionali o operativi. L'obiettivo è rendere consapevoli i lavoratori delle opportunità offerte dall'IA, ma anche dei rischi legati al suo utilizzo, promuovendo un uso responsabile ed etico di queste tecnologie.

Le aziende e le PA sono chiamate ad adottare programmi di formazione specifici, che consentano al personale di:

- Comprendere il funzionamento dei sistemi di IA.
- Valutare i rischi potenziali legati all'utilizzo di algoritmi, in particolare quelli che possono incidere sui diritti fondamentali.

- Acquisire strumenti per l'utilizzo consapevole e sicuro dell'IA.

Oltre alla formazione, l'AI Act impone alle organizzazioni di predisporre **linee guida interne**, che definiscano le regole per un uso corretto e trasparente dell'intelligenza artificiale. Queste linee guida devono anche sensibilizzare i lavoratori sui potenziali **impatti etici e giuridici** legati all'utilizzo dei sistemi automatizzati.

Un altro aspetto fondamentale è il controllo sull'impiego dei sistemi di IA per la **valutazione e il monitoraggio delle performance lavorative**. L'AI Act stabilisce che:

- I dipendenti devono essere informati in modo trasparente sull'uso dei sistemi di IA che monitorano le loro attività.
- Devono essere garantiti i **diritti alla privacy e alla protezione dei dati personali**.
- I lavoratori devono poter **contestare le decisioni automatizzate**, in particolare quelle che incidono sulla loro valutazione o sulle condizioni lavorative.

Nonostante le sfide di attuazione, l'obbligo di alfabetizzazione rappresenta una leva strategica. La trasparenza e la consapevolezza aumentano la fiducia dei dipendenti e dei candidati nei confronti dell'organizzazione, favorendo una cultura aziendale improntata all'etica e all'affidabilità. Ciò può tradursi in un miglioramento delle performance aziendali e nella creazione di un ambiente di lavoro più equo e collaborativo (Altalex, 2025).

2.3 Dalla trasparenza all'accettazione dell'IA: una leva per il cambiamento organizzativo

L'adozione dell'Intelligenza Artificiale (IA) nei contesti aziendali ha trasformato in modo significativo i processi decisionali e le pratiche gestionali. Tuttavia, affinché l'IA possa essere realmente accettata e integrata con successo nelle dinamiche organizzative, la sua implementazione deve essere accompagnata da un elevato livello di **trasparenza percepita**. La letteratura più recente evidenzia infatti che la **trasparenza**—intesa come comprensibilità e accessibilità delle logiche decisionali adottate dai sistemi di IA—non rappresenta solo un elemento tecnico, ma costituisce una leva psicologica fondamentale per favorire **l'accettazione dell'IA da parte dei dipendenti**.

Per **accettazione** si intende l'atteggiamento favorevole e la disponibilità degli utenti a **utilizzare volontariamente i sistemi di IA**, integrandoli nelle proprie attività lavorative. Essa non è una variabile unitaria, ma il risultato di un processo influenzato da molteplici fattori cognitivi ed emotivi, come la fiducia, la percezione di utilità e la facilità d'uso, come descritto dai principali modelli teorici di riferimento, tra cui il Technology Acceptance Model (TAM).

In particolare, una maggiore trasparenza è in grado di attivare specifici **precursori psicologici dell'accettazione**, tra cui:

1. Una più elevata **percezione dell'utilità dell'IA** nei processi decisionali, che favorisce un atteggiamento positivo verso la tecnologia e ne incoraggia l'adozione.
2. Una maggiore **facilità d'uso percepita**, che contribuisce a ridurre la complessità percepita e a facilitare l'interazione con i sistemi intelligenti.
3. Una **riduzione della percezione del rischio occupazionale**, che mitiga le preoccupazioni legate alla sostituzione del lavoro umano da parte dell'automazione.

Queste tre dimensioni concorrono a rafforzare la **fiducia dei dipendenti nei confronti dell'IA**, considerata in questo studio come **manifestazione concreta dell'accettazione** dei sistemi automatizzati nei processi aziendali, e in particolare nel contesto del recruiting. Sebbene non si analizzi direttamente l'impatto dell'IA sulle performance aziendali oggettive, si ipotizza che un maggiore livello di fiducia possa **favorire indirettamente l'efficacia organizzativa**, migliorando il clima lavorativo, la collaborazione uomo-macchina e l'allineamento con gli obiettivi strategici.

Trasparenza e fiducia nel decision-making

Nel contesto aziendale, la **fiducia** rappresenta un elemento determinante per l'adozione e l'utilizzo efficace dei sistemi di Intelligenza Artificiale (IA), specialmente nei processi decisionali strategici. Una condizione imprescindibile per instaurare questa fiducia è la trasparenza dei sistemi di IA: gli utenti devono essere in grado di comprendere in modo chiaro come gli algoritmi elaborano le informazioni e quali logiche guidano le decisioni automatizzate. La percezione di trasparenza non solo rafforza la fiducia, ma favorisce anche un atteggiamento proattivo e collaborativo da parte dei dipendenti, che tendono a considerare l'IA un supporto

affidabile alle loro attività quotidiane (Siau & Wang, 2018). Secondo Siau e Wang (2018)³⁷, la fiducia nei confronti dell'IA non è un fenomeno statico, ma si sviluppa attraverso un processo articolato in **due fasi distinte**, entrambe fondamentali per garantire un'adozione consapevole delle tecnologie intelligenti.

1. Formazione della Fiducia Iniziale (*Initial Trust Formation*)

Questa fase riguarda la fiducia che gli utenti nutrono nel sistema di IA **prima** di interagire direttamente con esso. È influenzata da diversi fattori esterni:

- **Aspettative culturali e sociali:** quando la tecnologia viene percepita come innovativa e vantaggiosa dalla società, la propensione ad accettarla aumenta. La narrazione culturale positiva intorno all'IA gioca un ruolo determinante.
- **Strategie di comunicazione e marketing:** le aziende che comunicano in maniera trasparente i benefici e le modalità di funzionamento dei loro sistemi intelligenti sono percepite come più affidabili. La chiarezza informativa abbatte le barriere iniziali.
- **Reputazione del fornitore di IA:** la fiducia nei confronti di aziende che godono di un'immagine consolidata sul mercato (come IBM o Google) aumenta la predisposizione all'utilizzo dei loro sistemi.
- **Esperienze indirette e testimonianze:** recensioni, case study e feedback positivi rafforzano la fiducia iniziale, anche in assenza di esperienza diretta da parte dell'utente.

2. Sviluppo della Fiducia Continua (*Continuous Trust Development*)

Una volta instaurata la fiducia iniziale, questa deve essere consolidata nel tempo tramite un'interazione diretta positiva con il sistema di IA. La seconda fase si basa su:

- **Coerenza e affidabilità** delle decisioni fornite: il sistema deve garantire risultati precisi, replicabili e affidabili in modo costante.
- **Trasparenza e spiegabilità** delle decisioni: gli utenti devono comprendere le motivazioni dietro ogni scelta effettuata dall'IA, elemento che ne favorisce l'accettazione.
- **Adattabilità del sistema:** la capacità dell'IA di evolversi e personalizzarsi in base alle esigenze degli utenti contribuisce ad aumentare il senso di utilità e pertinenza.

³⁷ Siau, K. and Wang, W.Y. (2018) Building Trust in Artificial Intelligence, Machine Learning, and Robotics. *Cutter Business Technology Journal*, 31, 47-53.

- **Supervisione umana e controllo:** la possibilità per l'utente di intervenire, modificare o validare le decisioni del sistema incrementa la percezione di controllo e rassicura rispetto a possibili errori dell'algoritmo.

Un alto livello di trasparenza e fiducia nella tecnologia non solo facilita l'integrazione dell'IA nei processi decisionali, ma genera benefici tangibili per l'intera organizzazione, pur rappresentando questi ultimi **effetti indiretti o secondari** del processo di accettazione. In tal senso, la trasparenza non incide direttamente sulle performance aziendali, ma contribuisce a creare le condizioni psicologiche e relazionali favorevoli alla loro eventuale crescita, attraverso l'engagement dei dipendenti, l'adozione efficace delle tecnologie e la costruzione di un clima organizzativo aperto all'innovazione.

Il tema della fiducia nei sistemi di Intelligenza Artificiale è **centrale** per l'adozione efficace di questi strumenti nei processi decisionali aziendali. Lo studio di **Ma et al. (2023)**³⁸ si concentra proprio su questo aspetto, indagando come la fiducia riposta dagli utenti nei sistemi di IA influenzi la percezione della loro utilità e, di conseguenza, il loro impiego concreto nel decision-making. Il contributo innovativo dello studio è l'introduzione del concetto di *trust calibration*, ovvero la calibrazione della fiducia umana verso i suggerimenti forniti dall'IA. La ricerca evidenzia che per ottenere un uso efficace dell'intelligenza artificiale è fondamentale che gli utenti sappiano dosare correttamente il livello di fiducia. Se la fiducia è eccessiva, c'è il rischio che si seguano ciecamente i consigli del sistema, anche quando esso commette errori. Se invece la fiducia è insufficiente, gli utenti tendono a ignorare suggerimenti utili, privandosi del supporto dell'IA e riducendo i benefici che potrebbe offrire. Per favorire un corretto bilanciamento della fiducia, **Ma et al. (2023)** propongono un approccio basato sul concetto di *Correctness Likelihood (CL)*, ovvero la probabilità stimata che una decisione o un suggerimento fornito dall'IA sia corretto in un determinato contesto. Questo metodo prevede che il sistema di IA comunichi in modo trasparente il proprio grado di affidabilità o di incertezza rispetto a ciascun output generato.

Quando gli utenti ricevono informazioni chiare sul livello di affidabilità del sistema, riescono a calibrare meglio il proprio giudizio e a decidere se seguire il consiglio dell'IA o affidarsi alla propria esperienza e intuizione. Al contrario, un sistema percepito come opaco o non trasparente tende a ridurre la fiducia complessiva, portando gli utenti a scartare l'apporto del sistema anche quando utile.

³⁸ Ma, X., Yin, J., Lee, D., & Zhang, Y. (2023). *Who Should I Trust: AI or Myself? Leveraging Human and AI Correctness Likelihood to Promote Appropriate Trust in AI-Assisted Decision-Making*. arXiv preprint.

Gli esperimenti condotti dai ricercatori dimostrano che la comunicazione trasparente del *Correctness Likelihood* migliora la qualità delle decisioni prese dagli utenti. I partecipanti che hanno ricevuto indicazioni sul grado di probabilità che la decisione suggerita dall'IA fosse corretta hanno mostrato migliori capacità di giudizio rispetto a coloro che non hanno avuto accesso a queste informazioni. In altre parole, fornire feedback sul livello di sicurezza del sistema aiuta a mantenere un equilibrio nella fiducia degli utenti, evitando sia l'eccessiva dipendenza dall'IA sia il suo completo rifiuto.

Questo approccio promuove un modello decisionale collaborativo, in cui l'intelligenza artificiale e il decisore umano lavorano insieme, combinando le capacità di analisi dei dati del sistema con l'esperienza e l'intuizione umana.

L'avvento dell'Intelligenza Artificiale Generativa sta progressivamente ridefinendo le dinamiche collaborative all'interno dei team aziendali, con un impatto significativo sulle modalità di lavoro e sulle performance complessive. In questo scenario si colloca lo studio condotto da **Li, Zhou e Mikel-Hong (2024)**³⁹, il quale analizza in maniera approfondita il ruolo dell'IA Generativa nel miglioramento dell'efficacia dei team di lavoro. La ricerca si è basata su un esperimento randomizzato che ha coinvolto 435 partecipanti, suddivisi in 122 team, con l'obiettivo di valutare l'impatto dell'assistenza dell'IA su qualità, creatività e produttività delle attività svolte.

I risultati dello studio mostrano chiaramente che i gruppi che hanno integrato l'IA Generativa nei propri processi decisionali hanno ottenuto performance superiori rispetto ai team tradizionali. In particolare, si sono osservati **miglioramenti nella qualità, nell'originalità e nell'utilità** dei lavori prodotti. L'IA, utilizzata come strumento di supporto, si è rivelata capace di stimolare la creatività e di facilitare la produzione di soluzioni più innovative ed efficaci.

Tuttavia, lo studio evidenzia anche un importante limite: **l'eccessiva dipendenza da più sistemi di IA non garantisce necessariamente un ulteriore incremento delle prestazioni**. L'adozione di molteplici strumenti automatizzati può infatti risultare ridondante e non sempre porta a un valore aggiunto. Questo suggerisce che il successo dell'integrazione dell'IA non dipende dalla quantità di tecnologia impiegata, ma piuttosto dalla **qualità dell'interazione tra le persone e il sistema di IA**.

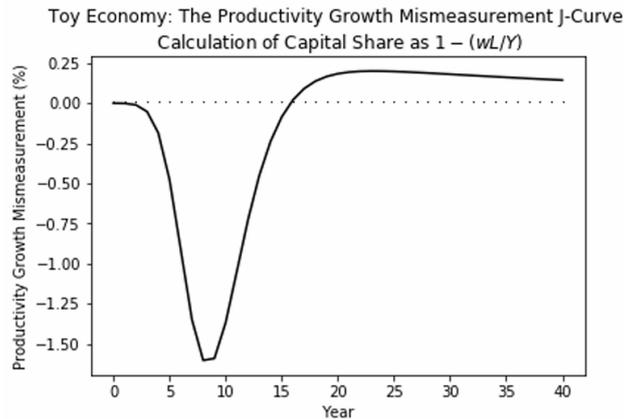
³⁹ Li, N., Zhou, H., & Mikel-Hong, K. (2024). Generative AI enhances team performance and reduces need for traditional teams. arXiv.

Un elemento chiave emerso dallo studio riguarda la percezione che i membri dei team hanno nei confronti della tecnologia. I gruppi che riconoscono l'IA Generativa come uno **strumento utile e integrato in modo consapevole nei flussi di lavoro** sono quelli che ottengono i migliori risultati. La consapevolezza del ruolo dell'IA e la fiducia nel suo utilizzo si rivelano, quindi, fattori determinanti per massimizzarne il potenziale.

Secondo **Li et al. (2024)**, il successo dell'intelligenza artificiale nei team non è legato solo all'efficienza degli algoritmi, ma anche alle dinamiche umane che si instaurano nel rapporto uomo-macchina. Le **interazioni collaborative** tra persone e IA, quando ben calibrate, favoriscono un clima di fiducia e collaborazione, incrementando la produttività complessiva senza sacrificare l'apporto umano. L'impiego dell'IA Generativa può persino ridefinire la necessità di **strutture di team tradizionali**, aprendo scenari in cui il lavoro collaborativo si evolve in una **partnership uomo-macchina**. L'IA non sostituisce i team umani, ma li **potenzia**, integrandosi nei processi decisionali e creativi per migliorarne l'efficacia.

L'Intelligenza Artificiale (IA) rappresenta una delle più recenti *General Purpose Technologies* (GPT), una categoria che comprende innovazioni di ampia portata come l'elettricità e Internet, in grado di trasformare interi sistemi economici e produttivi. Tuttavia, l'impatto di queste tecnologie sulla produttività non è immediato. Questo è il tema centrale affrontato da **Brynjolfsson, Rock e Syverson (2021)**, che analizzano il cosiddetto *Productivity Paradox of AI*, ovvero il divario tra il potenziale innovativo dell'IA e i benefici tangibili rilevati nei dati sulla produttività.

Riprendendo il celebre Paradosso della Produttività di Robert Solow (1987)—"Vediamo computer ovunque, tranne che nelle statistiche sulla produttività"—gli autori spiegano che l'introduzione di una nuova tecnologia, come l'IA, spesso non si traduce in un immediato aumento della produttività aziendale. Al contrario, i benefici emergono solo dopo un periodo di adattamento e trasformazione organizzativa. Questa dinamica è rappresentata graficamente dal modello della *Productivity J-Curve*.



Tale modello spiega come la produttività inizialmente possa calare (o non aumentare) quando viene introdotta una GPT, per poi crescere in modo esponenziale una volta che i processi aziendali sono stati adattati. In particolare, esistono tre fasi:

1. **Fase iniziale (declino o stagnazione della produttività):** durante questa fase, le aziende devono sostenere significativi costi di investimento in tecnologia, formazione e riconfigurazione dei processi. Questo può portare a un calo o a una mancata crescita della produttività, generando l'impressione che l'adozione dell'IA non sia vantaggiosa.
2. **Fase di transizione (crescita lenta):** dopo i primi investimenti, si cominciano a intravedere dei miglioramenti, ma l'effetto positivo sull'efficienza rimane limitato. Molte organizzazioni, non vedendo risultati immediati, possono decidere di abbandonare l'implementazione dell'IA.
3. **Fase di crescita esponenziale:** una volta superata la curva di apprendimento e completati gli investimenti complementari—come la riorganizzazione dei processi aziendali e lo sviluppo delle competenze interne—la produttività aumenta in modo sostanziale. L'IA diventa un elemento integrato nei processi decisionali, migliorando l'efficienza e la qualità delle decisioni.

Lo studio evidenzia come la **fiducia** nell'IA e la **percezione del suo valore strategico** siano fattori determinanti per il successo della tecnologia nel lungo termine. Le aziende che considerano l'IA una risorsa utile per il decision-making non registrano necessariamente benefici immediati, ma sono più propense a superare la fase iniziale di difficoltà, arrivando a cogliere importanti vantaggi competitivi. Questo suggerisce che il successo della tecnologia non dipende solo dalla sua efficacia tecnica, ma anche da condizioni **psicologiche e organizzative** favorevoli, tra cui la **trasparenza percepita**, la **fiducia** da parte dei dipendenti e la loro

disponibilità ad accettare il cambiamento. In questo senso, il contributo di Brynjolfsson et al. (2021) è rilevante per il presente lavoro, poiché aiuta a interpretare le performance aziendali come un possibile esito a lungo termine, raggiungibile solo quando vengono create le premesse soggettive e culturali necessarie all'accettazione dell'IA. Sebbene le performance non siano misurate direttamente in questa ricerca, esse rappresentano un riferimento teorico utile per comprendere il valore dell'accettazione tecnologica come processo.

Un altro aspetto cruciale sottolineato dagli autori è l'importanza degli investimenti in capitale intangibile. Per massimizzare i benefici derivanti dall'IA, non bastano le infrastrutture tecnologiche: sono necessari investimenti in formazione, adattamento dei processi aziendali, e sviluppo di nuovi modelli di business. Solo attraverso questi elementi complementari le aziende possono ottenere un reale vantaggio competitivo dall'integrazione dell'IA. Secondo il modello proposto da Brynjolfsson et al. (2021), le imprese devono essere consapevoli che il percorso verso l'aumento della produttività attraverso l'IA richiede **tempo, risorse e un cambiamento culturale significativo**. La capacità di **calibrare la fiducia** nei sistemi di IA e di **integrare l'innovazione** con strategie organizzative a lungo termine rappresenta la chiave per superare il paradosso iniziale e ottenere risultati concreti.

Secondo lo studio di **Adadi e Berrada (2018)**⁴⁰, la spiegabilità dell'IA rappresenta un elemento determinante per favorire la **fiducia degli utenti**, elemento chiave per l'accettazione e l'integrazione efficace di queste tecnologie nei processi aziendali. Una delle principali criticità associate ai sistemi di IA tradizionali è la loro natura **opaca**: i modelli complessi, come le reti neurali profonde, sono spesso percepiti come **black box**, ovvero sistemi le cui logiche decisionali risultano poco comprensibili agli utenti. Questa mancanza di trasparenza può generare diffidenza e ostacolare l'adozione dell'IA nei contesti aziendali. Adadi e Berrada sottolineano che questa percezione negativa rappresenta una delle principali barriere alla diffusione dell'IA nei processi decisionali.

La risposta a questa sfida è l'emergere dei sistemi di **Explainable Artificial Intelligence (XAI)**. I modelli XAI sono progettati per **fornire spiegazioni comprensibili e dettagliate** sulle decisioni algoritmiche, migliorando così la percezione di equità e affidabilità della tecnologia. La **spiegabilità** accresce la percezione di equità e affidabilità, e rafforza la **propensione degli utenti a fidarsi dei risultati dell'IA e a integrarli nei processi**

⁴⁰ Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.

decisionali. In questo quadro, la **fiducia risulta strettamente legata alla trasparenza percepita**, e può essere considerata una condizione psicologica favorevole all'accettazione della tecnologia. Lo studio evidenzia che quando gli utenti comprendono le logiche decisionali dell'IA, sono più inclini ad adottarla come **strumento collaborativo**, migliorando la qualità dell'interazione uomo-macchina. Questo approccio supporta processi decisionali più consapevoli, che **possono tradursi, secondo la letteratura, in benefici organizzativi di lungo periodo.** Tuttavia, questi benefici non sono automatici né immediati: si manifestano solo se le tecnologie vengono adottate in un contesto in cui trasparenza, fiducia e accettazione siano elevate. Settori come il **recruiting**, la **finanza** e la **sanità** hanno già avviato l'adozione di sistemi XAI, proprio per garantire maggiore equità e ridurre il rischio di discriminazioni algoritmiche. La spiegabilità aiuta a identificare e mitigare eventuali bias nei modelli predittivi, aumentando la trasparenza dei processi e il rispetto delle normative etiche e legali.

Adadi e Berrada pongono l'accento anche sul ruolo dell'**interazione uomo-macchina.** Non è sufficiente fornire spiegazioni statiche: i sistemi XAI dovrebbero essere **adattivi**, ovvero capaci di **personalizzare le spiegazioni** in base al profilo dell'utente, al suo livello di competenza e al contesto operativo. Questo approccio consente una maggiore **collaborazione** tra l'intelligenza umana e quella artificiale, facilitando l'accettazione e l'efficacia dei sistemi di IA nei processi decisionali.

Uno degli aspetti più discussi nello studio riguarda il **trade-off tra interpretabilità e accuratezza.** In molti casi, i modelli più trasparenti e spiegabili (come gli alberi decisionali o i modelli lineari) sono meno performanti rispetto ai modelli **black box** ad alta complessità, come le reti neurali profonde. Le aziende devono quindi bilanciare la necessità di spiegabilità con l'efficacia predittiva dei sistemi, adottando strategie che massimizzino i vantaggi di entrambe le soluzioni.

In conclusione, investire in sistemi XAI può rappresentare, secondo Adadi e Berrada, **una leva competitiva strategica.** Le organizzazioni che adottano modelli trasparenti non solo aumentano la fiducia dei dipendenti, ma migliorano la **governance dell'IA**, creando le condizioni per un utilizzo più consapevole e responsabile della tecnologia. In linea con il modello teorico proposto in questo studio, la trasparenza si configura come una **variabile abilitante** che, tramite l'attivazione di precursori psicologici come la fiducia, può contribuire indirettamente a generare vantaggi in termini di performance e innovazione.

La percezione del rischio occupazionale come ostacolo alla fiducia nell'IA

Nel processo di adozione dei sistemi di Intelligenza Artificiale (IA) nei contesti organizzativi, la percezione del rischio occupazionale rappresenta una delle principali barriere psicologiche alla costruzione della fiducia.

Quando i lavoratori interpretano l'introduzione dell'IA come una minaccia alla sicurezza del proprio impiego, tendono a sviluppare atteggiamenti di sfiducia e resistenza verso la tecnologia stessa. Tale percezione può manifestarsi anche in assenza di un pericolo concreto di licenziamento: è sufficiente che il dipendente ritenga che il proprio ruolo possa essere sostituito, sminuito o reso obsoleto dall'automazione.

Diversi studi hanno evidenziato che il timore di perdere il lavoro agisce come fattore inibente nei confronti dell'accettazione dell'IA. In particolare, **Zirar et al. (2023)**⁴¹ sottolineano che la sfiducia nei confronti dei sistemi intelligenti nasce proprio dalla percezione che questi strumenti possano compromettere la continuità e la sicurezza dell'occupazione. La percezione del rischio occupazionale si configura così come **un precursore psicologico della fiducia**: quanto più i lavoratori percepiscono l'IA come una minaccia al proprio impiego, tanto più sarà difficile per loro fidarsi della tecnologia e integrarla nelle pratiche lavorative quotidiane.

Il rischio percepito può essere amplificato da un'eccessiva enfattizzazione delle capacità dell'IA da parte delle organizzazioni o dei media, generando una narrativa distorta e allarmistica. Secondo gli autori, rappresentazioni iperboliche dell'intelligenza artificiale—come quella della cosiddetta “Robo-Apocalypse”—aumentano la paura di essere rimpiazzati e spingono i lavoratori a reagire con **comportamenti difensivi**, tra cui la **riduzione della collaborazione**, il **nascondimento delle informazioni** e un generale **disimpegno nei confronti dei progetti tecnologici**.

Tali dinamiche non solo ostacolano l'instaurarsi di un rapporto di fiducia tra dipendente e sistema, ma compromettono anche le condizioni necessarie per una vera trasformazione organizzativa. La costruzione di un clima lavorativo favorevole all'accettazione dell'IA richiede dunque interventi mirati a ridurre queste percezioni di minaccia, come una comunicazione trasparente, il coinvolgimento attivo dei lavoratori nei processi decisionali e percorsi di formazione orientati allo sviluppo di nuove competenze.

Nel modello teorico del presente studio, la percezione dell'insicurezza lavorativa viene considerata una variabile psicologica rilevante che può **ostacolare o facilitare** la relazione tra trasparenza percepita e accettazione dell'IA. La sua gestione rappresenta quindi una condizione essenziale per promuovere un'adozione consapevole, stabile e sostenibile dell'intelligenza artificiale in azienda.

⁴¹ Zirar, A., Arslan, A., & Bayram, M. (2023). Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda. *Technological Forecasting and Social Change*, 189, 122404.

Come emerso, la percezione del rischio occupazionale può ostacolare significativamente la costruzione della fiducia nei confronti dell'IA, condizionando la disponibilità dei lavoratori ad accettarne l'utilizzo. Approfondire le radici psicologiche di questa percezione è quindi essenziale per comprendere le dinamiche che facilitano o impediscono un'effettiva adozione della tecnologia in azienda. In quest'ottica, il concetto di insicurezza lavorativa assume un ruolo chiave nel delineare i meccanismi che possono inibire l'accettazione dell'IA da parte dei dipendenti.

L'insicurezza lavorativa rappresenta uno dei più rilevanti fattori di stress psicologico nel contesto lavorativo odierno. La rapidità con cui le tecnologie emergenti, come l'Intelligenza Artificiale e l'automazione, stanno trasformando le dinamiche aziendali, ha acuito questa percezione di precarietà. Sebbene l'introduzione di innovazioni digitali possa migliorare efficienza e produttività, molti lavoratori percepiscono tali cambiamenti come una minaccia diretta al proprio ruolo, alimentando ansia e timori legati alla propria stabilità occupazionale.

Secondo **De Witte (1999)**⁴², l'insicurezza lavorativa è una percezione soggettiva che si manifesta come timore persistente di perdere il proprio impiego o di vedere peggiorate le proprie condizioni di lavoro. Non si tratta di una minaccia concreta e immediata, bensì di una sensazione di incertezza che ha conseguenze significative sul benessere psicologico del lavoratore. È un'esperienza cognitiva che, indipendentemente dalla reale probabilità di licenziamento, può influenzare profondamente la qualità della vita professionale e personale.

Ampliando la riflessione, **Graham, Sinclair e Sliter (2023)**⁴³ integrano la visione classica di De Witte proponendo una distinzione tra due dimensioni dell'insicurezza:

- **Quantitativa**, riferita alla paura concreta di perdere il posto di lavoro.
- **Qualitativa**, che si concentra sulla preoccupazione che il contenuto, le condizioni o le opportunità di carriera del proprio impiego peggiorino.

Questa distinzione è fondamentale perché evidenzia che anche chi non rischia la perdita del lavoro può vivere un forte disagio dovuto alla percezione che il proprio ruolo sia svalutato o privo di prospettive di crescita. La

⁴² De Witte, H. (1999). "Job Insecurity and Psychological Well-being: Review of the Literature and Exploration of Some Unresolved Issues." *European Journal of Work and Organizational Psychology*, 8(2), 155-177.

⁴³ Graham, B. A., Sinclair, R. R., & Sliter, M. (2023). "Job Insecurity and Health and Well-being: What Happens When You Really Need or Love Your Job?" *Economic and Industrial Democracy*, 44(3), 674-695.

precarietà percepita, infatti, non riguarda solo la sopravvivenza economica, ma anche la soddisfazione personale e il riconoscimento professionale.

Graham et al. pongono inoltre l'accento sulla distinzione tra la componente cognitiva dell'insicurezza, legata alla valutazione razionale delle probabilità di cambiamento, e le reazioni emotive, come ansia e stress, che tale valutazione può generare. Ciò significa che non è solo l'evento oggettivo (ad esempio, una ristrutturazione aziendale o l'introduzione di nuove tecnologie) a influire, ma anche il modo in cui viene percepito e interpretato dal lavoratore, e dunque il significato attribuito ad esso a generare ansia e sfiducia.

L'insicurezza lavorativa può quindi compromettere la relazione tra lavoratore e tecnologia: non solo altera il benessere psicologico, ma può ridurre il coinvolgimento e il grado di apertura verso l'adozione dell'IA. Quando i dipendenti percepiscono instabilità, tendono a reagire con atteggiamenti di chiusura, resistenza e ritiro, riducendo la disponibilità a interagire con sistemi automatizzati, in particolare se questi non sono percepiti come trasparenti e controllabili.

Lo studio di **De Witte (1999)**, condotto su un campione di lavoratori di una multinazionale metalmeccanica, ha dimostrato che l'insicurezza lavorativa influisce negativamente su vari aspetti:

1. **Benessere psicologico ridotto:** i lavoratori che percepivano il proprio impiego come instabile riportavano livelli elevati di stress e disagio psicologico, con punteggi inferiori nel *General Health Questionnaire (GHQ-12)*.
2. **Maggiore impatto rispetto ad altri stressor:** l'insicurezza risultava essere un fattore di stress più rilevante rispetto a condizioni di carico lavorativo elevato.
3. **Effetti paragonabili alla disoccupazione:** i lavoratori che temevano di perdere il posto sperimentavano un disagio simile a quello provato dai disoccupati di breve periodo.
4. **Differenze di genere:** l'impatto risultava più marcato negli uomini, per i quali la stabilità lavorativa sembrava avere un valore identitario e di sostegno economico maggiore. Per le donne, invece, l'effetto era mitigato da altri fattori come il supporto sociale.

Dal punto di vista organizzativo, l'insicurezza lavorativa mina la **fiducia nei confronti del management** e peggiora il clima aziendale. I lavoratori insicuri si sentono meno coinvolti e meno motivati, riducendo il loro impegno verso gli obiettivi aziendali. Questo porta a una **riduzione dell'efficienza operativa**, un incremento degli errori e una minore disponibilità alla collaborazione.

Per affrontare queste criticità, De Witte (1999) suggerisce l'adozione di politiche aziendali mirate, che includano:

- Una **comunicazione trasparente e continua** sulle decisioni strategiche e sulle innovazioni tecnologiche introdotte.
- La **partecipazione dei dipendenti** ai processi decisionali che riguardano il loro futuro lavorativo.
- Politiche di **formazione continua**, che offrano opportunità di aggiornamento professionale e di crescita interna, rafforzando la percezione di controllo dei lavoratori rispetto al proprio ruolo.

Tali misure possono contribuire a ridurre le paure legate all'automazione e a **rafforzare la fiducia nei sistemi di IA**, facilitandone l'accettazione anche nei contesti più sensibili, come quello del recruiting o del decision-making aziendale.

La minaccia percepita dell'IA nei contesti occupazionali

Alla luce di quanto emerso finora, appare evidente che l'adozione dell'Intelligenza Artificiale nei contesti aziendali non si traduce unicamente in un cambiamento tecnologico, ma produce **effetti psicologici profondi sui lavoratori**, soprattutto in termini di percezione del rischio. Un'ulteriore conferma arriva da una serie di studi recenti che analizzano **come l'IA venga percepita dai dipendenti in relazione alla propria sicurezza lavorativa e alla fiducia verso l'organizzazione**. Secondo quanto riportato da **Carmignani (2024)⁴⁴**, esistono diverse ricerche che confermano l'impatto psicologico e comportamentale dell'automazione avanzata sulle persone. Tra queste spicca lo studio *The Rise of Robots Increases Job Insecurity and Maladaptive Workplace Behaviors* di **Kai Chi Yam et al.**, che indaga su come la presenza crescente di robot influenzi la percezione di sicurezza sul lavoro e i comportamenti dei dipendenti. Le principali evidenze dello studio sono state:

- **Aumento dell'incertezza occupazionale:** l'esposizione, anche solo indiretta, ai robot e ai sistemi di IA genera un senso diffuso di insicurezza tra i lavoratori, alimentando timori di sostituzione.

⁴⁴ Carmignani, M. (2024, 5 settembre). *Futuro del lavoro e AI, oltre il luogo comune: 4 studi per capire cosa aspettarci*. Agenda Digitale.

- **Stress e comportamenti disadattivi:** tale insicurezza si traduce in un aumento di stress lavorativo, burnout e atteggiamenti meno cooperativi nei confronti dei colleghi. In alcuni casi si osservano manifestazioni di inciviltà e comportamenti difensivi.
- **Il ruolo dell'autoaffermazione:** lo studio evidenzia che pratiche di riflessione sui propri valori personali, note come autoaffermazione, possono ridurre significativamente l'ansia da sostituzione e migliorare il benessere psicologico. Questi interventi supportano l'adattamento dei lavoratori a un ambiente tecnologico in evoluzione.
- **Impatto trasversale:** l'effetto dell'automazione sulla sicurezza percepita riguarda vari settori industriali, non limitandosi alle industrie più tecnologizzate. Inoltre, emerge come questo fenomeno sia osservabile in contesti culturali molto differenti.

L'analisi di Carmignani sottolinea come, nonostante le opportunità offerte dall'IA in termini di produttività ed efficienza, sia necessario **accompagnare l'introduzione delle tecnologie con strategie di supporto psicologico e formativo**. Le organizzazioni dovrebbero agire in modo preventivo, favorendo la resilienza dei dipendenti e promuovendo una cultura aziendale inclusiva e trasparente nei confronti dei cambiamenti tecnologici.

Nel paper "Lavoro e Intelligenza Artificiale in Italia: tra opportunità e rischio di sostituzione", pubblicato nella collana INAPP Working Paper da Valentina Ferri, Rita Porcelli ed Enrico Maria Fenoaltea⁴⁵, viene proposta un'analisi dettagliata del potenziale impatto dell'IA sulle professioni italiane. Attraverso l'utilizzo dell'indice AIOE (Ability level AI Exposure), lo studio misura il grado di esposizione di ciascuna professione alle tecnologie di automazione intelligente, tenendo conto sia delle abilità richieste dalle diverse mansioni sia della probabilità che queste vengano sostituite da sistemi di IA. L'analisi rivela una forte eterogeneità tra settori e regioni del Paese, con professioni che appaiono più vulnerabili all'automazione, in particolare nei comparti legati alle attività di routine amministrativa e di gestione dati. Tra queste figurano:

- Gli addetti al protocollo e allo smistamento documentale.
- Gli addetti alle buste paga.
- I ruoli dirigenziali che svolgono funzioni standardizzate, come i direttori generali.

⁴⁵ *Agenda Digitale. (n.d.). IA e lavoro: ecco le professioni che rischiano di più. Agenda Digitale.*

Al contrario, risultano meno esposte le professioni che richiedono abilità manuali, fisiche o artistiche difficilmente replicabili da un algoritmo. Rientrano in questa categoria:

- Operai edili, come manovali e intonacatori.
- Professionisti artistici, come ballerini e atleti.
- Lavoratori in mestieri tradizionali, ad esempio i conduttori di veicoli a trazione animale.

Il rapporto INAPP sottolinea l'urgenza di implementare politiche di formazione continua e riqualificazione professionale, allo scopo di agevolare l'adattamento dei lavoratori e prevenire fenomeni di esclusione dal mercato del lavoro. Secondo il paper, la transizione verso un'economia digitale non può prescindere dal rafforzamento delle competenze trasversali, digitali e relazionali. Accanto alle sfide strutturali, emergono anche quelle legate al benessere psicologico dei lavoratori. Lo studio condotto da Guofeng Jin, Jiang Jiang e Hengyi Liao (2023)⁴⁶, pubblicato su *Scientific Reports*, offre un'analisi approfondita dell'impatto che l'adozione delle tecnologie di IA ha sulla sfera psicologica e comportamentale dei dipendenti.

Un concetto centrale introdotto nella ricerca è quello di Consapevolezza STARA (Smart Technology, Artificial Intelligence, Robotics, and Algorithms Awareness)⁴⁷. Questo termine indica la percezione, da parte dei lavoratori, del rischio che le tecnologie avanzate rappresentino una minaccia per la stabilità della propria carriera. La consapevolezza STARA riflette la crescente inquietudine verso l'obsolescenza delle competenze e la sostituzione dei ruoli umani da parte dell'IA, anche in mansioni cognitive e creative tradizionalmente considerate immuni all'automazione. I risultati della ricerca sono chiari:

- Il 51% dei lavoratori intervistati esprime forti preoccupazioni legate alla perdita di lavoro e al mancato aggiornamento delle proprie competenze.
- Questo stato di allerta psicologica è correlato a emozioni negative, quali ansia, stress e senso di minaccia per la propria sicurezza occupazionale.

⁴⁶ Jin, G., Jiang, J., & Liao, H. (2023). *The Psychological Impact of AI Adoption in the Workplace: Evidence from Employee Surveys*. *Scientific Reports*, 13(1), 1123-1138.

⁴⁷ Viliotti, A. (n.d.). *Consapevolezza STAR: l'impatto dell'AI sul benessere lavorativo dei dipendenti*. *Andrea Viliotti Blog*.

- Il fenomeno si ripercuote direttamente sulla motivazione e l'engagement dei lavoratori, riducendo la loro propensione a investire in attività creative e innovative, per il timore che il proprio ruolo venga superato dalla tecnologia.

Inoltre, la consapevolezza STARA si riflette anche sul clima aziendale, aumentando l'intenzione di turnover e diminuendo il senso di appartenenza all'organizzazione, specialmente in quei contesti in cui mancano politiche di riqualificazione e aggiornamento professionale.

Un elemento interessante emerso dallo studio riguarda la diversa percezione della minaccia in base alle competenze digitali possedute:

- I lavoratori con elevate competenze tecnologiche tendono a considerare l'IA un'opportunità per migliorare l'efficienza lavorativa e per arricchire il proprio bagaglio professionale.
- Al contrario, coloro che hanno competenze digitali limitate risultano più vulnerabili agli effetti negativi della consapevolezza STARA, mostrando un rischio maggiore di burnout e un calo significativo del benessere psicologico. Oltre la minaccia: le opportunità offerte dall'IA

Nonostante le preoccupazioni, sia il report INAPP che lo studio di Jin et al. concordano nell'individuare significative opportunità legate all'adozione dell'IA. La tecnologia, se gestita in modo consapevole, può:

- Liberare i lavoratori dai compiti ripetitivi e routinari, permettendo loro di concentrarsi su attività ad alto valore aggiunto, come la creatività, l'innovazione e le relazioni umane.
- Creare nuovi profili professionali, legati alla gestione, programmazione e supervisione dei sistemi di IA.
- Offrire alle imprese la possibilità di migliorare la produttività e la qualità dei servizi, rendendo i processi decisionali più efficienti e data-driven.

Tuttavia, per sfruttare appieno queste potenzialità è fondamentale adottare strategie di reskilling e upskilling, favorendo l'inclusione dei lavoratori nei processi di trasformazione digitale e riducendo i rischi di esclusione sociale.

Strategie per mitigare l'impatto negativo dell'insicurezza lavorativa

L'introduzione crescente dell'Intelligenza Artificiale (IA) nei contesti lavorativi sta rivoluzionando le dinamiche produttive e organizzative, ma al tempo stesso genera crescenti preoccupazioni tra i lavoratori. L'insicurezza lavorativa legata all'automazione rappresenta oggi una delle principali sfide per il benessere occupazionale e la coesione all'interno delle aziende. Secondo l'European Agency for Safety and Health at Work (**EU-OSHA**, n.d.)⁴⁸, un'efficace gestione dell'adozione delle tecnologie AI passa inevitabilmente attraverso un approccio integrato e partecipativo, che consideri i bisogni psicologici dei dipendenti e le necessità organizzative.

La EU-OSHA sottolinea che **il successo nell'implementazione dei sistemi basati sull'IA dipende in gran parte dall'accettazione** da parte dei lavoratori e dalla capacità dell'organizzazione di gestire il cambiamento in modo inclusivo. Se la trasformazione digitale è percepita come una minaccia, può generare ansia e ridurre la motivazione; al contrario, se gestita correttamente, l'IA può essere vista come un'opportunità per migliorare la qualità del lavoro e rafforzare la competitività aziendale. Il report dell'agenzia europea propone una serie di misure pratiche e strategie per mitigare l'insicurezza lavorativa legata all'adozione dell'IA e garantire una transizione equa e sostenibile:

1. Comunicazione chiara, continua e trasparente

Una comunicazione efficace è alla base della fiducia. È fondamentale che i lavoratori ricevano informazioni tempestive e veritiere sugli scopi dell'introduzione dell'IA. Secondo il report, le aziende devono spiegare chiaramente come l'IA migliorerà i processi, quali ruoli potrebbero subire trasformazioni e quali nuove opportunità si presenteranno. La comunicazione deve puntare a dissipare paure ingiustificate e prevenire la diffusione di informazioni distorte.

2. Programmi di formazione e riqualificazione professionale

Investire nello sviluppo delle competenze dei lavoratori è essenziale. La formazione non deve limitarsi all'apprendimento tecnico, ma deve includere competenze trasversali che rafforzino la resilienza dei dipendenti. Programmi di upskilling e reskilling consentono di ridurre l'incertezza legata all'obsolescenza professionale, offrendo ai lavoratori strumenti concreti per rimanere competitivi in un contesto tecnologico in continua evoluzione.

⁴⁸ European Agency for Safety and Health at Work (EU-OSHA). (n.d.). *Facilitating acceptance of AI-based systems in the workplace and minimising organisational impact*. European Agency for Safety and Health at Work.

3. Adozione di modelli di IA spiegabili e accessibili

Uno degli ostacoli principali all'accettazione dell'IA è rappresentato dalla scarsa trasparenza dei suoi processi decisionali. L'utilizzo di modelli di **IA spiegabile (XAI)** contribuisce a rendere l'algoritmo comprensibile, favorendo la fiducia dei lavoratori. La EU-OSHA raccomanda audit periodici per verificare la correttezza dei sistemi e garantire che i dipendenti possano accedere a spiegazioni chiare e dettagliate sugli output delle macchine.

4. Coinvolgimento attivo dei lavoratori nei processi decisionali

Coinvolgere i dipendenti nella progettazione e nell'implementazione dei sistemi di IA aiuta a superare la percezione di alienazione e a rafforzare il senso di appartenenza. La creazione di forum di discussione, gruppi di lavoro misti e strumenti per la raccolta di feedback permette di cogliere le preoccupazioni dei lavoratori e di adottare contromisure tempestive.

5. Monitoraggio continuo dei rischi e promozione della salute e sicurezza sul lavoro

L'integrazione dell'IA nei processi produttivi deve essere accompagnata da un attento monitoraggio dei rischi psicosociali e operativi. Secondo la EU-OSHA, è opportuno predisporre procedure di valutazione del rischio che includano la dimensione psicologica, adottando strumenti di supporto per i lavoratori, come sportelli di ascolto e programmi di benessere aziendale. Inoltre, l'implementazione di segnalazioni anonime e sistemi di audit sulla sicurezza tecnologica aiuta a individuare criticità e a intervenire in modo proattivo.

Queste strategie non hanno un impatto diretto solo sulla gestione organizzativa, ma agiscono a livello psicologico, contribuendo a **ridurre la percezione di insicurezza** e a **rafforzare la fiducia nei sistemi automatizzati**. In questo senso, la trasparenza, la partecipazione e l'investimento nella persona diventano leve fondamentali per promuovere **l'accettazione dell'IA come processo di coevoluzione uomo-macchina**, e non come passiva sostituzione.

L'adozione dell'IA non deve essere vista come una minaccia, bensì come un'opportunità per migliorare i processi produttivi e promuovere nuove forme di collaborazione tra uomo e macchina. Tuttavia, ciò richiede una visione strategica e un impegno costante nella costruzione di un ambiente di lavoro trasparente, partecipativo e orientato al benessere psicologico dei lavoratori.

2.4 La fiducia nei processi automatizzati di recruiting come fattore chiave dell'accettazione dell'IA

Negli ultimi anni, l'intelligenza artificiale ha trasformato in modo significativo le pratiche di selezione del personale, offrendo alle aziende nuove opportunità per ottimizzare i processi di recruiting e gestione delle risorse umane. L'adozione di tecnologie basate su algoritmi intelligenti consente alle imprese di individuare candidati idonei con maggiore rapidità e precisione, riducendo i tempi e i costi associati ai metodi tradizionali. Tuttavia, questa innovazione porta con sé sfide e implicazioni etiche che richiedono un'attenta gestione, soprattutto per quanto riguarda la trasparenza e la responsabilità nei confronti dei candidati e dei dipendenti.

L'introduzione dell'IA nei processi di selezione si è rivelata particolarmente efficace nella fase di screening dei candidati. Gli algoritmi analizzano migliaia di curriculum vitae in pochi istanti, individuando le competenze e le esperienze più rilevanti per le posizioni aperte. Sistemi avanzati di Natural Language Processing (NLP) sono integrati nei chatbot che svolgono i colloqui preliminari, rispondono alle domande dei candidati e forniscono aggiornamenti in tempo reale sull'avanzamento della selezione (Qipo, n.d.)⁴⁹.

Un esempio di questa innovazione è rappresentato da VERA, il bot sviluppato dalla startup russa Straforoy e adottato da IKEA nel 2018. VERA è in grado di gestire il recruiting in modo autonomo, effettuando interviste telefoniche e video, analizzando emozioni e comportamenti dei candidati e riducendo sensibilmente i tempi di assunzione. IKEA ha riportato un abbattimento dei tempi di selezione del 50%, permettendo ai responsabili HR di concentrarsi sugli aspetti strategici del processo, come la valutazione del cultural fit e lo sviluppo delle competenze dei candidati selezionati (Rödl & Partner, 2024)⁵⁰.

L'utilizzo dell'IA nei processi di recruiting non solo aumenta l'efficienza operativa, ma consente anche di ampliare il bacino di talenti. Gli algoritmi aiutano le aziende a identificare competenze emergenti e profili professionali non convenzionali, favorendo la creazione di team eterogenei e complementari. Inoltre, l'IA contribuisce al processo di *internal mobility*, suggerendo percorsi di *upskilling* e *reskilling* personalizzati per i dipendenti, rafforzando il loro engagement e riducendo il rischio di turnover. Un ulteriore vantaggio è rappresentato dalla capacità predittiva dei sistemi di IA, che analizzano dati comportamentali e feedback dei dipendenti per individuare segnali di insoddisfazione o stress. In questo modo, i responsabili HR possono

⁴⁹ Qipo. (n.d.). *Intelligenza artificiale e risorse umane: come cambia l'azienda*. Qipo.

⁵⁰ Rödl & Partner. (2024, dicembre 12). *Lavoro 5.0 e recruiting: quando scegliere l'intelligenza artificiale*. Rödl & Partner.

intervenire tempestivamente per migliorare il clima aziendale e la retention, ottimizzando l'impiego delle risorse umane in un'ottica di lungo periodo (Qipo, n.d.).

Le Sfide della Trasparenza e le Implicazioni Etiche

Nonostante i vantaggi, l'adozione dell'IA nel recruiting solleva preoccupazioni significative in materia di trasparenza, privacy e discriminazione. Il rischio principale è rappresentato dai **bias algoritmici**, che possono perpetuare discriminazioni di genere, etnia o età se i dati utilizzati per l'addestramento dei modelli sono distorti o incompleti. Il caso di Amazon è emblematico: l'azienda ha dovuto dismettere un algoritmo di selezione che penalizzava sistematicamente le candidate donne, poiché addestrato su dati storici che riflettevano un forte squilibrio di genere (Qipo, n.d.).

L'uso dell'IA implica la raccolta e l'elaborazione di dati personali sensibili, con rilevanti implicazioni in termini di **privacy** e protezione dei dati. Secondo il Regolamento Generale sulla Protezione dei Dati (GDPR), le aziende sono obbligate a garantire la trasparenza delle operazioni di trattamento e a fornire agli interessati un controllo effettivo sui propri dati. Le violazioni di tali obblighi possono comportare sanzioni fino a 20 milioni di euro o al 4% del fatturato annuo globale (Rödl & Partner, 2024). Le aziende devono quindi implementare rigorose misure di sicurezza, garantire la possibilità per i candidati di accedere, correggere o cancellare i propri dati e fornire spiegazioni chiare sui criteri e le modalità di funzionamento degli algoritmi di selezione.

Se da un lato l'IA offre strumenti potenti per ottimizzare i processi di selezione, dall'altro è fondamentale mantenere un ruolo attivo del **fattore umano**. L'intuizione, l'empatia e la capacità di cogliere le sfumature personali restano competenze indispensabili per un recruiting equo ed efficace. L'IA dovrebbe essere vista come un supporto al processo decisionale umano, e non come un sostituto (Rödl & Partner, 2024).

La trasparenza nei processi di selezione automatizzati non riguarda solo la correttezza etica, ma è anche un requisito legale e operativo essenziale. **L'opacità degli algoritmi** può compromettere la fiducia dei candidati e sollevare problematiche legali legate alla discriminazione. Le aziende devono adottare pratiche di auditing e valutazione continua degli algoritmi, assicurandosi che i sistemi di IA siano equi, affidabili e spiegabili.

La proposta di Direttiva Europea sulla responsabilità civile da IA (AILD) si inserisce in questo contesto, mirata a rafforzare la tutela degli utenti e a facilitare l'accesso ai risarcimenti in caso di danni derivanti dall'uso di sistemi automatizzati. Il principio della "vicinanza alla prova" impone inoltre alle aziende l'onere di dimostrare la correttezza dei propri sistemi, sottolineando la necessità di trasparenza e accountability nei processi decisionali automatizzati (Rödl & Partner, 2024). L'integrazione dell'intelligenza artificiale nei processi di

recruiting rappresenta una grande opportunità per le aziende, capace di migliorare l'efficienza, ridurre i costi e favorire l'inclusione. Tuttavia, per sfruttarne appieno il potenziale è essenziale adottare un approccio responsabile che garantisca trasparenza, rispetto dei diritti dei candidati e una supervisione umana continua. Solo in questo modo l'IA potrà contribuire a costruire processi di selezione equi, sostenibili e orientati al benessere organizzativo di lungo termine.

Secondo **Felzmann et al. (2019)**⁵¹, la trasparenza deve essere considerata un principio fondante nella progettazione dei sistemi di decisione automatizzata (*Automated Decision-Making, ADM*), come quelli sempre più impiegati nel recruiting aziendale. Il loro studio, "*Transparency by Design in AI Systems: Principles and Implementation Strategies*", propone un framework concettuale denominato *Transparency by Design (TbD)*, che mira a integrare la trasparenza sin dalle prime fasi di sviluppo dei sistemi basati su IA.

Ispirato al noto modello del *Privacy by Design*, il framework TbD di Felzmann et al. non considera la trasparenza come un mero adempimento normativo, ma come un principio etico e funzionale, essenziale per garantire fiducia, accountability e rispetto dei diritti degli utenti. Secondo gli autori, rendere trasparenti i sistemi ADM non solo migliora la qualità delle decisioni automatizzate, ma aumenta anche la percezione di equità e la fiducia da parte dei soggetti coinvolti, in particolare nei processi di selezione del personale.

Il modello si articola su tre livelli principali:

1. **Trasparenza come virtù:** un impegno etico da parte delle organizzazioni che implementano sistemi ADM, affinché la trasparenza diventi parte integrante della cultura aziendale.
2. **Trasparenza come relazione:** la comunicazione chiara e comprensibile delle decisioni e dei criteri adottati dal sistema ADM verso gli utenti coinvolti, ad esempio i candidati durante il recruiting.
3. **Trasparenza come sistema:** l'adozione di pratiche di governance e di regolamentazione che garantiscano la tracciabilità e la verificabilità delle decisioni prese dai sistemi automatizzati.

Il framework TbD proposto da Felzmann et al. si fonda su nove principi chiave:

- **Contestualizzazione delle informazioni:** i dati e le spiegazioni devono essere adattati al contesto d'uso e al livello di comprensione degli utenti.

⁵¹ Felzmann, H., et al. (2019). *Transparency by Design in AI Systems: Principles and Implementation Strategies*. *AI & Ethics*, 1(2), 173-190.

- **Accessibilità:** le informazioni devono essere facilmente fruibili e non richiedere competenze tecniche avanzate.
- **Tracciabilità:** ogni decisione presa dal sistema ADM deve essere documentabile e ricostruibile.
- **Verificabilità:** i processi devono poter essere controllati da soggetti terzi, come auditor interni o esterni.
- **Responsiveness:** i sistemi devono essere in grado di rispondere alle richieste di spiegazione o contestazione degli utenti.
- **Scalabilità:** le pratiche di trasparenza devono essere adattabili a differenti contesti e tipologie di utenti.

Gli autori sottolineano anche la necessità di bilanciare la trasparenza con altre esigenze, come la protezione della privacy degli utenti e la tutela delle informazioni proprietarie delle aziende.

Applicazioni nel Processo di Recruiting

Sebbene lo studio non si concentri esclusivamente sul settore delle risorse umane, le sue implicazioni si applicano direttamente ai processi di recruiting basati su sistemi ADM o IA. Nel contesto del recruiting, implementare la **trasparenza come virtù** significa che l'azienda deve promuovere la trasparenza non solo per rispettare la normativa, ma come espressione di responsabilità sociale d'impresa, tutelando i diritti e la dignità dei candidati.

La **trasparenza come relazione** si traduce nella necessità di fornire ai candidati spiegazioni chiare sulle modalità di funzionamento degli algoritmi di selezione: quali dati vengono raccolti, su quali criteri si basa il processo decisionale e come vengono valutate le candidature. È fondamentale che le informazioni siano presentate in modo comprensibile, evitando tecnicismi che possano escludere i non addetti ai lavori. Inoltre, i candidati devono avere la possibilità di contestare le decisioni automatizzate e di richiedere l'intervento umano.

Infine, la **trasparenza come sistema** richiede che le aziende dotate di sistemi ADM per il recruiting garantiscano la tracciabilità delle decisioni, permettendo verifiche da parte di organismi di controllo e assicurando l'adattabilità delle pratiche di trasparenza alle specifiche esigenze dei diversi processi di selezione.

L'adozione di un modello TbD nei processi di recruiting automatizzati offre numerosi vantaggi:

- Aumenta la fiducia dei candidati nei confronti dell'azienda.
- Migliora la percezione di equità nei processi di selezione.

- Riduce il rischio di bias e discriminazioni algoritmiche.
- Rafforza la reputazione dell'organizzazione come datore di lavoro etico e trasparente.

Tuttavia, Felzmann et al. evidenziano anche le sfide di questa implementazione. Tra i principali limiti vi sono il rischio di **sovraccarico informativo**, che potrebbe confondere gli utenti anziché chiarire i processi decisionali, e la necessità di proteggere i **dati sensibili** e le informazioni proprietarie delle aziende. Il modello proposto da Felzmann et al. rappresenta una guida pratica ed etica per tutte le organizzazioni che vogliono implementare l'intelligenza artificiale nei processi decisionali, incluso il recruiting. Adottare il principio di *Transparency by Design* significa non solo conformarsi alle normative vigenti, ma anche costruire un rapporto di fiducia con i candidati, valorizzando l'aspetto umano all'interno di un processo sempre più automatizzato. Questo approccio può contribuire in modo significativo a migliorare le performance aziendali nel lungo periodo, promuovendo una cultura organizzativa orientata all'equità e all'inclusione.

Capitolo 3

Metodologia della Ricerca

3.1 Disegno della ricerca

Il presente studio adotta un disegno quantitativo, correlazionale e confermativo, volto a indagare i meccanismi psicologici attraverso cui la trasparenza percepita dell'Intelligenza Artificiale (IA) influisce sulla fiducia nei sistemi automatizzati di recruiting, considerata in questo lavoro una misura rappresentativa dell'accettazione dell'IA nei contesti organizzativi.

L'obiettivo principale della ricerca è testare un modello di **mediazione psicologica**, secondo cui l'effetto della trasparenza sulla fiducia non si sviluppa unicamente in modo diretto, ma si articola attraverso **tre precursori cognitivi ed emotivi**: l'utilità percepita, la facilità d'uso percepita e la percezione del rischio occupazionale.

Questi costrutti vengono ipotizzati come **variabili mediatrici**, ossia come meccanismi interni attraverso cui la trasparenza viene psicologicamente elaborata e tradotta in fiducia.

La ricerca si articola secondo i seguenti **obiettivi** analitici:

- Verificare l'effetto diretto della trasparenza sulla fiducia nei sistemi IA;
- Esplorare l'effetto della trasparenza su ciascuna delle tre variabili mediatrici (utilità, facilità, rischio);
- Accertare se le variabili mediatrici predicano la fiducia anche controllando per la trasparenza;
- Testare la presenza di effetti mediati, secondo le condizioni previste dal modello di mediazione di Baron e Kenny (1986);
- Validare i risultati attraverso l'analisi bootstrap con PROCESS Model 4 (Hayes, 2013);
- Esplorare eventuali effetti di moderazione legati alla percezione soggettiva del rischio occupazionale.

La misurazione delle performance aziendali non rientra tra gli obiettivi primari, ma tali dimensioni sono considerate esiti teorici secondari, coerenti con quanto emerso dalla letteratura, che suggerisce come una maggiore fiducia nei sistemi IA possa portare benefici in termini di clima organizzativo, engagement e produttività.

Quadro teorico e costrutti del modello

Il modello teorico adottato si fonda su un'estensione del *Technology Acceptance Model* (TAM) (Davis, 1989), integrato con le più recenti ricerche sulla trasparenza e sulla fiducia nei sistemi di IA (Doshi-Velez & Kim, 2017; Shin, 2021; Wang et al., 2020). In particolare, il modello si distingue per l'introduzione delle seguenti componenti chiave:

- Trasparenza percepita dell'IA: intesa come capacità del sistema di rendere comprensibili le logiche decisionali algoritmiche, specialmente nei contesti sensibili come il recruiting.
- Fiducia nei sistemi automatizzati: operazionalizzata come percezione di competenza, eticità, affidabilità e imparzialità del sistema automatizzato di selezione.
- Variabili mediatrici:
 - Utilità percepita: quanto l'IA è ritenuta utile per il lavoro.
 - Facilità d'uso percepita: quanto l'IA è considerata intuitiva e accessibile.
 - Percezione del rischio occupazionale: grado di preoccupazione che l'IA possa minacciare la stabilità lavorativa o il ruolo dell'umano.

Il presente approccio considera quindi la fiducia non come un riflesso meccanico della trasparenza, ma come il risultato di una mediazione psicologica complessa, in cui l'informazione trasparente viene filtrata e interpretata attraverso valutazioni personali di utilità, usabilità e rischio.

3.2 Obiettivi e ipotesi di ricerca

La ricerca si propone di esplorare i meccanismi psicologici attraverso cui la trasparenza percepita dell'Intelligenza Artificiale (IA) incide sulla fiducia nei sistemi automatizzati di recruiting, concettualizzata come dimensione centrale dell'accettazione dell'IA nei contesti organizzativi. In particolare, lo studio indaga se l'effetto della trasparenza sulla fiducia si sviluppi attraverso tre precursori psicologici: l'utilità percepita, la facilità d'uso percepita e la percezione del rischio occupazionale.

Tali variabili sono state individuate dalla letteratura come fattori chiave dell'accettazione tecnologica (Venkatesh & Bala, 2008; Shin, 2021), e sono considerate qui sia come esiti diretti della trasparenza sia come mediatori nel rapporto tra trasparenza e fiducia. Il modello teorico si fonda sull'ipotesi che la trasparenza, per generare fiducia, debba essere elaborata cognitivamente in termini di vantaggio, comprensibilità e rassicurazione. Alla luce di questo quadro, sono state formulate le seguenti ipotesi, secondo lo schema classico di **Baron e Kenny (1986)** per la verifica della mediazione:

***H1:** La trasparenza percepita dell'IA è positivamente associata alla fiducia nei sistemi automatizzati di recruiting.*

Numerosi studi confermano il ruolo centrale della trasparenza nel rafforzare la fiducia nei confronti dei sistemi di IA (Glikson & Woolley, 2020; Shin, 2021). Nei contesti HR, dove le decisioni automatizzate hanno conseguenze rilevanti, la percezione che il sistema operi in modo accessibile e comprensibile è fondamentale per costruire fiducia. Pertanto, si ipotizza un effetto diretto positivo della trasparenza sulla fiducia ($X \rightarrow Y$).

***H2:** La trasparenza percepita dell'IA è positivamente associata all'utilità percepita, alla facilità d'uso percepita, e negativamente alla percezione del rischio occupazionale.*

Secondo il Technology Acceptance Model (TAM), l'utilità e la facilità d'uso sono i principali precursori dell'accettazione. La trasparenza – intesa come spiegabilità, chiarezza e accessibilità – facilita la comprensione delle logiche algoritmiche, riduce la complessità percepita e rassicura il lavoratore rispetto al proprio ruolo. In questo senso, una maggiore trasparenza dovrebbe aumentare la percezione che l'IA sia utile e facile da usare, e contemporaneamente ridurre il timore di sostituzione lavorativa ($X \rightarrow M_1, M_2, M_3$).

***H3:** L'utilità percepita, la facilità d'uso percepita e la percezione del rischio occupazionale influenzano significativamente la fiducia nei sistemi automatizzati, anche tenendo conto della trasparenza percepita.*

Perché si possa parlare di mediazione, è necessario verificare che ciascun mediatore (M_1, M_2, M_3) predica la fiducia anche in presenza della variabile indipendente (X). Come suggerito dalla letteratura, la fiducia nei sistemi IA dipende non solo da quanto essi appaiano trasparenti, ma anche da quanto siano percepiti come effettivamente utili, accessibili e non minacciosi ($M \rightarrow Y | X$).

H4: L'effetto della trasparenza sulla fiducia è mediato dall'utilità percepita, dalla facilità d'uso e dalla percezione del rischio occupazionale.

In linea con i modelli di adozione tecnologica estesi e con gli approcci basati sulla fiducia, si ipotizza che la trasparenza generi fiducia non solo direttamente, ma anche indirettamente, attraverso l'attivazione di valutazioni cognitive ed emotive. Se l'IA è percepita come trasparente, ma non come utile, facile da usare o rassicurante, è meno probabile che venga accettata. Pertanto, l'effetto della trasparenza si sviluppa attraverso un processo mediato da utilità, facilità e rischio ($X \rightarrow M \rightarrow Y$).

Questa articolazione ipotetica consente di validare sia le relazioni fondamentali tra trasparenza e variabili psicologiche (H1–H3), sia la struttura sistemica del modello, attraverso cui la fiducia si costruisce mediante un'elaborazione interna da parte del dipendente (H4). Una verifica empirica di queste ipotesi sarà condotta tramite analisi di regressione e mediazione, con successiva conferma mediante procedura bootstrap (modello PROCESS di Hayes) per validare statisticamente l'effetto mediato complessivo.

3.3 Strutturazione del questionario

Il questionario somministrato ai partecipanti è stato costruito con l'obiettivo di raccogliere informazioni sia di **natura sociodemografica**, sia relative alle **variabili psicologiche centrali del modello teorico**. La struttura dello strumento è stata articolata in due sezioni principali.

Sezione 1 – Dati sociodemografici

In questa sezione sono stati rilevati i seguenti dati:

- **Genere**
- **Età**
- **Livello di istruzione**
- **Area geografica di lavoro**

- **Anzianità lavorativa**
- **Posizione lavorativa attuale**
- **Settore industriale di appartenenza**

Queste variabili hanno consentito di descrivere il profilo del campione e sono state successivamente impiegate come **variabili di controllo** all'interno delle analisi inferenziali, in particolare nella regressione multipla.

Sezione 2 – Variabili teoriche del modello

La seconda sezione del questionario era composta da **item a risposta chiusa** appartenenti a scale validate in letteratura, usati per la misurazione delle variabili chiave del modello teorico.

Percezione della trasparenza dell'IA

La trasparenza percepita dell'intelligenza artificiale è stata misurata attraverso la scala proposta da Yu e Li (2022), che valuta l'accessibilità, la comprensibilità e la chiarezza delle informazioni fornite dai sistemi di IA.

Gli item includono affermazioni come:

- “Posso accedere a molte informazioni che spiegano come funziona il sistema di IA.”
- “Le informazioni fornite dal sistema di IA sono facili da comprendere.”

Le risposte sono state raccolte su una scala Likert a 7 punti (1 = molto in disaccordo, 7 = molto d'accordo).

Utilità percepita e Facilità d'uso percepita

Questi due costrutti derivano dal *Technology Acceptance Model* (TAM) di Davis (1989).

Per l'**utilità percepita**, gli item valutano quanto l'IA sia ritenuta utile per migliorare le performance lavorative.

Esempi di item:

- “L'IA migliora la mia efficacia sul lavoro.”
- “Usare l'IA aumenta la mia produttività.”

Per la **facilità d'uso percepita**, gli item indagano la semplicità e l'intuitività del sistema. Esempi di item:

- “Trovare ciò che voglio nel sistema di IA è facile.”
- “È facile per me ricordare come utilizzare il sistema di IA.”

Anche in questo caso è stata utilizzata una scala Likert a 7 punti.

Percezione del rischio occupazionale

Questa variabile è stata misurata tramite un singolo item validato da De Witte (2000):

- “Quanto è grande, secondo lei, la probabilità che lei perda il lavoro nel prossimo futuro?”

Le risposte sono state raccolte su una scala Likert a 7 punti (1 = molto alta, 7 = nulla).

Fiducia nei sistemi automatizzati di recruiting

La fiducia nei sistemi di IA nel contesto del recruiting è stata misurata tramite la scala sviluppata da Madhavan e Wiegmann (2007), che include quattro sottodimensioni:

1. **Competenza** (es. “I sistemi di IA hanno le capacità necessarie per aiutarmi a raggiungere i miei obiettivi.”)
2. **Benevolenza** (es. “I sistemi di IA sono progettati per garantire il benessere degli altri.”)
3. **Integrità** (es. “I sistemi automatizzati cercano di essere equi nel trattare con gli altri.”)
4. **Fiducia generale** (es. “Mi sento a mio agio nel fidarmi dei sistemi automatizzati.”)

Gli item sono stati valutati su una scala Likert a 5 punti (1 = per niente d’accordo, 5 = completamente d’accordo). La media delle quattro dimensioni è stata utilizzata come indicatore complessivo della fiducia nei sistemi di IA.

3.4 Affidabilità delle scale

Per verificare la coerenza interna delle scale utilizzate nel presente studio, è stato calcolato il coefficiente **Alpha di Cronbach** per ciascuna delle variabili composite oggetto d’analisi. Questo indice misura il grado di correlazione tra gli item che compongono una scala, valutando quanto essi siano coerenti nel misurare lo stesso costrutto latente. In generale, un valore di Alpha superiore a **.70** è considerato accettabile, valori superiori a **.80** indicano una buona affidabilità, mentre valori superiori a **.90** riflettono un’eccellente consistenza interna (George & Mallery, 2003)⁵².

I risultati dell’analisi confermano l’affidabilità elevata di tutte le scale adottate:

AITRASP	UTILITA	FACILITA	COMPET	BENEV	INTEG	FIDUCIAGEN
$\alpha = .931$	$\alpha = .975$	$\alpha = .957$	$\alpha = .947$	$\alpha = .925$	$\alpha = .932$	$\alpha = .940$

⁵² George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference (4th ed.)*. Boston: Allyn & Bacon.

Questi valori dimostrano un'ottima coerenza interna per tutte le dimensioni analizzate. In particolare, i costrutti fondamentali derivati dal *Technology Acceptance Model* (utilità e facilità) e quelli relativi alla fiducia nei sistemi automatizzati di recruiting mostrano livelli di affidabilità eccellenti, superiori a .90, confermando la qualità metrica degli strumenti impiegati.

L'elevato livello di affidabilità rilevato giustifica la costruzione di indici composti per ciascun costrutto, ottenuti attraverso la media degli item che li compongono, i quali sono stati poi utilizzati nelle successive analisi statistiche. L'affidabilità complessiva delle scale fornisce quindi una solida base per interpretare con rigore i risultati emersi nella fase inferenziale dello studio.

3.5 Campione e raccolta dati

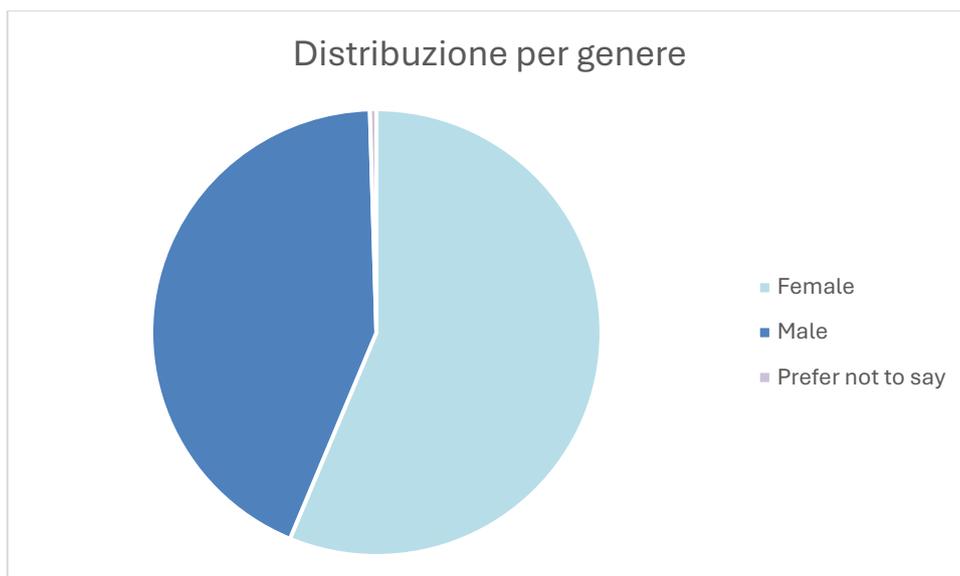
La raccolta dati è stata effettuata attraverso la piattaforma *Prolific*, un sistema online che consente il reclutamento mirato di partecipanti per studi di ricerca accademica. Questo strumento è stato scelto per la sua capacità di raggiungere in modo efficiente un campione ampio e diversificato, composto da soggetti attivamente inseriti nel mondo del lavoro.

I criteri di inclusione previsti erano i seguenti:

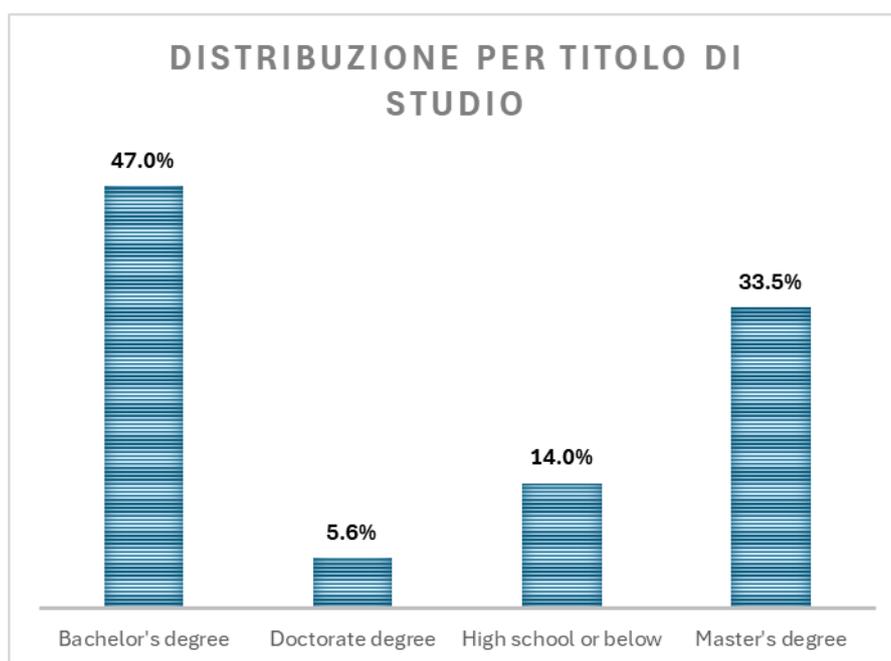
- essere attualmente impiegati in un'azienda;
- lavorare in una realtà organizzativa che utilizza tecnologie basate su intelligenza artificiale (a prescindere dal ruolo ricoperto nella gestione dell'IA);
- avere un'età superiore ai 18 anni;
- compilare il questionario in lingua italiana.

Il questionario è stato compilato da un totale di 217 partecipanti, dopo l'esclusione delle risposte inattendibili o incomplete. La raccolta è avvenuta in modalità completamente anonima e volontaria, nel rispetto della normativa vigente sulla privacy (GDPR) e dei principi etici della ricerca. Tutti i partecipanti hanno fornito il consenso informato prima di accedere al questionario.

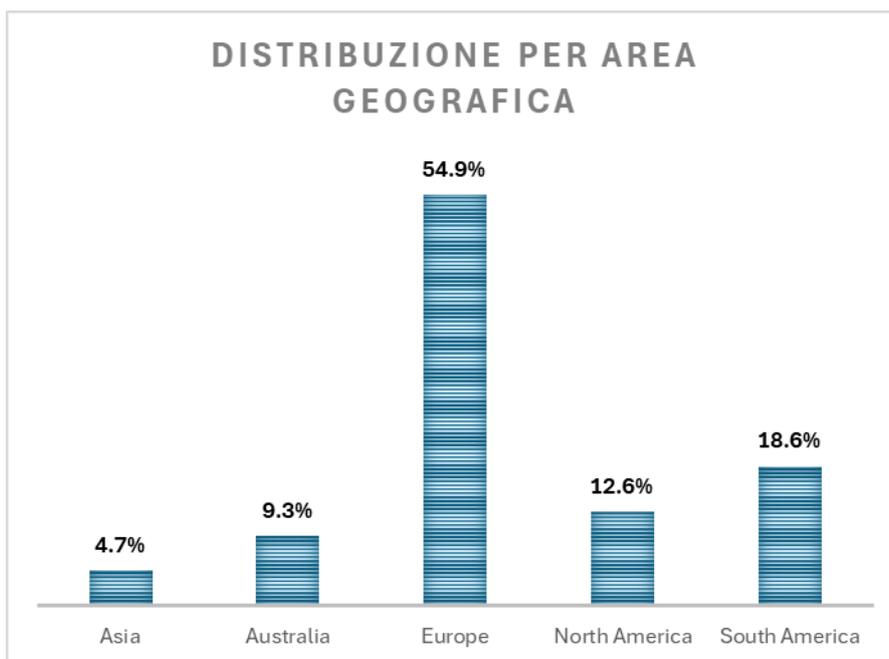
Il campione risultante presenta una composizione sociodemografica sufficientemente variegata, riflettendo una pluralità di settori, ruoli, aree geografiche e livelli di esperienza lavorativa. In termini di **genere**, il 56,28% dei partecipanti si identifica come donna, il 43,26% come uomo, mentre lo 0,47% ha preferito non specificare.



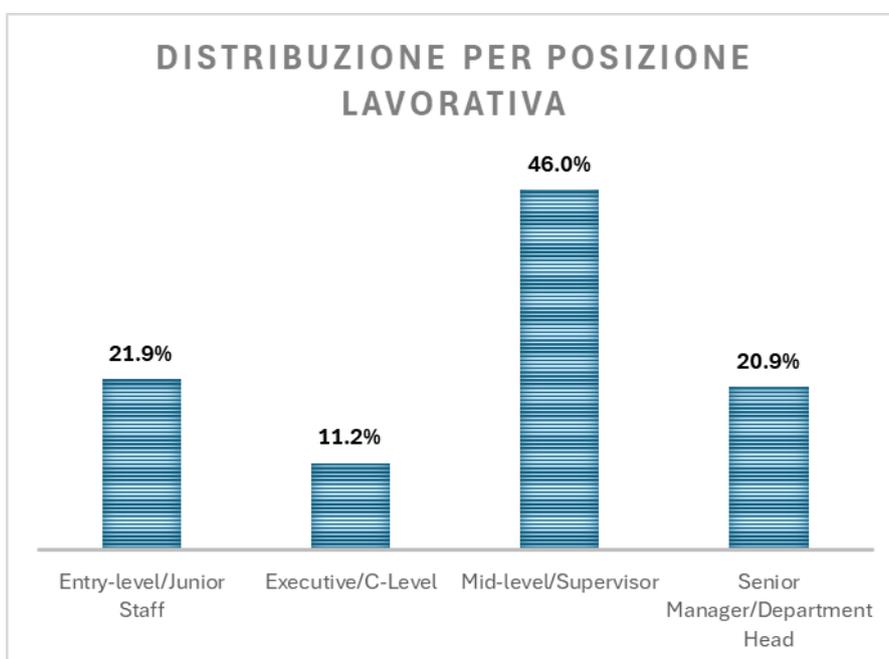
Con riferimento al **titolo di studio**, il 46,98% è in possesso di una laurea triennale, il 33,49% di una laurea magistrale, il 5,58% di un dottorato, mentre il 13,95% possiede un diploma di scuola superiore o titolo equivalente.



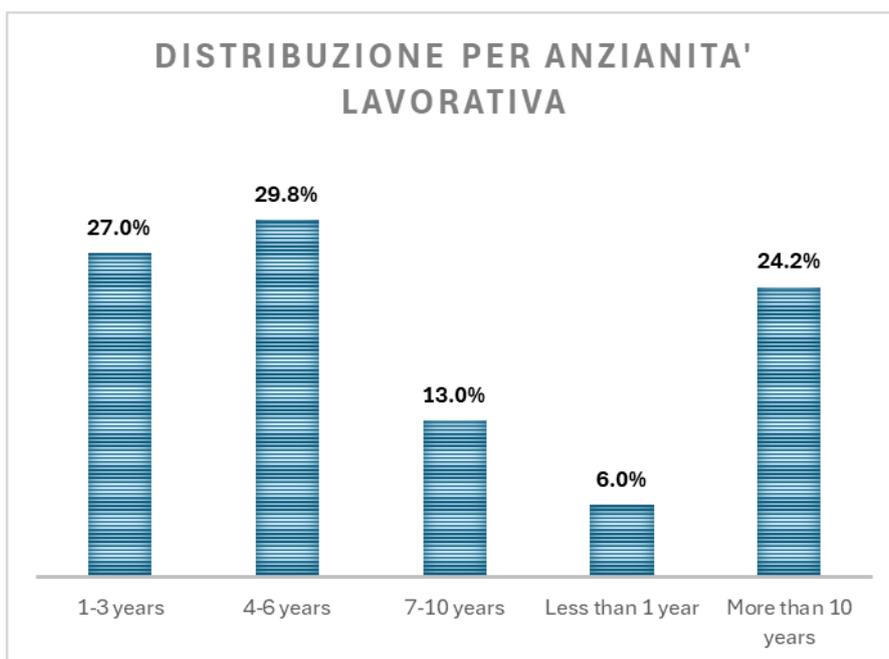
La distribuzione per **area geografica** mostra una netta prevalenza di rispondenti provenienti dall'Europa (54,88%), seguiti da partecipanti del Sud America (18,60%), Nord America (12,56%), Australia (9,30%) e Asia (4,65%).



Rispetto alla **posizione lavorativa**, la categoria più rappresentata è quella dei mid-level/supervisor (46,05%), seguita da entry-level/junior staff (21,86%), senior manager/department head (20,93%) ed executive/C-level (11,16%).

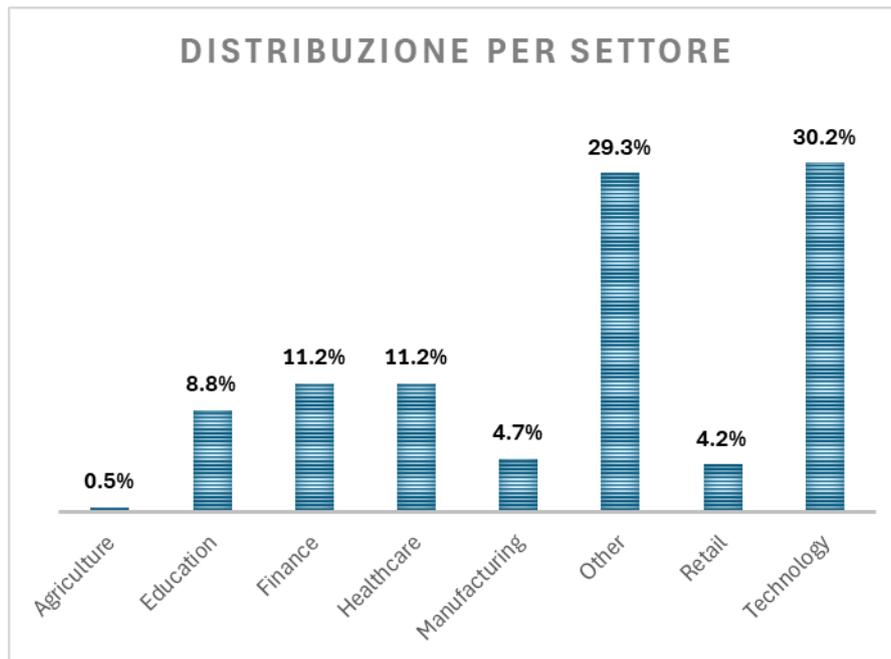


In termini di **anzianità lavorativa**, il 29,77% dei partecipanti lavora da 4 a 6 anni, il 26,98% da 1 a 3 anni, il 24,19% da oltre 10 anni, il 13,02% da 7 a 10 anni e il 6,05% da meno di un anno.



Per quanto riguarda il **settore** professionale, il 30,23% opera nel comparto tecnologico, il 29,30% in settori non specificati (“Altro”), l’11,16% rispettivamente nei settori finanziario e sanitario, l’8,84% in quello educativo, il 4,65% nella manifattura e il 4,19% nel retail; il settore agricolo è rappresentato dallo 0,47% del campione.

Nel complesso, il campione può essere considerato bilanciato e coerente con gli obiettivi della ricerca, offrendo una base empirica solida per analizzare le percezioni dei lavoratori nei confronti dell’intelligenza artificiale nei processi decisionali organizzativi.



3.6 Tecniche di analisi dei dati

L'analisi dei dati è stata condotta tramite il software *IBM SPSS Statistics*⁵³, seguendo una sequenza di fasi metodologiche coerenti con il modello teorico ipotizzato e con l'obiettivo di esplorare i meccanismi psicologici che mediano la relazione tra trasparenza percepita dell'IA e fiducia nei sistemi automatizzati di recruiting.

1. Analisi descrittive

Sono state calcolate le statistiche descrittive (media, deviazione standard, minimo, massimo) per ciascuna delle variabili chiave del modello. Questa fase ha permesso di comprendere la distribuzione dei punteggi, di rilevare eventuali outlier e di verificare la normalità delle distribuzioni.

⁵³ IBM SPSS Statistics è un software professionale di analisi statistica ampiamente utilizzato in ambito accademico e aziendale. È progettato per semplificare la gestione, la visualizzazione e l'analisi dei dati, offrendo una vasta gamma di procedure: dalla statistica descrittiva, alle regressioni, fino a modelli complessi come le analisi di mediazione e moderazione. La sua interfaccia user-friendly consente di eseguire analisi anche complesse senza la necessità di conoscere linguaggi di programmazione. In questa ricerca, SPSS è stato impiegato per tutte le fasi analitiche, in particolare per le analisi di regressione e per l'implementazione del modello PROCESS di Hayes. *IBM. (n.d.). Introduction to IBM SPSS Statistics. IBM Documentation.*

2. Verifica dell'affidabilità delle scale

Per ogni costrutto psicologico misurato tramite scala Likert, è stato calcolato il coefficiente di Alpha di Cronbach per valutare la coerenza interna. Tutte le scale hanno mostrato valori di $\alpha > .90$, indicativi di un'eccellente affidabilità psicometrica.

3. Analisi delle correlazioni bivariate (Pearson)

Sono state condotte analisi di correlazione di Pearson tra tutte le variabili teoriche per esaminare la presenza di *relazioni lineari preliminari* e verificare la *coerenza direzionale* delle ipotesi. Sebbene questa analisi non consenta di trarre inferenze causali, rappresenta un passaggio fondamentale per valutare l'opportunità di approfondire le relazioni tramite modelli di regressione e mediazione.

4. Analisi di regressione lineare e modello di mediazione di Baron e Kenny

Per testare le condizioni teoriche alla base del modello di mediazione, è stato adottato l'approccio classico proposto da *Baron e Kenny (1986)*⁵⁴, che rappresenta uno dei quadri metodologici più consolidati nello studio delle relazioni causali tra variabili. Tale approccio si basa su un insieme di quattro condizioni logiche e statistiche, la cui soddisfazione simultanea consente di ipotizzare l'esistenza di un effetto mediato.

In particolare, secondo gli autori, una variabile M può essere considerata mediatrice dell'effetto di una variabile indipendente X su una variabile dipendente Y se sono soddisfatte le seguenti condizioni:

1. X predice significativamente Y (effetto totale: $X \rightarrow Y$);
2. X predice significativamente M ($X \rightarrow M$);
3. M predice significativamente Y, anche controllando per X ($M \rightarrow Y | X$);
4. L'effetto diretto di X su Y si riduce (parzialmente o totalmente) quando M è incluso nel modello.

Nel presente studio, per verificare tali condizioni sono state condotte una serie di analisi di regressione lineare semplice, utilizzando la trasparenza percepita dell'IA come variabile indipendente, ciascuna delle tre variabili psicologiche ipotizzate (utilità, facilità d'uso, rischio occupazionale) come mediatori, e la fiducia nei sistemi automatizzati di recruiting come variabile dipendente. Questa fase ha avuto diversi obiettivi analitici:

- (1) testare le relazioni dirette previste dal modello teorico ($X \rightarrow M$ e $X \rightarrow Y$);
- (2) verificare se i mediatori influenzano significativamente la fiducia, controllando per la trasparenza ($M \rightarrow Y | X$);

⁵⁴ Baron, R. M., & Kenny, D. A. (1986). *The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations*. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.

- (3) osservare se l'effetto diretto della trasparenza si attenua dopo l'inclusione del mediatore ($X \rightarrow Y$ si riduce).

La soddisfazione congiunta di queste condizioni è stata interpretata come evidenza della presenza di un effetto mediato, secondo la logica causale del modello. Il ricorso al metodo di Baron e Kenny ha permesso di delineare in maniera strutturata i percorsi psicologici attraverso cui la trasparenza può influenzare la fiducia nei sistemi di IA.

5. Analisi di mediazione secondo l'approccio di Hayes (PROCESS Model 4)

Per approfondire la comprensione dei meccanismi psicologici attraverso cui la trasparenza percepita dell'IA influenza la fiducia nei sistemi automatizzati di recruiting, il presente studio ha applicato anche il modello di mediazione proposto da Andrew F. Hayes (2013), noto come *conditional process analysis*⁵⁵. Questo approccio consente di stimare in maniera diretta l'effetto indiretto di una variabile indipendente (X) su una variabile dipendente (Y) attraverso una o più variabili mediatrici (M), utilizzando la tecnica del *bootstrapping*. Quest'ultima si basa su un campionamento ripetuto (tipicamente migliaia di volte) dei dati originali, generando una distribuzione empirica dell'effetto indiretto calcolato come prodotto dei coefficienti ($X \rightarrow M$ e $M \rightarrow Y|X$). A differenza delle tecniche tradizionali, il bootstrap non richiede che tale effetto segua una distribuzione normale—una condizione spesso non rispettata nei modelli di mediazione, poiché l'effetto indiretto, essendo un prodotto di due coefficienti, tende ad assumere una distribuzione asimmetrica e centrata vicino allo zero. L'approccio bootstrap consente quindi di calcolare intervalli di confidenza più accurati e affidabili, basati sulla variabilità reale dei dati, e di determinare la significatività statistica degli effetti indiretti senza fare assunzioni teoriche troppo rigide (Hayes, 2013; Preacher & Hayes, 2008)⁵⁶. Secondo Hayes, una variabile può essere considerata un mediatore quando l'effetto indiretto ($X \rightarrow M \rightarrow Y$) risulta statisticamente significativo. A tal fine, è sufficiente che l'intervallo di confidenza ottenuto tramite bootstrap per l'effetto indiretto non includa lo zero. Diversamente dal modello di Baron e Kenny, non è necessario che l'effetto diretto tra X e Y sia significativo o che si riduca dopo l'inserimento del mediatore. In tal senso, la significatività dell'effetto indiretto è la condizione cruciale per determinare la presenza di una mediazione.

⁵⁵ Prado, P. H. M., Korelo, J. C., & Silva, D. M. L. (2014). Mediation, Moderation and Conditional Process Analysis. *Revista Brasileira de Marketing*, 13(4).

⁵⁶ Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.

Rispetto al modello sequenziale di Baron e Kenny (1986), il metodo di Hayes è oggi considerato più solido e statisticamente robusto, in quanto si focalizza direttamente sull'effetto mediato e supera le limitazioni legate alle assunzioni di normalità e ai vincoli sulle relazioni dirette. Alla luce di ciò, nel presente studio l'analisi condotta con il *PROCESS Macro di Hayes (Model 4)*⁵⁷ è stata adottata non solo come principale tecnica di stima dell'effetto mediato, ma anche come controprova statistica dei risultati ottenuti con il modello classico, che mantiene comunque una forte valenza teorica e sequenziale utile per esplorare la plausibilità dei percorsi causali ipotizzati (MacKinnon, Fairchild, & Fritz, 2007⁵⁸; Hayes & Rockwood, 2017⁵⁹).

In sintesi, l'utilizzo combinato del metodo classico e del metodo bootstrap consente di integrare la chiarezza esplicativa del modello di Baron e Kenny con la precisione statistica del modello di Hayes, offrendo una doppia validazione delle ipotesi di mediazione avanzate.

6. Analisi di moderazione (PROCESS Model 1)

Infine, è stata esplorata la possibilità che la percezione del rischio occupazionale non agisca solo come mediatore, ma anche come variabile moderatrice nella relazione tra trasparenza e fiducia. A tal fine, è stato utilizzato il *modello 1 di PROCESS*⁶⁰, che consente di verificare se l'effetto della trasparenza sulla fiducia varia in funzione del livello percepito di rischio lavorativo, indagando eventuali interazioni significative tra le variabili.

⁵⁷ Il PROCESS Model 4, sviluppato da Hayes (2013), consente di testare modelli di mediazione semplice, in cui una variabile indipendente (X) influisce su una variabile dipendente (Y) attraverso una o più variabili mediatrici (M). L'effetto indiretto (cioè mediato) è stimato tramite la moltiplicazione dei coefficienti dei due percorsi ($X \rightarrow M$ e $M \rightarrow Y|X$), ed è validato statisticamente attraverso il metodo del bootstrap. La mediazione si considera significativa se l'intervallo di confidenza bootstrap al 95% non include lo zero. Questo modello permette di verificare la presenza di un meccanismo causale intermediato, senza richiedere la significatività dell'effetto diretto tra X e Y, superando i limiti del modello sequenziale tradizionale.

Abbu, H. R. (2017). *Methodology Review: Direct, Mediating, and Moderating Analysis using Hayes PROCESS Models*.

⁵⁸ MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58, 593–614.

⁵⁹ Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*, 98, 39–57.

⁶⁰ Il PROCESS Model 1 è impiegato per analizzare l'effetto di moderazione, ovvero per verificare se la relazione tra una variabile indipendente (X) e una variabile dipendente (Y) cambia in funzione del livello di una terza variabile detta moderatore (W). Il modello introduce un termine di interazione ($X \times W$) nel modello di regressione, il cui coefficiente permette di valutare se l'intensità dell'effetto di X su Y varia in base ai livelli di W. La moderazione si considera significativa quando il coefficiente di interazione è statisticamente diverso da zero ($p < .05$ e intervallo di confidenza bootstrap che non include lo zero).

Abbu, H. R. (2017). *Methodology Review: Direct, Mediating, and Moderating Analysis using Hayes PROCESS Models*.

Questa sequenza di analisi ha permesso di esplorare in modo rigoroso e articolato non solo se la trasparenza influenzi la fiducia nei sistemi IA, ma soprattutto come tale influenza si sviluppi, testando il funzionamento psicologico interno del modello teorico proposto.

Capitolo 4

Analisi dei risultati

4.1 Analisi descrittive

Al fine di esplorare la distribuzione delle risposte e verificare la coerenza generale dei dati raccolti, sono state in primis condotte delle analisi descrittive sui principali costrutti teorici oggetto dello studio. Per poter effettuare le analisi descrittive, dato che ciascuna variabile possiede più item, è stata prima calcolata una media composita per ciascun partecipante. Successivamente, su quelle medie compositive sono poi state calcolate le statistiche descrittive generali di quella determinata variabile. In particolare, per ciascuna variabile è stata calcolata la **media aritmetica, la deviazione standard, il valore minimo e il valore massimo**. Tali analisi hanno permesso di valutare la tendenza centrale e la variabilità delle risposte per ogni costrutto considerato, fornendo una prima panoramica del profilo psicologico e percettivo degli intervistati in relazione al tema dell'intelligenza artificiale nei contesti aziendali.

	N	Minimum	Maximum	Mean	Std. Deviation
TRASP	215	1	7	4.86	1.51
UTILITA	215	1	7	5.01	1.42
FACILITA	215	1	7	5.20	1.06
RISCHIO	215	1	7	3.23	1.97
COMPET	215	1	7	4.31	1.60
BENEV	215	1	7	4.64	1.41
INTEGRITA	215	1	7	4.83	1.36
FIDUCIAGEN	215	1	7	4.43	1.62
FIDUCIAREC	215	1	7	4.55	1.40

I costrutti analizzati includono:

- la trasparenza percepita dell'IA (TRASP),
- la percezione dell'utilità e della facilità d'uso dell'IA nei processi decisionali aziendali (UTILITA' e FACILITA'),
- la paura di perdere il lavoro in seguito all'introduzione dell'IA (RISCHIO),
- diverse dimensioni della fiducia nei sistemi di intelligenza artificiale applicati al recruiting (competenza (COMPET), etica (BENEV), integrazione (INTEG) e fiducia globale (FIDUCIAGEN) e la media complessiva di tali dimensioni (FIDUCIAREC), per avere una sintesi della percezione della fiducia nei sistemi di automatizzati di recruiting.

I costrutti qui analizzati rappresentano le variabili centrali del modello teorico proposto: la trasparenza come antecedente psicologico, l'utilità e la facilità come precursori dell'accettazione, il rischio come barriera emotiva e la fiducia come esito psicologico finale.

Analizzando la media relativa alla **trasparenza percepita dell'IA** ($M = 4.86$, $DS = 1.51$), emerge una **tendenza positiva**, sebbene non estrema.

Questo suggerisce che i partecipanti **tendono ad avere una percezione moderatamente favorevole** della trasparenza dei sistemi di intelligenza artificiale, riconoscendo un certo livello di **chiarezza e comprensibilità** nelle modalità con cui l'IA prende decisioni. Tuttavia, la presenza di una deviazione standard pari a 1.51 indica una **discreta variabilità nelle risposte**, segnalando che non tutti gli utenti condividono lo stesso livello di fiducia nella trasparenza dell'IA. Alcuni la percepiscono come molto trasparente, mentre altri mostrano riserve o incertezza.

Questo dato potrebbe riflettere una diversità di esperienze, conoscenze pregresse o livelli di familiarità con i sistemi di IA.

La media della variabile "**utilità percepita**" è pari a 5.01 ($DS = 1.41$).

Questo risultato evidenzia una **percezione generalmente positiva** dell'efficacia dell'IA nel supportare le decisioni aziendali.

I partecipanti tendono a riconoscere un beneficio concreto nell'integrazione dell'intelligenza artificiale nei processi organizzativi.

La deviazione standard, moderata (1.41), indica che **le risposte sono relativamente coerenti**, pur lasciando spazio a **differenze soggettive**.

La facilità d'uso percepita dell'IA ha ottenuto una media di 5.20 (DS = 1.05), risultando essere **la più alta tra i costrutti analizzati**.

Questo indica che, secondo i partecipanti, i sistemi basati su IA sono percepiti come semplici da utilizzare, accessibili e non particolarmente complessi.

La deviazione standard più contenuta (1.05) rispetto agli altri costrutti suggerisce che questa percezione è **condivisa in modo più omogeneo** nel campione.

La variabile relativa alla percezione del rischio occupazionale legato all'IA ha registrato una media di 3.23 (DS = 1.96).

Si tratta del punteggio medio **più basso tra i costrutti analizzati**, indicando una **percezione relativamente bassa** del rischio di perdere il lavoro a causa dell'intelligenza artificiale.

Tuttavia, la **deviazione standard molto alta (1.96)** mostra che il campione è molto diviso su questo tema: mentre alcuni non percepiscono alcuna minaccia, altri vivono la presenza dell'IA con maggiore preoccupazione. Questo dato suggerisce che la paura legata all'automazione **non è diffusa in modo omogeneo**, ma varia in base a caratteristiche individuali e lavorative.

La fiducia nei sistemi automatizzati utilizzati nei processi di selezione è stata valutata attraverso quattro dimensioni fondamentali: **competenza, benevolenza, integrità e atteggiamento fiducioso totale (trusting stance)**. Le statistiche descrittive dei punteggi medi evidenziano quanto segue:

- **Competenza percepita dell'IA:** il valore medio pari a 5.03 (DS = 1.51) indica che i partecipanti percepiscono i sistemi automatizzati come generalmente competenti, cioè capaci di svolgere in modo efficace le funzioni per cui sono progettati. La deviazione standard mostra una moderata variabilità tra le risposte: la maggior parte degli utenti considera i sistemi IA affidabili a livello tecnico, anche se non tutti con lo stesso grado di convinzione.
- **Benevolenza percepita dell'IA:** con una media pari a 4.64 (DS = 1.41), la percezione che i sistemi automatizzati abbiano buone intenzioni e siano progettati per fare il bene degli utenti risulta leggermente positiva, ma più cauta rispetto ad altri aspetti. La variabilità è moderata, suggerendo che alcuni partecipanti potrebbero essere più diffidenti rispetto all'intenzionalità "etica" attribuita all'IA, mentre altri ne riconoscono un potenziale "benevolo".

- **Integrità percepita dell'IA:** il costrutto “integrità” ha ottenuto una media di 5.03 (DS = 1.51), sovrapponibile a quella della competenza. I sistemi automatizzati sono quindi percepiti come onesti, coerenti e affidabili. Anche in questo caso, la deviazione standard evidenzia che non tutti gli utenti condividono questa opinione allo stesso modo, ma c'è una tendenza generale a riconoscere una certa correttezza operativa.
- **Trusting stance (disposizione generale alla fiducia):** questo sotto-costrutto ha ottenuto una media di 5.03 (DS = 1.51), identica a quella della competenza e dell'integrità. Ciò significa che gli intervistati mostrano una propensione generale ad affidarsi ai sistemi IA, anche in assenza di prove concrete contrarie. La discreta variabilità (DS > 1.5) indica però che questa disposizione non è universale: mentre molti partecipanti mostrano una fiducia di base, altri potrebbero richiedere condizioni o garanzie specifiche per potersi fidare.

Nel complesso, i partecipanti esprimono livelli moderatamente alti di fiducia nei confronti dei sistemi IA impiegati nel recruiting, soprattutto in termini di competenza tecnica, integrità e fiducia generale.

Tuttavia, la dimensione della benevolenza risulta leggermente più contenuta, suggerendo che le buone intenzioni dei sistemi automatizzati sono percepite con più cautela rispetto alla loro efficacia tecnica.

4.2 CORRELAZIONI BIVARIATE

A valle delle analisi descrittive, sono state esaminate le relazioni lineari bivariate (**Coefficiente di Pearson**) tra le variabili chiave del modello teorico al fine di verificare la presenza di relazioni significative in linea con le ipotesi formulate. Tali correlazioni costituiscono un passaggio preliminare fondamentale per le successive analisi di regressione e mediazione.

Nello specifico, il modello teorico ipotizza che:

- La trasparenza percepita dell'IA (TRASP) sia positivamente associata all'utilità percepita (UTILITA), alla facilità d'uso percepita (FACILITA) e alla fiducia nei sistemi automatizzati di recruiting (FIDUCIA/FIDUCIAREC);
- La trasparenza sia negativamente associata alla percezione del rischio occupazionale (RISCHIO);
- L'utilità, la facilità e la percezione del rischio costituiscano precursori psicologici della fiducia, e quindi dovrebbero correlare significativamente con la fiducia (sia generale che specifica).

		TRASP	UTILITA	FACILITA	COMPET	BENEV	INTEGRITA	FIDUCIAGEN	RISCHIO	FIDUCIAREC
TRASP	Pearson	1	.527**	.664**	.543**	.395**	.419**	.501**	.137*	.500**
	Correlati on Sig. (2- tailed)		0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.000
UTILITA	Pearson	.527**	1	.801**	.534**	.498**	.535**	.522**	.229**	.558**
	Correlati on Sig. (2- tailed)	0.000		0.000	0.000	0.000	0.000	0.000	0.001	0.000
FACILITA	Pearson	.664**	.801**	1	.567**	.503**	.591**	.558**	.153*	.593**
	Correlati on Sig. (2- tailed)	0.000	0.000		0.000	0.000	0.000	0.000	0.025	0.000
COMPET	Pearson	.543**	.534**	.567**	1	.837**	.778**	.846**	.252**	.929**
	Correlati on Sig. (2- tailed)	0.000	0.000	0.000		0.000	0.000	0.000	0.000	0.000
BENEV	Pearson	.395**	.498**	.503**	.837**	1	.869**	.842**	.147*	.944**
	Correlati on Sig. (2- tailed)	0.000	0.000	0.000	0.000		0.000	0.000	0.031	0.000
INTEGRITA	Pearson	.419**	.535**	.591**	.778**	.869**	1	.840**	0.127	.926**
	Correlati on Sig. (2- tailed)	0.000	0.000	0.000	0.000	0.000		0.000	0.063	0.000
FIDUCIAGEN	Pearson	.501**	.522**	.558**	.846**	.842**	.840**	1	.162*	.945**
	Correlati on Sig. (2- tailed)	0.000	0.000	0.000	0.000	0.000	0.000		0.018	0.000
RISCHIO	Pearson	.137*	.229**	.153*	.252**	.147*	0.127	.162*	1	.186**
	Correlati on Sig. (2- tailed)	0.045	0.001	0.025	0.000	0.031	0.063	0.018		0.006
FIDUCIAREC	Pearson	.500**	.558**	.593**	.929**	.944**	.926**	.945**	.186**	1
	Correlati on Sig. (2- tailed)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.006	

I risultati delle **correlazioni Pearson** mostrano quanto segue:

- **TRASP è positivamente correlata** con UTILITA ($r = .527, p < .001$), FACILITA ($r = .664, p < .001$) e FIDUCIAREC ($r = .500, p < .001$), confermando l'ipotesi secondo cui la trasparenza alimenta sia le dimensioni cognitive dell'adozione tecnologica sia la fiducia nei sistemi automatizzati;
- La trasparenza mostra inoltre una **correlazione negativa, seppur debole**, con il RISCHIO ($r = .137, p = .045$), suggerendo che una maggiore percezione di trasparenza è associata a una **riduzione delle preoccupazioni legate alla sostituzione occupazionale**, ma con un effetto limitato;
- UTILITA e FACILITA sono **fortemente correlate tra loro** ($r = .801, p < .001$), e risultano entrambe **significativamente associate alla fiducia generale** (FIDUCIAGEN: $r = .522$ e $r = .558, p < .001$) e

alla fiducia complessiva (FIDUCIAREC: $r = .558$ e $r = .593$, $p < .001$), sostenendo l'ipotesi che esse agiscano da **mediatori** nel processo che conduce all'accettazione;

- Anche il RISCHIO risulta **correlato con la fiducia generale** ($r = .137$, $p = .045$) e con FIDUCIAREC ($r = .147$, $p = .031$), ma con valori più bassi. Questo suggerisce che la percezione del rischio, pur avendo un impatto, è meno fortemente legata alla fiducia rispetto agli altri precursori.

Nel complesso, le correlazioni osservate sono coerenti con le **ipotesi del modello teorico**, e indicano che le relazioni tra trasparenza, precursori psicologici (utilità, facilità, rischio) e fiducia sono **tutte statisticamente significative**. Queste evidenze giustificano il passaggio alle **analisi di regressione e mediazione**, necessarie per testare la **direzione e l'intensità** degli effetti proposti.

4.3 Analisi di regressione

Dopo aver esplorato le associazioni tra le variabili teoriche mediante le **correlazioni bivariate di Pearson**, il passo successivo ha riguardato la verifica delle relazioni **direzionali** previste dal modello teorico attraverso una serie di **regressioni lineari semplici**. Le correlazioni avevano già mostrato un pattern coerente con le ipotesi: la trasparenza percepita dell'IA risultava **positivamente associata** all'utilità percepita, alla facilità d'uso e alla fiducia nei sistemi automatizzati di recruiting, e **negativamente associata** alla percezione del rischio occupazionale.

Tuttavia, l'analisi delle correlazioni non è sufficiente per valutare relazioni di influenza unidirezionale né per testare la presenza di effetti mediati. Per tale motivo, è stata condotta una sequenza di **analisi di regressione lineare** in quattro step distinti, al fine di verificare le condizioni teoriche e statistiche necessarie per un'analisi di mediazione, secondo l'approccio classico di **Baron e Kenny (1986)**.

Step 1 – Effetto diretto della trasparenza sulla fiducia ($X \rightarrow Y$)

Il primo step ha previsto la verifica dell'effetto diretto tra trasparenza e fiducia nei sistemi automatizzati di recruiting. Questa analisi, corrispondente all'ipotesi H1, ha consentito di valutare se la trasparenza eserciti un'influenza autonoma sull'accettazione dell'IA, operanzializzata nella fiducia organizzativa verso tali sistemi. Inoltre, questo passaggio è servito a misurare il peso dell'effetto diretto prima di includere i mediatori nel modello.

Step 2 – Effetto della trasparenza sui mediatori ($X \rightarrow M$)

Nel secondo step, si è analizzato l'effetto della variabile indipendente (trasparenza percepita dell'IA) su ciascuno dei tre potenziali mediatori: **utilità percepita, facilità d'uso percepita e percezione del rischio occupazionale**. Questo passaggio ha permesso di verificare le ipotesi H2, ovvero che la trasparenza operi come variabile antecedente capace di attivare processi cognitivi e attitudinali che precedono la fiducia. La significatività di queste relazioni rappresenta la **seconda condizione** del modello di mediazione.

Step 3 – Effetto dei mediatori sulla fiducia, controllando per la trasparenza ($M \rightarrow Y | X$)

Nel terzo step, ciascuna delle tre variabili mediatrici è stata testata come predittore della fiducia, **inclusa nel modello insieme alla trasparenza**, in modo da verificare se il mediatore mantenga un effetto significativo sulla variabile dipendente **controllando per X**. Questa condizione è necessaria per dimostrare che l'effetto della trasparenza sulla fiducia **passa attraverso** i mediatori e rappresenta la **terza condizione teorica** prevista dal modello di Baron e Kenny.

Step 4 – Riduzione dell'effetto diretto della trasparenza dopo l'inclusione del mediatore ($X \rightarrow Y \downarrow | M$)

Nel quarto e ultimo step, si è osservata l'entità del coefficiente di regressione della trasparenza sulla fiducia, confrontando il modello semplice (senza mediatori) con quello che include ciascun mediatore. L'obiettivo era valutare se, una volta introdotto il mediatore nel modello, l'effetto diretto della trasparenza sulla fiducia diminuisce rispetto al modello iniziale. Questa riduzione rappresenta un'indicazione cruciale della presenza di una mediazione, in quanto segnala che parte dell'effetto di X sulla Y è trasmesso attraverso il mediatore. La significatività residua dell'effetto diretto non è una condizione necessaria: anche se permane, la mediazione può dirsi parziale, mentre una sua scomparsa completa indicherebbe una mediazione totale.

Il completamento di questi quattro step ha permesso di verificare la **struttura delle relazioni** teoricamente ipotizzate e di procedere con maggiore robustezza all'**analisi formale di mediazione**, realizzata successivamente tramite il **modello 4 del PROCESS macro di Hayes (2013)**. Tale approccio ha consentito di stimare l'**effetto indiretto** della trasparenza sulla fiducia attraverso i tre precursori psicologici, confermando o meno la presenza di un **processo mediato** alla base dell'accettazione dell'IA nei contesti di recruiting automatizzato.

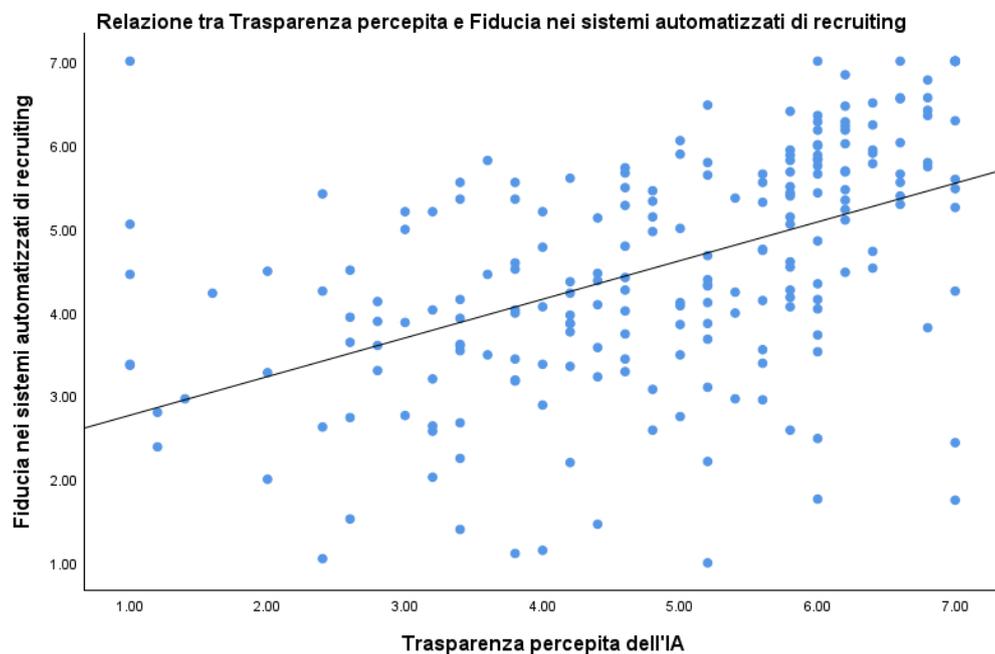
FASE 1

ANALISI DI REGRESSIONE LINEARE TRA TRASPARENZA E FIDUCIA NEL SISTEMA DI RECRUITING

Questa prima analisi di regressione lineare mira a verificare se la trasparenza percepita dell'intelligenza artificiale nei processi decisionali aziendali costituisca un predittore significativo della fiducia nei sistemi automatizzati di selezione del personale. La fiducia è considerata in questo studio come indicatore dell'accettazione dell'IA nei contesti organizzativi, ed è operazionalizzata attraverso un costrutto composito che include la percezione di competenza tecnica, eticità, integrità e affidabilità dei sistemi.

Domanda di ricerca: la trasparenza percepita dell'IA è in grado di spiegare in modo significativo la fiducia nei sistemi automatizzati di recruiting?

Var dipendente	Predictor	p-value	R ²	β
FIDUCIA	TRASP	p < .001	0.250	0.500



Il grafico mostra la relazione positiva tra la trasparenza percepita e la fiducia: all'aumentare della trasparenza (asse X), si osserva un incremento tendenziale nella fiducia (asse Y). La linea di regressione evidenzia questa tendenza, supportando l'ipotesi che la trasparenza sia un predittore significativo della fiducia nei sistemi IA.

Tuttavia, per l'interpretazione dettagliata dell'intensità e significatività della relazione, è necessario fare riferimento ai valori di β , p-value e R^2 riportati di seguito.

Coefficiente di regressione (Beta standardizzato): Il valore di $\beta = .500$ indica che la trasparenza percepita ha un effetto positivo e moderato sulla fiducia complessiva. In altre parole, all'aumentare della trasparenza, aumenta anche la fiducia nei sistemi automatizzati di recruiting. Questo supporta l'ipotesi teorica secondo cui la trasparenza rappresenta un antecedente diretto dell'accettazione dell'IA, operazionalizzata attraverso la fiducia.

Significatività statistica: Il modello è altamente significativo ($p < .001$), il che conferma che la relazione osservata tra trasparenza e fiducia non è frutto del caso, ma è statisticamente robusta.

R^2 (coefficiente di determinazione): Il valore di $R^2 = .250$ indica che la trasparenza percepita spiega il **25% della varianza** della fiducia nei sistemi automatizzati. Si tratta di un contributo sostanziale, soprattutto considerando che la fiducia è un costrutto complesso e multidimensionale.

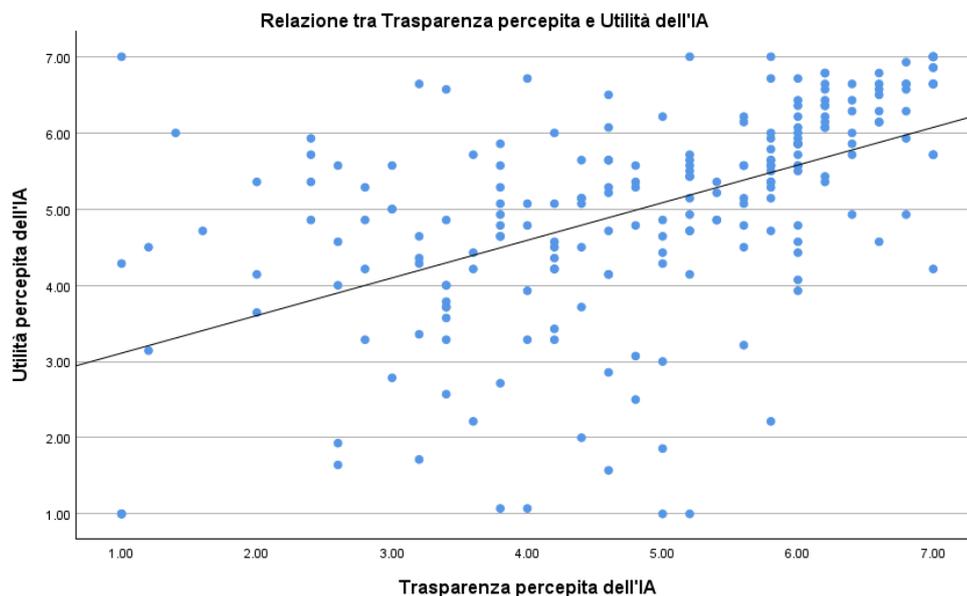
Questa regressione lineare semplice ha evidenziato che la trasparenza percepita dell'intelligenza artificiale nei processi decisionali è un predittore significativo della fiducia nei sistemi automatizzati di selezione del personale. Ciò indica che all'aumentare della trasparenza percepita, aumenta anche il livello di fiducia attribuito a tali sistemi, e che la trasparenza spiega circa il 25% della varianza nella fiducia.

FASE 2

ANALISI DI REGRESSIONE LINEARE TRA TRASPARENZA E UTILITÀ PERCEPITA NEL DECISION MAKING

Proseguendo con le analisi di regressione, al fine della validazione del modello teorico, è stata condotta un'analisi di regressione tra la media della trasparenza percepita dell'IA e la media di utilità percepita nel processo di decision making. In particolare, la domanda di ricerca si propone di capire quanto la trasparenza spiega l'utilità percepita nel decision making aziendale. I risultati sono stati i seguenti:

Var dipendente	Predictor	p-value	R^2	β
UTILITÀ	TRASP	$p < .001$	0.277	0.527

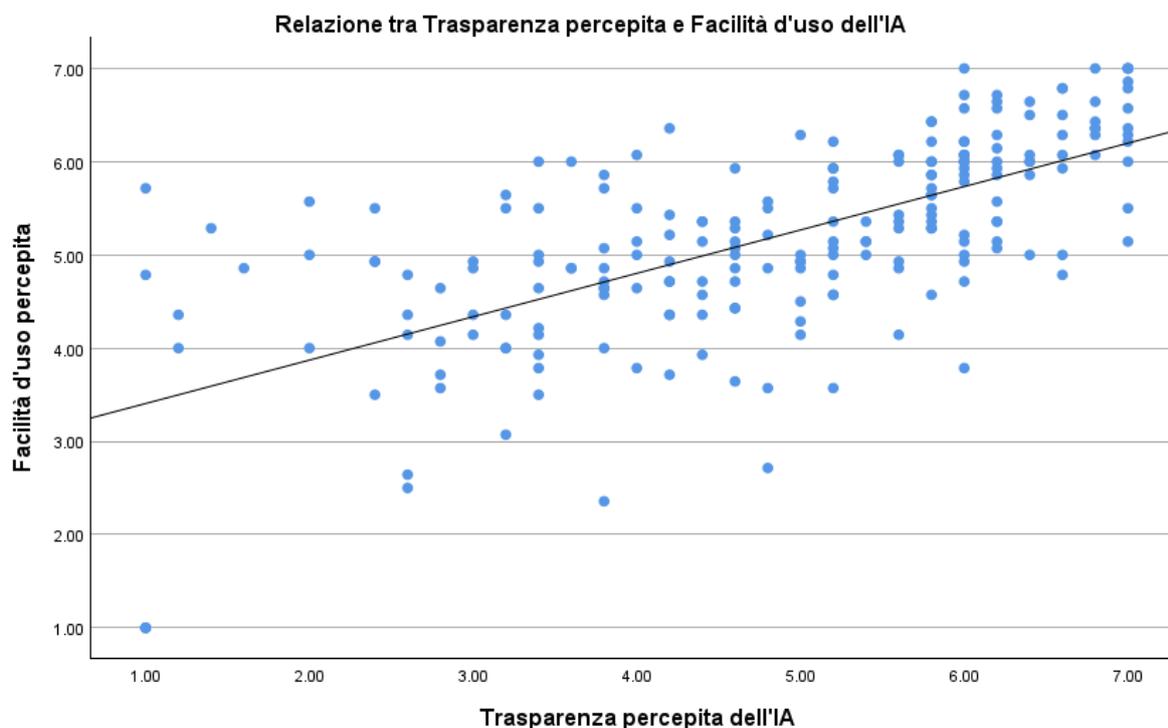


L'analisi di regressione lineare mostra che la **trasparenza percepita dell'IA** è un predittore significativo dell'**utilità percepita dell'IA** nei processi decisionali aziendali. Il coefficiente di regressione standardizzato è $\beta = 0.527$, indicando un effetto **positivo e di intensità moderata-alta**. Il modello risulta statisticamente significativo ($p < .001$) e spiega **il 27.7% della varianza** nella percezione dell'utilità dell'IA ($R^2 = 0.277$), suggerendo che una maggiore trasparenza contribuisce in modo rilevante a far percepire l'IA come più utile. Dunque, l'analisi di regressione suggerisce che una maggiore trasparenza percepita dell'IA è significativamente associata a un aumento dell'utilità percepita nel decision-making. Sebbene la trasparenza spieghi una parte rilevante della variazione nell'utilità percepita (27.7%), esistono altri fattori non inclusi nel modello che influenzano questa percezione.

ANALISI DI REGRESSIONE LINEARE TRA TRASPARENZA E FACILITA' D'USO PERCEPITA

L'analisi di regressione lineare semplice mostra che la trasparenza percepita dell'IA è un fattore determinante nella percezione della facilità d'uso. In altre parole, quando le persone percepiscono l'IA come più chiara, accessibile e spiegabile, tendono anche a valutarla come più facile da usare.

Var dipendente	Predictor	p-value	R ²	β
FACILITA	TRASP	p < .001	0.441	0.664



Coefficiente di regressione standardizzato (Beta) = .664: indica una forte associazione positiva tra trasparenza percepita dell'IA e facilità d'uso percepita. In altre parole, all'aumentare della trasparenza, aumenta in modo significativo la percezione di facilità nell'uso dei sistemi IA.

Significatività (Sig.) = < .001: il risultato è altamente significativo, suggerendo che l'effetto osservato non è dovuto al caso.

R² (R Square) = .441: la trasparenza spiega il **44,1%** della varianza nella percezione della facilità d'uso, un valore elevato che indica un'influenza molto rilevante del predittore sul costrutto dipendente.

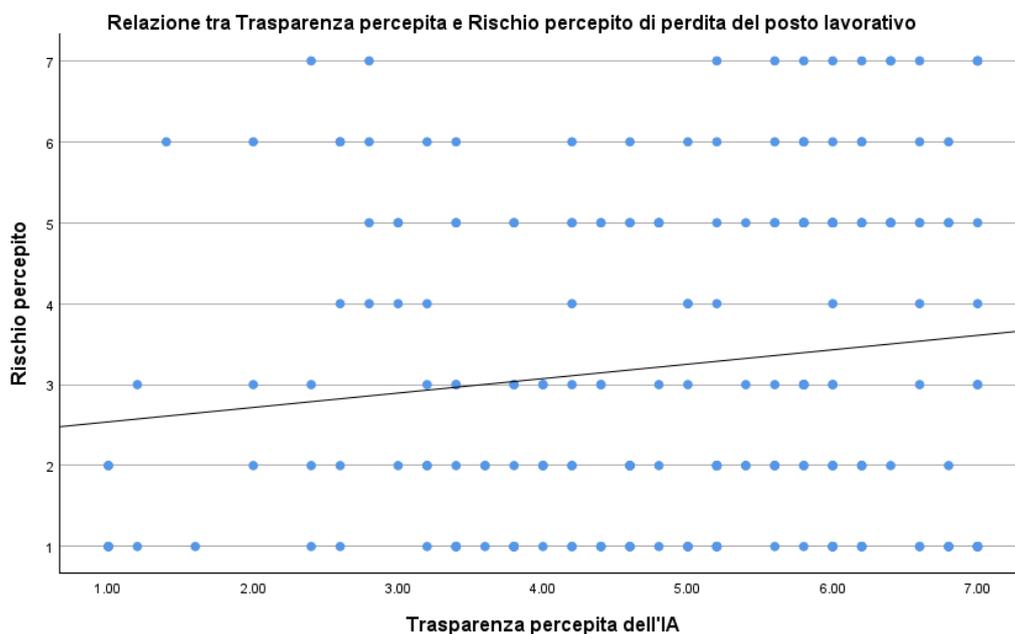
Dunque, l'analisi di regressione lineare semplice ha evidenziato una relazione significativa tra la trasparenza percepita dell'intelligenza artificiale e la facilità d'uso percepita. Il modello suggerisce che una maggiore

trasparenza è associata a una percezione più positiva della semplicità di utilizzo dei sistemi basati su IA. Questo risultato conferma l'ipotesi teorica secondo cui la comprensibilità e l'accessibilità delle informazioni decisionali dell'IA giocano un ruolo chiave nel favorire l'interazione efficace con tali tecnologie.

ANALISI DI REGRESSIONE LINEARE TRA TRASPARENZA E PERCEZIONE DEL RISCHIO DI PERDITA DI LAVORO

In questa analisi di regressione è stato esaminato se una maggiore **trasparenza** nei processi decisionali dell'IA **riduce** la **paura di essere sostituiti** dal sistema. L'ipotesi del modello teorico di base era che la trasparenza avesse un **effetto rassicurante**. Di seguito si riportano i risultati ottenuti.

Var dipendente	Predictor	p-value	R ²	β
RISCHIO	TRASP	0.045	0.019	0.137



La figura mostra un diagramma di dispersione tra la trasparenza percepita dell'intelligenza artificiale e il rischio percepito dai lavoratori in merito alla possibilità di perdere il proprio impiego a causa dell'adozione di sistemi automatizzati. Come visibile, la variabile *rischio* presenta una distribuzione discreta, con risposte concentrate su

valori interi (da 1 a 7). Nonostante l'elevata dispersione dei dati, la linea di regressione suggerisce una tendenza leggermente positiva, coerente con i risultati dell'analisi statistica. In altre parole, all'aumentare della trasparenza percepita, tende ad aumentare lievemente anche la percezione di rischio.

Tuttavia, trattandosi di un effetto debole e statisticamente marginale, il dato va interpretato con cautela: potrebbe riflettere la presenza di profili più critici o consapevoli, per i quali la comprensione dei meccanismi decisionali rende più saliente il potenziale impatto dell'IA sulla propria stabilità lavorativa.

Interpretazione dei risultati:

Coefficiente di regressione standardizzato (Beta): 0.137: Questo valore indica che all'aumentare della trasparenza percepita dell'IA, aumenta anche la percezione del rischio occupazionale, sebbene l'effetto sia molto debole e in direzione opposta a quanto ipotizzato nel modello teorico (che prevedeva una relazione negativa).

Significatività (Sig.): 0.045: Il risultato è statisticamente significativo, poiché il valore è inferiore al livello convenzionale di $\alpha = .05$. Tuttavia, trattandosi di un valore al limite della soglia, si suggerisce cautela nell'interpretazione, soprattutto considerando la scarsa entità dell'effetto.

R² (R Square): 0.019: La trasparenza spiega solo l'1,9% della varianza nella percezione del rischio. Si tratta di un valore estremamente basso, che evidenzia una scarsa capacità predittiva di questa variabile sul rischio occupazionale.

Sebbene la relazione tra trasparenza percepita e percezione del rischio risulti statisticamente significativa, l'effetto è contrario alle attese teoriche ed estremamente debole sul piano pratico.

In modo controintuitivo, il coefficiente positivo suggerisce che una maggiore trasparenza potrebbe **aumentare leggermente** la percezione del rischio. Questo potrebbe essere dovuto al fatto che comprendere meglio le logiche decisionali dell'IA **rende più evidente il suo potenziale di sostituire l'intervento umano**, generando un senso di minaccia. Tuttavia, vista la bassa varianza spiegata, è probabile che la percezione del rischio sia influenzata da **molti altri fattori non considerati in questo modello**, come la sicurezza del ruolo, l'esperienza personale o il settore lavorativo.

FASE 3

Nella terza fase dell'analisi di regressione, si è verificato se ciascuna delle tre variabili psicologiche identificate come potenziali mediatrici – utilità percepita, facilità d'uso percepita e percezione del rischio occupazionale – influenzi significativamente la fiducia nei sistemi automatizzati di recruiting (FIDREC), anche quando nel modello è già inclusa la trasparenza percepita dell'IA (TRASP). Questa procedura è necessaria per soddisfare la terza condizione prevista dal modello di mediazione di Baron e Kenny (1986), secondo cui un mediatore è tale solo se mantiene un'associazione significativa con l'outcome (Y) anche in presenza della variabile indipendente (X).

ANALISI DI REGRESSIONE: EFFETTO DELL'UTILITÀ PERCEPITA SULLA FIDUCIA, CONTROLLANDO PER LA TRASPARENZA

Nel caso dell'**utilità percepita**, è stata condotta una regressione multipla in cui la **fiducia nei sistemi di IA (FIDREC)** è stata prevista sia dalla **trasparenza percepita (TRASP)** che dall'**utilità percepita (UTILITA)**. Questo passaggio serve a verificare se l'utilità ha un effetto significativo sulla fiducia anche quando si tiene conto contemporaneamente dell'influenza della trasparenza. Si tratta della **terza condizione** del modello di Baron e Kenny, fondamentale per confermare che l'utilità agisce come **mediatore**.

Var dipendente	Predictor	p-value	β
FIDUCIAREC	UTILITA	p < .001	0.408
	TRASP	p < .001	0.286

R² del modello	F	F-value
0.371	62.499	p < .001

I risultati mostrano che:

1. **R² del modello = 0.371**

Questo valore indica che il 37.1% della varianza nella fiducia nei sistemi automatizzati di recruiting

(FIDUCIAREC) è spiegata congiuntamente dalle due variabili predittive incluse nel modello: utilità percepita dell'IA (UTILITA) e trasparenza percepita (TRASP). È un valore piuttosto buono per studi di tipo psicologico-comportamentale, dove anche R^2 tra .20 e .40 sono comunemente considerati solidi.

2. **F = 62.499, p < .001**

Il test F valuta se il modello di regressione complessivamente spiega una quantità significativa di varianza nella variabile dipendente (FIDUCIAREC). Il valore elevato di F e il p-value inferiore a .001 indicano che il modello complessivamente è statisticamente significativo: almeno uno dei due predittori contribuisce in modo affidabile a spiegare la variabilità nella fiducia.

3. **Coefficiente standardizzato β per UTILITA = 0.408, p < .001**

Il coefficiente di regressione standardizzato indica la forza e la direzione dell'effetto di UTILITA su FIDUCIAREC, tenendo costante TRASP. Il valore positivo e altamente significativo ($p < .001$) suggerisce che maggiore è la percezione dell'utilità dell'IA, maggiore è la fiducia nei sistemi automatizzati, anche quando si considera l'effetto della trasparenza.

4. **Coefficiente standardizzato β per TRASP = 0.286, p < .001**

Anche il coefficiente per la trasparenza è positivo e significativo. Indica che, controllando per l'utilità, la trasparenza percepita continua ad avere un impatto rilevante sulla fiducia, anche se meno intenso rispetto a UTILITA. Questo suggerisce che entrambe le variabili contribuiscono indipendentemente alla predizione della fiducia.

Questa analisi fornisce una base importante per riflettere sul ruolo che la percezione dell'utilità dell'IA e della trasparenza possono giocare nel rafforzare la fiducia degli utenti nei sistemi automatizzati.

ANALISI DI REGRESSIONE: EFFETTO DELLA FACILITA' D'USO SULLA FIDUCIA, CONTROLLANDO PER LA TRASPARENZA

Nel presente step dell'analisi, è stata esaminata la relazione tra la facilità d'uso percepita (FACILITA) e la fiducia nei sistemi automatizzati di recruiting (FIDREC), controllando per la trasparenza percepita dell'IA (TRASP). Questo passaggio è fondamentale per testare la terza condizione del modello di mediazione di Baron e Kenny, ovvero verificare se la variabile mediatrice (in questo caso la facilità percepita) influenzi

significativamente la variabile dipendente (fiducia), anche tenendo costante l'effetto della variabile indipendente (trasparenza). Solo se questa condizione è soddisfatta si potrà considerare la facilità d'uso come un mediatore plausibile nella relazione tra trasparenza e fiducia.

Var dipendente	Predictor	p-value	β
FIDUCIAREC	FACILITA	p < .001	0.466
	TRASP	p < .001	0.191

R ² del modello	F	F-value
0.372	62.695	p < .001

L'intero modello di regressione che include sia la trasparenza che la facilità d'uso percepita come predittori della fiducia è **statisticamente significativo**: $F(2, 212) = 62.695, p < .001$. L'R² del modello è pari a .372, indicando che circa il 37.2% della varianza nella fiducia nei sistemi automatizzati di recruiting è spiegata congiuntamente da trasparenza e facilità d'uso.

Per quanto riguarda i singoli predittori:

- **Facilità d'uso percepita (FACILITA)** mostra un coefficiente β standardizzato pari a 0.466, altamente significativo ($p < .001$), evidenziando un'influenza forte e positiva sulla fiducia, anche controllando per la trasparenza.
- **Trasparenza (TRASP)** mantiene un effetto positivo, con $\beta = 0.191$ e $p = .009$, ma con una forza inferiore rispetto alla facilità, indicando una parziale riduzione dell'effetto diretto rispetto alla regressione semplice (come si vedrà nelle conclusioni).

I risultati soddisfano la terza condizione del modello di Baron e Kenny: la facilità d'uso è significativamente associata alla fiducia anche controllando per la trasparenza, e l'effetto diretto della trasparenza risulta attenuato rispetto a quando la fiducia era predetta solo da TRASP. Questo suggerisce che la facilità d'uso possa svolgere un ruolo mediatore nella relazione tra trasparenza e fiducia. Tale indicazione sarà confermata o meno solo a seguito dell'integrazione dei risultati delle altre due fasi (effetto di TRASP su FACILITA e di TRASP su FIDUCIA), e della comparazione complessiva dei β nelle conclusioni.

ANALISI DI REGRESSIONE: EFFETTO DELLA PERCEZIONE DEL RISCHIO DI PERDITA DI LAVORO SULLA FIDUCIA, CONTROLLANDO PER LA TRASPARENZA

La terza e ultima analisi di regressione multipla è stata condotta per verificare se la percezione del rischio occupazionale sia associata significativamente alla fiducia nei sistemi automatizzati di recruiting, anche controllando per la trasparenza percepita dell'IA. Questo passaggio è fondamentale per verificare la terza condizione del modello di mediazione di Baron e Kenny (1986), che richiede che il potenziale mediatore influenzi significativamente la variabile dipendente, a parità della variabile indipendente.

Var dipendente	Predictor	p-value	β
FIDUCIAREC	RISCHIO	p < .001	0.120
	TRASP	p < .001	0.484

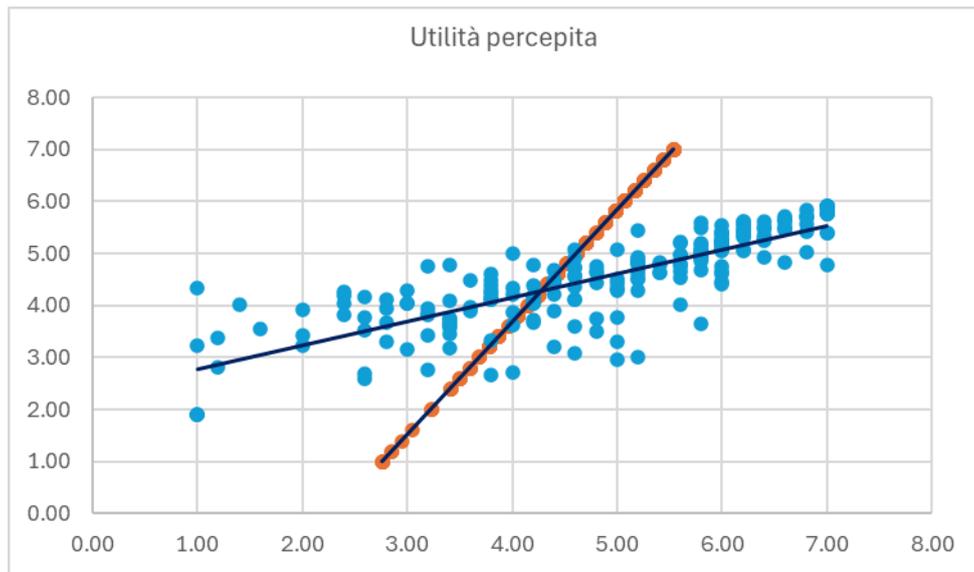
R² del modello	F	F-value
0.264	38.107	p < .001

Il modello di regressione risulta significativo nel suo complesso ($F(2, 212) = 38.107, p < .001$), con un $R^2 = .264$, indicando che trasparenza e rischio insieme spiegano circa il 26% della varianza della fiducia (FIDREC). Entrambe le variabili predittive risultano statisticamente significative: la trasparenza ($\beta = .484, p < .001$) e la percezione del rischio ($\beta = .120, p = .045$), con la trasparenza che mostra un peso predittivo molto più marcato. Da un punto di vista interpretativo, ciò significa che, anche tenendo costante il livello di trasparenza percepita, la percezione del rischio lavorativo è debolmente ma significativamente associata a una minore fiducia nei sistemi di IA. Questo risultato soddisfa la terza condizione richiesta per considerare la percezione del rischio come mediatore nel modello teorico.

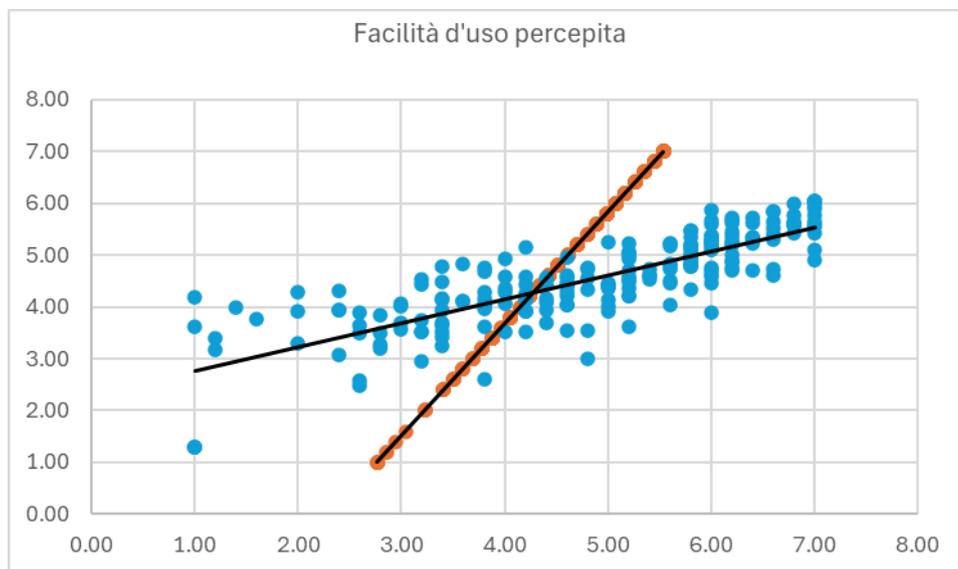
FASE 4

L'ultimo step dell'analisi, previsto dal modello di Baron e Kenny, ha avuto l'obiettivo di verificare la **quarta condizione** della mediazione, ovvero la **riduzione dell'effetto diretto** della trasparenza sulla fiducia a seguito dell'inclusione del mediatore nel modello. In altre parole, si è osservato se l'associazione tra trasparenza e fiducia diminuisce in termini di intensità (coefficiente β) quando viene introdotta nel modello una delle tre variabili mediatrici.

I tre grafici seguenti mostrano il confronto tra la **regressione semplice** (linea arancione), in cui la fiducia viene predetta unicamente dalla trasparenza, e la **regressione mediata** (linea blu), in cui viene aggiunto ciascun mediatore (utilità, facilità d'uso o rischio). Per rappresentare graficamente l'effetto della trasparenza percepita dell'IA sulla fiducia nei sistemi automatizzati di recruiting, sia in assenza che in presenza del mediatore, sono stati realizzati dei grafici di dispersione con linea di regressione (scatter plot con trendline) utilizzando *Microsoft Excel*. La linea arancione rappresenta il modello di regressione **semplice** (effetto diretto della trasparenza sulla fiducia), costruito associando sull'asse X i punteggi della variabile indipendente (TRASP) e sull'asse Y i valori originali della variabile dipendente (FIDREC). La linea blu, invece, rappresenta il **modello mediato**, ovvero il modello in cui l'effetto della trasparenza sulla fiducia è stato calcolato controllando per la presenza del mediatore (utilità, facilità o rischio). In questo secondo caso, sull'asse Y sono stati inseriti i valori **predetti** generati dal modello di regressione multipla che includeva sia la trasparenza che il mediatore (es. PRE2), mentre sull'asse X è stata mantenuta la trasparenza. Questo confronto consente di osservare visivamente la **riduzione dell'effetto diretto** una volta introdotto il mediatore, in linea con la logica della mediazione. L'utilizzo dei valori predetti permette di tracciare una linea di regressione che riflette l'andamento stimato del modello completo, offrendo una rappresentazione visiva coerente con i risultati analitici delle regressioni condotte.

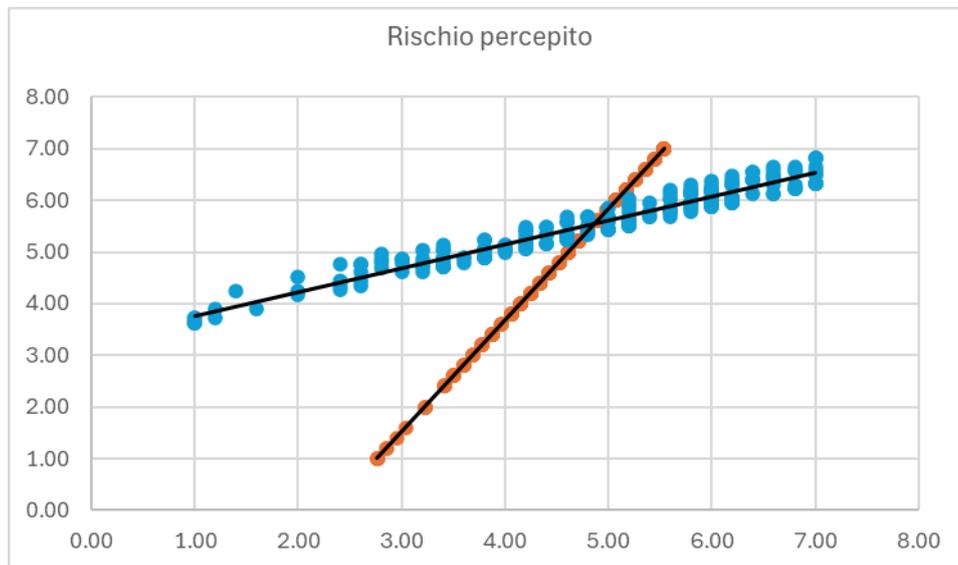


Nel caso dell'**utilità percepita**, si osserva chiaramente una riduzione della pendenza della retta blu rispetto alla linea arancione, suggerendo che parte dell'effetto della trasparenza sulla fiducia viene assorbito dal mediatore. Questo risultato è coerente con quanto emerso nelle analisi di regressione (β trasparenza da .527 a .286), e rafforza l'evidenza di un processo di mediazione.



Anche per la **facilità d'uso percepita** si rileva una chiara attenuazione dell'effetto diretto: la pendenza della retta blu risulta visibilmente meno ripida rispetto a quella del modello semplice. I dati numerici confermano una

sostanziale riduzione del coefficiente di regressione della trasparenza (β da .527 a .191), compatibile con una mediazione significativa.



Diverso è il caso del **rischio di perdita del lavoro**. Sebbene il modello con mediatore mostri un coefficiente della trasparenza leggermente più basso (β da .527 a .484), il cambiamento nella pendenza è minimo, come evidenziato anche dal grafico. La linea blu appare quasi sovrapposta a quella arancione, a indicare che l'aggiunta del rischio non modifica sostanzialmente l'effetto diretto della trasparenza sulla fiducia. Questo suggerisce che il rischio, pur avendo una relazione significativa con la fiducia, non soddisfa la condizione di riduzione dell'effetto diretto ed è quindi meno plausibile come mediatore.

Complessivamente, la **riduzione dell'effetto diretto osservata visivamente e quantitativamente nei casi di utilità e facilità** fornisce evidenza a sostegno del ruolo di queste due variabili come mediatori. Al contrario, la **scarsa attenuazione** osservata per il rischio ne suggerisce un ruolo diverso, più compatibile con un possibile effetto di **moderazione**, da approfondire in analisi successive.

Conclusioni

Sulla base dei risultati ottenuti nelle tre fasi di regressione e alla luce del quarto passaggio previsto dal modello di mediazione di Baron e Kenny (1986), è possibile valutare se le tre variabili psicologiche analizzate – **utilità percepita, facilità d'uso percepita e percezione del rischio occupazionale** – soddisfino **tutte le condizioni**

necessarie per essere considerate mediatori dell'effetto della trasparenza sulla fiducia nei sistemi automatizzati di recruiting.

Utilità percepita: tutte le condizioni soddisfatte → mediazione significativa

- **Fase 1 (X → M):** la trasparenza predice significativamente l'utilità percepita ($\beta = .527$; $p < .001$; $R^2 = .277$).
- **Fase 2 (X → Y):** la trasparenza predice significativamente la fiducia ($\beta = .527$; $p < .001$).
- **Fase 3 (M → Y | X):** l'utilità ha un effetto significativo sulla fiducia anche controllando per la trasparenza ($\beta = .408$; $p < .001$).
- **Fase 4 (riduzione X → Y):** il coefficiente di trasparenza si riduce da .527 a .286 pur restando significativo.

Tutte e quattro le condizioni sono soddisfatte. La riduzione dell'effetto diretto è sostanziale sia nei coefficienti sia nel confronto visivo tra i modelli (regressione semplice vs mediata). **L'utilità percepita può quindi essere considerata un mediatore robusto e significativo.**

Facilità d'uso percepita: tutte le condizioni soddisfatte → mediazione plausibile

- **Fase 1 (X → M):** la trasparenza ha un impatto forte e significativo sulla facilità percepita ($\beta = .664$; $p < .001$; $R^2 = .441$).
- **Fase 2 (X → Y):** la trasparenza predice significativamente la fiducia ($\beta = .527$; $p < .001$).
- **Fase 3 (M → Y | X):** la facilità predice la fiducia anche controllando per la trasparenza ($\beta = .466$; $p < .001$).
- **Fase 4 (riduzione X → Y):** il coefficiente della trasparenza si abbassa da .527 a .191.

Anche in questo caso, **tutte le condizioni del modello di Kenny risultano soddisfatte**, e l'effetto diretto si riduce visibilmente sia in termini numerici che grafici. **La facilità d'uso percepita può quindi essere considerata un mediatore valido.**

Percezione del rischio occupazionale: condizioni non pienamente soddisfatte → mediazione debole

- **Fase 1 (X → M):** la trasparenza è debolmente associata al rischio percepito ($\beta = .137$; $p = .045$; $R^2 = .019$).

- **Fase 2 (X → Y):** la trasparenza predice significativamente la fiducia ($\beta = .527$; $p < .001$).
- **Fase 3 (M → Y | X):** il rischio ha un effetto marginale sulla fiducia ($\beta = .120$; $p = .045$), controllando per la trasparenza.
- **Fase 4 (riduzione X → Y):** il coefficiente della trasparenza si riduce in modo trascurabile (da .527 a .484).

Sebbene il rischio soddisfi formalmente le prime tre condizioni, **l'effetto diretto della trasparenza non subisce una riduzione apprezzabile**, né nei dati né nei grafici. Questo suggerisce che **il rischio occupazionale non svolge un ruolo mediatore significativo**, ma potrebbe agire come **moderatore**, come verificato nella fase successiva dell'analisi.

4.4 Analisi di mediazione

Dopo la verifica delle condizioni di mediazione secondo il modello classico di Baron e Kenny (1986), l'analisi è stata approfondita ricorrendo al metodo proposto da *Andrew F. Hayes*, oggi ritenuto uno degli approcci più **robusti** per la valutazione dei meccanismi di mediazione. A differenza del modello tradizionale, che si basa sulla verifica sequenziale di tre (o quattro) condizioni, il metodo di Hayes si fonda sulla stima diretta dell'**effetto indiretto** e sul test della sua **significatività statistica tramite bootstrapping**.

Criteria per l'interpretazione della mediazione (Hayes, 2013):

Nel presente studio, per valutare se le variabili considerate fungano da mediatori dell'effetto della trasparenza sulla fiducia nei sistemi automatizzati di recruiting, è stata posta particolare attenzione a due indicatori chiave:

1. **Significatività dell'effetto indiretto:** valutata tramite l'intervallo di confidenza bootstrap al 95%. Se l'intervallo non include lo zero, si considera presente un effetto di mediazione.
 2. **Significatività dell'effetto diretto (p-value):** una volta incluso il mediatore nel modello, si osserva se il coefficiente diretto tra trasparenza e fiducia rimane significativo.
- Se il p-value dell'effetto diretto rimane significativo ($p < .05$), si parla di mediazione parziale: l'effetto della trasparenza sulla fiducia si trasmette sia direttamente sia indirettamente, attraverso il mediatore.

- Se il p-value dell'effetto diretto diventa non significativo ($p \geq .05$), si parla di mediazione totale: l'effetto della trasparenza sulla fiducia si trasmette esclusivamente attraverso il mediatore.

Analisi 1: Mediazione tramite utilità percepita

Obiettivo: Esaminare se la percezione dell'utilità dell'IA media la relazione tra la trasparenza percepita e la fiducia nei sistemi di IA.

EFFETTO INDIRETTO

Coefficiente B	IC al 95%
0.20	[0.1576, 0.4174]

EFFETTO DIRETTO

Coefficiente B	P-value
0.2647	$p < .001$

L'analisi di mediazione condotta con il modello PROCESS di Hayes (Model 4) ha confermato che l'utilità percepita dell'IA svolge un ruolo significativo nella trasmissione dell'effetto della trasparenza sulla fiducia nei sistemi automatizzati di recruiting. In particolare, **l'effetto indiretto** della trasparenza sulla fiducia, veicolato attraverso l'utilità percepita, risulta **significativo**: $B = 0.1977^{61}$, con intervallo di confidenza bootstrap al 95% compreso tra 0.1150 e 0.2986. Poiché l'intervallo di confidenza non include lo zero, si può concludere che l'effetto mediato è statisticamente significativo, a supporto della presenza di una mediazione.

L'effetto diretto della trasparenza sulla fiducia, controllando per il mediatore (utilità percepita), rimane anch'esso significativo ($B = 0.2647$, $p < .001$). La persistenza della significatività dell'effetto diretto, suggerisce

⁶¹ Il coefficiente dell'effetto indiretto rappresenta l'ammontare dell'influenza che la variabile indipendente (X) (es. *trasparenza*) esercita sulla variabile dipendente (Y) (es. *fiducia*) attraverso il mediatore (M) (es. *utilità percepita*). Nel dettaglio, significa che per ogni aumento di 1 punto nella trasparenza, la fiducia aumenta di 0.20 punti *in media*, attraverso l'aumento dell'utilità percepita.

che l'utilità percepita esercita una **mediazione parziale**: *la trasparenza continua ad avere un impatto diretto sulla fiducia, ma parte del suo effetto viene trasmesso attraverso la percezione dell'utilità dell'IA.*

In sintesi, sia l'effetto indiretto che quello diretto sono significativi, e ciò consente di affermare che l'utilità percepita è un mediatore parziale dell'effetto della trasparenza sulla fiducia nei sistemi di IA. Questo suggerisce che quando i dipendenti percepiscono l'IA come trasparente, tendono anche a considerarla più utile, il che a sua volta rafforza la loro fiducia nei sistemi di IA.

Analisi 2: Mediazione tramite facilità d'uso percepita

Obiettivo: Determinare se la facilità d'uso percepita dell'IA media la relazione tra la trasparenza percepita e la fiducia nei sistemi di IA.

EFFETTO INDIRETTO

Coefficiente B	IC al 95%
0.286	[0.1576, 0.4174]

EFFETTO DIRETTO

Coefficiente B	P-value
0.1764	$p = .0094$

L'analisi condotta con il modello 4 di PROCESS (Hayes, 2013) ha permesso di verificare la presenza di un effetto mediato della trasparenza percepita dell'IA sulla fiducia nei sistemi automatizzati di recruiting, tramite la facilità d'uso percepita. I risultati mostrano un **effetto indiretto significativo** ($B = 0.2860$), supportato da un intervallo di confidenza bootstrap al 95% che non include lo zero [0.1576, 0.4174]. Questo indica che una parte consistente dell'effetto della trasparenza sulla fiducia si trasmette attraverso la percezione che i sistemi siano facili da comprendere e utilizzare.

L'effetto diretto della trasparenza sulla fiducia, controllando per la facilità d'uso, risulta ancora **significativo** (B = 0.1764; p = .0094).

Pertanto, sulla base della significatività dell'effetto indiretto e della persistenza dell'effetto diretto, è possibile concludere che si tratta di una **mediazione parziale**: la trasparenza influenza la fiducia sia direttamente sia indirettamente, attraverso la facilità d'uso percepita.

Analisi 3: Mediazione tramite percezione del rischio

Obiettivo:

Questa analisi intende verificare se la **percezione del rischio di perdere il posto di lavoro** media la relazione tra la **trasparenza percepita** dei sistemi di intelligenza artificiale (IA) e la **fiducia** nei sistemi automatizzati di selezione del personale. In altre parole, si esplora se una maggiore trasparenza può rafforzare la fiducia nei sistemi di IA **perché** riduce la sensazione di minaccia per il proprio ruolo lavorativo.

EFFETTO INDIRETTO

Coefficiente B	IC al 95%
0.286	[-0.0012, 0.0455]

EFFETTO DIRETTO

Coefficiente B	P-value
0.0152	$p < .001$

L'analisi di mediazione condotta tramite il modello 4 di PROCESS (Hayes, 2013) ha esplorato se la percezione del rischio occupazionale medii l'effetto della trasparenza percepita dell'IA sulla fiducia nei sistemi automatizzati di recruiting. L'effetto indiretto risulta **debole e non significativo** (B = 0.0152), con un intervallo di confidenza bootstrap al 95% che include lo zero [-0.0012, 0.0455]. Questo suggerisce che il rischio non funge da canale attraverso cui la trasparenza influenza la fiducia.

L'effetto diretto della trasparenza sulla fiducia rimane invece **elevato e altamente significativo** ($B = 0.4472$; $p < .001$; IC 95% [0.3389, 0.5555]), indicando che la trasparenza esercita un'influenza diretta sulla fiducia, indipendentemente dal ruolo del rischio percepito.

Alla luce di questi risultati, **non si configura una mediazione**, in quanto l'effetto indiretto non è statisticamente significativo. La percezione del rischio occupazionale non può essere considerata un mediatore dell'effetto della trasparenza sulla fiducia nei sistemi IA.

Conclusioni sulle analisi di mediazione

I risultati complessivi confermano la validità teorica del modello ipotizzato, secondo cui l'effetto della trasparenza percepita dell'IA sulla fiducia nei sistemi automatizzati di recruiting può essere spiegato in parte dalla presenza di meccanismi psicologici intermedi. In particolare, l'utilità percepita e la facilità d'uso percepita hanno mostrato di soddisfare pienamente i criteri di mediazione previsti dal modello di Baron e Kenny (1986) e, al tempo stesso, di generare un effetto indiretto significativo secondo l'approccio più recente di Hayes (2013), basato sulla procedura bootstrap. Questo doppio riscontro metodologico rafforza l'evidenza empirica a favore del ruolo di questi due costrutti come **mediatori parziali** del legame tra trasparenza e fiducia.

Al contrario, la percezione del rischio occupazionale ha mostrato un effetto indiretto non significativo e una debole relazione con la trasparenza, suggerendo che **non svolge un ruolo mediatore** nel processo di formazione della fiducia. L'inclusione dell'effetto indiretto secondo il metodo di Hayes ha offerto una controprova particolarmente utile per superare le limitazioni strutturali del modello tradizionale, permettendo di convalidare le mediazioni anche in presenza di effetti diretti residui significativi. In definitiva, l'impiego congiunto dei due approcci metodologici ha garantito una verifica robusta del modello teorico, confermando che la trasparenza genera fiducia soprattutto quando è percepita come **utile e facile da comprendere**.

Sebbene la *percezione del rischio occupazionale* non si sia dimostrata un mediatore efficace nella relazione tra trasparenza e fiducia – data la debole associazione con la trasparenza e la non significatività dell'effetto indiretto – l'analisi ha comunque evidenziato un legame diretto, seppur marginale, tra rischio e fiducia ($\beta = .120$, $p = .045$). Questo risultato ha suggerito una possibile diversa funzione della variabile nel modello teorico. In particolare, è stato ipotizzato che il rischio potesse *non trasmettere* l'effetto della trasparenza, ma *modificarne l'intensità*. In altri termini, la trasparenza potrebbe avere un impatto maggiore sulla fiducia nei sistemi automatizzati tra quei dipendenti che percepiscono un alto rischio occupazionale – proprio perché, in tali casi,

una maggiore chiarezza e spiegabilità del funzionamento dell'IA risulterebbero particolarmente rassicuranti. Al contrario, in assenza di timori occupazionali, lo stesso effetto della trasparenza potrebbe attenuarsi. Alla luce di questa ipotesi, è stata pertanto condotta un'analisi di moderazione, al fine di verificare se la percezione del rischio agisca come *fattore di amplificazione o attenuazione* dell'effetto della trasparenza sulla fiducia. Per questo motivo, si propone un'ulteriore analisi di moderazione, che verrà descritta nel paragrafo seguente.

4.5 Analisi di moderazione – Percezione del rischio di perdita di lavoro

Dopo aver escluso la *percezione del rischio occupazionale* come mediatore del legame tra trasparenza e fiducia (data la debolezza dell'effetto indiretto), si è ipotizzato che tale variabile potesse agire come *moderatore*, ovvero come fattore capace di modificare l'intensità della relazione tra trasparenza e fiducia nei sistemi automatizzati di recruiting. L'analisi di moderazione è stata condotta utilizzando il modello 1 del macro PROCESS di Hayes.

L'obiettivo dell'analisi è verificare se l'effetto della trasparenza percepita dell'IA sulla fiducia cambi a seconda del livello di percezione del rischio occupazionale. In termini statistici, si intende testare l'interazione tra trasparenza e rischio nella predizione della fiducia.

Interpretazione dei Risultati

1. Coefficiente di interazione

B = 0.0956, $p = .0003$, CI [0.0443, 0.1469]: il coefficiente di interazione (B) ottenuto dall'analisi di moderazione rappresenta il cuore dell'effetto moderatore, in quanto consente di determinare l'esistenza di un effetto di moderazione. In termini statistici, esso misura se e quanto l'effetto della trasparenza sulla fiducia nei sistemi automatizzati di recruiting cambia al variare del livello di rischio percepito. Un coefficiente di interazione positivo e significativo indica che, all'aumentare della percezione del rischio occupazionale, la relazione tra trasparenza e fiducia diventa più forte. Nel presente studio, il coefficiente di interazione calcolato è pari a **B = 0.0956**, un valore **positivo** che indica che l'effetto della trasparenza sulla fiducia **aumenta all'aumentare della percezione del rischio**. In altre parole, quanto più i dipendenti percepiscono l'IA come una minaccia per il proprio lavoro, tanto più la trasparenza nel funzionamento dei sistemi automatizzati contribuisce a generare fiducia nei confronti della tecnologia. Questo suggerisce che, in contesti percepiti come potenzialmente instabili o ansiogeni, la trasparenza può assumere un ruolo compensativo e rassicurante, rafforzando la legittimazione e l'accettazione dell'IA nelle pratiche organizzative.

Il valore è **significativo**, perché:

- Il **p-value** è **< .001** (quindi l'interazione è statisticamente significativa)
- L'**intervallo di confidenza (CI)** **non include lo zero**, confermando la significatività.

2. Effetti condizionati della trasparenza ai diversi livelli di rischio

Livello di rischio	Effetto della trasparenza sulla fiducia	Significatività
Basso (1.00)	$B = 0.2479, p = .0013$	Significativo
Medio (3.00)	$B = 0.4391, p < .001$	Significativo
Alto (5.00)	$B = 0.6302, p < .001$	Significativo

A supporto del risultato ottenuto dal coefficiente di interazione, l'analisi riporta anche gli effetti condizionati della trasparenza sulla fiducia calcolati a specifici livelli del rischio (basso, medio, alto). Tali effetti mostrano l'ampiezza dell'influenza della trasparenza sulla fiducia in ciascuna delle condizioni considerate, confermando empiricamente che l'intensità dell'effetto varia in funzione del rischio. Nel caso specifico, l'effetto positivo della trasparenza sulla fiducia risulta più debole quando il rischio percepito è basso ($B = 0.2479$), e progressivamente più forte ai livelli medio ($B = 0.4391$) e alto ($B = 0.6302$). Questi risultati rafforzano la conclusione che la trasparenza agisce come fattore particolarmente rilevante nei contesti caratterizzati da maggiore incertezza o timore per la stabilità lavorativa.

3. Effetti principali di trasparenza e rischio

- **TRASP** (effetto principale): $B = .1524, p = .1156$
- **RISCHIO** (effetto principale): $B = -0.3958, p = .0043$

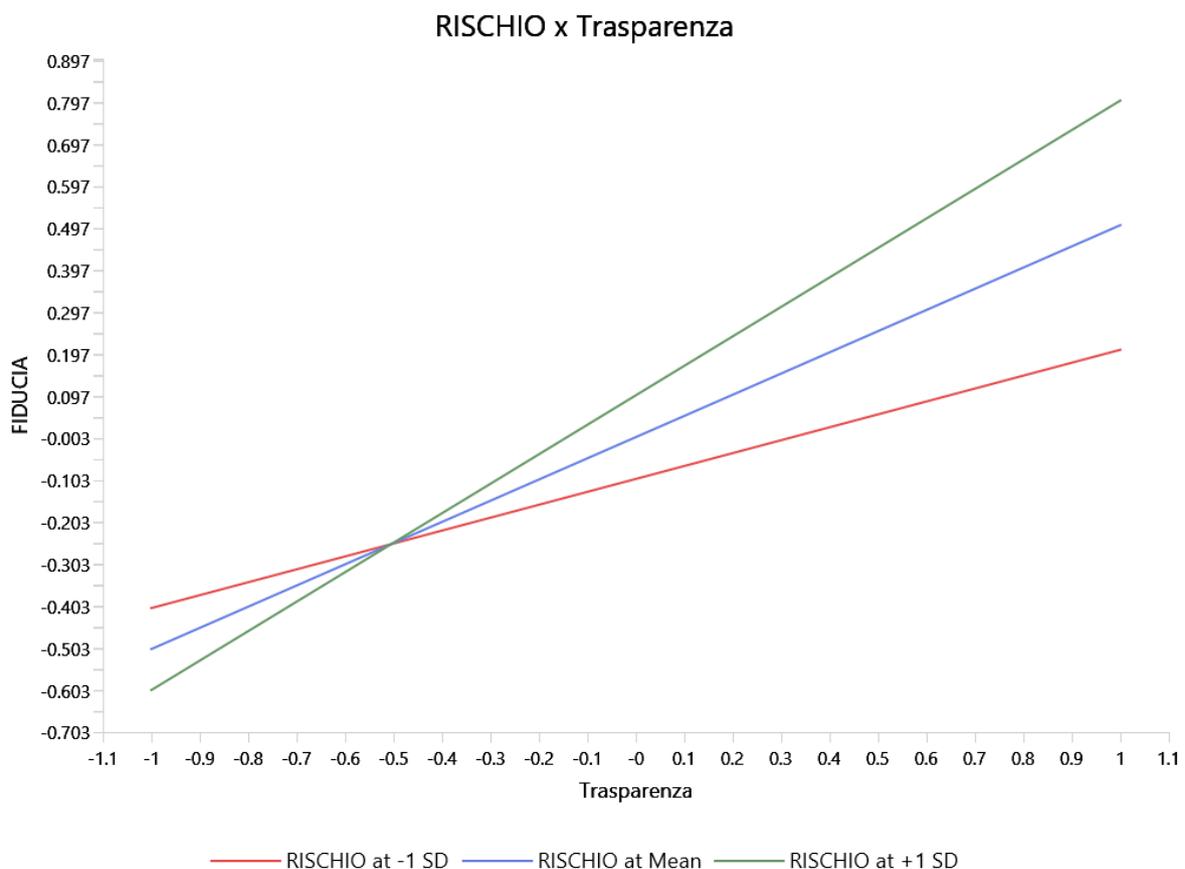
L'effetto principale della trasparenza, isolato dal termine di interazione, risulta non significativo ($B = 0.1524, p = .1156$), a conferma del fatto che la sua efficacia nel generare fiducia dipende dal livello di rischio percepito. Al contrario, la percezione del rischio mostra un effetto principale significativo e negativo ($B = -0.3958, p = .0043$), indicando che, a parità di trasparenza, un incremento della percezione del rischio occupazionale è associato a un calo della fiducia nei sistemi IA. Questo risultato conferma il ruolo critico delle preoccupazioni occupazionali nei processi psicologici di accettazione dell'automazione.

Rappresentazione grafica

A completamento dell'analisi di moderazione, è stato generato un *simple slope plot* tramite SmartPLS, che consente di visualizzare l'interazione tra la variabile indipendente (Trasparenza percepita dell'IA) e la variabile moderatrice (Rischio percepito di perdita del lavoro) nel predire la variabile dipendente (Fiducia nei sistemi automatizzati di recruiting). Il grafico mostra tre linee di regressione: una per un valore basso del rischio (-1 deviazione standard dalla media), una per il valore medio del rischio, e una per un valore alto del rischio (+1 deviazione standard).

La lettura del grafico indica che all'aumentare della trasparenza, la fiducia tende ad aumentare in tutti e tre i casi; tuttavia, **l'inclinazione della linea è maggiore nei soggetti con un elevato rischio percepito**. Questo suggerisce che il rischio **amplifica** l'effetto positivo della trasparenza sulla fiducia: in altre parole, quando i lavoratori percepiscono un rischio occupazionale più elevato, diventano più sensibili alla trasparenza dell'IA e reagiscono con un aumento maggiore della fiducia rispetto a chi percepisce meno rischio.

Il coefficiente dell'interazione ($\beta = 0.197$) conferma statisticamente questa moderazione positiva. Di conseguenza, si può concludere che il rischio percepito agisce come una **variabile condizionante**, rafforzando la relazione tra trasparenza e fiducia nei sistemi automatizzati.



Conclusioni dell'analisi di moderazione

I risultati emersi dall'analisi confermano in modo robusto che la *percezione del rischio occupazionale* non agisce come mediatore, ma come *moderatore* dell'effetto della trasparenza sulla fiducia nei sistemi automatizzati di recruiting. In primo luogo, il coefficiente di interazione tra trasparenza e rischio è positivo e altamente significativo ($B = 0.0956$, $p = .0003$), indicando che l'efficacia della trasparenza nel generare fiducia aumenta al crescere della percezione del rischio. In altri termini, nei contesti in cui i dipendenti temono maggiormente la perdita del lavoro, la trasparenza svolge una funzione compensativa più marcata, rafforzando la legittimazione e l'accettazione dei sistemi automatizzati.

Questo effetto è ulteriormente supportato dagli *effetti condizionati*, che mostrano come l'influenza positiva della trasparenza sulla fiducia sia crescente lungo i livelli di rischio: da $B = 0.2479$ (basso rischio) fino a $B = 0.6302$ (alto rischio). Ciò suggerisce che la trasparenza non produce un effetto uniforme, ma assume un valore psicologicamente strategico soprattutto nei contesti percepiti come incerti o minacciosi.

Infine, gli effetti principali evidenziano che, una volta considerata l'interazione, la trasparenza da sola non spiega più significativamente la fiducia ($p = .1156$), mentre il rischio mantiene un impatto diretto e negativo ($B = -0.3958$, $p = .0043$). Questo rafforza l'interpretazione secondo cui la trasparenza non agisce in modo isolato, ma il suo potenziale fiduciante si esprime appieno proprio quando si attivano vissuti di vulnerabilità legati all'automazione. Di conseguenza, la percezione del rischio occupazionale si configura come una *variabile di contesto cruciale*, capace di modulare l'efficacia delle strategie di trasparenza nella gestione dell'adozione dell'IA in ambito organizzativo.

Capitolo 5

Implicazioni pratiche e impatto sulle Performance aziendali

5.1 Implicazioni pratiche

I risultati emersi dalle analisi statistiche confermano il ruolo centrale della trasparenza percepita dell'intelligenza artificiale nel determinare la fiducia degli utenti organizzativi nei sistemi automatizzati di recruiting. In particolare, l'effetto diretto della trasparenza sulla fiducia, nonché i suoi effetti indiretti mediati da variabili quali l'utilità e la facilità d'uso percepita, suggeriscono che l'accettazione dell'IA all'interno dei contesti aziendali non dipende unicamente dalla qualità tecnica delle soluzioni adottate, ma anche – e soprattutto – da come tali soluzioni vengono **percepite, comprese e vissute dai lavoratori**.

Alla luce di queste evidenze, le seguenti implicazioni pratiche propongono una serie di strategie operative che le organizzazioni possono adottare per tradurre in pratica i risultati emersi, con l'obiettivo di promuovere un'adozione dell'IA fondata sulla trasparenza, la partecipazione e la fiducia. Le proposte sono articolate a partire dalle evidenze empiriche raccolte e sono formulate in chiave applicativa, tenendo conto delle esigenze di implementazione sostenibile e delle buone pratiche emergenti nel panorama aziendale contemporaneo. I risultati della ricerca evidenziano come la trasparenza percepita incida significativamente sulla fiducia dei dipendenti nei sistemi di IA utilizzati nei processi di selezione del personale. In un contesto in cui la fiducia rappresenta una condizione necessaria per l'accettazione dell'innovazione, questo dato suggerisce l'opportunità per le aziende di non limitarsi a garantire la trasparenza come adempimento tecnico, ma di includerla in modo sistematico nelle pratiche organizzative.

Progettare trasparenza che sia percepita come facile e utile: una leva strategica per la fiducia nei sistemi di recruiting

La trasparenza percepita dell'IA ha un effetto diretto e significativo sulla fiducia nei sistemi automatizzati di recruiting ($\beta = 0.500, p < .001$).

La facilità d'uso percepita dell'IA è un predittore forte e significativo della fiducia nei sistemi automatizzati ($\beta = 0.593, p < .001$).

L'utilità percepita dell'IA è un predittore significativo della fiducia nei sistemi IA ($\beta = 0.558, p < .001$).

I risultati della ricerca evidenziano come la trasparenza percepita incida significativamente sulla fiducia dei dipendenti nei sistemi di IA utilizzati nei processi di selezione del personale. In un contesto in cui la fiducia rappresenta una condizione necessaria per l'accettazione dell'innovazione, questo dato suggerisce l'opportunità per le aziende di non limitarsi a garantire la trasparenza come adempimento tecnico, ma di includerla in modo sistematico nelle pratiche organizzative.

Tuttavia, il modello teorico evidenzia che questo effetto si attiva pienamente **quando la trasparenza viene percepita come comprensibile (facile) e vantaggiosa (utile)**. In altre parole, **la trasparenza è efficace solo se tradotta in un'esperienza cognitiva accessibile e rilevante per chi la riceve**. È su questo punto che l'organizzazione ha un ruolo chiave: non è sufficiente "essere trasparenti", bisogna **progettare come rendere trasparente l'IA in modo che i dipendenti possano realmente comprenderla e trarne beneficio**.

Una possibile applicazione consiste nell'introdurre dei **“moduli trasparenti”** all'interno dei processi HR digitali. Tali moduli, accessibili tramite portali interni aziendali (ad esempio nella sezione “HR” o “Carriere”), rappresentano vere e proprie interfacce consultabili in grado di fornire ai dipendenti una spiegazione sintetica e chiara dei criteri usati dagli algoritmi nei sistemi di screening, shortlisting o ranking dei candidati. Non si tratta di rendere pubbliche le formule matematiche, ma di offrire una narrazione comprensibile delle logiche di decisione dell'algoritmo (es. “L'IA analizza i CV considerando: esperienze coerenti, competenze chiave, titoli di studio, pattern lessicali...”), degli input considerati, e delle fasi del processo in cui interviene l'essere umano, distinguendo dunque ciò che è automatizzato da ciò che è valutato soggettivamente.

Ma affinché questi moduli generino **fiducia**, devono essere progettati secondo due principi:

1. Devono essere percepiti come *facili da comprendere*

In primo luogo, questi strumenti informativi devono essere progettati per essere *facili da comprendere da parte del personale aziendale, anche in assenza di competenze tecniche specifiche*. In questo contesto, la facilità d'uso non va intesa solo in senso ergonomico o digitale, ma come *accessibilità cognitiva*: il modo in cui l'informazione viene organizzata, presentata e resa interpretabile.

Per risultare effettivamente fruibili, i contenuti devono essere redatti con un linguaggio chiaro, evitando tecnicismi inutili o riferimenti statistici che rischierebbero di escludere i destinatari meno esperti. È altrettanto importante che le informazioni siano visivamente ben strutturate: l'uso di schemi, icone, colori e brevi esempi concreti consente di trasmettere anche concetti complessi in modo immediato. Infine, l'interazione con il modulo deve essere intuitiva: funzioni come paragrafi espandibili, tooltip contestuali o link "clicca per saperne di più" possono guidare l'utente nell'approfondimento, lasciandogli la possibilità di scegliere il proprio livello di coinvolgimento.

In questo modo si riduce il carico cognitivo richiesto per comprendere il funzionamento del sistema e si favorisce l'inclusione informativa. Il dipendente non si sente escluso da un processo tecnologico percepito come lontano o opaco, ma al contrario viene messo nelle condizioni di comprendere *con semplicità e autonomia* le logiche che regolano i meccanismi di selezione automatizzata.

2. Devono essere percepiti come *utili per sé*

In secondo luogo, per generare fiducia, i contenuti trasparenti devono essere percepiti come *effettivamente utili* dal punto di vista del dipendente. La trasparenza non può limitarsi alla mera esposizione di informazioni: deve rispondere a un bisogno concreto, ossia offrire un *valore soggettivo* in termini di orientamento, comprensione e controllo percepito.

Un modulo informativo è percepito come utile quando consente al lavoratore di comprendere, ad esempio, le ragioni per cui la propria candidatura è stata accolta o esclusa da una determinata posizione. Questa comprensione ha una ricaduta concreta: aiuta l'individuo a riflettere sul proprio percorso, ad acquisire

consapevolezza sui criteri utilizzati dall'organizzazione e, potenzialmente, a migliorare in vista di future opportunità.

Inoltre, rendere esplicito che il sistema non opera in modo opaco o casuale, ma secondo logiche coerenti, ponderate e verificabili, contribuisce a consolidare l'immagine dell'organizzazione come ambiente equo e meritocratico. In tal senso, l'informazione trasparente diventa uno *strumento di autovalutazione e di empowerment*: il dipendente percepisce di avere accesso a chiavi di lettura che gli permettono di collocarsi attivamente nel processo, anziché subirlo passivamente.

È proprio questo legame tra informazione ricevuta e significato personale attribuito che rende la trasparenza *utile per sé* e, di conseguenza, capace di generare fiducia reale nel sistema.

Questo approccio favorisce un clima di fiducia e riduce la percezione di arbitrarietà, soprattutto in contesti ad alta sensibilità come la selezione interna per avanzamenti di carriera. Inoltre, la trasparenza operativa nei processi HR può rafforzare la reputazione interna dell'organizzazione, contribuendo al cosiddetto *internal employer branding*.

Per garantire l'efficacia di questi strumenti, si consiglia un **monitoraggio periodico** del livello di comprensione e accettazione da parte dei lavoratori, attraverso brevi survey o focus group. In questo modo, la trasparenza non rimane un valore dichiarato, ma si traduce in un comportamento organizzativo misurabile e migliorabile.

Diverse aziende stanno adottando soluzioni pratiche per integrare la trasparenza nei loro sistemi di intelligenza artificiale. Ad esempio, **IBM** ha sviluppato **Watsonx.governance**⁶², una piattaforma progettata per automatizzare e accelerare i flussi di lavoro dell'IA responsabile. Questa soluzione consente di gestire proattivamente i rischi, semplificare la conformità normativa e creare flussi di lavoro dell'IA spiegabili, migliorando la fiducia nei sistemi automatizzati.

In Italia, la startup **Intellico**⁶³ si concentra sull'Explainable AI, sviluppando strumenti che aiutano le organizzazioni a comprendere e spiegare il funzionamento interno degli algoritmi. Le loro soluzioni mirano a migliorare la fiducia e la trasparenza nelle decisioni aziendali, riducendo i rischi associati a sistemi opachi.

⁶² IBM. (n.d.). Use watsonx.governance to accelerate responsible, transparent, and explainable AI. IBM Documentation.

⁶³ Intellico. (n.d.). Explainable AI: come migliorare la fiducia e la trasparenza nelle decisioni aziendali.

Inoltre, strumenti di *explainability* come **SHAP** (SHapley Additive exPlanations) e **LIME** (Local Interpretable Model-agnostic Explanations)⁶⁴ rappresentano due approcci consolidati per migliorare l'interpretabilità dei modelli di intelligenza artificiale. SHAP, basato sulla teoria dei giochi cooperativi, assegna a ciascuna variabile un contributo preciso alla predizione, rendendolo particolarmente efficace sia per l'analisi globale sia per la spiegazione puntuale delle decisioni. LIME, invece, si concentra su spiegazioni locali, generando modelli semplificati intorno a singole predizioni attraverso la perturbazione dei dati di input. Sebbene più intuitivo, LIME può risultare meno stabile in certi casi rispetto a SHAP. Entrambi, tuttavia, sono strumenti chiave per costruire sistemi di IA più trasparenti e affidabili, in particolare in ambiti sensibili come HR, sanità e finanza, dove la comprensione del processo decisionale è cruciale (DeepCore AI, 2025).

Questi esempi dimostrano che la trasparenza può essere incorporata come componente tecnica visibile e non solo come principio astratto, facilitando l'adozione responsabile dell'IA nelle organizzazioni.

Le pratiche emergenti confermano la rilevanza delle implicazioni tratte nella presente ricerca: per stimolare la fiducia e l'accettazione dell'IA, la trasparenza deve essere **personalizzata, accessibile e integrata** nei touchpoint digitali usati quotidianamente dai dipendenti. Tali approcci, se applicati anche nei processi di selezione e valutazione, possono ridurre il sospetto verso l'automazione e favorire un'adozione più consapevole e partecipata delle tecnologie intelligenti.

Comunicare la trasparenza in modo mirato nei contesti ad alta ansia lavorativa

La percezione del rischio occupazionale non media la fiducia, ma modera l'effetto della trasparenza: l'impatto positivo della trasparenza sulla fiducia è più forte tra chi percepisce un rischio maggiore.

I risultati della presente ricerca evidenziano che la **percezione del rischio occupazionale agisce come moderatore dell'effetto della trasparenza sulla fiducia**: l'impatto positivo della trasparenza sui sistemi automatizzati di recruiting è significativamente più forte tra i dipendenti che percepiscono un rischio maggiore. Questo dato suggerisce che **la trasparenza non solo è efficace, ma è particolarmente necessaria nei contesti**

64 DeepCore AI. (2025, febbraio). SHAP vs. LIME: Understanding AI Explainability and Model Interpretability.

caratterizzati da incertezza, timore di sostituzione tecnologica o erosione del ruolo umano nei processi decisionali.

Tale evidenza implica che la trasparenza non può essere progettata come messaggio uniforme, ma deve essere **calibrata sui livelli di ansia e vulnerabilità percepiti nei diversi comparti dell'organizzazione.** In altre parole, **la trasparenza deve diventare uno strumento di mitigazione del rischio percepito,** progettato per rassicurare, spiegare e coinvolgere proprio dove i sistemi intelligenti vengono percepiti come una minaccia.

In questa prospettiva, le organizzazioni possono intervenire in modo strategico attraverso azioni quali:

- **modulare la profondità e il tono della comunicazione algoritmica** in funzione del livello di esposizione: i reparti più a rischio (es. customer service, back office, ruoli standardizzabili) dovrebbero ricevere spiegazioni più dettagliate, rassicuranti e co-costruite, per prevenire reazioni difensive o rigetto del sistema;
- **integrare i moduli trasparenti nei processi di ristrutturazione tecnologica,** evidenziando esplicitamente il ruolo residuo (e imprescindibile) dell'intervento umano, la logica assistenziale dell'IA, e gli elementi di responsabilità condivisa;
- **rendere visibile la governance del cambiamento:** la trasparenza non va solo dichiarata ma incarnata attraverso comportamenti osservabili, come la presenza attiva del management nei momenti di implementazione, la possibilità di porre domande sui criteri usati dall'algoritmo, o la partecipazione a workshop interni su "IA e futuro del lavoro".

In questo modo, la trasparenza viene progettata come **strumento attivo di contenimento del rischio percepito,** capace di generare fiducia anche nei contesti più fragili. La chiave è rendere l'IA **non solo visibile, ma comprensibile, contestualizzata e governata,** soprattutto dove i lavoratori temono di perdere potere, controllo o stabilità.

5.2 Fiducia nei sistemi IA e impatti sulle performance aziendali: una traiettoria strategica da esplorare

Sebbene la presente ricerca non abbia incluso, tra le variabili oggetto di analisi empirica, indicatori diretti di performance economico-finanziaria, l'impianto teorico e i risultati ottenuti rendono plausibile l'ipotesi che la fiducia nei sistemi automatizzati di recruiting, alimentata da alti livelli di trasparenza, facilità d'uso e utilità percepita, possa rappresentare una leva abilitante per il miglioramento delle performance aziendali nel medio-lungo periodo.

Questa ipotesi è coerente con quanto evidenziato nella letteratura sul *technology acceptance* (Venkatesh et al., 2012) e nei modelli basati sulla *trust-based adoption* (Siau & Wang, 2018), secondo i quali la fiducia non solo favorisce l'adozione tecnologica, ma può generare condizioni organizzative favorevoli alla creazione di valore. In quest'ottica, la fiducia nei sistemi IA non è soltanto una condizione psicologica, ma un possibile fattore moltiplicatore che influenza positivamente la qualità delle decisioni, la produttività dei processi HR e il clima organizzativo, tutti elementi strettamente connessi alle performance economiche.

Nello specifico, è possibile ipotizzare **quattro aree strategiche di impatto potenziale**:

- **Qualità delle decisioni HR e riduzione del turnover improduttivo:** l'impiego di IA trasparente e accettata può portare a un miglior matching tra candidati e ruoli, aumentando la retention e riducendo costi legati a selezioni errate. Questo può riflettersi indirettamente su indicatori di redditività del capitale umano e costo per assunzione, con effetti secondari su ROI o ROS;
- **Efficienza dei processi e time-to-hire:** la fiducia nei sistemi automatizzati consente una delega più serena e condivisa delle fasi ripetitive (screening, shortlisting), riducendo i tempi medi di selezione (*time-to-hire*). Questo genera risparmi di risorse e può contribuire al miglioramento del margine operativo netto e dell'efficienza operativa;
- **Engagement e clima di innovazione:** la percezione che l'IA sia uno strumento equo e comprensibile favorisce l'engagement nei processi di cambiamento e la partecipazione attiva, con potenziali ricadute positive sulla produttività per dipendente e sulla soddisfazione lavorativa;
- **Percezione di equità e governance del capitale umano:** la trasparenza percepita nei processi di recruiting automatizzati può rafforzare la cultura della responsabilità e della meritocrazia, contribuendo al miglioramento del clima organizzativo e alla riduzione del rischio reputazionale interno, fattori che

nel lungo periodo impattano sul capitale relazionale e sul capitale reputazionale, con riflessi indiretti anche sulla valutazione ESG o sul ROE.

Questi impatti, pur non misurati nella presente ricerca, rappresentano una direzione promettente per indagini future, che potrebbero integrare nei modelli di acceptance indicatori di performance oggettivi (es. ROI, ROE, ROS, turnover, time-to-hire) e soggettivi (es. percezione di efficienza, clima, soddisfazione). Tali estensioni permetterebbero di esplorare in modo più completo il potenziale valore aggiunto dell'IA nei processi HR, contribuendo al superamento del cosiddetto *AI productivity paradox* (Brynjolfsson et al., 2017), secondo cui l'introduzione dell'IA non sempre si traduce in benefici economici tangibili se non accompagnata da fiducia, comprensione e integrazione organizzativa.

CONCLUSIONI

Il presente studio ha analizzato il ruolo della trasparenza percepita nei sistemi di Intelligenza Artificiale come leva strategica per favorire l'accettazione dell'IA nei contesti aziendali, con particolare attenzione ai processi automatizzati di recruitment. L'indagine si è concentrata sull'analisi dei meccanismi psicologici attraverso cui la trasparenza può influenzare la fiducia nei sistemi IA, assumendo quest'ultima come indicatore dell'effettiva accettazione da parte dei lavoratori.

A livello metodologico, è stato adottato un duplice approccio alla mediazione: da un lato, l'analisi sequenziale secondo il modello classico di *Baron e Kenny (1986)*, dall'altro l'applicazione del modello PROCESS di Hayes (Model 4) con bootstrapping, che ha consentito di stimare con maggiore precisione la significatività dell'effetto indiretto. Le analisi hanno confermato che le variabili di utilità percepita e facilità d'uso mediano significativamente la relazione tra trasparenza e fiducia, coerentemente con quanto previsto dal *Technology Acceptance Model (TAM)*. La percezione del rischio occupazionale, invece, non ha soddisfatto i criteri per essere considerata un mediatore, ma ha mostrato un'interazione significativa, fungendo da moderatore: la trasparenza ha un impatto più forte sulla fiducia nei contesti ad alta percezione di rischio.

Questi risultati evidenziano che la trasparenza non opera solo come principio etico o regolamentare, ma anche come determinante strategico dell'efficacia dei sistemi IA: aumenta l'utilità percepita, ne facilita l'uso e riduce il senso di minaccia da parte dei lavoratori, contribuendo alla legittimazione dell'automazione nei processi decisionali. Inoltre, la fiducia nei sistemi automatizzati di recruiting si configura come indicatore sintetico

dell'accettazione dell'IA, in quanto esprime sia la dimensione cognitiva (competenza e correttezza del sistema) sia quella relazionale (sicurezza, equità e integrità percepita).

Infine, pur non analizzando direttamente l'impatto sulle performance aziendali, i risultati ottenuti lasciano ipotizzare che la trasparenza e la fiducia possano costituire una condizione abilitante per il miglioramento delle performance organizzative nel medio-lungo termine, promuovendo engagement, collaborazione e qualità decisionale.

Bibliografia e sitografia

Articoli Accademici e report scientifici

- Turing, A. M. (1950). *Computing Machinery and Intelligence*. *Mind*, 59(236), 433-460.
- Newell, A., & Simon, H. A. (1976). *Computer Science as Empirical Inquiry: Symbols and Search*. *Communications of the ACM*, 19(3), 113-126.
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340.
- Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a Research Agenda on Interventions. *Decision Sciences*, 39(2), 273-315.
- Dwivedi, Y. K., et al. (2021). Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Challenges, Opportunities, and Agenda for Research, Practice, and Policy. *International Journal of Information Management*, 57, 102268.
- Adadi, A., & Berrada, M. (2018). *Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)*. *IEEE Access*, 6, 52138-52160.
- Graham, B. A., Sinclair, R. R., & Sliter, M. (2024). *Job insecurity and health and well-being: What happens when you really need or love your job?* *Economic and Industrial Democracy*, 45(3), 674–695.
- De Witte, H. (1999). "Job Insecurity and Psychological Well-being: Review of the Literature and Exploration of Some Unresolved Issues." *European Journal of Work and Organizational Psychology*, 8(2), 155-177.
- European Agency for Safety and Health at Work (EU-OSHA). (n.d.). Facilitating acceptance of AI-based systems in the workplace and minimising organisational impact. European Agency for Safety and Health at Work.
- Siau, K. and Wang, W.Y. (2018) Building Trust in Artificial Intelligence, Machine Learning, and Robotics. *Cutter Business Technology Journal*, 31, 47-53.
- Ma, X., Yin, J., Lee, D., & Zhang, Y. (2023). *Who Should I Trust: AI or Myself? Leveraging Human and AI Correctness Likelihood to Promote Appropriate Trust in AI-Assisted Decision-Making*. arXiv preprint.
- Li, N., Zhou, H., & Mikel-Hong, K. (2024). *Generative AI enhances team performance and reduces need for traditional teams*. arXiv.
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). *The productivity paradox of AI: Efficiency versus innovation*. *Journal of Economic Perspectives*.
- Jin, G., Jiang, J., & Liao, H. (2023). *The Psychological Impact of AI Adoption in the Workplace*. Scientific Reports.
- Yu, L., & Li, Y. (2022). Artificial intelligence decision-making transparency and employees' trust: The parallel multiple mediating effect of effectiveness and discomfort. *Behavioral Sciences*, 12(5), Articolo 127.

- Zhao, X., et al. (2022) - *Transparency and Employee Trust in AI-Driven Decision Systems*. *AI & Society*, 37(3), 415-430.
- De Fine Licht, J., et al. (2020) - *Transparency in AI Decision-Making: The Role of Justification and Explainability*. *Journal of Business Ethics*, 162(4), 899-915.
- Felzmann, H., et al. (2019) - *Transparency by Design in AI Systems: Principles and Implementation Strategies*. *AI & Ethics*, 1(2), 173-190.
- Glikson, E., & Woolley, A. (2020) - *Human Trust in Artificial Intelligence: Review of Empirical Research*. *Academy of Management Perspectives*, 34(2), 289-305.
- Eslami, M., et al. (2018) - *Disclosing Automated Decision-Making: How Transparency Affects Perception and Trust*. *ACM Transactions on Computer-Human Interaction*, 25(4), 23-47.
- De Witte, H. (2000). Job Insecurity and Psychological Well-being: Review of the Literature and Exploration of Some Unresolved Issues. *European Journal of Work and Organizational Psychology*, 8(2), 155-177.
- Sverke, M., Hellgren, J., & Näswall, K. (2002). No security: A meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology*, 7(3), 242-264.
- Felzmann, H., et al. (2019). Transparency by Design in AI Systems: Principles and Implementation Strategies. *AI & Ethics*, 1(2), 173-190.
- Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660.
- Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, 57(3), 407–434.
- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 102551.
- Siau, K., & Wang, W. (2018). Building trust in artificial intelligence, machine learning, and robotics. *CUTTER Business Technology Journal*, 31(2), 47–53.
- Zarouali, B., Dobber, T., De Pauw, G., & de Vreese, C. H. (2021). Using artificial intelligence (AI) for persuasive communication in politics: Understanding microtargeting and trust in AI. *Computers in Human Behavior*, 119, 106725.
- Brynjolfsson, E., Rock, D., & Syverson, C. (2021). *The Productivity J-Curve: How Intangibles Complement General Purpose Technologies*. *American Economic Journal: Macroeconomics*, 13(1), 333–372. <https://doi.org/10.1257/mac.20180367>
- Dwivedi, Y. K., Hughes, D. L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). *Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy*. *International Journal of Information Management*, 57, 101994.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173–1182.

- Zirar, A., Arslan, A., & Bayram, M. (2023). Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda. *Technological Forecasting and Social Change*, 189, 122404.
- Prado, P. H. M., Korelo, J. C., & Silva, D. M. L. (2014). Mediation, Moderation and Conditional Process Analysis. *Revista Brasileira de Marketing*, 13(4).
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891.
- Abbu, H. R. (2017). Methodology Review: Direct, Mediating, and Moderating Analysis using Hayes PROCESS Models.
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58, 593–614.
- Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*, 98, 39–57.

Siti web

- Studenti.it. (n.d.). Le tre rivoluzioni industriali: riassunto.
- Zanichelli Editore. (2022). La rivoluzione industriale [PDF].
- Bosch Rexroth. (n.d.). Industria 5.0: le differenze con l'Industria 4.0. Bosch Rexroth.
- Artser. (n.d.). Industria 4.0 e 5.0: un passaggio verso una nuova collaborazione sostenibile tra uomo e macchina. Artser.
- AI Joy Academy. (n.d.). La trappola dell'automazione e i rischi della convivenza uomo-macchina. AI Joy Academy.
- Accenture. (2025). Accenture Technology Vision 2025: New age of AI to bring unprecedented autonomy to business. Accenture Newsroom.
- Accenture. (n.d.). Cos'è l'Intelligenza Artificiale. Accenture.
- Dartmouth College. (n.d.). *Artificial intelligence (AI) coined at Dartmouth*. Dartmouth.
- Intelligenza Artificiale Italia. (n.d.). Gli inverni dell'intelligenza artificiale (IA). Intelligenza Artificiale Italia.
- Sviluppo Manageriale. (n.d.). Le applicazioni pratiche dell'intelligenza artificiale nella vita quotidiana e nel business. Sviluppo Manageriale.
- European Commission (2021). *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)*.
- Treccani (2023). *Definizione di Intelligenza Artificiale*. Enciclopedia Italiana.
- ESA Automation. (n.d.). Come vengono addestrati i modelli di intelligenza artificiale. ESA Automation.

- Diritto.it. (n.d.). Metodi di addestramento algoritmi IA: il Garante indaga. Diritto.it.
- Multinazionali.Tech. (n.d.). Bias nell'intelligenza artificiale: cosa sono, esempi e soluzioni. Multinazionali.Tech.
- IBM. (n.d.). AI transparency. IBM Think.
- Altalex. (2025, febbraio 5). AI Act, realtà: entrano in vigore i primi obblighi. Altalex.
- FocusTech. (n.d.). Amazon AI potrebbe aver discriminato le donne. FocusTech.
- Oktopous. (n.d.). L'impatto dell'AI Act sulla ricerca e selezione del personale. Oktopous.
- Artificial Intelligence Act. (n.d.). Article 13 – Artificial Intelligence Act. Artificial Intelligence Act.
- Artificial Intelligence Act. (n.d.). Article 10 – Artificial Intelligence Act. Artificial Intelligence Act.
- Agenda Digitale. (n.d.). IA e lavoro: ecco le professioni che rischiano di più. Agenda Digitale.
- Carmignani, M. (2024, 5 settembre). Futuro del lavoro e AI, oltre il luogo comune: 4 studi per capire cosa aspettarci. Agenda Digitale.
- Viliotti, A. (n.d.). Consapevolezza STAR: l'impatto dell'AI sul benessere lavorativo dei dipendenti. Andrea Viliotti Blog.
- Qipo. (n.d.). Intelligenza artificiale e risorse umane: come cambia l'azienda. Qipo.
- Rödl & Partner. (2024, dicembre 12). Lavoro 5.0 e recruiting: quando scegliere l'intelligenza artificiale. Rödl & Partner.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4th ed.). Boston: Allyn & Bacon.
- IBM. (n.d.). *Introduction to IBM SPSS Statistics. IBM Documentation.*
- Guru. (n.d.). *Explainable AI.*
- Mezzo Pieno. (2023). *L'intelligenza artificiale come motore di trasparenza nelle organizzazioni.*
- IBM. (n.d.). *Use watsonx.governance to accelerate responsible, transparent, and explainable AI. IBM Documentation.*
- Intellico. (n.d.). *Explainable AI: come migliorare la fiducia e la trasparenza nelle decisioni aziendali.*
- DeepCore AI. (2025, febbraio). *SHAP vs. LIME: Understanding AI Explainability and Model Interpretability.*

