

UNIVERSITÀ DEL PIEMONTE ORIENTALE  
DIPARTIMENTO DI GIURISPRUDENZA E SCIENZE POLITICHE,  
ECONOMICHE E SOCIALI

CORSO DI LAUREA MAGISTRALE IN ECONOMIA E  
MANAGEMENT

TESI DI LAUREA

**Intelligenza Artificiale, trasformazione dei mercati e fiducia nelle  
raccomandazioni algoritmiche: un'analisi economico-  
comportamentale**

Relatore:

Chiar.mo Prof. Graziano Abrate

Correlatrice:

Chiar.ma Prof.ssa Clementina Bruno

Candidato:

Alessio Failla

ANNO ACCADEMICO 2024/2025



## INDICE

INTRODUZIONE .....	9
CAPITOLO 1 – L’Intelligenza Artificiale come leva di trasformazione economica e strategica dei mercati .....	13
1.1 – La nuova economia dei dati.....	13
1.1.1. Dal capitale fisico al capitale informativo .....	13
1.1.2. La natura economica del dato: bene non rivale ma escludibile.....	14
1.1.3. I dati come capitale intangibile: investimenti e valore macroeconomico .....	15
1.1.4. Creazione di valore e politiche per la condivisione dei dati.....	16
1.1.5. Implicazioni macroeconomiche e prospettive di crescita.....	17
1.1.6. Verso un paradigma di crescita data-driven .....	18
1.2 – L’Intelligenza Artificiale come fattore di produzione e vantaggio competitivo .....	19
1.2.1 L’IA come nuovo motore della produttività.....	19
1.2.2 L’Intelligenza Artificiale come capitale tecnologico .....	19
1.2.3 Evidenze empiriche: l’impatto dell’IA sulla produttività.....	20
1.2.4 Il paradosso della produttività e i ritardi di adozione .....	21
1.2.5 L’IA come fonte di vantaggio competitivo .....	22
1.2.6 Politiche economiche e prospettive di sviluppo .....	22
1.3 – Strategie di mercato <i>data-driven</i> .....	23
1.3.1 Dalla personalizzazione alla competizione algoritmica .....	23
1.3.2 Architetture: scegliere ciò che serve (e solo ciò che serve).....	24
1.3.3 Dati e metriche: ciò che misuri è ciò che muovi .....	24
1.3.4 <i>Hyper-personalization</i> : dall’algoritmo alla capacità organizzativa .....	25
1.3.5 Dalla <i>accuracy</i> al valore: raccomandazioni <i>value-aware</i> e <i>profit-aware</i> .....	26
1.3.6 Decisioni <i>data-driven</i> e performance d’impresa .....	27
1.3.7 Dalla teoria alla pratica: applicazioni <i>business</i> e valutazione sperimentale.....	27
1.3.8 Sfide tecniche e organizzative: <i>cold start</i> , sparsità, <i>echo chambers</i> e <i>governance</i> .....	28
1.3.9 Strategie <i>data-driven</i> : quadro operativo.....	28
1.3.10 Dalla strategia alle implicazioni competitive e di fiducia .....	29

1.4 – Implicazioni economiche per la concorrenza e il valore del consumatore.....	30
1.4.1 Dalla centralità dei dati alla trasformazione dei mercati.....	30
1.4.2 Le nuove dinamiche della concorrenza nell’economia dei dati .....	31
1.4.3 I sistemi di raccomandazione e il valore del consumatore .....	32
1.4.4 Implicazioni di welfare e redistribuzione del valore .....	34
1.4.5 Politiche economiche e regolazione della concorrenza digitale.....	35
1.4.6 Verso un equilibrio tra efficienza, concorrenza e fiducia .....	36
1.5 – Rischi e limiti dell’economia algoritmica .....	38
1.5.1 Dal progresso tecnico alla vulnerabilità sistemica .....	38
1.5.2 Rischi economici e strutturali dell’automazione algoritmica.....	39
1.5.3 Bias, opacità e problemi di accountability nei sistemi AI.....	41
1.5.4 Impatti etici e sociali: privacy, manipolazione e disuguaglianze .....	42
1.5.5 Verso una governance responsabile dell’IA.....	43
1.5.6 Equilibrio tra innovazione, fiducia e sostenibilità.....	44
CAPITOLO 2 – Letteratura accademica e modelli teorici sulla fiducia nei sistemi di raccomandazione AI	
.....	46
2.1 Le origini del dibattito sull’avversione agli algoritmi.....	46
2.1.1. Perché parlare di “avversione” (e perché non basta).....	46
2.1.2. Che cosa intendiamo per “algoritmo” in questa letteratura.....	47
2.1.3. Il paradigma dell’“algorithm aversion”: l’errore come rottura della fiducia .....	47
2.1.4. Perché l’errore dell’algoritmo pesa di più: norme implicite, aspettative e agency .....	48
2.1.5. Il controcanto: l’algorithm appreciation e la preferenza per l’algoritmo .....	49
2.1.6. “Human vs algorithm trust”: il ruolo del compito, della posta in gioco e della cornice di presentazione .....	50
2.1.7. Ricomporre il puzzle: perché avversione e apprezzamento possono coesistere .....	52
2.1.8. Implicazioni per la tesi e ponte verso i capitoli successivi.....	53
2.2. Il modello “Word-of-Machine” (Longoni & Cian, 2022).....	54
2.2.1. Perché serve un modello “contestuale” della fiducia nelle raccomandazioni algoritmiche ...	54
2.2.2. Definizione del “Word-of-Machine effect” e intuizione di fondo .....	54
2.2.3. Il meccanismo: credenze di competenza e “appropriatezza” della fonte .....	55
2.2.4. Evidenze empiriche: sintesi ragionata delle principali evidenze sperimentali.....	56

2.2.5. Confini dell'effetto: quando la macchina perde terreno anche nel dominio utilitaristico (Studio 5).....	58
2.2.6. Un secondo confine: dalla sostituzione all'"augmented intelligence" (Studio 6).....	59
2.2.7. Interventi di attenuazione: cambiare le credenze (Studio 7a–7b) .....	60
2.2.8. Implicazioni teoriche: un ponte tra avversione/apprezzamento e motivazioni di consumo... 60	
2.2.9. Dal modello teorico al disegno sperimentale: misurazione, ipotesi e continuità con il Capitolo 3 .....	61
2.2.10. Limiti e spunti per le sezioni successive .....	62
2.3. Consumo edonistico e utilitaristico: cornice teorica e implicazioni per le raccomandazioni algoritmiche.....	63
2.3.1. Perché distinguere tra edonismo e utilitarismo nelle decisioni di consumo.....	63
2.3.2. Definizioni operative: edonico e utilitaristico come attributi, come scopi e come contesti... 64	
2.3.3. Il nodo della "giustificazione": perché il piacere richiede una ragione .....	65
2.3.4. Preferenze che si capovolgono: valutazione separata e valutazione congiunta .....	66
2.3.5. Pagare in tempo o pagare in denaro: il prezzo psicologico della giustificazione.....	67
2.3.6. Acquisire o rinunciare: quando la cornice di scelta rende l'edonico più "pesante".....	68
2.3.7. Evidenze "di mercato": il valore di scambio e il divario acquirente–venditore.....	69
2.3.8. Una visione integrativa: il modello decisionale e la "self-attribution" nelle scelte edoniche 69	
2.3.9. Locus of choice e causality: quando delegare riduce (o non riduce) la soddisfazione .....	71
2.3.10. Implicazioni per la fiducia nei sistemi di raccomandazione: una continuità verticale verso i Capitoli 2 e 3 .....	72
2.4. Fiducia nei sistemi AI: dall'"appropriate reliance" nell'automazione alla fiducia nelle raccomandazioni algoritmiche.....	73
2.4.1. Introduzione: perché la fiducia è la variabile-ponte tra potenziale tecnologico e impatto economico.....	73
2.4.2. Che cos'è la fiducia: vulnerabilità, incertezza e aspettative di esito .....	75
2.4.3. La tradizione "trust in automation": misuse, disuse e progettazione dell'affidamento appropriato.....	76
2.4.4. Dall'automazione tradizionale all'AI: perché la fiducia diventa più difficile (e più centrale) .....	77
2.4.5. Le dimensioni percettive della fiducia: competenza e calore come grammatica universale del giudizio .....	78

2.4.6. Fiducia in AI e fiducia nei recommendation agents: quando la tecnologia diventa “attore sociale” .....	79
2.4.7. Determinanti della fiducia nell’AI: affidabilità, trasparenza, tangibilità, antropomorfismo e comportamenti di immediatezza.....	80
2.4.8. Calibrazione della fiducia: undertrust, overtrust e qualità dell’esito decisionale.....	81
2.4.9. Fiducia, spiegazioni e percezione di imparzialità: il problema della benevolenza nei mercati digitali.....	82
2.4.10. Verso un modello integrato: fiducia come funzione di performance, percezioni sociali e architettura di interazione .....	83
2.5 Modello concettuale e ipotesi di ricerca: dalla teoria sulla fiducia alla verifica empirica .....	86
2.5.1. Costrutti chiave: fiducia, competenza, calore e propensione a seguire il consiglio.....	87
2.5.2 Variabili esplicative: fonte del consiglio e dominio decisionale.....	88
2.5.3 Meccanismi teorici: avversione, apprezzamento e “word-of-machine” .....	89
2.5.4 Ipotesi di ricerca .....	90
2.5.5 Sintesi del modello e ponte verso il Capitolo 3.....	91
CAPITOLO 3 – Analisi empirica: fiducia nelle raccomandazioni AI tra decisioni di investimento e di donazione.....	92
3.1 Disegno della ricerca .....	92
3.1.1 Logica generale: dal modello concettuale alla verifica empirica .....	93
3.1.2 Disegno sperimentale: un 3×2 tra-soggetti.....	93
3.1.3 Manipolazioni: Costruzione dello scenario e coerenza tra le condizioni .....	94
3.1.4 Variabili misurate: variabili dipendenti principali e dimensioni percettive .....	95
3.1.5 Strategia di analisi: test delle ipotesi e lettura dei meccanismi .....	97
3.1.6 Come si passa dal disegno ai dati .....	98
3.2 Struttura del questionario e misure.....	98
3.2.1 Scenari sperimentali .....	99
3.2.2 Scale e costrutti .....	101
3.2.3 Costruzione degli indici, affidabilità e validità delle misure.....	104
3.3 Campione e procedura .....	108
3.3.1 Reclutamento e criteri di inclusione.....	109
3.3.2 Descrizione del campione .....	110
<b>Tabella 3.6 – Caratteristiche del campione e altre variabili di controllo (N = 402) .....</b>	<b>111</b>

3.3.3 Procedura.....	113
3.4 Analisi dei dati.....	115
3.4.1 Preparazione dei dati .....	115
3.4.2 Statistiche descrittive.....	116
3.4.3 Analisi principali: ANOVA fattoriale 3×2.....	121
3.4.4 Il peso delle variabili percettive nei due domini decisionali .....	132
3.4.5 Analisi di moderazione: il ruolo dell'uso dell'IA .....	135
3.5.1 Discussione delle ipotesi alla luce della letteratura.....	148
3.5.2 Implicazioni manageriali.....	157
3.5.4 Limiti e sviluppi futuri .....	159
Conclusioni.....	161
BIBLIOGRAFIA.....	165
APPENDICE .....	168



## INTRODUZIONE

Negli ultimi anni l'intelligenza artificiale è passata, con rapidità sorprendente, da promessa tecnologica a componente ordinaria dell'infrastruttura decisionale dei mercati contemporanei. La sua diffusione non riguarda soltanto i settori ad alta intensità tecnologica, né si limita alle funzioni più visibili dell'automazione industriale o dell'analisi predittiva. L'IA entra sempre più spesso nei punti in cui si orientano scelte, si selezionano opzioni, si costruiscono preferenze: suggerisce prodotti, ordina contenuti, personalizza offerte, classifica profili di rischio, indirizza investimenti, raccomanda consumi, filtra opportunità. In questo senso, non è solo uno strumento di efficienza. È una tecnologia che interviene direttamente nella mediazione tra informazione e decisione.

Per l'economia e per il management, il tema è cruciale. Nei mercati data-driven il vantaggio competitivo non dipende unicamente dalla disponibilità di dati o dalla potenza degli algoritmi che li elaborano. Conta anche la capacità di trasformare quell'elaborazione in un consiglio ritenuto credibile, appropriato e utile da chi lo riceve. Una raccomandazione, infatti, non genera valore soltanto perché è tecnicamente accurata. Lo genera se viene presa sul serio, se viene percepita come legittima, se riesce a orientare davvero il comportamento dell'utente. La questione della fiducia nelle raccomandazioni algoritmiche, quindi, non è un problema accessorio rispetto all'innovazione tecnologica. È una delle condizioni attraverso cui tale innovazione può tradursi in adozione, in efficacia organizzativa e, più in generale, in trasformazione economica.

Da qui nasce il problema che questa tesi affronta. Quando un consiglio proviene da un sistema di intelligenza artificiale, da un consulente umano oppure da una configurazione ibrida tra i due, il destinatario lo valuta sempre allo stesso modo? Oppure il giudizio cambia a seconda del tipo di decisione che viene proposta? E ancora: ciò che conta davvero è la natura della fonte in sé, oppure il modo in cui quella fonte viene percepita — in termini di competenza, calore, affidabilità, vicinanza, credibilità? La domanda può essere formulata in modo ancora più semplice: non basta chiedersi se l'IA “funzioni” meglio o peggio dell'umano; occorre capire quando, in quali contesti e per quali utenti una raccomandazione algoritmica venga considerata convincente e degna di essere seguita.

La rilevanza del tema si colloca esattamente all'incrocio tra due livelli analitici che spesso, nella discussione pubblica e talvolta anche in quella accademica, tendono a rimanere separati. Il primo è il livello macro e strategico della trasformazione dei mercati. Qui l'attenzione si concentra sul valore economico dei dati, sulla crescente centralità dei modelli predittivi, sulla capacità dell'IA di modificare produttività, organizzazione e concorrenza. Il secondo è il livello micro-comportamentale. Qui il fuoco si sposta su come gli individui reagiscono alla tecnologia, su quali criteri usano per valutare una fonte di consiglio, su come si formi la fiducia nei confronti di sistemi percepiti come intelligenti ma non umani. È proprio nell'incontro tra questi due piani che si colloca la tesi.

Il Capitolo 1 sviluppa il primo versante di questo percorso. L'intelligenza artificiale viene ricostruita come fattore della trasformazione economica dei mercati contemporanei, dentro un quadro in cui dati, capacità computazionale e apprendimento automatico incidono sui modelli di business, sulle strategie di personalizzazione, sulla costruzione del vantaggio competitivo e sulla ridefinizione dei processi decisionali. In questa prospettiva, la raccomandazione algoritmica non appare come un semplice output tecnico, ma come una leva di coordinamento economico e di orientamento delle scelte.

Il Capitolo 2 compie invece uno spostamento di scala. Dal piano della trasformazione economica si passa a quello della ricezione soggettiva del consiglio. Vengono discussi i principali filoni teorici relativi alla fiducia negli algoritmi: il dibattito tra *algorithm aversion* e *algorithm appreciation*, la letteratura sulla persuasività della fonte non umana e sul *word-of-machine*, il tema della *trust in automation* e dell'*appropriate reliance*, nonché il ruolo delle dimensioni percettive di competenza e calore. L'ipotesi di fondo che attraversa questa parte del lavoro è che la fiducia non sia una risposta fissa alla tecnologia, ma una costruzione situata, legata al tipo di compito, alla fonte del consiglio e al significato che il soggetto attribuisce alla decisione da prendere.

La tesi nasce precisamente da questo intreccio. Da un lato, i mercati digitali rendono sempre più centrale il consiglio generato da sistemi intelligenti. Dall'altro, la letteratura suggerisce che l'accettazione di tali consigli non dipenda soltanto dalla loro accuratezza, ma anche dalla percezione della fonte. Ciò che ancora merita di essere chiarito, però, è il modo in cui questi elementi si combinano concretamente quando si confrontano diverse fonti di consiglio in contesti decisionali differenti. Non è sufficiente, infatti, opporre in astratto "macchina" e "umano". Occorre osservare come la valutazione cambi quando la decisione

richiama una logica più analitica e strumentale, come nel caso dell'investimento, oppure una logica più valoriale e prosociale, come nel caso della donazione. Ed è altrettanto importante capire se la configurazione ibrida, in cui l'IA elabora e l'umano interpreta o media, rappresenti una forma di sintesi particolarmente efficace.

La domanda di ricerca che guida il lavoro può dunque essere formulata così: come cambia la fiducia nelle raccomandazioni al variare della fonte del consiglio — IA, umano, ibrido — e del contesto decisionale in cui tale consiglio è inserito? Questa domanda ne contiene almeno altre due, più specifiche. La prima riguarda i meccanismi: quali dimensioni percettive, tra competenza, calore e affidabilità, aiutano a spiegare la maggiore o minore disponibilità a seguire il consiglio? La seconda riguarda le differenze individuali: la familiarità con l'IA modifica il modo in cui la fonte viene valutata?

Per rispondere a questi interrogativi, la tesi adotta un'impostazione empirica di tipo sperimentale. Senza entrare qui nei dettagli metodologici, che verranno presentati compiutamente nel Capitolo 3, la logica del disegno è quella di mettere a confronto diverse fonti di consiglio in diversi contesti decisionali, osservando come cambino alcune dimensioni centrali: la fiducia nel consiglio, la competenza percepita della fonte, il calore/empatia percepita, l'intenzione di seguire il consiglio e l'affidabilità attribuita alla fonte. Ciò che interessa non è soltanto sapere quale configurazione ottenga il valore più alto in assoluto, ma comprendere se il giudizio sulla fonte dipenda dal contesto e se tale dipendenza venga riorientata da caratteristiche dei partecipanti, come il loro livello di utilizzo dell'IA.

Il contributo atteso del lavoro si colloca esattamente qui. La tesi non punta a dimostrare una superiorità generale dell'IA né a riaffermare, in chiave difensiva, la centralità indiscriminata dell'intervento umano. Un risultato del genere sarebbe forse rassicurante nella sua semplicità, ma poco utile sul piano teorico e manageriale. Il contributo che questo lavoro intende offrire è diverso: mostrare che la fiducia nelle raccomandazioni è contestuale, relazionale e mediata. Contestuale, perché cambia con il tipo di compito. Relazionale, perché il giudizio sulla fonte non dipende solo dalla correttezza tecnica, ma anche da dimensioni come vicinanza, sensibilità e legittimazione. Mediata, perché tra fonte ed esito finale si collocano percezioni come competenza, calore e affidabilità che trasformano una raccomandazione in qualcosa di più o meno credibile. In questa prospettiva, la configurazione ibrida appare fin dall'inizio come una pista teoricamente promettente: non come soluzione universalmente

ottimale, ma come possibile punto di mediazione tra razionalità algoritmica e accettabilità umana.

L'architettura della tesi riflette questa logica. Il primo capitolo ricostruisce il ruolo dell'IA nella trasformazione economica dei mercati, mettendo al centro dati, piattaforme, personalizzazione e modelli decisionali. Il secondo capitolo sviluppa il quadro teorico sul rapporto tra algoritmi, fiducia e fonti di consiglio, fino alla formulazione delle ipotesi di ricerca. Il terzo capitolo presenta il disegno empirico, la struttura del questionario, il campione, le analisi statistiche e la discussione dei risultati, fino alle implicazioni manageriali e ai limiti dello studio. Le Conclusioni, infine, riportano il percorso della tesi al suo significato complessivo: comprendere che cosa accade quando la decisione economica non è più solo assistita dall'informazione, ma mediata da una fonte — umana, algoritmica o ibrida — che deve prima di tutto conquistare fiducia.

Questa è, in fondo, la posta in gioco del lavoro. Nei mercati contemporanei, il problema non è soltanto quanto sia sofisticato un algoritmo, ma se la sua raccomandazione riesca a essere riconosciuta come appropriata da chi la riceve. L'innovazione tecnologica, da sola, non basta. Deve incontrare aspettative, percezioni e forme di legittimazione. È in questo spazio — tra trasformazione dei mercati e psicologia economica della fiducia — che si colloca la tesi.

Al fine della redazione del presente elaborato, mi sono avvalso dell'uso di strumenti di intelligenza artificiale esclusivamente per correttezza formale, quali la riformulazione e la revisione linguistica del testo. I contenuti sostanziali e le analisi qui presentate sono interamente frutto di ricerche, elaborazioni e riflessioni personali.

## **CAPITOLO 1 – L’Intelligenza Artificiale come leva di trasformazione economica e strategica dei mercati**

### **1.1 – La nuova economia dei dati**

#### **1.1.1. Dal capitale fisico al capitale informazionale**

Negli ultimi due decenni, l’economia globale ha assistito a una profonda trasformazione che ha spostato il centro della creazione di valore dal capitale fisico e finanziario a quello informazionale, ossia alla capacità di raccogliere, elaborare e sfruttare dati su larga scala. In tale contesto, i dati non rappresentano più un semplice sottoprodotto delle attività economiche, bensì una risorsa strategica in grado di generare vantaggi competitivi, innovazione e crescita economica.

L’Organizzazione per la Cooperazione e lo Sviluppo Economico (OECD, 2022) definisce i dati come un fattore di produzione intangibile, stoccabile e potenzialmente riutilizzabile, che oggi rappresenta una componente essenziale dei modelli di business digitali. Tuttavia, la crescente importanza dei dati nei modelli di business non si riflette ancora nelle statistiche macroeconomiche e nei conti nazionali. Gli investimenti in dati non vengono contabilizzati allo stesso modo degli investimenti in capitale fisico o umano.

La transizione verso una data-driven economy ha importanti implicazioni per la struttura e il funzionamento dei mercati e delle politiche economiche. Come osservano Corrado, Haskel, Iommi e Jona-Lasinio (OECD, 2022), il crescente stock di dati prodotti da aziende e consumatori ha cambiato il funzionamento dei processi produttivi e il funzionamento della concorrenza tra le imprese. Le fonti tradizionali di creazione di valore lasciano il posto a forme di creazione di valore basate sull’analisi, la condivisione e la monetizzazione delle informazioni.

In questo scenario, i dati sono un fattore di produzione, una risorsa strategica e anche una merce soggetta a regolamentazione, ponendo questioni inedite in termini di proprietà, privacy e politica della concorrenza. Per comprendere meglio tali dinamiche, è necessario definire le caratteristiche economiche dei dati e il modo in cui queste (non rivalità, escludibilità) guidano incentivi e politiche; come verrà mostrato nel Capitolo 2, anche la fiducia degli utenti nei sistemi intelligenti gioca un ruolo nel determinare se i benefici economici attesi si concretizzano effettivamente.

### **1.1.2. La natura economica del dato: bene non rivale ma escludibile**

La teoria economica dei dati è stata significativamente arricchita dal contributo di Jones e Tonetti (2020), che definiscono, in primo luogo, che i dati sono un bene non rivale, ossia le informazioni possono essere consumate simultaneamente da diversi attori economici senza che il consumo di uno riduca la capacità di consumo di un altro. Questo concetto rivoluzionario è ciò che distingue i dati da tutti gli altri beni tradizionali presenti sul mercato e, inoltre, ne evidenzia la capacità di generare rendimenti crescenti di scala: più aziende/istituzioni condividono e utilizzano i dati, maggiore è il potenziale di aumento della produttività. Le caratteristiche di non rivalità implicano che il valore sociale è spesso maggiore del valore privato appropriato dalle aziende che lo generano. Tuttavia, allo stesso tempo, i dati sono escludibili, il che suggerisce che esistono mezzi legali o tecnici per limitarne l'accesso tramite protezione, brevetti o diritti di proprietà. Esiste quindi un fallimento del mercato naturale per cui è difficile conciliare un'allocatione socialmente efficiente con la protezione della privacy dei dati: poiché l'eccessiva restrizione nell'uso dei dati disincentiva la diffusione della conoscenza e l'innovazione, mentre fornire un accesso troppo libero può compromettere i diritti dei consumatori (Jones & Tonetti, 2020). In termini reali, ciò significa tre potenziali regimi di proprietà dei dati: (i) proprietà aziendale, in cui le aziende non sono disposte a cedere i dati per paura di compromettere il loro vantaggio competitivo, limitando così l'efficienza sociale; (ii) proprietà pubblica o regolamentata, in cui il governo definisce limiti e condizioni d'uso, bilanciando privacy e innovazione; (iii) proprietà dei consumatori, in cui i consumatori determinano a chi desiderano vendere i propri dati e se desiderano venderli, a seconda di quanto sono disposti a sacrificare in termini di privacy rispetto al benessere collettivo.

Jones e Tonetti (2020) illustrano che, in un equilibrio decentralizzato, l'attribuzione dei diritti di proprietà ai consumatori tende a generare risultati che si avvicinano all'ottimo sociale, incendiando una diffusione più ampia dei dati nel rispetto delle preferenze individuali per la riservatezza. Il passo successivo, dopo una corretta assegnazione dei diritti e degli incentivi, è vedere come i dati diventano capitale e le implicazioni macroeconomiche: questo avviene nel capitolo successivo e sul piano delle strategie d'impresa nel capitolo 1.3.

### 1.1.3. I dati come capitale intangibile: investimenti e valore macroeconomico

Una delle sfide più importanti della nuova economia dei dati è come misurare il valore del capitale informativo. Corrado et al. (OECD, 2022) forniscono un quadro di valutazione che suggerisce di considerare i dati come un bene immateriale alla stregua della conoscenza, software e del capitale organizzativo. In quanto, secondo tale approccio, gli investimenti effettuati nella raccolta, nell'archiviazione e nell'analisi dei dati, possono essere considerati un accumulo di capitale, con rendimenti valutati in base ai guadagni in termini di efficienza produttiva e alle capacità innovative.

Secondo le stime dell'OECD (2022), dal 2010 al 2018, i dati e la “data intelligence” hanno rappresentato investimenti nei paesi europei più importanti (Francia, Germania, Italia, Spagna e Regno Unito) compresi tra il 5 e il 6,5% del valore aggiunto nel settore privato. Queste cifre suggeriscono che il capitale informativo sta diventando rilevante dal punto di vista macroeconomico nella stessa misura in cui lo era il capitale fisico durante la fase di sviluppo industriale.

Da una prospettiva microeconomica, Farboodi e Veldkamp (2021) determinano che la raccolta e l'elaborazione dei dati creano un ciclo endogeno di investimento/crescita. Nel loro modello, le aziende acquisiscono dati attraverso le interazioni con i consumatori, che poi utilizzano tali dati per sviluppare nuovi prodotti di migliore qualità e prendere decisioni più efficaci. Ciò crea un ciclo auto-rinforzante: l'aumento dell'utilizzo dei dati accresce la produttività, stimola ulteriori investimenti in tecnologie di analisi e, di conseguenza, amplifica il vantaggio competitivo delle imprese data-intensive (Farboodi & Veldkamp, 2021).

Questo processo ha dato origine a un'economia “data-intensive”, in cui la capacità di estrarre valore dai dati si configura come una nuova forma di capitale organizzativo. La conseguenza macroeconomica è un aumento della dispersione di produttività tra imprese, poiché quelle dotate di competenze e infrastrutture digitali superiori riescono a sfruttare meglio le esternalità positive del dato (OECD, 2022).

Pertanto, non è sufficiente limitarsi ad “accumulare” dati, ma è necessario possedere la capacità di rielaborarli algebricamente e diffonderli. Questi status collegano le strategie e le policy a livello micro e conducono al livello successivo.

#### **1.1.4. Creazione di valore e politiche per la condivisione dei dati**

Nell'ambito dell'economia dei dati, quando si focalizza l'attenzione sul concetto di "generare un valore", si tiene conto da un lato ai diversi elementi inerenti al numero di informazioni raccolte e dall'altro alla capacità di creare una conoscenza che sia vantaggiosa e profittevole.

Secondo quanto riportato da Farboodi e Veldkamp (2021), i dati primari producono valore soltanto dal momento in cui vengono uniti a processi di algoritmi analitici che sono in grado di estrapolare insight strategici, pattern e correlazioni (Farboodi & Veldkamp, 2021).

Per farla breve, il valore economico dei dati non è da ricercare nella loro mera disponibilità bensì in un'estensione degli stessi definita proprio come intelligenza derivata.

L'OECD (2022) riprende il concetto sopra citato distinguendo e descrivendo tre precisi stati del ciclo del valore dei dati: acquisizione e archiviazione; analisi e modellizzazione ed infine monetizzazione e condivisione. Suddetta rete data-driven, nasce dunque dall'interazione fra imprese, consumatori e istituzioni pubbliche, il cui equilibrio e regolazione reciproca diviene cruciale per impedire la manifestazione di possibili inefficienze sistemiche.

Esistono inoltre delle specifiche politiche che, promuovendo la portabilità dei dati, la collaborazione tra sistemi e la genesi di mercati dei dati, permettono di ridurre i costi di transazione e agevolare la concorrenza, con il fine ultimo di ampliare le opportunità di innovazione.

Ciononostante, Jones e Tonetti (2020) evidenziano che, in assenza di stimoli adeguati, le imprese tendono a "trattenere" i propri dati per paura di una nuova ed emergente innovazione, ostacolando così la diffusione di conoscenza e l'efficienza paretiana.

Come si evince dal principio europeo di data sovereignty, un quadro normativo in grado di attribuire ai consumatori la supervisione dei propri dati, può permettere di raggiungere un equilibrio più efficace tra tutela della privacy e crescita economica.

Tenendo conto di tale prospettiva, i dati assumono quindi una duplice funzione: la condivisione tra bene economico e bene relazionale condiziona il livello di fiducia tra gli interpreti del mondo digitale. Per tale motivo, in assenza di fiducia e sistemi di governance

appropriati, i flussi di valore rimangono sottovalutati e trascurati. Quest'ultimo aspetto in particolare, che sarà affrontato più dettagliatamente nel capitolo 2, preannuncia il ruolo cruciale della propensione da parte degli individui, alla condivisione dei dati e all'approvazione delle scelte algoritmiche.

### **1.1.5. Implicazioni macroeconomiche e prospettive di crescita**

Dal punto di vista macroeconomico, l'espansione del capitale informativo ha effetti osservabili sulla produttività del lavoro e sulla resilienza economica. Secondo Corrado et al. (OECD, 2022), una diffusione, anche parziale, degli asset in dati potrebbe aumentare la crescita della produttività fino a 0,5 punti percentuali annui, se l'intervento pubblico incoraggiasse l'uso dei dati e l'acquisizione di competenze digitali. Tuttavia, allo stesso tempo, la supremazia informativa di una manciata di grandi aziende digitali può aumentare le disuguaglianze e le distorsioni della concorrenza, rendendo necessario un intervento politico che democratizzi l'uso delle informazioni.

In definitiva, questa nuova economia basata sui dati rappresenta una transizione dall'economia dell'informazione, in cui l'intelligenza artificiale, le infrastrutture digitali e il capitale umano qualificato lavorano assieme per ridefinire i confini della produttività e dell'innovazione. Pertanto, l'obiettivo dei governi e delle imprese è quello di trovare un equilibrio che soddisfi tre obiettivi primari: crescita economica alimentata dall'uso delle informazioni; protezione della privacy e dei dati dei consumatori; uguaglianza nell'accesso e nella condivisione del valore delle informazioni.

Pertanto, l'equilibrio non è solo una questione macroeconomica, ma si riflette anche nelle azioni strategiche e nelle strutture organizzative delle aziende, come illustrato nel capitolo 1.2 (L'IA come fattore di produzione e vantaggio competitivo) e nel capitolo 1.3 (Processo decisionale basato sui dati).

### 1.1.6. Verso un paradigma di crescita data-driven

La “nuova economia dei dati” è una struttura complessa e multilivello che considera i dati come capitale produttivo, bene pubblico e oggetto di scambio economico. Come evidenziato da Jones e Tonetti (2020) e dall’OECD (2022), la sfida principale consiste nel riconciliare la non rivalità del dato con le esigenze di escludibilità e protezione, definendo istituzioni e mercati in grado di massimizzare i benefici collettivi derivanti dalla sua circolazione.

In prospettiva, il valore economico dei dati non risiede soltanto nella loro quantità, ma soprattutto nella capacità di trasformarli in decisioni migliori, innovazioni più rapide e servizi personalizzati, grazie all’intelligenza artificiale e alle tecnologie di apprendimento automatico (Farboodi & Veldkamp, 2021). Pertanto, la crescita futura dipende dalla capacità di sfruttare il capitale intangibile dei dati, sviluppare infrastrutture digitali interoperabili e sistemi normativi che supportino una condivisione responsabile delle informazioni.

Pertanto, la nuova economia dei dati non è solo una fase tecnologica, ma rappresenta un nuovo paradigma socioeconomico basato sull’informazione come mezzo di produzione, vantaggio competitivo e bene collettivo.

Tutti questi aspetti saranno integrati nei capitoli 1.2-1.5 nei fattori di produzione/vantaggio competitivo (1.2), strategie data-driven (1.3), concorrenza e welfare (1.4) e rischi/limiti (1.5), completando l’inquadramento teorico-empirico.

#### Riferimenti bibliografici – (APA 7<sup>a</sup> ed.)

Corrado, C., Haskel, J., Iommi, M., & Jona-Lasinio, C. (2022). *The value of data in digital-based business models: Measurement and economic policy implications* (OECD Economics Department Working Papers No. 1723). OECD Publishing. <https://doi.org/10.1787/d960a10c-en>

Farboodi, M., & Veldkamp, L. (2021). *A growth model of the data economy* (NBER Working Paper No. 28427). National Bureau of Economic Research. <https://www.nber.org/papers/w28427>

Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the economics of data. *American Economic Review*, 110(9), 2819–2858. <https://doi.org/10.1257/aer.20191330>

## **1.2 – L’Intelligenza Artificiale come fattore di produzione e vantaggio competitivo**

### **1.2.1 L’IA come nuovo motore della produttività**

Nel panorama economico contemporaneo, l’Intelligenza Artificiale (IA) è divenuta un elemento cardine dei processi produttivi e delle strategie competitive delle imprese. Essa non rappresenta più un semplice strumento di automazione, bensì un fattore di produzione autonomo, il quale è capace di generare valore attraverso l’elaborazione di informazioni, l’apprendimento dai dati nonché il miglioramento continuo delle decisioni operative (Brynjolfsson, Rock & Syverson, 2017).

Come sottolinea Mazzotta (2024), solo in termini di produttività totale dei fattori (TFP) è possibile comprendere fino a che punto l’IA sia un fattore produttivo, poiché essa modifica simultaneamente l’efficienza del lavoro, la qualità del capitale e la capacità di innovazione.

In un’Europa in cui la produttività si è stabilizzata e in cui le vulnerabilità strutturali rendono più difficile la digitalizzazione, l’adozione di sistemi intelligenti costituisce una leva strategica per il rilancio della crescita.

Nel seguito l’IA è trattata come capitale tecnologico (1.2.2), se ne documentano le evidenze empiriche (1.2.3) e si discutono i ritardi di adozione (1.2.4), con ricadute sulle strategie (1.2.5) e sulle policy (1.2.6).

### **1.2.2 L’Intelligenza Artificiale come capitale tecnologico**

Nei modelli economici classici, la crescita deriva da una dotazione più ampia di capitale fisico e umano. Tuttavia, con l’introduzione dell’IA, nel nuovo mondo del lavoro è emersa una nuova forma di capitale: il capitale algoritmico, ossia l’insieme dei sistemi e delle applicazioni basati su algoritmi, reti neurali e machine learning che potenziano la produzione e i processi decisionali (Greci, 2023).

Brynjolfsson et al. (2017) distinguono tra “capitale informazionale” (dati, piattaforme digitali, infrastrutture e opportunità) e capitale algoritmico (la capacità di un’impresa di trasformare i dati in previsioni, decisioni e innovazioni). Il capitale algoritmico deve essere coltivato attraverso investimenti, manutenzione e continuo aggiornamento, analogamente al capitale fisico, ma a differenza di quest’ultimo è soggetto a rendimenti crescenti di scala, poiché ogni unità aggiuntiva di dato migliora le prestazioni dell’intero sistema.

Secondo Mazzotta (2024), l'intelligenza artificiale impiegata nei sistemi produttivi europei rappresenta un moltiplicatore di forza per la produttività: consente alle imprese di utilizzare input meno produttivi in modo più efficiente, riduce la produzione di scarti e favorisce una distribuzione più efficace delle risorse. L'IA opera tramite database, strumenti predittivi e analisi automatizzate, generando una relazione di complementarità continua tra capitale umano e capitale tecnologico, i quali rafforzano reciprocamente i propri contributi alla crescita economica.

Pertanto, in relazione a questo quadro teorico, è necessario determinare in che misura e se tali previsioni trovino riscontro nei dati. La sezione seguente esamina le evidenze attualmente disponibili in Europa.

### **1.2.3 Evidenze empiriche: l'impatto dell'IA sulla produttività**

Le analisi empiriche condotte da Mazzotta (2024) su un campione di paesi dell'Unione Europea, basate sui dati di EUKLEMS, AI Watch e dell'AI Index, analizzano il livello di sviluppo dell'IA nei diversi settori nel tempo e la crescita della produttività settoriale, riscontrando una forte correlazione positiva nel lungo periodo. In un modello econometrico OLS, ad esempio, l'autore identifica le regioni con livelli settoriali più avanzati di sviluppo dell'IA come quelle con livelli superiori di produttività totale dei fattori (TFP), corroborando così l'idea che l'IA rappresenti un input universale di efficienza.

L'aumento della produttività derivante da tali sviluppi dell'IA si manifesta in due modi: l'incremento dell'automazione nelle attività routinarie libera risorse umane da mansioni a basso valore aggiunto, consentendo loro di essere riallocate altrove, mentre il miglioramento del processo decisionale deriva da una conoscenza ampia e in tempo reale. Come osservato da Greci (2023), la differenza tra l'IA e le precedenti innovazioni tecnologiche risiede nella sua capacità di "imparare a imparare": l'IA non si limita a sostituire il lavoro fisico svolto dall'uomo, ma lo integra e lo potenzia in modi finora impossibili, valorizzando tali sforzi attraverso un'elaborazione personalizzata e intensiva.

Inoltre, l'IA genera una riallocazione disomogenea delle risorse tra i settori: quelli ad alta intensità di conoscenza e tecnologia riportano una maggiore utilità derivante dagli sviluppi innovativi dell'IA (ICT, finanza, servizi professionali), mentre i settori a basso contenuto

tecnologico registrano aumenti più contenuti. Ciò riflette differenze nei livelli di esposizione, nella capacità di assorbimento tecnologico e nella qualità del capitale umano (Mazzotta, 2024).

Tuttavia, nonostante i benefici derivanti dagli sviluppi dell'IA evidenziati a livello aggregato, tali guadagni tendono a manifestarsi con un certo ritardo, a causa dei costi di allineamento, delle ristrutturazioni organizzative e delle specificità del lavoro: si tratta del paradosso della produttività, discusso nella sezione successiva.

#### **1.2.4 Il paradosso della produttività e i ritardi di adozione**

Brynjolfsson, Rock e Syverson (2017) collegano il “paradosso della produttività dell'IA” a una prospettiva tipica dell'economia internazionale. Essi sottolineano che l'intelligenza artificiale non può essere immediatamente ottimizzata dal punto di vista operativo per generare benefici. Difatti, l'IA rappresenta inizialmente un investimento costoso, poiché la sua implementazione richiede cambiamenti organizzativi preliminari, investimenti in capitale umano altamente qualificato e un adeguato livello di competenze tecnologiche. Pertanto, in termini economici, l'IA è un investimento che produce ritorni ritardati sulla produttività: nella fase iniziale i costi sono elevati e l'efficacia è ridotta a causa delle difficoltà di implementazione e di integrazione dei sistemi. Solo dopo un certo periodo, quando un'organizzazione riesce ad adattare in modo efficace i cambiamenti introdotti e a integrarli con l'IA, emergono i benefici in termini di produttività.

Questi risultati si allineano anche all'analisi di Mazzotta (2024) nel contesto europeo, secondo cui i paesi con livelli più elevati di digitalizzazione e politiche orientate allo sviluppo delle competenze (come Germania e Paesi Bassi) riescono a cogliere prima gli effetti positivi, mentre gli Stati membri caratterizzati da deficit di competenze rimangono indietro.

Da questa prospettiva, il vantaggio competitivo non deriva semplicemente dalla disponibilità di una nuova tecnologia, ma dalla capacità di riadattare processi e competenze. Tale questione verrà approfondita nella prossima sezione, che affronta le valutazioni data-informed del Capitolo 1.3.

### **1.2.5 L'IA come fonte di vantaggio competitivo**

Inoltre, in termini di produttività, l'IA è fonte di un vantaggio competitivo sostenibile. Secondo Greci (2023), l'apprendimento automatico e le tecniche di big data offrono alle aziende l'opportunità di acquisire conoscenze predittive sugli interessi dei consumatori e sulle fluttuazioni del mercato, facilitando lo sviluppo di iniziative più articolate. In questo modo, promuove l'efficienza allocativa e la capacità di innovazione anticipatoria. Le fonti di vantaggio competitivo direttamente collegate all'IA includono: efficienza operativa (riduzione dei costi di produzione e miglioramento della qualità per output attraverso l'automazione intelligente), personalizzazione (beni e servizi su misura per le esigenze individuali dei consumatori) e apprendimento organizzativo (la capacità di acquisire conoscenze nel tempo e adattare infinitamente algoritmi e analisi predittive per il processo decisionale). Brynjolfsson et al. (2017) affermano che l'IA consente ai dati di produrre azioni: le macchine di apprendimento diventano un circolo virtuoso positivo che, con i dati giusti, produce vantaggi quasi impossibili da ottenere per altri.

### **1.2.6 Politiche economiche e prospettive di sviluppo**

L'integrazione dell'IA solleva sfide di policy. L'OECD (2022) ritiene che lo sviluppo decentralizzato dell'IA debba essere promosso attraverso la sovvenzione della riqualificazione digitale, dei progetti di ricerca e sviluppo e delle iniziative di open data, al fine di evitare una distribuzione ineguale dei benefici e promuovere un aumento generalizzato della produttività.

Per quanto riguarda Mazzotta (2024), l'approccio è più incentrato sull'uomo. La tecnologia non sostituisce le capacità umane, ma le potenzia.

Laddove non è possibile fidarsi dell'efficacia e della trasparenza dei sistemi di IA, nessun investimento pubblico o privato si tradurrà in un'implementazione effettiva. Si veda il Capitolo 2 per il nesso tra fiducia e adozione dell'IA.

## Riferimenti bibliografici – Capitolo 1.2 (APA 7<sup>a</sup> ed.)

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics* (NBER Working Paper No. 24001). National Bureau of Economic Research. <https://doi.org/10.3386/w24001>

Greci, G. M. (2023). *Effetti dell'adozione dell'intelligenza artificiale sulle attività d'impresa*.

Mazzotta, A. (2024). *L'impatto dell'intelligenza artificiale sulla produttività* [Tesi di laurea, LUISS Guido Carli].

Corrado, C., Haskel, J., Iommi, M., & Jona-Lasinio, C. (2022). *The value of data in digital-based business models: Measurement and economic policy implications* (OECD Economics Department Working Papers No. 1723). OECD Publishing. <https://doi.org/10.1787/d960a10c-en>

## 1.3 – Strategie di mercato *data-driven*

### 1.3.1 Dalla personalizzazione alla competizione algoritmica

Negli ultimi vent'anni, la personalizzazione si è evoluta da miraggio tecnologico a infrastruttura competitiva. Per le aziende basate sui dati, l'apprendimento dai dati non è più un'attività secondaria, ma la ragione fondamentale alla base dell'intera sequenza di creazione di valore, dalla raccolta di segnali comportamentali (clickstream, tempi di permanenza, interazioni transazionali) alla produzione di previsioni individualizzate su bisogni, desideri e propensione all'acquisto (Roy & Dutta, 2022; Alfaifi, 2024). Pertanto, i sistemi di raccomandazione (Recommender Systems, RS) fungono da interfaccia di abbinamento tra varietà dell'offerta e eterogeneità della domanda, riducendo al minimo le spese di ricerca e alleviando il sovraccarico di informazioni e operando contemporaneamente internalizzando metriche di business quali conversione, retention e lifetime value (Roy & Dutta, 2022, pp. 2-4; Alfaifi, 2024).

Questo capitolo sostiene il posizionamento di mercato basato sui dati attraverso il quale le aziende sfruttano gli RS e l'analisi predittiva per ottenere un vantaggio competitivo

difendibile. La discussione è organizzata in quattro parti: (i) scelte architettoniche e pattern progettuali dei RS; (ii) progettazione delle feature e delle metriche come leve strategiche; (iii) iper-personalizzazione come strategia di differenziazione dinamica; e (iv) obiettivi economici come parte dei modelli (orientati al valore e orientati al profitto) fino alla valutazione sperimentale (A/B testing) e al disegno organizzativo orientato al data-driven decision-making (DDD).

### 1.3.2 Architetture: scegliere ciò che serve (e solo ciò che serve)

In teoria, esistono molte famiglie di RS (basate sui contenuti, collaborative, ibridi). In pratica, la soluzione più adatta è relativa:

- Laddove gli elementi possono essere descritti (testo, immagini) e il catalogo cambia continuamente, le RS basate sui contenuti aiutano cold start degli item.
- Laddove l'interazione utente-prodotto è frequente, il collaborative filtering rivela strutture implicite e collegamenti sorprendenti.
- In realtà, le soluzioni ibride (spesso potenziate da deep learning) dominano poiché si basano su modelli contestuali/sequenziali che aiutano a determinare quando e perché un suggerimento è valido (Roy & Dutta, 2022; Alfaifi, 2024).

La regola empirica manageriale è quella di smettere di soffrire di "ansia da SOTA": non si tratta dell'algorithm migliore, ma piuttosto di quello più adatto ai vincoli (spiegabilità in domini regolati, latenza in tempo reale, esigenze di formazione, integrazione IT) e agli obiettivi (acquisizione, cross-selling, fidelizzazione). L'architettura adeguata consente tempi di apprendimento più rapidi, una scalabilità appropriata e una più semplice aderenza ai requisiti normativi, tutti e tre legati al valore operativo (Alfaifi, 2024).

### 1.3.3 Dati e metriche: ciò che misuri è ciò che muovi

Nelle strategie *data-driven*, l'architettura dei dati è di per sé una decisione strategica. Senza i segnali appropriati (feedback loop espliciti/impliciti, consapevolezza contestuale,

metadati affidabili), il modello "ottimizza al buio". E, cosa ancora peggiore, ottimizza solo le metriche che gli vengono fornite.

Le metriche classiche (precision@k, nDCG) mostrano la rilevanza tecnica; le metriche esperienziali (diversità, novità, serendipità, copertura) informano sulla qualità del percorso utente; le metriche economiche (margine atteso, LTV, propensione al churn) informano sul valore reale (Roy & Dutta, 2022; De Biasio, 2024).

L'obiettivo è rilevanza  $\times$  esperienza  $\times$  economia. Se questi vengono misurati esclusivamente sulla base dei clic e delle conversioni, allora l'algoritmo si concentra sul breve termine (quelli più comuni esposti più frequentemente hanno una saturazione più elevata). Se imponi vincoli e obiettivi diversi, ottieni un elenco più sostenibile: qualcosa di altamente performante ma senza il rischio di ridurre la varietà o di suscitare sfiducia (Alfaifi, 2024; De Biasio, 2024).

#### **1.3.4 Hyper-personalization: dall'algoritmo alla capacità organizzativa**

La letteratura manageriale più recente identifica l'iper-personalizzazione come l'erogazione di micro-esperienze contestuali in tempo reale, abilitate dall'uso di dati di prima parte, indicatori in-session e algoritmi che prevedono percorsi cross-channel e cross-customer. Tuttavia, sebbene l'articolo di Hao (2023) non sia stato sottoposto a revisione paritaria, esso sfida i professionisti con una proposta strategica secondo cui l'iper-personalizzazione non è un algoritmo, ma una capacità organizzativa che collega raccomandazioni, offerte, prezzi, promozioni e creatività lungo tutto il customer journey e attraverso tutti i canali. Dal punto di vista economico, essa consente (i) estrazione di surplus tramite una più fine corrispondenza delle preferenze, (ii) costi di ricerca inferiori e (iii) fidelizzazione attraverso esperienze percepite come su misura.

Da un punto di vista tecnico, la spinta verso l'iper-personalizzazione si fonda su metodologie sequenziali (a livello di sessione), sensibili al contesto e multimodali sempre più interconnesse, con meccanismi di re-ranking vincolato (esposizione ai fornitori, diversità, rotazioni) e un'ottimizzazione multi-obiettivo (rilevanza  $\times$  qualità  $\times$  esperienza utente). Le soluzioni predittive funzionano meglio quando risolvono il dilemma esplorare/sfruttare e applicano politiche adattive per evitare il deterioramento e l'affaticamento pubblicitario (Roy & Dutta, 2022; Alfaifi, 2024).

L'iper-personalizzazione sarà sostenibile a lungo termine solo quando l'ottimo privato (KPI d'impresa) sarà allineato con l'ottimo dell'utente (utilità percepita): il ponte tra i due è rappresentato dal disegno value-aware dei modelli.

### 1.3.5 Dalla *accuracy* al valore: raccomandazioni *value-aware* e *profit-aware*

L'assunto secondo cui *maggior accuratezza*  $\Rightarrow$  *maggior valore d'impresa* è contingente; una raccomandazione fortemente rilevante, per un utente, potrebbe non essere così profittevole (cannibalizzando le vendite più redditizie con offerte a basso margine); il deterioramento del benessere a lungo termine (ad esempio saturazione, dipendenza dai canali, affidamento su contenuti controversi) porta anche a una diminuzione della fidelizzazione.

In questo contesto si colloca l'analisi di De Biasio (2024) sulla sottoclasse delle raccomandazioni value-aware ed economic/profit-aware, in cui prezzi, costi, margini, elasticità e valore a lungo termine vengono incorporati nella funzione obiettivo sia in fase di in-processing (adattando la loss e i criteri di ranking) sia in fase di post-processing (riordinamento della lista top-k), al fine di ottimizzare il valore economico previsto, dove la rilevanza prevista riceve un trattamento equo. I risultati mostrano raccomandazioni economicamente più profittevoli con livelli simili (o piccole perdite nel valore percepito) e vantaggi per l'elaborazione in-processing rispetto all'elaborazione post-processing (De Biasio, 2024, Parte II, capp. 5.2–5.5).

De Biasio rileva inoltre che viene aggiunto un elemento sociale e comportamentale maggiore laddove l'equità e l'esposizione possono diversificare ma non compromettere drasticamente le raccomandazioni aziendali, purché siano trattate come un vincolo o obiettivo in termini di perdita (De Biasio, 2024, §2.3.5; §6). Ciò significa che la prospettiva basata sui dati integra un approccio multicriterio (profitto, rilevanza, equità, sicurezza) e multi-orizzonte (breve/lungo termine).

Le imprese *frontier* cambiano l'approccio alle raccomandazioni da una "macchina dei clic" a una macchina dei margini (e relazionale) attraverso una co-progettazione di modelli, metriche e incentivi.

### 1.3.6 Decisioni *data-driven* e performance d'impresa

L'evidenza empirica sostiene il legame tra pratiche di data-driven decision-making (DDD) e performance. Gli autori di "Strength in Numbers" riferiscono che, sulla base di un panel di 179 grandi aziende quotate, quelle che implementano un processo decisionale basato sui dati hanno livelli di produttività e output superiori del 5-6% rispetto al controfattuale; tali impatti sono persistenti nel tempo sull'efficienza degli asset, sul ROE e sulla capitalizzazione di mercato; le stime IV suggeriscono un effetto causale, e non un problema di selezione inversa (Brynjolfsson, Hitt & Kim, SSRN). Esiste una controparte teorica a questa scoperta: l'azienda è un input della funzione di produzione dell'IA; mentre la tecnologia rende disponibili i segnali, ma routine, gli incentivi e le capacità analitiche ne determinano il rendimento marginale.

L'IA come fattore di produzione (Capitolo 1.2) dispiega il suo pieno potenziale solo in assetti DDD-compliant: diritti decisionali decentralizzati, sistemi informativi verticali, cultura sperimentale (test-and-learn) e A/B testing a ciclo stretto (SSRN; Lizzardi, 2023).

### 1.3.7 Dalla teoria alla pratica: applicazioni *business* e valutazione sperimentale

Collegandoci con riferimenti all'uso della RS in servizi come lo streaming, i social media e il turismo, nonché nella sanità, nell'istruzione, nell'e-commerce e nel mercato del lavoro, Lizzardi (2023) ci mostra sia i benefici tangibili (ricavi, cross selling, fidelizzazione) che quelli intangibili (esperienza utente, lealtà, reputazione) di applicazioni RS molto diverse tra loro, non solo dimostrando ma anche sottolineando che la qualità dell'abbinamento e alcuni parametri diversi dalla sola accuratezza sono fattori chiave nell'impatto economico (Lizzardi, 2023).

Nella valutazione, la letteratura distingue tra offline (ad esempio, split hold-out, split temporale, valutazione controfattuale) e online (ad esempio, test A/B e interleaving). Per citarne alcune, ad esempio, Roy & Dutta (2022) e Alfaifi (2024) citano metriche di accuratezza (RMSE/MAE per *rating prediction*; precision/recall/nDCG per *top-k*) e metriche di ranking e utilità (novità, diversità, serendipity, coverage) (Alfaifi, 2024; Roy & Dutta, 2022). Da un punto

di vista value-aware, De Biasio (2024) analizza le metriche economiche (margine atteso, LTV, profit@k) e di benessere sociale (diversità/fairness) (De Biasio, 2024).

*What you measure is what you move*: senza KPI composite (rilevanza × valore × esperienza × equità), i sistemi ottimizzano surrogati (CTR, tempo) con possibili danni collaterali su *trust* e lungo periodo.

### 1.3.8 Sfide tecniche e organizzative: *cold start*, sparsità, *echo chambers* e *governance*

In primo luogo, la sparsità e il cold start: questi aspetti possono essere affrontati con una progettazione adeguata dei metadati, utilizzando embeddings e modelli session-based per i nuovi utenti. In secondo luogo, i loop di popolarità e ed eco-chambers: questi dovrebbero essere mitigati da vincoli di diversità e coverage, exploration controllata o indicatori di serendipity. In terzo luogo, il bias e l'opacità: con schede modello e spiegazioni comprensibili (e persino controlli di screening per entrambi), è possibile ridurre al minimo il bias e l'opacità. Questi rischi, come discusso nel Capitolo 1.5, possono tradursi in costi reputazionali, regolatori e competitivi. Tutte le misure citate in precedenza possono sembrare procedure pesanti o formalità burocratiche, ma in realtà, sono di una garanzia di continuità del valore (Roy & Dutta, 2022; Alfaifi, 2024;). A livello organizzativo, ciò comporta solide piattaforme di dati su cui i modelli vengono addestrati e versionati (compresi i relativi processi MLOps, come audit trail e rollback), ruoli dedicati per sostenere i vantaggi dell'RS nel tempo.

### 1.3.9 Strategie *data-driven*: quadro operativo

Allineare architettura e obiettivi scegliendo CBF/CF/ibridi in base ai dati di input, alle esigenze di latenza/spiegabilità e ai KPI finali. (Alfaifi, 2024; Roy & Dutta, 2022).

Progettare il segnale prima del modello controllando ciò che si registra (strumentazione, metadati, contesto, coerenza della registrazione e KPI). (Alfaifi, 2024; Roy & Dutta, 2022).

Ottimizzare per criteri multipli: combinando rilevanza, esperienza ed economia (margine/LTV) con vincoli di equità/diversità. Applicare tali criteri meglio nella loss o

attraverso un nuovo posizionamento, se fattibile. (Alfaifi, 2024; De Biasio, 2024; Roy & Dutta, 2022).

Testare sempre: offline per selezionare, online per decidere (test A/B, interleaving, test controfattuali). Prendere nota delle ipotesi e valutare gli effetti collaterali (saturazione, affaticamento pubblicitario, fiducia). (Alfaifi, 2024; Lizzardi, 2023; Roy & Dutta, 2022).

Governa in esercizio: monitoraggio post-implementazione, guard-rails (cioè anti-saturazione), risposta agli incidenti/rollback e revisione periodica delle metriche, mentre gli incentivi interni e le responsabilità devono essere allineati (DDD). (Brynjolfsson, Hitt, & Kim; Lizzardi, 2023).

### 1.3.10 Dalla strategia alle implicazioni competitive e di fiducia

Le strategie di mercato data-driven convertono l'AI da "motore di predizione" a motore di decisione. Il vantaggio competitivo che si viene a creare non deriva da un algoritmo specifico, ma piuttosto dall'uso coerente di: architetture adeguate, segnali adeguati progettati in modo efficiente, obiettivi economici multicriterio, sperimentazione continua e governance in esercizio.

Come illustrato nel capitolo 1.4, queste scelte hanno conseguenze competitive (concentrazione informativa, potere di raccomandazione). Mentre, come illustrato nel capitolo 2, l'efficacia di queste scelte dipende dalla fiducia: se non viene percepita l'equità, trasparenza e del controllo, l'adozione ristagna e non emerge alcun valore.

#### **Riferimenti (citati nel testo, stile APA 7<sup>a</sup> ed.)**

- Alfaifi, Y. H. (2024). *Recommender Systems Applications: Data Sources, Features, and Challenges*. *Information*, 15(660). Open access.
- Brynjolfsson, E., Hitt, L., & Kim, H. (SSRN). *Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?* (analisi su 179 imprese; +5–6% produttività).
- De Biasio, A. (2024). *Value-Aware Recommendation: Algorithms and Applications*. PhD Thesis, Università di Padova (parti II; capp. 3–6; profit-aware; fairness/diversity).

- Lizzardi, M. J. (2023). *Recommender Systems and Business Applications*. Università Politecnica delle Marche (cap. 4–5, applicazioni e valore).
- Roy, D., & Dutta, M. (2022). *A systematic review and research perspective on recommender systems*. *Journal of Big Data*, 9(59).

## 1.4 – Implicazioni economiche per la concorrenza e il valore del consumatore

### 1.4.1 Dalla centralità dei dati alla trasformazione dei mercati

Negli ultimi anni, i dati digitali e gli algoritmi di intelligenza artificiale si sono affermati come nuove infrastrutture economiche dei mercati globali. Nei capitoli precedenti si è analizzato come i dati rappresentino una forma immateriale di capitale e l'IA un fattore di produzione che contribuisce ad amplificare la produttività e all'efficienza (Mazzotta, 2024; Brynjolfsson, Rock & Syverson, 2019). In questo capitolo si analizzano gli effetti complessivi di questa trasformazione sulla concorrenza e sul benessere collettivo.

Per Coyle e Li (2021), i dati sono più di un semplice input per la produzione, ma un bene strategico che definisce nuove barriere all'ingresso, concentra il potere economico e ridefinisce le forme della concorrenza. Infatti, l'OECD (2022) sostiene che l'economia digitale tende a sviluppare strutture di mercato oligopolistiche. In queste strutture di mercato un piccolo numero di attori internazionali con risorse avanzate per la raccolta, l'analisi e l'implementazione dei dati supera di gran lunga chiunque altro detenendo un vantaggio competitivo difficilmente replicabile.

Di conseguenza, la concorrenza tradizionale basata sul prezzo lascia spazio a una competizione fondata su informazione, predizione e personalizzazione.

Le aziende che dominano i flussi informativi non solo funzionano in modo più efficiente, ma modellano le preferenze dei consumatori attraverso sistemi di raccomandazione e algoritmi predittivi (Zhang, 2018). Pertanto, il nuovo mezzo di concorrenza non risiede nell'efficienza produttiva, ma nella capacità di influenzare le scelte di consumo. Questo cambiamento incide in modo significativo sul welfare e su come il valore viene distribuito.

Questo capitolo esamina quindi i meccanismi attraverso cui l'economia dei dati modifica l'equilibrio tra concorrenza e innovazione rispetto al benessere dei consumatori (e decostruisce il ruolo delle politiche pubbliche nel correggere questo squilibrio del benessere sociale).

#### **1.4.2 Le nuove dinamiche della concorrenza nell'economia dei dati**

L'economia digitale si distingue per la presenza di forti effetti di rete, di rendimenti di scala crescenti e la concentrazione del potere informativo.

Come spiegano Coyle e Li (2021), in un'economia basata sui dati, “il vincitore tende a prendere tutto”. Più un'impresa raccoglie dati, più gli algoritmi diventano più precisi ed efficaci, e la maggior precisione ed efficacia di questi algoritmi attirano ancora più utenti che sono spinti a generare nuovi dati. Questo è un processo di apprendimento cumulativo che favorisce un circolo virtuoso, il quale genera un vantaggio competitivo che si autoalimenta.

Secondo l'OECD (2022), si crea un ciclo di apprendimento cumulativo. In questo ciclo i pochi grandi operatori digitali sono chiamati Big Tech.

Questi operatori consolidano il loro dominio grazie a tre fattori principali:

1. Accumulo di dati proprietari, che riduce la contendibilità del mercato;
2. Integrazione verticale tra piattaforme, motori di ricerca e servizi complementari;
3. Maggiore capacità predittiva, che rende i servizi più personalizzati e riduce l'elasticità della domanda.

Queste caratteristiche producono un paradosso della concorrenza algoritmica, poiché la stessa tecnologia che aumenta l'efficienza del mercato può ridurre la concorrenzialità con la concentrazione del potere economico. La disponibilità di grandi volumi di dati non assicura un accesso uguale per tutte le imprese. Al contrario, favorisce gli incumbents e barriere cognitive e tecnologiche all'ingresso (OECD, 2022).

Inoltre, Coyle e Li (2021) affermano che questa concentrazione di informazioni non crea un problema di dimensione economica, ma anche un'asimmetria informativa strutturale: le piattaforme fanno molte cose sui consumatori, mentre i consumatori fanno poco sui mezzi attraverso i quali le piattaforme elaborano le informazioni che influenzano il processo

decisionale. Tale squilibrio riduce la trasparenza del mercato e diminuisce la capacità dei consumatori di valutare le alternative e i prezzi.

Inoltre, la concorrenza si sposta progressivamente dal livello dei beni a quello degli ecosistemi: le aziende competono attraverso reti integrate di servizi (es. cloud, e-commerce, pubblicità, media) in cui la raccolta e l'applicazione dei dati fungono da collante strategico.

Si sviluppano così nuovi mercati multilaterali interdipendenti in cui il funzionamento di una piattaforma influisce sul benessere degli attori esterni attraverso le esternalità di rete (OECD, 2022).

Come verrà ulteriormente sviluppato nel capitolo 1.5, tali squilibri informativi e di potere sono motivo di preoccupazione dal punto di vista etico e normativo: laddove esiste una concentrazione algoritmica, questa può tradursi in uno sfruttamento dei dati personali o in tentativi di discriminazione algoritmica con impatti negativi sul benessere e sulla fiducia dei consumatori.

### **1.4.3 I sistemi di raccomandazione e il valore del consumatore**

Un ambito emblematico in cui si manifestano queste dinamiche sopra descritte è tuttavia quello dei sistemi di raccomandazione (RS), che rappresentano oggi uno dei mezzi più diffusi attraverso cui le piattaforme digitali generano ed estraggono valore economico. Tuttavia, secondo Zhang (2018), gli RS non si limitano a suggerire prodotti, ma fungono piuttosto da intermediari cognitivi tra domanda e offerta, influenzano attivamente le preferenze e i comportamenti degli utenti.

Da un punto di vista economico, i sistemi di raccomandazione influiscono sul benessere in due direzioni:

1. L'impatto sull'efficienza allocativa, dove i costi di ricerca sono ridotti, la corrispondenza tra preferenze e prodotto aumenta e l'incremento della soddisfazione del consumatore;

2. L'impatto sul potere di mercato, dove le imprese orientano le scelte degli utenti e estraggono il surplus di informazioni, aumentando al contempo il loro vantaggio competitivo.

Zhang (2018) osserva che, all'interno di un modello teorico di welfare algoritmico, questi impatti sul benessere sono positivi o negativi a seconda della trasparenza e della varietà dei sistemi di raccomandazione. Se gli algoritmi sono costruiti per favorire una maggiore varietà di opzioni, il welfare collettivo aumenta poiché le persone hanno un maggiore accesso a prodotti più eterogenei e idee alternative. Se i sistemi sono costruiti per la massimizzazione del click-through o del profitto immediato, gli algoritmi sono orientati verso una diminuzione della varietà: si rinforzano le preferenze esistenti e si creano bolle di consumo (Zhang 2018).

L'OECD (2022) conferma che la personalizzazione algoritmica rappresenta un'arma a doppio taglio, se da un lato aumenta il valore percepito dal consumatore, dall'altro può restringere la concorrenza riducendo la visibilità dei concorrenti minori. Pertanto, le piattaforme che controllano i propri sistemi di raccomandazione controllano ciò che gli utenti vogliono e non vogliono, diventando regolatori privati del mercato digitale.

Inoltre, come esaminato nel capitolo 1.3, la personalizzazione basata sui dati è attualmente una leva cruciale della concorrenza e, senza un'adeguata valutazione del benessere dell'impatto sociale, ciò è pericoloso. La distribuzione del valore generato dai sistemi di raccomandazione non è neutrale: la maggior parte del surplus di informazioni rimane all'interno delle piattaforme, mentre all'utente viene concesso solo un beneficio limitato in termini di convenienza e varietà.

Pertanto, Zhang (2018) afferma che il benessere collettivo può essere massimizzato solo attraverso la governance pubblica dei criteri di raccomandazione, che deve stabilire un equilibrio tra la **ottimizzazione** algoritmica e la diversità delle informazioni.

In definitiva, i sistemi di raccomandazione presentano la tensione estrema nell'economia algoritmica, in quanto, pur facilitando l'efficienza e la soddisfazione degli utenti, accentuano anche la concentrazione del potere informativo.

#### 1.4.4 Implicazioni di welfare e redistribuzione del valore

Man mano che le interazioni sul mercato digitale diventano sempre più basate sui dati e determinate da algoritmi, sorge una domanda cruciale: chi beneficia della creazione di valore dell'IA e dei sistemi di raccomandazione?

Secondo Coyle e Li (2021), sebbene l'economia dei dati sia un fattore determinante per l'aumento del PIL globale, la distribuzione del valore delle informazioni è estremamente diseguale. La concentrazione della proprietà dei dati nelle mani di pochi grandi operatori digitali si traduce in una asimmetria informativa nella cattura del surplus, poiché la maggior parte dei guadagni derivanti dalla monetizzazione dei dati e dalle attività di machine learning resta interna alle piattaforme, mentre i consumatori ottengono solo una frazione di quel valore sotto forma di servizi gratuiti o personalizzati.

Infatti, secondo l'OECD (2022), il valore sociale e il valore privato dei dati mostrano una disparità a causa di questo fenomeno. Quando i dati vengono utilizzati, dovrebbero creare esternalità positive in tutta l'economia, poiché migliorano l'efficienza dei mercati e la qualità del processo decisionale pubblico e privato. Tuttavia, senza sforzi di redistribuzione o normative specifiche in materia di condivisione dei dati, tali benefici vengono assorbiti solo da alcuni agenti economici.

Da un punto di vista microeconomico, Zhang (2018) chiarisce la situazione: nei sistemi di raccomandazione il surplus delle informazioni viene catturato principalmente attraverso le piattaforme che controllano il matching tra utenti e prodotti. Tuttavia, la redistribuzione del valore dipende dal design dell'algoritmo: se l'algoritmo massimizza la soddisfazione individualizzata rispetto alla diversità collettiva, potrebbe in ultima analisi ridurre il benessere complessivo.

Inoltre, il valore cognitivo si sovrappone al valore economico, poiché determinando ciò che verrà esposto, i sistemi di raccomandazione non solo influenzano il potenziale di consumo, ma anche ciò che plasma le opinioni attraverso la pluralità informativa. Da ciò deriva una forma di paternalismo algoritmico, in cui il consumatore delega parte della propria autonomia decisionale a un sistema opaco, riducendo la propria capacità di autodeterminazione (OECD, 2022).

Come osserva Coyle & Li (2021), non si tratta solo di una questione di valore economico, ma anche istituzionale; le misure del benessere sociale (PIL o produttività del

lavoro) non tengono conto del beneficio sociale dei dati o delle conseguenze distributive dell'economia algoritmica. In definitiva, da qui emerge la necessità di definire nuovi indicatori di benessere digitale in relazione agli standard di esperienza utente, alla diversità e alla fiducia nei sistemi algoritmici. Ciò verrà ripreso nel capitolo 2, dove la fiducia nella tecnologia intelligente verrà analizzata come una variabile nella valutazione della fattibilità economica e sociale dell'intelligenza artificiale.

#### **1.4.5 Politiche economiche e regolazione della concorrenza digitale**

La regolamentazione della concorrenza nell'economia dei dati è una delle maggiori sfide per gli attori pubblici. Secondo l'OECD (2022), i sistemi antitrust tradizionali e la politica della concorrenza non sono sufficientemente adeguati alle particolari caratteristiche dei mercati digitali: rendimenti crescenti, effetti di rete, integrazione verticale e asimmetrie informative.

Pertanto, è duplice l'obiettivo che deve essere perseguito dalle autorità di regolazione:

In primo luogo deve promuovere l'innovazione e l'efficienza senza frenare la dinamica tecnologica ed evitare che il potere economico si concentri eccessivamente, poiché tale concentrazione potrebbe ridurre la concorrenza e compromettere allo stesso tempo il benessere collettivo.

Vengono proposte in tal senso diverse linee d'azione, dall'OECD (2022) e dalla Commissione Europea:

- Portabilità/interoperabilità dei dati: la possibilità per i consumatori di trasferire i dati tra piattaforme favorisce la concorrenza e riduce le barriere all'ingresso.

- Obblighi di condivisione dei dati per le imprese dominanti: imporre alle piattaforme con posizione di mercato significativa di condividere in modo sicuro una parte dei loro dataset con attori più piccoli e pubblici.

- Trasparenza algoritmica: richiedere la spiegabilità delle decisioni automatizzate mentre, al fine di garantire il controllo sociale sugli effetti dell'IA, i criteri di raccomandazione dovrebbero essere pubblicati.

- Tutela dei consumatori digitali: le raccomandazioni personalizzate non possono trasformarsi in differenziazione dei prezzi o nello sfruttamento delle vulnerabilità cognitive degli utenti.

Inoltre, secondo Coyle e Li (2021), la politica della concorrenza nell'economia dei dati dovrebbe evolvere verso un modello proattivo e adattivo, che consideri il valore dei dati come infrastruttura pubblica.

L'obiettivo non dovrebbe essere solo quello di creare spazi di condivisione dei dati, ma anche di istituzioni digitali, per regolamentare e gestire in modo efficace e sostenibile l'uso dei dati.

Un elemento centrale riguarda la costruzione di un mercato dei dati regolamentato nel quale il diritto alla privacy e la libertà economica coesistano. L'OECD (2022) sostiene che l'obiettivo, oltre a limitare gli abusi di una posizione dominante, è garantire una contendibilità dinamica: i nuovi operatori dovrebbero entrare e competere sulla base dei loro algoritmi invece che semplicemente sulla quantità di informazioni di cui dispongono.

Questa prospettiva impone di rivedere come si misura la concorrenza. Infatti, gli indicatori tradizionali, come la quota di mercato o il prezzo di vendita, risultano inadeguati nei contesti digitali, dato che molto spesso la moneta di scambio non è il denaro ma l'informazione. Dato per cui bisogna valutare la concorrenza sulla base dell'accessibilità, della trasparenza algoritmica e della pluralità delle informazioni.

La politica pubblica dovrà integrare strumenti pratici di alfabetizzazione algoritmica e educazione digitale affinché i consumatori comprendano al meglio i sistemi che influenzano le loro scelte economiche. Al fine che una società partecipi in modo equo all'economia dei dati, la stessa deve essere informata e consapevole.

#### **1.4.6 Verso un equilibrio tra efficienza, concorrenza e fiducia**

Il presente capitolo ha mostrato come l'economia dei dati e l'intelligenza artificiale ridefiniscano i confini della concorrenza e la distribuzione del valore economico. I dati sono diventati una nuova forma di capitale e al contempo una fonte di potere economico e politico. Tale potere ha la capacità di influenzare le decisioni dei consumatori e di determinare nuovi squilibri di welfare.

Tre diverse analisi (Coyle e Li (2021), OECD (2022), Zhang (2018)) mettono in luce una tensione strutturale tra efficienza allocativa e concentrazione del vantaggio informativo in mano a pochi attori. Così, mentre l'IA e i sistemi di raccomandazione aumentano la produttività

e migliorano l'esperienza dei consumatori, allo stesso tempo riducono la concorrenza e ridistribuiscono il valore nelle mani delle piattaforme dominanti.

Pertanto, la risposta normativa a tali sfide deve promuovere un nuovo approccio normativo che riconosca che i dati hanno una doppia (o tripla) natura: risorsa privata, infrastruttura collettiva, bene relazionale. La condivisione dei dati, la responsabilità algoritmica e la partecipazione dei consumatori come soggetto attivo nella creazione del valore informativo dovrebbero essere incoraggiate attraverso una risposta normativa.

In prospettiva, la sostenibilità dell'economia digitale dipenderà dalla capacità delle istituzioni di bilanciare tre obiettivi fondamentali:

1. Efficienza, ossia il miglior utilizzo delle risorse informative.
2. Equità competitiva, cioè l'accesso equivalente ai dati e ai mercati.
3. Fiducia, come condizione necessaria per la cooperazione tra imprese, consumatori e autorità pubbliche.

Come verrà discusso nel capitolo 2, la fiducia diventerà in realtà la variabile più importante per comprendere le relazioni tra esseri umani e sistemi intelligenti, per assicurare che l'IA sia anche uno strumento di benessere collettivo e coesione sociale oltre che di efficienza economica.

### **Riferimenti bibliografici (APA 7<sup>a</sup> edizione)**

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). *Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics*. NBER Working Paper No. 24001.

Coyle, D., & Li, W. (2021). *The Data Economy: Market Size and Global Trade*. Bennett Institute for Public Policy, University of Cambridge.

Mazzotta, A. (2024). *L'impatto dell'Intelligenza Artificiale sulla produttività*. LUISS Guido Carli, Tesi di Laurea Magistrale.

OECD. (2022). *The Value of Data in Digital Business Models*. OECD Publishing.

Zhang, J. (2018). *Welfare Properties of Recommender Systems*. Carnegie Mellon University, Heinz College Research Paper.

## **1.5 – Rischi e limiti dell'economia algoritmica**

### **1.5.1 Dal progresso tecnico alla vulnerabilità sistemica**

L'ultimo decennio è stato caratterizzato da una trasformazione digitale, la cui espressione più avanzata è rappresentata dall'intelligenza artificiale. Tuttavia, l'OECD (2024) sottolinea che il progresso tecnologico non è mai neutrale, poiché ogni innovazione porta con sé nuovi rischi di natura economica, sociale e istituzionale.

A tal fine, l'automazione algoritmica incoraggia una dipendenza sempre maggiore da operazioni sempre più complesse e spesso oscure, che migliorano i processi decisionali, ma aumentano il rischio di vulnerabilità sistemiche difficili da controllare.

Con il concetto di *vulnerabilità sistemica* si intende la possibilità in cui un errore tecnico, informativo o decisionale si propaghi rapidamente attraverso l'interconnessione dei sistemi digitali, producendo così effetti economici amplificati.

Secondo il World Economic Forum (2023-2024), siamo giunti ad una fase nell'economia mondiale in cui il rischio tecnologico è un rischio macroeconomico che influenza le operazioni di mercato e la percezione di fiducia da parte dei consumatori.

Ortiz (2021) ci suggerisce che le organizzazioni possono essere soggette a tre diverse forme di vulnerabilità.

Il primo tipo è tecnico: la sempre maggiore complessità dei modelli di machine learning e le difficoltà di anticiparne il comportamento in condizioni estreme.

Il secondo tipo è organizzativo: la carenza di competenze adeguate per il controllo, la verifica e la validazione dei sistemi all'interno delle imprese.

Il terzo è sistemico: nasce dalle forti interdipendenze tra infrastrutture digitali, piattaforme e catene di fornitura globali.

Si viene quindi a sviluppare un paradosso secondo cui, da un lato, l'intelligenza artificiale aumenta l'efficienza, ma dall'altro, si diventa sempre più dipendenti da strutture automatizzate e difficili da comprendere o da sostituire.

Pomo (2023) sostiene che, quando vengono prese decisioni difficili e affidate a sistemi di raccomandazione o predizione algoritmica, la delega di tali decisioni provoca un deficit di controllo umano che può tradursi in errori di valutazione, discriminazioni e perdita di responsabilità.

Pertanto, la vulnerabilità sistemica si estende oltre le organizzazioni private anche agli enti pubblici che affidano agli algoritmi processi di selezione, l'allocazione delle risorse e la valutazione delle prestazioni. Infatti, senza un quadro regolatorio adeguato, il rischio a cui si va incontro è la *tecnocrazia automatizzata* nella quale le decisioni appaiono oggettive ma riproducono logiche di mercato opache o pregiudizi impliciti.

In definitiva, l'implementazione dell'intelligenza artificiale nei processi produttivi e nel processo decisionale rappresenta la transizione da un'economia basata sul capitale fisico a un'economia basata sull'informazione automatizzata, dove la conoscenza è mediata attraverso algoritmi e la trasparenza del processo decisionale diventa più difficile da ottenere. Tale transizione ha un impatto sulla fiducia, come discusso nel capitolo 2, perché le tecnologie intelligenti devono essere credibili prima che la fattibilità economica possa essere realizzata come benessere collettivo.

### **1.5.2 Rischi economici e strutturali dell'automazione algoritmica**

Uno dei problemi più significativi dell'economia algoritmica è che concentra il potere economico e informativo nelle mani di poche piattaforme globali. L'OECD (2024) osserva che l'intelligenza artificiale e i big data rafforzano le dinamiche di lock-in competitivo. Tali imprese che hanno a disposizione ampi volumi di dati ed elevate capacità di calcolo, acquisiscono grazie ad essi un vantaggio cumulativo e autoalimentante, associato a rendimenti difficilmente replicabili.

Si tratta quindi di fenomeni di rendimenti crescenti della scala informativa, come visto nel Capitolo 1.4, ma laddove quest'ultimo era collegato alla competitività nei mercati digitali

basata sul valore, si tratta di monopolio del capitale cognitivo. Le piattaforme dominanti possono condizionare ciò che consumatori/aziende fanno, influenzare prezzi, le preferenze e persino la percezione di scelta (Căpușneanu 2025).

McKinsey (2016) ha ipotizzato anni fa che l'automazione intelligente avrebbe creato un tessuto produttivo polarizzato: coloro che potevano lavorare con l'intelligenza artificiale e la previsione avrebbero ottenuto livelli di produttività più elevati, mentre coloro che non potevano farlo ne avrebbero sofferto, lasciando dietro di sé crescenti disparità tra settori e regioni. Tale previsione si è in gran parte avverata. Secondo il WEF (2023-2024), oggi l'implementazione diffusa dell'IA genera un duplice effetto:

- riduzione dei costi marginali di produzione per le imprese leader,
- mentre le imprese che non riescono a investire in IA e infrastrutture digitali vedono ridursi sempre di più la loro capacità di competere.

La conseguenza macroeconomica è la formazione di una curva di disuguaglianza digitale (Ortiz 2021) che implica che la produttività e i profitti si verificano dove le tecnologie sono concentrate, mentre il resto dell'economia ristagna. Inoltre, l'intelligenza artificiale sostituisce non solo il lavoro manuale, ma anche parte delle funzioni cognitive intermedie, poiché il rischio occupazionale si sposta ora sulla classe media (OECD 2024). Un altro rischio strutturale è il cosiddetto lock-in algoritmico: infatti, le aziende che utilizzano sistemi di intelligenza artificiale proprietarie diventano dipendenti da coloro che creano le tecnologie necessarie e perdono flessibilità strategica. Tale esposizione alla tecnologia, se non regolamentata, può creare vulnerabilità per l'economia paragonabili a quelle delle catene di approvvigionamento fisiche, sebbene più rapide e con minore trasparenza. In terzo luogo, quando si parla di produttività e occupazione, la produttività in rapporto all'occupazione viene adeguata a causa dell'automazione algoritmica. A differenza delle rivoluzioni industriali precedenti, l'automazione algoritmica consente di produrre di più senza assumere più lavoratori, rompendo il meccanismo storico per cui il progresso tecnologico generava anche nuova occupazione (Ortiz 2021). In particolare, i processi digitali consentono di scalare la produzione senza un aumento proporzionale del personale. Efficienza significa politiche compensative relative alla formazione digitalizzata, al welfare adattivo e alla redistribuzione dei guadagni di produttività, come raccomandato dall'OECD (2024). Pertanto, oggi, qualsiasi efficienza relativa all'IA comporta un rischio di iniquità sistemica, a meno che non venga fornita una governance economica per redistribuire tutto il valore creato.

### 1.5.3 Bias, opacità e problemi di accountability nei sistemi AI

Un'altra criticità centrale dell'economia algoritmica riguarda i bias decisionali e la mancanza di trasparenza nei modelli di intelligenza artificiale. Pomo (2023) mostra come, nei sistemi di raccomandazione, le distorsioni derivanti dai dati di addestramento o dalle metriche di ottimizzazione possano produrre effetti discriminatori anche in assenza di intenzionalità.

I bias algoritmici si presentano a tre livelli:

1. Dati, quando le informazioni storiche riflettono pregiudizi sociali o economici;
2. Modello, dove l'architettura è progettata con un effetto moltiplicatore di alcune risposte correlazionali rispetto ad altre;
3. Il terzo e ultimo livello è rappresentato dalla Decisione, secondo cui un sistema produce risultati che favoriscono gruppi o comportamenti specifici, riducendo così la diversità delle scelte (Pomo 2023).

La difficoltà principale risiede nell'opacità dei processi decisionali. La maggior parte degli algoritmi è basata su reti neurali profonde e quindi agisce come *scatole nere*, producendo risultati, ma nemmeno i creatori sanno spiegare come. Questa mancanza di trasparenza solleva problemi di accountability: chi è responsabile dell'errore/danno di una decisione automatizzata?

Pertanto, secondo l'OECD (2024), la responsabilità legale relativa ai sistemi di intelligenza artificiale non ha ancora tracciato un percorso chiaro. La difficoltà di individuare un soggetto giuridico responsabile rende complessa l'applicazione delle norme su responsabilità civile e tutela dei consumatori. Allo stesso tempo, si evidenzia una ridotta fiducia nei sistemi intelligenti, dove la mancanza di trasparenza alimenta una riluttanza culturale nei confronti dell'implementazione dell'intelligenza artificiale in settori sensibili come sanità, finanza e governo. Secondo *AI Marketing: Exploring Ethical Frontiers* (2024), la mancanza di trasparenza negli algoritmi va oltre una questione tecnica e diventa un problema etico e comunicativo. La maggior parte delle persone non sa come vengono scelti i propri contenuti o le offerte personalizzate, il che porta a una manipolazione cognitiva, generando un'illusione di libertà di scelta. Pertanto, come verrà discusso nel Capitolo 2, la fiducia nelle raccomandazioni automatizzate dipende in larga misura dalla percezione di equità e controllo da parte dell'utente. Per questo motivo, trasparenza e esplicabilità non hanno solo un valore etico, ma anche un

motore economico del vantaggio competitivo. Le aziende che promuovono la spiegabilità dei propri algoritmi diventano più legittimate a livello sociale e più stabili nel lungo periodo.

#### **1.5.4 Impatti etici e sociali: privacy, manipolazione e disuguaglianze**

Oltre alle implicazioni economiche e strutturali, l'economia algoritmica pone anche questioni etiche e sociali su privacy, autonomia decisionale ed equità distributiva. *Exploring Ethical Frontiers of AI in Marketing (2024)* afferma che i servizi personalizzati basati sui dati sono intrinsecamente rischiosi e aprono la strada all'iper-sorveglianza; raccogliere e analizzare costantemente le intenzioni degli utenti attraverso un processo basato sui big data offusca il confine tra utilità ed intrusione. L'etica dell'economia algoritmica ruota attorno all'equilibrio tra personalizzazione e libertà. Gli utenti possono apprezzare gli algoritmi che prevedono ciò di cui hanno bisogno e desiderano, ma così facendo diventano oggetto di una profilazione costante e incessante. In ambito marketing, questo si traduce in una forma di nudging algoritmico dove le scelte vengono orientate attraverso suggerimenti mirati che non sono sempre così trasparenti come sembrano. La libertà di prendere decisioni esiste in superficie; ciò che esiste invece è una libertà assistita basata sul modo in cui vengono creati i sistemi digitali (Pomo, 2023).

Un altro problema riguarda la disuguaglianza informativa. Gli algoritmi che ottimizzano l'engagement tendono a privilegiare contenuti polarizzanti o economicamente vantaggiosi, creando così echo chambers e rafforzando dinamiche di esclusione sociale (OECD, 2024). Il pericolo, osserva il World Economic Forum (2023-2024), è che l'intelligenza artificiale non faccia altro che replicare e amplificare i bias preesistenti, aprendo nuove linee di frattura tra chi controlla i dati e chi li fornisce inconsapevolmente.

In termini di privacy, la sfida è garantire che gli utenti possano effettivamente dare un consenso informato. In casi normali, non comprendono la portata dell'utilizzo dei loro dati o le inferenze che vengono tratte dagli algoritmi. Come osserva Căpușneanu (2025), l'asimmetria cognitiva tra aziende e consumatori in questo ambito mina la trasparenza del mercato e limita il controllo che il consumatore può esercitare sulla propria esperienza di consumo.

La conseguenza etica di gran lunga più allarmante riguarda la dimensione relazionale della fiducia. Quando le raccomandazioni automatizzate diventano il principale punto di contatto tra le persone e il mercato, la fiducia si sposta dagli attori umani alle macchine. Questo

cambiamento implica una riconsiderazione antropologica del concetto stesso di decisione economica. Questo aspetto sarà affrontato nel Capitolo 2.

### 1.5.5 Verso una governance responsabile dell'IA

Considerati i pericoli delineati, emerge la necessità di una governance multilivello dell'IA che concili innovazione, tutela dei diritti e sostenibilità economica. L'OECD (2024) sottolinea che la regolamentazione dell'IA non può limitarsi a norme tecniche, ma dovrebbe includere anche principi etici, strumenti di monitoraggio e meccanismi di responsabilità condivisa.

Possono essere riassunti in cinque ambiti i principali suggerimenti per una governance responsabile:

- 1. Trasparenza e spiegabilità:** garantire non solo l'implementazione di modelli intrinsecamente interpretabili, ma anche la comprensione dei meccanismi alla base del processo decisionale negli algoritmi. Una documentazione accessibile aiuta sia gli utenti che gli enti regolatori a comprendere i criteri che guidano il processo decisionale automatizzato (Pomo, 2023).
- 2. Responsabilità distribuita:** è essenziale chiarire chi risponde delle decisioni prese lungo l'intera filiera dell'intelligenza artificiale, coinvolgendo sviluppatori, imprese private e istituzioni pubbliche. In questa prospettiva, l'OECD (2024) raccomanda l'adozione di standard di *algorithmic auditing*, affidati anche a soggetti terzi, che includano verifiche sia tecniche sia non tecniche sugli impatti etici e sull'equità dei processi decisionali automatizzati
- 3. Data governance aperta:** per prevenire monopoli informativi, è necessario promuovere politiche di condivisione dei dati sicure e responsabili. Garantendo l'interoperabilità e la portabilità dei dati, è possibile livellare il campo di gioco ed evitare asimmetrie informative tra le grandi piattaforme e gli altri operatori del mercato.
- 4. Alfabetizzazione digitale:** per un utilizzo sostenibile dell'IA, sono necessari anche utenti informati e consapevoli. Le iniziative educative possono migliorare la comprensione dei sistemi automatizzati da parte delle persone, consentendo un'interazione critica con essi (WEF, 2023-2024).

**5. Sostenibilità e inclusione:** infine, lo sviluppo di sistemi intelligenti dovrebbe riflettere fin dall'inizio obiettivi di equità sociale e sostenibilità. A questo proposito, è fondamentale l'allineamento con i *Sustainable Development Goals* dell'ONU.

Il World Economic Forum (2023-2024) sottolinea che la fiducia nei sistemi intelligenti dipende dalla loro legittimazione sociale. Ciò implica un nuovo patto tra tecnologia e democrazia in cui le decisioni automatizzate siano soggette a forme di supervisione collettiva. A questo proposito, la regolamentazione non dovrebbe essere considerata un ostacolo all'innovazione, ma un'infrastruttura istituzionale di affidabilità. Ortiz (2021) suggerisce di considerare la governance dell'IA come un bene pubblico globale, che funziona meglio attraverso la cooperazione tra Stati, imprese e comunità scientifiche. Ciò sembra coincidere con la visione dell'OECD (2024), secondo cui un prerequisito fondamentale per la competitività stessa dei mercati digitali è la creazione di fiducia.

#### **1.5.6 Equilibrio tra innovazione, fiducia e sostenibilità**

L'analisi presentata in questo capitolo ha dimostrato come l'economia algoritmica rappresenti, pur essendo uno dei principali motori di crescita e innovazione, una seria minaccia economica, etica e sociale. L'intelligenza artificiale aumenta l'efficienza produttiva, ma allo stesso tempo rafforza le asimmetrie informative e la concentrazione economica. Le sfide principali riguardano la trasparenza, la responsabilità e l'equa distribuzione del valore informativo. Se non governata, l'automazione intelligente può trasformarsi in una macchina asimmetrica controllata da una manciata di attori, mentre molti sono sottoposti al processo decisionale algoritmico. Tuttavia, gli ultimi lavori dell'OECD (2024) e del World Economic Forum (2023-2024) indicano che è possibile immaginare una diversa traiettoria di sviluppo tecnologico in cui progresso e responsabilità non si escludono a vicenda e svolgono un ruolo cruciale nell'efficienza degli usi digitali; l'efficienza tecnico-economica delle pratiche digitali non è sufficiente se non si basa sulla fiducia dell'utente e sulla trasparenza del sistema.

In quest'ottica, le condizioni strutturali di una fiducia algoritmica sostenibile richiedono almeno due trasformazioni complementari. Da un lato, un cambiamento culturale, basato sull'insegnamento di competenze essenziali e su una rivisitazione dell'etica della tecnologia; dall'altro, un rafforzamento istituzionale, nel senso della definizione di standard comuni per la governance e il controllo dei sistemi intelligenti. Come verrà discusso nel secondo capitolo, la

fiducia non è solo una variabile psicologica, ma anche una variabile economica per l'efficienza e l'adozione. Solo se percepiti come equi e affidabili, i sistemi di raccomandazione e le piattaforme di intelligenza artificiale potranno esprimere il loro potenziale produttivo. Da questo punto di vista, responsabilità e trasparenza non sono un vincolo all'innovazione, ma una condizione abilitante dell'innovazione.

In sintesi, l'economia algoritmica richiede un nuovo equilibrio tra efficienza, fiducia e sostenibilità. Regolamentare l'IA significa riconoscere che ogni decisione algoritmica è anche una decisione sociale e il futuro dei mercati digitali dipenderà dalla capacità collettiva di orientare la tecnologia verso il bene comune.

La sostenibilità dell'economia algoritmica dipende dalla fiducia collettiva nelle macchine decisionali, un aspetto di cui parleremo nel Capitolo 2. Trasparenza, spiegabilità, controllo umano significativo ed equità percepita sono le condizioni istituzionali necessarie. Tuttavia, la loro efficacia è mediata da fattori psicologico-comportamentali quali norme di giustizia, percezione di agency e responsabilità, tolleranza per l'errore umano vs algoritmico, preferenze per consigli utilitaristici o edonistici. Senza queste condizioni, i vantaggi economici dell'IA si scontrano con la resistenza cognitiva (avversione/sfiducia nell'algoritmo), che ne riduce la diffusione e l'impatto sul benessere. Al contrario, quando i sistemi sono percepiti come affidabili/benevoli, tuttavia, si verifica un apprezzamento algoritmico che consente l'innovazione a lungo termine. Il passaggio dalla conformità tecnica alla legittimità sociale dell'IA avviene quindi nel contesto della psicologia economica della fiducia, che sarà il filo conduttore dello sviluppo teorico e delle ipotesi di ricerca presentate nel capitolo seguente.

### **Riferimenti bibliografici (APA 7<sup>a</sup> ed.)**

- Căpușneanu, S. (2025). *Reshaping the Digital Economy with Big Data: A Meta-Analysis*. MDPI.
- McKinsey & Company. (2016). *Industry 4.0: How to Navigate Digitization of the Manufacturing Sector*.
- OECD. (2024). *Artificial Intelligence, Data and Competition – Note by Austria*. Directorate for Financial and Enterprise Affairs, Competition Committee.
- Ortiz, J. H. (Ed.). (2021). *Industry 4.0: Current Status and Future Trends*. IntechOpen. <https://doi.org/10.5772/intechopen.97805>

Pomo, G. (2023). *Towards Responsible AI in Recommender Systems*. Politecnico di Bari.  
*Exploring Ethical Frontiers of AI in Marketing*. (2024). ScienceDirect.  
World Economic Forum. (2023–2024). *Centre for the Fourth Industrial Revolution Network Report*.

## **CAPITOLO 2 – Letteratura accademica e modelli teorici sulla fiducia nei sistemi di raccomandazione AI**

### **2.1 Le origini del dibattito sull'avversione agli algoritmi**

#### **2.1.1. Perché parlare di “avversione” (e perché non basta)**

Dopo aver mostrato nel Capitolo 1 che l'Intelligenza Artificiale e l'economia dei dati stanno cambiando le regole della concorrenza, i modelli di business e i processi decisionali, diventa naturale porsi una domanda che è contemporaneamente sia empirica che teorica: le persone si fidano delle decisioni algoritmiche? E, soprattutto, in quali condizioni le persone sono disposte a seguire le raccomandazioni derivate computazionalmente, quando quell'output influenza scelte con conseguenze economiche e personali?

La questione della fiducia negli algoritmi viene spesso inquadrata attraverso una formula tanto intuitiva quanto parziale: le persone preferiscono gli esseri umani alle macchine. La narrazione, per quanto diffusa, non è sufficiente. La letteratura recente, infatti, riporta un quadro più complesso: in alcuni casi riscontriamo sfiducia (persino rifiuto) nei confronti dell'output algoritmico, mentre in altri riscontriamo un'attrazione verso gli algoritmi rispetto agli esseri umani (Dietvorst et al., 2015; Logg et al., 2019). Questa ambiguità concettuale è al centro di questa tesi perché, seguendo il modello qui proposto, la fiducia non è meramente un problema psicologico; è la condizione di possibilità affinché i guadagni economici dell'IA (in termini di produttività, efficienza allocativa e migliore abbinamento domanda/offerta) si concretizzino anziché rimanere una possibilità astratta.

La Sezione 2.1 ricostruisce dunque le origini del dibattito attuale e chiarisce come siamo passati da un dibattito quasi di senso comune (avversione generalizzata agli algoritmi) a una

visione più complessa in cui avversione e apprezzamento coesistono e variano in base (tra le altre) alle seguenti variabili: esperienza passata con l'errore, confronto con il proprio giudizio, competenza, natura del compito, rischiosità percepita e presentazione dell'algoritmo (Castelo et al., 2019; Logg et al., 2019).

### 2.1.2. Che cosa intendiamo per “algoritmo” in questa letteratura

Un primo passaggio, non sempre adeguatamente considerato, riguarda il significato operativo di algoritmo. Nei tre contributi centrali di questa sezione (Dietvorst et al., 2015; Logg et al., 2019; Castelo et al., 2019), l'algoritmo non coincide necessariamente con una “black box” (Logg et al., 2019, p. 91) sofisticata. Piuttosto, denota una procedura evidence-based in grado di produrre una previsione, una stima o una raccomandazione da un set di dati secondo regole di calcolo. Questa è una distinzione importante: molte reazioni psicologiche non riguardano la sofisticatezza tecnica, ma ciò che l'algoritmo rappresenta socialmente e simbolicamente (Logg et al., 2019). Inoltre, questi studi non considerano l'algoritmo "in astratto", ma lo considerano nel "mondo reale" di decisioni predittive, giudizi incerti e raccomandazioni in ambiti che variano in termini di oggettività e rilevanza personale. La fiducia, in altre parole, non è una caratteristica individuale fissa: è una risposta che si sviluppa nel contesto di una decisione.

### 2.1.3. Il paradigma dell’“algorithm aversion”: l’errore come rottura della fiducia

Il punto di svolta del dibattito recente è rappresentato dal lavoro di Dietvorst, Simmons e Massey (2015), che rende empiricamente misurabile quello che è stato un dibattito alquanto impressionistico in studi precedenti: le persone, pur sapendo che gli algoritmi sono più accurati degli esseri umani, tendono a evitarli quando li vedono sbagliare (Dietvorst et al., 2015). Il cuore dell’argomentazione è, a pensarci bene, psicologicamente potente: in molti contesti quotidiani siamo abituati all’errore umano; lo accettiamo, lo perdoniamo, lo consideriamo normale. L'errore della macchina, tuttavia, **sembra avere la peggio**: mina la nostra fiducia molto più rapidamente e in modo molto più profondo, anche quando l'algoritmo rimane complessivamente superiore.

Dietvorst et al. (2015) esplorano questo concetto in un esperimento complesso: i partecipanti devono fare previsioni basate su dati reali su un caso e a un certo punto sono

incentivati a scegliere se legare il loro incentivo alla previsione di un giudice umano o a un modello statistico. La variabile importante è la manipolazione dell' "esperienza": in alcune condizioni i partecipanti vedono la performance dell' algoritmo (e i suoi errori), in altre la performance umana, in altre entrambe, e in una condizione di controllo nessuna delle due (Dietvorst et al., 2015).

La scoperta di base è solo apparentemente controintuitiva: vedere l'algoritmo all'opera riduce la propensione a sceglierlo anche quando performa meglio dell'essere umano (Dietvorst et al., 2015). Questo aspetto è centrale per questa tesi poiché suggerisce che il problema non è soltanto "informativo" (mancanza di evidenza sulla qualità), ma cognitivo-valutativo: l'errore viene pesato diversamente a seconda della fonte. In termini di economia comportamentale, possiamo parlare di un aggiornamento asimmetrico: infatti, la penalizzazione dell' algoritmo per un errore è più severa della penalizzazione dell'essere umano per lo stesso errore. Dietvorst et al. (2015) mostrano appunto che la perdita di fiducia in un previsore algoritmico si verifica a un ritmo più rapido rispetto alla perdita di fiducia in un previsore umano quando entrambi commettono lo stesso errore. Questo è fondamentale perché può portare a decisioni economicamente dannose (Dietvorst et al., 2015). Optare per la previsione umana inferiore rispetto all'algoritmo superiore aumenta l'inefficienza decisionale, poiché può ridurre l'accuratezza e il benessere atteso nelle decisioni (Dietvorst et al., 2015). Nel quadro della presente tesi, ciò significa che la "razionalità" nell'adottare un algoritmo non prevale automaticamente; infatti, la fiducia può rappresentare un vincolo e talvolta un collo di bottiglia nel rapporto tra tecnologia e risultati economici.

#### **2.1.4. Perché l'errore dell'algoritmo pesa di più: norme implicite, aspettative e agency**

A questo punto, è opportuno chiedersi: *perché* l'errore della macchina viene giudicato più duramente? La letteratura non indica una sola ragione, ma suggerisce un intreccio di aspettative implicite. La prima è l'aspettativa di precisione (quasi) meccanica che caratterizza gli algoritmi. Anche se si ammette, in linea di principio, che i modelli siano imperfetti, si può comunque aggrapparsi a un'aspettativa latente di perfezione. Quando un algoritmo fallisce, non si limita a minare un risultato: mina un modello di perfezione. In questo modo, l'errore ha conseguenze simboliche piuttosto che statistiche (Dietvorst et al., 2015).

Secondo: l'errore umano può essere facilmente spiegato in termini narrativi (ad esempio, stanchezza, distrazione, intuizione imperfetta). L'errore algoritmico non è sempre, tuttavia, facilmente comprensibile. È l'incomprensibilità dell'errore che crea un senso di impotenza: *se non capisco perché sbaglia*, come posso affidarmi a esso per guidare le mie decisioni? Anche se non esplicitato, il tema della comprensione è un filo conduttore del concetto di fiducia.

In terzo luogo, c'è una questione di agency e controllo. In parte, affidarsi a un algoritmo è un atto di delega. E la delega può essere in contrasto con il bisogno psicologico di sentirsi autori della decisione. In una situazione di incertezza, le persone potrebbero preferire una decisione umanamente imperfetta che percepiscono come propria (o come imputabile a un altro umano) rispetto a una decisione meccanicamente migliore ma percepita come estranea.

Queste implicazioni forniscono già collegamenti verso i capitoli successivi della tesi; nel Capitolo 2.4, ad esempio, la fiducia viene scomposta nei suoi elementi (affidabilità, prevedibilità, comprensibilità, intenzionalità percepita) e collegata a modelli di fiducia nell'automazione; e nel Capitolo 2.3, la distinzione tra consumo utilitaristico ed edonistico si riferisce all'idea che la delega all'algoritmo potrebbe non essere percepita in egual misura nei casi in cui la decisione è "funzionale" rispetto a una legata all'identità, al gusto o all'emozione.

### **2.1.5. Il controcanto: l'algorithm appreciation e la preferenza per l'algoritmo**

Mentre Dietvorst et al. (2015) hanno sperimentalmente realizzato il concetto di avversione, Logg, Minson e Moore (2019) hanno reso il quadro più complesso dimostrando che, in molte condizioni, è vero il contrario: le persone sono più propense ad accettare un consiglio se credono che sia stato dato da un algoritmo piuttosto che da un essere umano, anche se il consiglio è identico (Logg et al., 2019).

Il contributo di Logg et al. (2019) è metodologico prima ancora che teorico. Gli autori osservano che la maggior parte delle prove a sostegno dell'avversione proviene da studi in cui le persone devono scegliere tra un algoritmo e se stesse (o il proprio giudizio). Tuttavia, la psicologia decisionale suggerisce che le persone sono spesso eccessivamente sicure del proprio giudizio e, più in generale, sottovalutano i consigli esterni (Logg et al., 2019). Se così fosse, il rifiuto dell'algoritmo potrebbe non essere avversione alla macchina, ma eccessiva fiducia in sé stessi.

Per essere più precisi, Logg et al. (2019) adottano il paradigma del Judge–Advisor System: le persone formulano un giudizio iniziale, ricevono un consiglio, poi formulano un giudizio finale. Il fattore critico è il peso del Consiglio. In diversi esperimenti, gli autori dimostrano che, quando il consiglio è lo stesso, sebbene l'etichetta della fonte sia la diversa, ha un peso maggiore quando proviene dall'etichetta di un algoritmo rispetto a quella di un essere umano (Logg et al., 2019). In altre parole, l'aura di competenza cognitiva porterà l'algoritmo a essere trattato come un consulente migliore. Questo risultato non esclude tuttavia Dietvorst et al. (2015), poiché suggerisce che il rapporto con l'algoritmo non è unidirezionale: l'atteggiamento dipende dalla situazione e dalla cornice comparativa. Logg et al. (2019), dopotutto, hanno anche individuato contesti in cui l'apprezzamento per l'algoritmo si attenua: confrontarsi con il proprio giudizio (piuttosto che con due consulenti esterni) e con elevati livelli di competenza può rendere i professionisti meno inclini a esternalizzare le decisioni (Logg et al., 2019). La domanda rilevante per questa tesi non è quindi solo "algoritmo sì o no", ma occorre mappare quando l'algoritmo viene percepito come alleato e quando viene percepito come minaccia all'autonomia o all'identità decisionale.

### **2.1.6. “Human vs algorithm trust”: il ruolo del compito, della posta in gioco e della cornice di presentazione**

La terza tessera del mosaico è rappresentata dal contributo di Castelo, Bos e Lehmann (2019), che sposta l'attenzione dal solo errore (Dietvorst et al., 2015) e dalla sola fonte dei consigli (Logg et al., 2019) a una serie più ampia di fattori determinanti: caratteristiche del compito, familiarità, conseguenze e rappresentazione sociale dell'algoritmo (Castelo et al., 2019).

Il punto di partenza è empiricamente solido: molte persone sono scettiche nei confronti degli algoritmi non perché non siano a conoscenza dei progressi dell'IA, ma perché associano alle macchine un profilo di capacità sbilanciato: forti nelle competenze cognitive (calcolo, coerenza, analisi), deboli nelle competenze umane (intuizione, emozione, sensibilità). Questa mappatura implicita ha un effetto predittivo: si tende a fidarsi maggiormente dell'algoritmo per compiti percepiti come oggettivi, misurabili e quantificabili e a preferire gli esseri umani per i

compiti percepiti come soggettivi, legati al gusto, alle relazioni o alla sfera personale (Castelo et al., 2019).

La nuova scoperta importante è che, per (Castelo et al., 2019), l'oggettività non è una proprietà fissa del compito, ma una percezione manipolabile. Quando un compito viene riformulato in modo tale da apparire più quantificabile, siamo più propensi a fidarci dell' algoritmo. Questa scoperta è particolarmente interessante nel caso dei sistemi di raccomandazione: molte raccomandazioni (film, musica, prodotti lifestyle) oscillano tra gusto ed efficienza, e la loro accettazione dipende anche da come la piattaforma presenta ciò che sta facendo (*ti suggeriamo questo perché...*).

Tuttavia, oltre all'oggettività, Castelo et al. (2019) sottolineano due variabili aggiuntive: familiarità e conseguenze. In generale, l'esperienza ripetuta con algoritmi quotidiani (es. mappe, consigli di intrattenimento) può portare a considerare la delega come normale; laddove la posta in gioco è più alta (salute, finanza, diagnosi), la soglia di fiducia tende ad aumentare e le persone potrebbero essere più diffidenti (Castelo et al., 2019). Questo elemento è rilevante anche in prospettiva analitica, poiché suggerisce che la fiducia negli algoritmi non dipende soltanto dalle loro prestazioni percepite, ma anche dalla rilevanza delle conseguenze associate alla decisione. Quando la posta in gioco aumenta, la delega algoritmica tende infatti a essere valutata con maggiore cautela, rendendo la consequenzialità della scelta una dimensione cruciale del contesto decisionale. Infine, Castelo et al. (2019) introducono un elemento interessante: l'umanizzazione dell'algoritmo può ridurre lo scetticismo, soprattutto nei compiti più soggettivi. Se le persone si abituano all'idea che gli algoritmi possano fare cose umane (creare arte, riconoscere emozioni), la distanza simbolica tra uomo e macchina si riduce, insieme all'avversione (Castelo et al., 2019). Questo punto prepara in modo naturale il terreno per il Capitolo 2.4, dove il framework warmth–competence può essere utilizzato per esplorare perché la percezione del "calore" (intenzioni benevole) sia importante quanto la competenza tecnica nel generare fiducia.

### 2.1.7. Ricomporre il puzzle: perché avversione e apprezzamento possono coesistere

A questo punto, il quadro generale è meno contraddittorio di quanto sembri. I tre filoni non si contraddicono a vicenda; descrivono invece **momenti diversi** del rapporto tra persone e algoritmi.

- Dietvorst et al. (2015) descrivono una dinamica di rottura: una volta che l'algoritmo entra nel dominio dell'esperienza concreta e rivela l'inevitabilità degli errori, la fiducia può sgretolarsi più velocemente di quanto accadrebbe con un essere umano.

- Logg et al. (2019) descrivono una dinamica di inferenza: quando l'algoritmo è visto come un consulente cognitivo e confrontato con un essere umano esterno, l'algoritmo può guadagnare credibilità e pesare di più nelle revisioni di giudizio.

- Castelo et al. (2019) aggiungono una dinamica di contesto: la fiducia non è "per l'algoritmo" in generale, ma dipende dal tipo di compito, dalla posta in gioco e dalla cornice con cui la tecnologia viene presentata.

Ne deriva un'idea compatibile con l'approccio economico-comportamentale: la fiducia può essere vista come capitale relazionale che può essere accumulato e consumato. In alcuni contesti l'algoritmo parte avvantaggiato (aura di competenza e coerenza), in altri parte svantaggiato (percezione di inadeguatezza emotiva; mancanza di identità). L'errore agisce come una scossa che esaurisce rapidamente quel capitale, soprattutto se la persona non ha i mezzi per dargli un senso o normalizzarlo.

Questa inquadratura contestualizza correttamente anche questo argomento di interesse all'interno di questo progetto di tesi: ciò che è interessante per questo progetto non è solo la preferenza dichiarata per gli esseri umani o gli algoritmi, ma come tale preferenza influenza l'adozione effettiva e quindi i benefici economici dei mercati digitali. Una piattaforma può avere un SR molto ben progettato, ma se gli utenti sono sproporzionatamente sensibili agli errori algoritmici (Dietvorst et al., 2015) o rifiutano l'algoritmo nei domini soggettivi (Castelo et al., 2019), la conseguenza positiva per il loro benessere può essere attenuata o addirittura annullata. D'altro canto, in contesti ad alta incertezza e bassa competenza, l'algoritmo potrebbe essere visto come una scorciatoia cognitiva accettata e persino preferita (Logg et al., 2019).

### 2.1.8. Implicazioni per la tesi e ponte verso i capitoli successivi

Sul piano della tesi, questa sezione svolge tre funzioni.

Contestualizza il problema della fiducia come una variabile economico-comportamentale che condiziona l'adozione di tecnologie *basate sui dati* e, quindi, l'efficienza dei mercati algoritmici descritti nel Capitolo 1.

**Costruisce un lessico concettuale** utile per le ipotesi: infatti, gli atteggiamenti verso gli algoritmi non sono (solo) una questione di avversione/apprezzamento, ma si muovono lungo dimensioni come esperienza ed errore (Dietvorst et al., 2015), fonte di consiglio e confronto (Logg et al., 2019), oggettività e presentazione (Castelo et al., 2019).

**Prepara direttamente il Capitolo 2.2:** se la fiducia dipende dal modo in cui l'algoritmo viene ascoltato come fonte di consiglio, allora diventa una transizione naturale al concetto di Word-of-Machine, ovvero l'effetto delle raccomandazioni algoritmiche sul comportamento e sulle preferenze, e in quali condizioni tale effetto è considerato legittimo.

In altre parole, il dibattito su avversione e apprezzamento non è solo una discussione teorica: prepara, invece, il terreno per valutare perché e quando una raccomandazione artificiale diventa persuasiva o respingente. È proprio su questa zona di confine (tra calcolo e psicologia economica, tra efficienza e legittimazione) che si innesta il prosieguo di questa analisi.

#### Riferimenti bibliografici (APA 7<sup>a</sup> ed.)

- Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). *Let the machine decide: When consumers trust or distrust algorithms*. NIM Marketing Intelligence Review, 11(2), 24–29.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). *Algorithm aversion: People erroneously avoid algorithms after seeing them err*. Journal of Experimental Psychology: General, 144(1), 114–126.
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). *Algorithm appreciation: People prefer algorithmic to human judgment*. Organizational Behavior and Human Decision Processes, 151, 90–103.

## **2.2. Il modello “Word-of-Machine” (Longoni & Cian, 2022)**

### **2.2.1. Perché serve un modello “contestuale” della fiducia nelle raccomandazioni algoritmiche**

Nel dibattito sull'accettazione delle decisioni algoritmiche, la letteratura più nota (vedi sezione 2.1) tende a presentare la relazione tra individui e algoritmi come una sorta di braccio di ferro stabile: da un lato c'è la diffidenza (algorithm aversion), dall'altro c'è una potenziale apertura, o addirittura una preferenza per la macchina (algorithm appreciation). È una cornice, se dovessimo attenerci ad essa, che ci porterebbe a leggere la fiducia negli algoritmi come una disposizione psicologica pressoché invariabile (mi fido/non mi fido degli algoritmi). Eppure, in un contesto economico, la fiducia negli algoritmi sembra essere condizionata, a quanto pare, dal tipo di scelta, dal contesto del suo utilizzo e dalle poste in gioco che comporta. È proprio in questa direzione che Longoni e Cian (2022) contribuiscono proponendo un modello capace di spiegare quando la raccomandazione AI viene percepita come più convincente e quando, al contrario, genera resistenza. L'idea chiave del loro modello è che le persone non valutano l'algoritmo in astratto: valutano la coerenza percepita tra quello che ritengono essere il livello di competenza della fonte (umana o artificiale) e la natura della decisione. Questo sviluppo è di fondamentale importanza per la presente tesi, perché sposta l'attenzione dal se gli algoritmi siano accettati al perché e in quali condizioni lo siano. Apre un terreno di esplorazione che può essere direttamente collegato ai mercati digitali e ai sistemi di raccomandazione di cui si è discusso nel Capitolo 1 (in particolare nelle sezioni sulle strategie data-driven e sulla creazione di valore attraverso la personalizzazione). Inoltre, l'approccio di Longoni e Cian (2022) è particolarmente utile per il disegno empirico sviluppato nel Capitolo 3: se la fiducia si muove in modo sistematico lungo una dimensione (utilitaristica vs. edonica), allora diventa possibile formulare ipotesi testabili e misurabili non solo riguardo all'esito di una raccomandazione (accettazione/rifiuto), ma anche riguardo ai meccanismi psicologici intermedi che la conducono (competenza percepita, appropriatezza della fonte, aspettative sul valore della raccomandazione).

### **2.2.2. Definizione del “Word-of-Machine effect” e intuizione di fondo**

Longoni e Cian (2022) definiscono l'effetto word-of-machine come il fenomeno per cui la preferenza (o la resistenza) alle raccomandazioni generate dall'IA dipende dal tipo di compromesso attivato nella decisione: quando la scelta è percepita come utilitaristica, si tende

preferire l'IA; quando invece è percepita come edonistica, la preferenza si sposta verso il consigliere umano. Questa regolarità non è presentata come una questione di gusto soggettivo, ma come il risultato di una convinzione diffusa (*lay belief*) sulle rispettive capacità di esseri umani e macchine. Intuitivamente:

- L'IA è associata alla logica, al calcolo, alla coerenza e alla capacità di elaborare una grande quantità di informazioni e scegliere la soluzione più razionale;

- Gli esseri umani sono associati alla sensibilità, all'esperienza, all'empatia, alle sfumature e alle preferenze non formalizzabili. Di conseguenza, se la decisione da prendere è legata alla massimizzazione della funzionalità, dell'efficienza, delle proprietà tecniche e della convenienza, la macchina è nel suo elemento. Se, d'altra parte, la decisione da prendere è legata al piacere, al gusto, all'estetica, all'atmosfera e all'esperienza, si suppone che gli esseri umani siano più adatti a interpretare ciò che è veramente importante. Questa concettualizzazione è immediatamente rilevante per i mercati digitali: esistono molte piattaforme che combinano dimensioni utilitaristiche ed edonistiche (si pensi al turismo digitale, alla ristorazione, alle piattaforme di streaming, persino la scelta di un prodotto funzionale che comunque trasmette identità). Per questo, come verrà approfondito nella sezione 2.3 (Consumo edonico vs. utilitaristico), la distinzione non va intesa come una rigida tassonomia dei beni, ma piuttosto come una lente motivazionale: ciò che cambia è il tipo prevalente di motivazione che il consumatore attiva nel momento della scelta.

### **2.2.3. Il meccanismo: credenze di competenza e “appropriatezza” della fonte**

La proposta teorica di Longoni e Cian (2022) si basa sulla semplice, ma potente ipotesi: le persone non valutano solo la qualità oggettiva di una raccomandazione, ma valutano anche *chi* la formula e se quel *chi* è considerato competente in relazione al tipo di valore in questione.

Nel loro modello, i passaggi causali si susseguono nel seguente ordine:

1. Attivazione dell'obiettivo (goal): utilitaristico o edonico;
2. Percezione di competenza relativa: l'intelligenza artificiale appare più competente nel dominio utilitaristico, gli esseri umani in quello edonico;

3. Preferenza e maggiore reliance: la raccomandazione proveniente dalla fonte percepita come più adatta viene accettata più facilmente;
4. Conseguenze percettive sul consumo: non cambia solo la scelta: cambia anche il modo in cui gli attributi del bene vengono valutati dopo l'esperienza.

Questo quarto e ultimo punto è particolarmente interessante: suggerisce che la fonte conta non solo ex ante (la decisione) ma anche ex post (l'interpretazione dell'esperienza). In termini di economia comportamentale, la raccomandazione agisce come un frame che orienta l'attenzione e la valutazione degli attributi, con possibili implicazioni su soddisfazione, fidelizzazione e il valore di lungo periodo del cliente. È importante notare che, nel testo di Longoni e Cian (2022), le percezioni di competenza sono inquadrare come credenze culturali e non come proprietà naturali della tecnologia stessa (l'intelligenza artificiale non è naturalmente incapace di comprendere le dimensioni edoniche, così come gli esseri umani non sono naturalmente incapaci di comprendere le dimensioni tecniche). Il problema sta nelle aspettative che le persone hanno da una macchina rispetto a un essere umano; e tali aspettative, proprio perché condivise, diventano regolarità psicologiche che influenzano i mercati. Tutto questo prepara anche il terreno per la sezione 2.4, che scomporrà la fiducia in dimensioni più articolate (competenza, calore, intenzionalità percepita, ecc.). Qui, tuttavia, l'accento è su una forma specifica di appropriatezza: la fiducia non riguarda quanto credo che l'IA sia brava, ma in cosa credo che l'IA sia brava, e se questo combaci con ciò che mi interessa.

#### **2.2.4. Evidenze empiriche: sintesi ragionata delle principali evidenze sperimentali**

Un tratto distintivo del lavoro di Longoni e Cian (2022) è la costruzione progressiva dell'argomento empirico. Gli autori non si limitano a mostrare un effetto in un singolo esperimento, ma lo testano su contesti diversi, con misure differenti e soprattutto con passaggi pensati per chiarire *meccanismo* e *confini* dell'effetto.

(a) La preferenza per i consulenti umani o IA varia in base all'obiettivo

Nella prima serie di risultati, l'obiettivo è dimostrare che la preferenza variabile per un consulente IA rispetto a un consulente umano è influenzata dall'attivazione di un obiettivo utilitaristico o edonico; come previsto, la raccomandazione algoritmica è favorita in presenza di un obiettivo utilitaristico, mentre la raccomandazione umana prevale in presenza di un obiettivo edonico.

Tuttavia, ciò che è interessante, ai fini di questa tesi, è che non si tratta semplicemente di una misurazione dell'atteggiamento ("Mi piace l'IA"), ma un comportamento di scelta della fonte. Questo avvicina l'analisi alla logica dei mercati digitali, dove la fiducia si traduce spesso in un'azione semplice ma economicamente decisiva: cliccare su una raccomandazione, seguire un suggerimento, acquistare un articolo suggerito, accettare una classifica.

(b) Effetti sulle percezioni degli attributi post-consumo

Una parte importante del contributo consiste nel mostrare che la fonte della raccomandazione influenza il modo in cui gli individui percepiscono gli attributi del bene, una volta che gli viene presentato; lo considerano di natura più utilitaristica quando generato dall'intelligenza artificiale e più edonica quando generato dall'uomo. Anche questo aspetto è rilevante per questa tesi: le raccomandazioni influenzano la percezione del bene piuttosto che essere una misura neutrale del suo valore, che può a sua volta influenzare valutazioni, recensioni e passaparola.

(c) Importanza relativa degli attributi e preferenza per la fonte

Un altro passaggio empirico mette in relazione l'importanza che gli individui attribuiscono agli attributi utilitaristici/edonici e la loro preferenza per i consulenti umani rispetto a quelli basati sull'intelligenza artificiale. Gli autori mostrano che la preferenza per l'intelligenza artificiale cresce all'aumentare dell'importanza assegnata agli attributi utilitaristici e diminuisce quando si dà maggiore peso ad attributi edonistici; inoltre, le analisi di soglia (Floodlight/Johnson-Neyman) consentono agli autori di individuare intervalli in cui l'effetto è statisticamente significativo.

Questa sezione è particolarmente rilevante per il Capitolo 3, in quanto suggerisce un modo per manipolare il contesto (edonico vs. utilitaristico) e anche per misurare *quanto* il soggetto valorizzi certe dimensioni e vedere se questa importanza modera la fiducia nella raccomandazione.

(d) Il ruolo delle percezioni di competenza e l'esclusione dell'alternativa "complessità"

Un risultato importante è che le differenze di scelta non sono (solo) spiegate dall'idea che l'intelligenza artificiale sia migliore nel gestire informazioni complesse, ma da un'ipotesi più specifica sulla competenza per i due domini di valore. Longoni e Cian (2022) testano l'ipotesi alternativa della complessità attraverso la manipolazione del numero degli attributi e, contestualmente, verificano che ciò non spiega l'interazione tra goal e tipo di consigliere;

piuttosto, sono le percezioni di competenza a mediare l'effetto, con una struttura di moderated mediation.

Questo risultato è interessante proprio perché apporta un'importante chiarificazione a livello teorico: il consumatore non pensa *Scelgo l'IA perché ci sono tante informazioni*, ma *Scelgo l'IA perché quando cerco funzionalità, la macchina sembra più competente in questo tipo di valutazione*. Questo porta a una connessione più sofisticata tra fiducia e valore: la fiducia non è un residuo emotivo, ma una valutazione di adeguatezza tra fonte e valore.

### **2.2.5. Confini dell'effetto: quando la macchina perde terreno anche nel dominio utilitaristico (Studio 5)**

Finora, il quadro potrebbe suggerire una regola quasi lineare: utilitaristico → IA; edonico → umano. Longoni e Cian (2022), tuttavia, dimostrano che l'effetto ha confini precisi e può, in determinate occasioni, persino invertirsi.

Un primo confine è la necessità di matching con preferenze uniche. Quando si ha la sensazione che la raccomandazione debba essere su misura in modo da tenere conto dell'unicità del decisore, la competenza del raccomandante umano torna a pesare anche nelle decisioni il cui contenuto è utilitaristico. Nello Studio 5, facendo comprendere ai partecipanti che la raccomandazione deve essere su misura in base all'unicità della loro decisione, la scelta dell'IA diminuisce anche negli obiettivi utilitaristici, al punto da invertire l'effetto (ovvero, una drastica riduzione della scelta dell'IA nel caso utilitaristico rispetto alla condizione di controllo).

Questo risultato è immediatamente rilevante per questa tesi, in quanto fa riferimento a due argomenti già esplorati nel Capitolo 1:

1. le implicazioni economiche della personalizzazione e dell'hyper-personalization (Capitolo 1.3), spesso inquadrate come un vantaggio naturale degli algoritmi;
2. la natura relazionale dei dati (Capitolo 1.1), in cui fiducia e controllo delle informazioni emergono come condizioni di sostenibilità.

La scoperta di Longoni e Cian (2022) suggerisce una tensione: proprio quando la personalizzazione diventa davvero personale (ovvero, quando l'individuo sente che il consiglio dovrebbe rispondere alla sua specificità), può emergere una preferenza per l'umano, in quanto la macchina viene percepita come incapace di cogliere caratteristiche idiosincratiche, il contesto di vita e le sfumature non formalizzabili. Questa scoperta ha tuttavia un'implicazione manageriale non banale: infatti, una maggiore personalizzazione non garantisce una maggiore fiducia nella macchina; a volte, man mano che aumenta la posta in gioco dell'identità o il senso di unicità, diventa strategico integrare segnali di presenza umana o di controllo umano.

### **2.2.6. Un secondo confine: dalla sostituzione all'“augmented intelligence” (Studio 6)**

Se lo studio 5 indica un caso in cui l'IA sta perdendo terreno, lo studio 6 suggerisce una soluzione progettuale per riconquistare la fiducia anche in contesti edonistici: il passaggio dall'intelligenza artificiale, vista come un sostituto degli umani, all'augmented intelligence, vista come un supporto per gli umani. Longoni e Cian (2022) tracciano una distinzione tra due architetture implicite: l'IA come sostituto del consulente umano (la macchina prende il posto dell'umano); l'IA come amplificatore del consulente umano (la macchina supporta, ma l'umano rimane il decisore finale). Quando la decisione è inquadrata come ibrida (umano-IA), l'effetto word-of-machine tende a ridursi o a scomparire: infatti, l'AI non viene più percepita come intrusa nel dominio edonistico, perché qui l'umano funge da garanzia di sensibilità e di esperienza; simmetricamente, nel dominio utilitaristico l'ibridazione può rafforzare l'efficacia percepita dell'umano grazie al supporto algoritmico.

A livello teorico, qui avviene un passaggio importante: in realtà, la fiducia non dipende solo da *chi* consiglia, ma anche da come è distribuito il controllo decisionale. Ciò è in linea con la letteratura sulla fiducia nell'automazione e sulle condizioni di appropriate reliance (che riprenderemo nella sezione 2.4), poiché la fiducia non è una delega totale: è spesso molto più sostenibile come collaborazione.

In questa prospettiva diventa utile richiamare anche la riflessione divulgativa di Hao (2020), che discute la necessità di sistemi capaci di deferire all'umano quando appropriato, evidenziando così il valore della complementarità tra giudizio umano e calcolo algoritmico (Hao, 2020).

Questa riflessione è coerente con l'idea che un maggior valore non sia necessariamente il risultato di una maggiore automazione. Nella maggior parte dei contesti di mercato, la vera sfida è come organizzare questi sistemi in modo che il ruolo della macchina e quello dell'uomo possano essere coordinati in modo ottimale, sapendo quando spingere l'algoritmo e quando frenarlo (cioè, stabilire *quando* la macchina debba consigliare e *quando* debba lasciare spazio a un intervento umano).

### **2.2.7. Interventi di attenuazione: cambiare le credenze (Studio 7a–7b)**

Un aspetto particolarmente rilevante, anche da un punto di vista applicativo, riguarda la possibilità di intervenire sulle convinzioni che determinano l'effetto. Longoni e Cian (2022) non si limitano a segnalare un fenomeno: testano un protocollo di debiasing (consider-the-opposite) per mitigare l'associazione stereotipata tra IA/utilitarismo e umano/edonismo. Nei risultati, invitare gli individui a considerare la possibilità che le aspettative iniziali sulla competenza della fonte possano essere errate può indebolire l'effetto: le persone valutano gli attributi delle due fonti in modo più uniforme e la macchina non è così pesantemente penalizzata nel dominio edonistico. Un elemento ancora più interessante dal punto di vista del mercato è che gli autori delineano una versione scalabile dell'intervento, integrata in uno scambio in stile chatbot, che potrebbe essere compatibile con contesti digitali reali. La fiducia non è solo un dato psicologico preesistente, ma può essere costruita (o distrutta) attraverso scelte di comunicazione, di interfaccia e del framing. Questo si collega verticalmente al Capitolo 1.5, dove le minacce dell'economia algoritmica sono legate anche all'opacità e al senso di manipolazione. Infatti, nei casi in cui i consumatori percepiscono il sistema come freddo, cioè privo di empatia o irrispettoso delle loro preferenze, i livelli di fiducia calano; ma quando i consumatori vengono incoraggiati a sfidare gli stereotipi e a considerare la tecnologia come un alleato piuttosto che un sostituto, l'accettazione può aumentare.

### **2.2.8. Implicazioni teoriche: un ponte tra avversione/apprezzamento e motivazioni di consumo**

Il valore teorico del word-of-machine effect, nel contesto di questa tesi, risiede non solo nella sua funzione chiarificatrice ma anche nella capacità di spiegare evidenze empiriche che a prima vista appaiono contraddittorie. Dato che alcuni studi evidenziano avversione e altri

apprezzamento, si ritiene che questa divergenza dipenda in parte dal contesto decisionale. In effetti, in domini in cui gli attributi utilitaristici giocano un ruolo centrale, l'IA può essere preferita; mentre in domini dove prevalgono aspetti edonistici, la stessa IA può suscitare delle resistenze.

A questo proposito, Longoni e Cian (2022) offrono una spiegazione situazionale che tiene conto dei risultati della sezione 2.1 senza ridurla a un singolo schema. Inoltre, il loro modello anticipa perfettamente la sezione 2.3: se edonistico e utilitaristico non sono categorie definite ma possono essere attivate come motivazioni, allora il comportamento nei confronti della raccomandazione dell'IA può dipendere anche dal modo in cui la piattaforma incornicia la scelta (testi, immagini, metriche, classifiche, argomenti di vendita). Una stessa decisione (es. quale ristorante scegliere) può essere resa più o meno utilitaristica (prezzo, distanza, tempo) o più o meno edonistica (esperienza, atmosfera, estetica), con effetti prevedibili sulla fonte considerata più affidabile.

In linea con l'approccio economico-comportamentale, l'algoritmo non agisce da solo: infatti, il suo contributo al valore economico emerge solo quando le sue prestazioni tecniche si combinano con le aspettative, le convinzioni e gli schemi cognitivi degli utenti; inoltre, in assenza di questo allineamento, anche soluzioni algoritmiche superiori rischiano di non essere adottate. In un mercato data-driven, la piattaforma può persino modellare la domanda orientando l'attenzione su determinati attributi (utilitari) e ignorandone altri (edonici), incoraggiando indirettamente le convinzioni dei soggetti ad allinearsi al suo sistema di raccomandazione.

### **2.2.9. Dal modello teorico al disegno sperimentale: misurazione, ipotesi e continuità con il Capitolo 3**

Per il disegno empirico sviluppato nel Capitolo 3, il modello word-of-machine suggerisce le seguenti linee guida operative:

**Manipolare o rendere saliente il goal** (utilitaristico vs edonico) e osservare le variazioni nella fiducia e nell'aderenza alla raccomandazione;

**Misurare il meccanismo:** utilizzando scale che misurano la competenza percepita della fonte (e, seguendo il punto 2.4, anche il calore/empatia e l'intenzionalità percepita);

**Considerare moderatori** come l'esigenza di personalizzazione unica (boundary condition) e la presenza di un modello ibrido (IA + umano) come potenziale leva di design;

Interventi di comunicazione (ad esempio, suggerimenti, spiegazioni, inquadramento) come leve manageriali ed eventualmente come condizioni sperimentali.

In particolare, la tesi può basarsi sulla distinzione tra:

- fiducia come convinzione ("questa fonte è competente"),
- fiducia come comportamento (decido di seguire la raccomandazione),
- fiducia come valutazione ex post (attribuisco un valore utilitaristico/edonico all'esperienza anche a causa della fonte). Questa distinzione tripartita mira a evitare un errore facile da commettere: trattare la fiducia come una variabile unidimensionale. Invece, come in Longoni e Cian (2022), la fiducia emerge come un costrutto dinamico, incarnato nel processo decisionale.

#### **2.2.10. Limiti e spunti per le sezioni successive**

Sebbene il modello sia solido, ci sono almeno tre implicazioni cautelative utili per una discussione critica:

- **Contesto culturale e cambiamento tecnologico:** le convinzioni sull'intelligenza artificiale e sugli esseri umani possono cambiare nel tempo, anche con l'evoluzione della tecnologia e la padronanza delle attività umane. Gli autori stessi sottolineano che si tratta di lay beliefs radicate culturalmente.

- Eterogeneità individuale: non tutti attribuiscono lo stesso peso al piacere e alla funzionalità, ; la moderazione in base all'importanza degli attributi suggerisce che anche le differenze psicografiche contano.

- Design e governance: l'esito non dipende solo dalla tecnologia, ma anche dal modello decisionale (sostituzione vs ibrido), dalle comunicazioni (interventi correttivi) e, di conseguenza, dai modelli di trasparenza e responsabilità (temi che si intrecciano con i rischi dell'economia algoritmica già discussi nel Capitolo 1.5).

In definitiva, l'effetto word-of-machine è un naturale proseguimento per le sezioni successive: da un lato, prepara il terreno per un'esplorazione approfondita del consumo edonistico/utilitaristico (2.3), dall'altro porta inevitabilmente alla fiducia come architettura multidimensionale (2.4). In questa tesi, infatti, la fiducia non sarà considerata un semplice atteggiamento pro o contro l'AI, ma invece una condizione economico-psicologica che determina la possibilità stessa che le raccomandazioni algoritmiche producano valore per il consumatore e per l'impresa.

### **Riferimenti bibliografici essenziali (APA 7<sup>a</sup> edizione)**

Hao, K. (2020, August 5). *AI is learning when it should and shouldn't defer to a human*. MIT Technology Review. [journals.sagepub.com](https://journals.sagepub.com)

Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The “word-of-machine” effect. *Journal of Marketing*, 86(1), 91–108. <https://doi.org/10.1177/0022242920957347>

## **2.3. Consumo edonistico e utilitaristico: cornice teorica e implicazioni per le raccomandazioni algoritmiche**

### **2.3.1. Perché distinguere tra edonismo e utilitarismo nelle decisioni di consumo**

Nella discussione sull'interazione del consumatore con i sistemi intelligenti nell'attuale dibattito accademico, la distinzione tra consumo edonistico e utilitaristico non è una

considerazione descrittiva, ma un vero e proprio quadro interpretativo per comprendere il comportamento economico. In termini molto concreti, la stessa scelta (l'acquisizione di un bene, l'accettazione di un consiglio, la delega di una decisione) non viene vissuta allo stesso modo a seconda che il risultato desiderato sia una gratificazione immediata, una soddisfazione sensoriale o simbolica, piuttosto che la risoluzione di problemi, migliorare l'efficienza, il risparmio sui costi o raggiungere un risultato funzionale (Botti & McGill, 2011). Questa differenza di “finalità” agisce come un filtro psicologico che influenza (i) il tipo di informazioni che il consumatore considera rilevanti, (ii) la facilità con cui il consumatore è in grado di giustificare la scelta, (iii) l'intensità emotiva associata all'esito e (iv) ) la sua disponibilità a lasciare spazio a fonti esterne di influenza, comprese, al giorno d'oggi, le raccomandazioni basate su algoritmi.

Questa sezione ha quindi un obiettivo specifico, in linea con la prospettiva economico-comportamentale di questa tesi: ricostruire i principali risultati della letteratura su edonismo e utilitarismo non come categorie statiche di prodotti, ma come modalità motivazionali e cornici decisionali che rimodellano la struttura della scelta stessa. Questa ricostruzione svolgerà in ultima analisi un ruolo in due direzioni diverse in questa tesi. Da un lato, contribuirà a costruire il quadro concettuale necessario per comprendere in quali contesti e per quali ragioni un individuo può reagire con maggiore fiducia o resistenza a una tecnologia che "raccomanda" (Capitolo 2, sezioni 2.4-2.5). Dall'altro, offre una base teorica per la distinzione sperimentale tra decisioni edonistiche e utilitaristiche che sarà utilizzata nell'analisi dello studio empirico (Capitolo 3), dove la fiducia nelle raccomandazioni dell'IA viene osservata proprio in funzione della natura del tipo di compito decisionale.

### **2.3.2. Definizioni operative: edonico e utilitaristico come attributi, come scopi e come contesti**

Una prima semplificazione, diffusa nel linguaggio comune, consiste nell'equiparare edonico a *superfluo* e utilitaristico a *necessario*. La letteratura, tuttavia, dimostra che questa equivalenza è fragile e fuorviante. Okada (2005) sostiene che edonismo e utilitarismo non vanno intesi come estremi di un'unica scala, ma piuttosto come costrutti sintetici che riassumono un insieme di significati e associazioni: un bene può essere sia edonico che utilitaristico, a seconda di come il consumatore lo percepisce nel contesto della scelta (Okada, 2005). L'idea di utilizzare la distinzione come costrutto teorico è stata ulteriormente sviluppata

da Khan, Dhar e Wertenbroch (2004), che la definiscono come un quadro che comprende la relazione tra preferenze esperienziali e obiettivi funzionali, e che le scelte edoniche sono più "affettivamente ricche" mentre le scelte utilitaristiche sono più "cognitivamente giustificabili" (Khan et al., 2004). Da questo punto di vista, ci sono tre livelli di interpretazione, da discutere insieme:

Livello degli attributi: alcuni beni sono generalmente considerati edonistici (incentrati sul piacere, sulla fantasia, sulla gratificazione sensuale o simbolica); altri sono generalmente considerati utilitaristici (incentrati sulla funzionalità, sui compiti o sui problemi). Questa è una distinzione utile, ma non esaustiva, poiché non tiene conto delle differenze individuali o contestuali.

Livello degli scopi: lo stesso bene può diventare edonico o utilitaristico a seconda del perché viene consumato. Botti e McGill (2011) insistono su questo punto: il fattore critico è la motivazione, che può essere terminale (fare qualcosa per il gusto di farlo, per il piacere dell'esperienza) o strumentale (fare qualcosa al servizio di un altro obiettivo). È qui che edonismo e utilitarismo acquisiscono pieno significato comportamentale (Botti e McGill, 2011).

Livello del contesto: la forma concreta della decisione, come valutare le alternative in sequenza o simultaneamente, acquistare o non acquistare e pagare in tempo o denaro, influenza la salienza delle dimensioni edonistiche e utilitaristiche e produce effetti sistematici e prevedibili (Okada, 2005; Dhar & Wertenbroch, 2000).

Questi tre livelli sono particolarmente rilevanti per la tesi perché le raccomandazioni algoritmiche intervengono proprio a livello di contesto; presentano alternative, rendono più evidenti alcune opzioni, trasformano i confronti impliciti in confronti espliciti e spesso riducono lo sforzo di ricerca. Pertanto, l'effetto dell'IA non può essere interpretato senza considerare come l'edonismo e l'utilitarismo influenzano la psicologia della scelta.

### **2.3.3. Il nodo della “giustificazione”: perché il piacere richiede una ragione**

Le scelte edonistiche sono più inclini a essere associate a sensi di colpa/autocritica e a un conseguente bisogno di giustificazione. Okada (2005) sviluppa la sua argomentazione da

una prospettiva intuitiva ma teoricamente rigorosa: divertirsi non è sempre psicologicamente gratuito; le persone trovano il piacere meno giustificabile razionalmente e socialmente della funzionalità. Da qui emergono le seguenti implicazioni: (i) le preferenze possono cambiare a seconda di quanto il contesto rende necessaria una giustificazione; (ii) i consumatori possono selezionare modalità di pagamento/scambio (tempo vs denaro), il che rende il piacere più giustificabile (Okada, 2005). Quest'ultima implicazione è direttamente rilevante per il legame tra edonismo/utilitarismo e le dinamiche della fiducia e dell'accettazione dei consigli esterni. Quando una decisione richiede giustificazione, il consumatore tende a cercare appigli razionali, segnali di legittimità, oppure scuse socialmente accettabili. Nel caso di un sistema di raccomandazione, questo significa che la semplice qualità predittiva dell'algoritmo potrebbe non essere sufficiente; deve anche svolgere il ruolo di motivare/spiegare/legittimare la raccomandazione; conciliando la tensione tra ciò che si vuole e ciò che si dovrebbe fare – una sfida che tornerà a farsi sentire nel Capitolo 2.4, nella discussione sulla fiducia nei sistemi intelligenti e sull'esigenza di trasparenza/interpretabilità.

#### **2.3.4. Preferenze che si capovolgono: valutazione separata e valutazione congiunta**

Un contributo particolarmente significativo di Okada (2005) è lo studio delle **reversal** di preferenza per scelte edoniche e utilitaristiche a seconda che le opzioni fossero valutate separatamente o congiuntamente: se un'opzione veniva valutata isolatamente, l'opzione edonica veniva ricordata come più attraente e la sensazione della scelta era più favorevole; se le scelte edoniche e utilitaristiche venivano presentate fianco a fianco, il semplice fatto dell'esistenza dell'alternativa utilitaristica rende più evidente la natura discrezionale della scelta edonica e ne rende più difficile la giustificazione, aumentando così la probabilità di scegliere invece l'opzione utilitaristica (Okada, 2005). Non si tratta tuttavia di un mero risultato sperimentale; descrive un meccanismo generale attraverso il quale la struttura del confronto influenza la moralità percepita di determinate scelte.

Applicata all'economia digitale, l'implicazione è immediatamente chiara: piattaforme e sistemi di raccomandazione istituzionalizzano questa forma di valutazione congiunta. Presentano alternative per il confronto, evidenziano differenze di prezzo e prestazioni, classificano le scelte (migliori scelte) e suggeriscono confronti (chi ha acquistato questo ha scelto anche...). In altre parole, aumentano la probabilità che il consumatore rifletta sulla scelta come un compromesso esplicito che ha un vantaggio naturale nella giustificazione: il criterio

utilitaristico del miglior valore. Ciò suggerisce una linea di ragionamento che verrà esplorata nuovamente nel quadro sperimentale del Capitolo 3: che la stessa raccomandazione algoritmica potrebbe essere più efficace se formulata in modo utilitaristico rispetto a quando deve essere giustificata in un contesto edonico, richiedendo un adattamento alla comunicazione di tale raccomandazione algoritmica.

### **2.3.5. Pagare in tempo o pagare in denaro: il prezzo psicologico della giustificazione**

Okada (2005) dimostra inoltre che la giustificazione si applica anche al lato del pagamento del bene: l'autore prosegue suggerendo che tempo e denaro non sono valute equivalenti, ma psicologicamente diverse: il denaro ha un valore percepito fisso e misurabile, mentre il tempo può essere una risorsa più sfocata, rendendolo quindi più razionalizzabile. Le implicazioni comportamentali sono che, se i consumatori si aspettano di dover giustificare un acquisto (in questo caso, comune per i beni edonici), potrebbero preferire pagare in tempo (impegno, sforzo, ricerca di sconti) piuttosto che in denaro; mentre pagare in denaro potrebbe rappresentare un ostacolo psicologico minore nel caso di un bene utilitaristico (Okada, 2005). Questo risultato è particolarmente rilevante per la tesi perché si riferisce a un ambito dei sistemi di raccomandazione solitamente sottovalutato: la *riduzione dei* costi di ricerca. Gli algoritmi fanno risparmiare tempo, ma questo risparmio di tempo non è percepito necessariamente come un beneficio netto in ogni contesto. Se il tempo dedicato alla ricerca è in realtà una forma di sforzo giustificato (Me lo sono meritato, Mi sono impegnato, non è un capriccio impulsivo), allora l'automazione potrebbe privare una parte della legittimità psicologica alla base del consumo edonico. Più direttamente, un sistema di raccomandazione che riduce lo sforzo può aumentare l'efficienza, ma vanificare la giustificabilità dell'edonico. Più pragmaticamente, un recommender che minimizza lo sforzo potrebbe aumentare l'efficienza, ma diminuire la giustificabilità dell'edonico. Questa tensione, tra efficienza algoritmica e psicologia della giustificazione, diventa un ponte naturale verso le sezioni sulla fiducia (2.4) e, cosa ancora più importante, sulla distinzione edonico/utilitaristico che guiderà le ipotesi del Capitolo 3.

### 2.3.6. Acquisire o rinunciare: quando la cornice di scelta rende l'edonico più “pesante”

Se Okada (2005) rivela l'effetto di un confronto esplicito e la necessità di giustificazione, Dhar e Wertenbroch (2000) svelano un altro aspetto non meno importante: la distinzione tra scelte di acquisizione e scelte di rinuncia (*forfeiture*). In teoria, si tratta della stessa scelta: scegliere un'opzione per un vantaggio edonistico rispetto a un'altra per un vantaggio utilitaristico. Ma la letteratura sulla *reference dependence* suggerisce che il cambiamento di direzione rispetto allo stato attuale (guadagnare qualcosa vs. perdere qualcosa) trasforma l'esperienza soggettiva della scelta. Dhar e Wertenbroch (2000) suggeriscono che nelle scelte di rinuncia, gli individui elaborino mentalmente più intensamente le conseguenze (attraverso prefattuali spontanei) e che, in questo processo, le conseguenze edoniche sono più facili da prevedere e quindi acquisiscono un peso complessivo maggiore (Dhar e Wertenbroch, 2000). Ne consegue un effetto sistematico: la preferenza relativa per l'edonico aumenta quando la scelta viene effettuata in un frame di rinuncia, mentre l'utilitaristico tende a prevalere quando la scelta viene effettuata in un frame di acquisizione.

Un altro aspetto della rilevanza metodologica di questo studio è che questa asimmetria non è dovuta a una generale "preferenza edonica", bensì a uno specifico meccanismo cognitivo: la diversa intensità di elaborazione mentale tra acquisizione e rinuncia. Gli autori dimostrano inoltre che l'effetto può essere ridotto da compiti che sopprimono o controllano la differenza di elaborazione, rafforzando l'interpretazione causale (Dhar & Wertenbroch, 2000). Questo è rilevante perché implica che le raccomandazioni algoritmiche possano funzionare in modo diverso a seconda del tipo di prodotto e del tipo di decisione che l'utente sente di prendere: sto scegliendo un nuovo prodotto o mi sto negando un'alternativa migliore? Sto *perdendo* un'esperienza migliore ignorando la raccomandazione? Inoltre, Dhar e Wertenbroch (2000) estendono la logica oltre il riferimento alla proprietà effettiva, dimostrando che il frame di acquisizione/rinuncia può essere indotto semplicemente attraverso una manipolazione del **reference point** a livello degli attributi (ad esempio, descrivendo l'opzione di riferimento come migliore o peggiore sia in termini edonistici che utilitaristici). Questo consente di trattare l'effetto come una forma di **loss aversion differenziale** tra attributi edonici e utilitaristici, per cui quando il riferimento è alto, la scelta assume il sapore della rinuncia e l'edonico "pesa" di più (Dhar e Wertenbroch, 2000).

### **2.3.7. Evidenze “di mercato”: il valore di scambio e il divario acquirente–venditore**

Un aspetto spesso ignorato dell'edonismo e dell'utilitarismo è che questi costrutti non solo spiegano le preferenze dichiarate in laboratorio, ma hanno anche conseguenze di mercato. Dhar e Wertenbroch (2000) associano l'asimmetria acquisizione/rinuncia alla dinamica del divario tra acquirente e venditore: se le considerazioni edoniche sono più importanti nella rinuncia, allora i proprietari di prodotti "edonici" dovrebbero essere meno disposti a separarsene, chiedendo prezzi di vendita relativamente più alti del valore di mercato (Dhar e Wertenbroch, 2000). Nella loro indagine sul mercato delle auto usate, gli autori trovano un risultato coerente con questa ipotesi: all'aumentare del valore edonico dell'auto percepito, aumenta anche il rapporto tra il prezzo di riserva richiesto dall'acquirente e il prezzo di riferimento di mercato (Dhar e Wertenbroch, 2000). Questo tipo di evidenza supporta il quadro economico-comportamentale che questa tesi intende adottare: infatti, edonico e utilitaristico non sono etichette, ma variabili che possono influenzare (i) la disponibilità a pagare, (ii) la disponibilità ad accettare, (iii) la dinamica dei prezzi e, in ultima analisi, (iv) la distribuzione del benessere tra consumatori e aziende, una considerazione che entra in gioco quando i sistemi di raccomandazione iniziano a influenzare la domanda e la visibilità dei prodotti.

Per le piattaforme digitali, questo sta a significare che l'IA non è neutrale: raccomandando prodotti con una forte componente edonica (intrattenimento, moda, esperienze), può rafforzare l'attaccamento, la riluttanza a rinunciare e la rigidità del valore percepito. Per i prodotti utilitaristici, può migliorare l'efficienza e ridurre gli sprechi con un minore attrito emotivo. Si tratta di un'ipotesi ampia, tuttavia, rappresenta un ponte logico verso la parte empirica della tesi, dove sarà importante prestare attenzione non solo a "quanto" le persone si fidano, ma anche a "cosa" stanno decidendo e al quadro interpretativo in cui si trovano.

### **2.3.8. Una visione integrativa: il modello decisionale e la “self-attribution” nelle scelte edoniche**

Khan, Dhar e Wertenbroch (2004) suggeriscono una sintesi teorica che aiuta a mettere ordine tra molte distinzioni affini: la dicotomia edonico/utilitaristico si affianca ad altre – vizi/virtù, desideri/doveri, affettivo/cognitivo, lussi/necessità – che vengono spesso utilizzate in modo intercambiabile ma non identico. Il contributo è importante per due ragioni.

In primo luogo, dimostrano che la scelta tra alternative edoniche e utilitaristiche non può essere compresa facendo appello alla razionalità cognitiva e relegando la componente emozionale in secondo piano.

Gli autori sostengono che la ricerca comportamentale ha a lungo enfatizzato l'errore cognitivo rispetto all'analisi normativa, ma è stata lenta ad abbracciare il lato emotivo che è al centro di molte decisioni dei consumatori (Khan et al., 2004).

In secondo luogo, la proposta più interessante è la formulazione di un modello di autoattribuzione della scelta edonica, in cui la scelta edonica è interpretata come un processo attraverso il quale gli individui costruiscono inferenze su se stessi e sui loro obiettivi di livello superiore (identità, valori, immagine personale) in modo tale che la scelta edonica diventi non solo una decisione di consumo, ma un processo decisionale finalizzato all'auto-significazione (Khan et al., 2004).

In questa prospettiva, la scelta edonica è raramente, se non mai, *solo* piacere, ma anche narrativa, giustificazione, coerenza identitaria o persino compensazione morale.

Questo modello ci permette di riformulare i risultati di cui sopra in modo unitario:

- l'inversione tra valutazione separata e congiunta (Okada, 2005) può essere vista come un cambiamento nella facilità di costruire un'autonarrazione coerente e difendibile;
- l'asimmetria acquisizione/rinuncia (Dhar & Wertenbroch, 2000) può essere vista come un cambiamento nella vividezza dell'immaginazione e quindi nel peso assegnato alla dimensione esperienziale nella costruzione del valore;
- la preferenza per il pagamento puntuale (Okada, 2005) può essere vista come un modo per *meritare* la scelta edonistica e per ridurre l'attrito identitario tra desiderio e obbligo.

Questa lettura integrativa è particolarmente rilevante nel mettere in relazione l'edonismo/utilitarismo con la questione della fiducia: se una scelta edonistica implica il sé e richiede giustificazione, allora il coinvolgimento di un agente esterno, che sia un algoritmo o qualsiasi altro mezzo di assistenza, può essere inteso come un'interferenza nell'autodefinizione della scelta.

### 2.3.9. Locus of choice e causality: quando delegare riduce (o non riduce) la soddisfazione

Mentre le sezioni precedenti hanno dimostrato l'impatto del contesto e della giustificazione sulla probabilità di prendere decisioni edoniche rispetto a quelle utilitaristiche, Botti e McGill (2011) introducono un ulteriore fattore: la distinzione tra scegliere e non scegliere (ovvero ricevere una decisione esterna) ha un impatto asimmetrico sulla soddisfazione e sul benessere nei due ambiti.

Nel loro articolo, gli autori forniscono prove – attraverso quattro studi – che la soddisfazione per un risultato è significativamente maggiore se la scelta è stata fatta autonomamente (anziché assegnata da un esperto o determinata dal caso) quando l'obiettivo della decisione è edonico, mentre la stessa distinzione non ha importanza (o conta meno) se l'obiettivo è utilitaristico (Botti e McGill, 2011). Le ragioni di questo effetto sono legate alla nozione di personal causality: nelle decisioni edoniche (terminali), l'individuo percepisce più intensamente di essere la causa dell'esito; quindi, la scelta autonoma aumenta il valore percepito dell'esperienza. Nelle decisioni utilitaristiche (strumentali), un secondo fine sposta questa prospettiva verso elementi esterni (il compito, la necessità, l'obiettivo). In altre parole, "sembra meno una scelta" (Botti & McGill, 2011, p. 1068) se la scelta è strumentale (Botti & McGill, 2011).

Questa scoperta è direttamente rilevante per la tesi perché i sistemi di raccomandazione sono, di fatto, strumenti che spostano parzialmente il *locus of choice*: non sempre decidono, ma spesso limitano la decisione, selezionano opzioni “per te”, ordinano alternative e spesso possono avere un impatto sulla scelta maggiore di quanto il consumatore pensi. Pertanto, un'ipotesi che potrebbe essere esplicitata (anche in relazione alle ipotesi del punto 2.5 e al lavoro empirico del capitolo 3) è che affidarsi (cioè delegare una decisione) a un sistema di raccomandazione algoritmico si tradurrà in una maggiore riduzione della soddisfazione (e potenzialmente anche della fiducia) quando la scelta è edonica, poiché la personal causality e la paternità dell'esperienza sono state ridotte. Un'altra implicazione più sottile emerge qui: se nelle decisioni utilitaristiche la differenza tra scegliere e non scegliere è ridotta, l'algoritmo può essere più facilmente accettato come "strumento" senza che ciò minacci l'identità decisionale dell'individuo. Questa asimmetria, se dimostrata vera, aiuta a spiegare perché in alcuni ambiti (ad esempio, finanza, ottimizzazione della spesa, scelte funzionali) ci si può fidare della tecnologia più rapidamente che in altri (intrattenimento, gusto, esperienze simboliche): una linea di interrogativi che è direttamente rilevante per la letteratura

sull'avversione/apprezzamento algoritmico (Capitolo 2.1) e per il modello Word-of-Machine (Capitolo 2.2).

### **2.3.10. Implicazioni per la fiducia nei sistemi di raccomandazione: una continuità verticale verso i Capitoli 2 e 3**

A questo punto, la distinzione tra edonismo e utilitarismo può essere tradotta in alcune implicazioni coerenti con l'impianto:

1. **La fiducia non dipende solo dall'accuratezza.** Nei contesti utilitaristici, l'accuratezza e la minimizzazione dello sforzo possono essere sufficienti a generare accettazione. Nei contesti edonistici, entrano in gioco giustificazione, identità e personal causality; pertanto, la fiducia può richiedere segnali aggiuntivi (spiegazioni, comportamento comunicativo, senso di comprensione, controllo).
2. **Le piattaforme favoriscono la valutazione congiunta.** Se il confronto esplicito favorisce l'utilitarismo (Okada, 2005), l'architettura digitale può orientare la domanda in modo sistematico verso scelte "difendibili" anche quando il consumatore desidererebbe l'edonico. Questo produce implicazioni di welfare e di benessere soggettivo: scegliere "il giusto" potrebbe non coincidere con scegliere "ciò che ti rende più felice".
3. **Delegare è psicologicamente pericoloso nel contesto edonico.** Se la scelta autonoma aumenta la soddisfazione solo quando sono in gioco obiettivi edonistici (Botti e McGill, 2011), l'IA può essere vista come una barriera laddove l'esperienza del consumatore è maggiormente legata all'identità. Questa nozione si collega in modo naturale alla parte successiva sulla fiducia nell'automazione e ai fattori psicologici (Capitolo 2.4) e alla formulazione di ipotesi (2.5).
4. **La cornice di acquisizione/rinuncia modifica il peso dell'edonico.** I frame percepiti come di rinuncia possono rafforzare l'edonico (Dhar & Wertenbroch, 2000). Per i

sistemi di raccomandazione, una raccomandazione identica può essere accettata o rifiutata a seconda del frame: "se non scegli questo, ti perdi X".

In breve, il consumo edonico e utilitaristico costituisce un terreno privilegiato per studiare la fiducia nelle raccomandazioni algoritmiche perché rende visibile la natura *non neutrale* della scelta economica: la decisione è anche giustificazione, agency, identità e gestione emotiva. Come verrà mostrato più dettagliatamente nel Capitolo 2 (sezione 2.4), la fiducia nelle tecnologie intelligenti è una condizione necessaria affinché **i benefici economici** dell'IA si realizzino pienamente; **ma tale fiducia non è uniforme e dipende, tra le altre cose, dalla natura edonica o utilitaristica del compito decisionale.**

#### **Riferimenti bibliografici essenziali (APA 7<sup>a</sup> ed.)**

Botti, S., & McGill, A. L. (2011). The locus of choice: Personal causality and satisfaction with hedonic and utilitarian decisions. *Journal of Consumer Research*, 37(6), 1065–1078. [OUP Academic](#)

Dhar, R., & Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. *Journal of Marketing Research*, 37(1), 60–71. [SAGE Journals+1](#)

Khan, U., Dhar, R., & Wertenbroch, K. (2004). *A behavioral decision theoretic perspective on hedonic and utilitarian choice* (Working Paper 2004/66/MKT).

Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of Marketing Research*, 42(1), 43–53.

## **2.4. Fiducia nei sistemi AI: dall'“appropriate reliance” nell'automazione alla fiducia nelle raccomandazioni algoritmiche**

### **2.4.1. Introduzione: perché la fiducia è la variabile-ponte tra potenziale tecnologico e impatto economico**

Nel contesto delle discussioni sulla trasformazione digitale, la fiducia è vista come una variabile "soft", ausiliaria rispetto a variabili più facilmente misurabili e modellabili: accuratezza, efficienza, riduzione dei costi, incremento di produttività. In realtà, l'oggetto di analisi è l'adozione di tecnologie intelligenti e, soprattutto, di affidamento su decisioni automatizzate/semi-automatizzate, la fiducia gioca un ruolo strutturale: poiché è il requisito che consente al potenziale tecnico di realizzarsi come comportamento e al comportamento di generare valore economico. Intuitivamente, un utente che non si fida di un sistema non lo usa; ma anche un utente che si fida troppo lo usa in modo inappropriato, con risultati a volte peggiori dell'assenza di tecnologia. In questo senso, la fiducia non è solo un antecedente dell'adozione; ma è anche un meccanismo regolazione dell'affidamento (*reliance*) e quindi determinante per la qualità dell'interazione uomo-macchina (Lee & See, 2004).

Questo sottocapitolo si colloca esattamente all'intersezione tra economia e psicologia comportamentale che questa tesi intende esplorare: da un lato, la fiducia come mezzo per rendere le transazioni più efficienti, ridurre i costi di ricerca, diminuire l'incertezza e facilitare l'adozione di sistemi basati sui dati; dall'altro, la fiducia come costrutto percettivo costruito da euristiche, costrutti sociali, aspettative e inferenze ricavate dalle "intenzioni" dell'agente tecnologico. Tale doppia natura appare particolarmente rilevante nel caso dei sistemi di raccomandazione: strumenti che operano come "assistenti virtuali" ma possono anche essere visti come estensioni del venditore/piattaforma e quindi agenti non neutrali e autoreferenziali (Wang & Benbasat, 2005).

In questo contesto, la Sezione 2.4 costruisce una teoria della fiducia in quattro fasi: (i) la fiducia è definita come la disponibilità ad assumersi rischi di fronte all'incertezza; (ii) recupera la tradizione della "trust in automation" e l'idea di *appropriate reliance*; (iii) questa viene aggiornata alle specificità dell'IA contemporanea (opacità, non determinismo, autonomia, forme di embodiment); (iv) la fiducia nell'IA è collegata alla fiducia nei recommendation agents, mostrando come dimensioni apparentemente "interpersonali" (benevolenza, integrità) siano attribuite anche agli artefatti tecnologici. In linea con il resto del Capitolo 2, questi fondamenti teorici forniranno quindi un senso al perché, secondo le Sezioni 2.1 e 2.2, il rapporto delle persone con l'algorithm oscilla tra avversione e apprezzamento.

#### 2.4.2. Che cos'è la fiducia: vulnerabilità, incertezza e aspettative di esito

Una definizione funzionale particolarmente utile in un contesto tecnologico definisce la fiducia come un atteggiamento basato sull'azione: l'aspettativa che un agente (umano o artificiale) faciliti il raggiungimento di un obiettivo in una situazione di incertezza e vulnerabilità. Lee e See (2004) propongono una formulazione particolarmente efficace: la fiducia è l'atteggiamento secondo cui l'agente faciliterà il raggiungimento di un obiettivo in una situazione di incertezza e vulnerabilità; e questa formulazione, proprio per queste ragioni, si applica alla guida della fiducia in sistemi complessi in cui la piena comprensione del sistema non è praticamente raggiungibile (Lee e See, 2004). Questa impostazione offre due vantaggi per l'analisi dei sistemi di intelligenza artificiale. In primo luogo, fonda la fiducia a livello di processo decisionale economico piuttosto che a livello di "simpatia". La fiducia implica l'esposizione al rischio (rinuncia del controllo, la delega di un compito, l'accettazione di una raccomandazione) sulla base dell'aspettativa di un risultato positivo. In secondo luogo, rende la fiducia concettualmente compatibile con l'automazione e con l'AI, evitando l'obiezione secondo cui la fiducia sarebbe riservata a entità dotate di coscienza e intenzionalità, poiché in termini comportamentali ciò che conta non è la vera intenzione dell'agente, bensì la percezione dell'utente: infatti, la fiducia è una costruzione cognitiva che organizza l'azione in condizioni di informazione incompleta.

È proprio qui che emerge un'ulteriore importante considerazione: nel rapporto con la tecnologia, la fiducia non può essere ridotta a un giudizio unidimensionale. Lee e See (2004) discutono come le definizioni di fiducia oscillino tra credenze, atteggiamenti, intenzioni e comportamenti, e sulla possibilità di distinguere tra livelli: credenze (beliefs) che alimentano atteggiamenti (attitudes), che generano intenzioni (intentions), fino a produrre un comportamento di reliance (Lee & See, 2004). Quindi la fiducia non coincide con l'uso; piuttosto agisce come una forza che orienta l'uso e ne determina l'appropriatezza. Questa distinzione appare decisiva perché se il problema fosse semplicemente aumentare la fiducia, la soluzione sarebbe la comunicazione persuasiva. In realtà, ciò che interessa in termini economici e ai fini di questa tesi è la calibrazione tra fiducia percepita e capacità reali del sistema; cioè un equilibrio dinamico che eviti sia la sottoutilizzazione (disuse) sia l'iperaffidamento (misuse). In altre parole, l'obiettivo non è aumentare la fiducia, ma piuttosto renderla adeguata all'affidabilità dell'IA e alla vulnerabilità del compito, qualcosa che, come vedremo, ritroveremo sia nella letteratura sull'automazione sia nei modelli più recenti di fiducia nell'IA (Lee & See, 2004; Glikson & Woolley, 2020).

### 2.4.3. La tradizione “trust in automation”: misuse, disuse e progettazione dell’affidamento appropriato

Lee e See (2004) costruiscono il loro lavoro sullo sfondo di un'automazione sofisticata – aviazione, navigazione, controllo di processo, guida – in cui l'interazione affidabile con l'uomo ha portato grandi benefici ma anche grandi disastri. Punto di partenza: l'automazione non uniformemente vantaggiosa; l'uso quando non si dovrebbe (misuse) o il mancato utilizzo (disuse) di capacità affidabili da parte degli esseri umani può causare il fallimento dell'automazione. Queste due situazioni rappresentano gli estremi di un reliance inappropriato, con costi economici e (in alcuni casi) umani (Lee e See, 2004).

Il concetto di *appropriate reliance* rappresenta un ponte straordinariamente utile verso l'intelligenza artificiale contemporanea. Anche ora, infatti, non si tratta di creare sistemi più accurati, ma piuttosto di progettare interazioni che inducano le persone a utilizzarli nel modo giusto; ad affidarsi alle competenze e a intervenire sulle incompetenze, per mantenere un controllo adeguato in situazioni ad alto rischio. Lee e See (2004) osservano che la fiducia alimenta la reliance in situazioni in cui il sistema è troppo complesso e l'ambiente troppo unico per poter adottare un approccio analitico completo alla situazione; pertanto, la fiducia diventa un'euristica: una scorciatoia cognitiva che permette di decidere se delegare o meno (Lee & See, 2004).

È particolarmente rilevante la natura interdisciplinare del concetto: Lee e See (2004) ricostruiscono la fiducia attraverso prospettive organizzative, sociologiche, interpersonali, psicologiche e persino neurologiche, proprio per mostrare che la fiducia non può basarsi su un singolo meccanismo, ma piuttosto emerge dall'interazione tra contesto, caratteristiche dell'automazione e processi cognitivi. Infatti, nella loro sintesi, le caratteristiche dell'automazione possono influenzare la fiducia attraverso vari processi – analitici, analogici e affettivi – e la progettazione deve tenere conto di questa pluralità (Lee e See, 2004).

Emerge qui un primo snodo, utile a questa tesi in una chiave economico-gestionale: la fiducia è un interruttore che regola l'intensità dell'impiego della tecnologia, ma al contempo è anche il prodotto dell'esperienza e della comunicazione del sistema. Se l'interfaccia rende chiari quali sono i limiti, le incertezze e le condizioni d'uso, la calibrazione è più facile; laddove il

sistema appare infallibile o viene comunicato come tale, aumenta il rischio di essere utilizzato impropriamente (misuse). Questa dicotomia è estremamente rilevante nell'intelligenza artificiale: le moderne tecniche di utilizzo dell'intelligenza artificiale —specialmente quelle basate su machine learning— possono apparire "magiche" nelle dimostrazioni, ma si deteriorano rapidamente quando vengono inserite in un nuovo contesto. La fiducia diventa quindi un problema di governance dell'interazione: cioè come distribuire controllo, responsabilità e comprensione tra utente e sistema.

#### **2.4.4. Dall'automazione tradizionale all'AI: perché la fiducia diventa più difficile (e più centrale)**

Se la fiducia nell'automazione era già complicata, la fiducia nell'intelligenza artificiale presenta ulteriori complicazioni. Glikson e Woolley (2020) sottolineano un aspetto cruciale, secondo cui l'IA attuale non è semplicemente una forma più sofisticata di automazione deterministica, quanto piuttosto una tecnologia che apprende, si adatta e produce comportamenti non del tutto prevedibili e spesso opachi per l'utente. In altre parole, la relazione input-output è meno chiara e la logica del processo decisionale è spesso difficile da spiegare, sia perché il processo è multilivello, sia perché l'IA può incorporare bias e dipendere dalla qualità del training data (Glikson e Woolley, 2020). La revisione di Glikson e Woolley (2020) offre un beneficio, sia metodologico sia teoretico, integrando le evidenze empiriche provenienti da diversi campi e offrendo al contempo un quadro che distingue le diverse forme di rappresentazione dell'IA (robotica, virtuale, embedded) e il livello percepito di machine intelligence, evidenziando che questi fattori sono antecedenti della fiducia. Non è indifferente, ad esempio, se si percepisce l'IA come un robot fisico, un chatbot o un algoritmo "invisibile" incorporato in un sistema in cui i segnali disponibili, le aspettative e i mezzi di attribuzione di agency differiscono (Glikson & Woolley, 2020).

Un'ultima distinzione, di grande utilità per la tesi, è quella tra fiducia cognitiva e fiducia emotiva. Gli autori, richiamando la tradizione che distingue componenti cognitive (affidabilità e valutazione delle competenze) da quelle emotive (sensazioni di sicurezza, comfort nella relazione e antropomorfismo) della fiducia, dimostrano che le leve della fiducia non sono le stesse: trasparenza, affidabilità e segnali di immediatezza comportamentale influenzano maggiormente la fiducia cognitiva, mentre l'antropomorfismo e alcuni segnali sociali influenzano maggiormente la fiducia emotiva (Glikson & Woolley, 2020). A questo punto, sta

diventando chiaro perché la fiducia sia la variabile ponte tra potenziale economico e impatto; più l'IA è potente e autonoma, più può creare valore. Ma allo stesso tempo, più è autonoma e opaca, maggiore è l'incertezza soggettiva e dunque cresce il bisogno di fiducia. È un paradosso pratico: la tecnologia più avanzata richiede, per essere adottata, un capitale psicologico e istituzionale più robusto.

#### **2.4.5. Le dimensioni percettive della fiducia: competenza e calore come grammatica universale del giudizio**

Per inquadrare la fiducia come un fenomeno psicologico, e non semplicemente come una reazione tecnica all'accuratezza, introduciamo un modello di psicologia sociale: le persone valutano gli altri (individui, gruppi, a volte persino agenti non umani) lungo due dimensioni fondamentali: competenza e calore (Cuddy et al., 2008, pp. 63–65). Cuddy, Fiske e Glick (2008) dimostrano che queste dimensioni sono universali nella percezione sociale: la competenza riguarda la capacità di perseguire le intenzioni; il calore si riferisce all'intenzione percepita, inclusi tratti come moralità, affidabilità, sincerità, benevolenza e affidabilità (Cuddy et al., 2008).

Sebbene questo modello sia stato creato per valutazioni interpersonali e intergruppo, si applica straordinariamente bene alle interazioni uomo-IA, per due motivi. Primo: molte interazioni con l'AI sono socialmente simulate: interfacce conversazionali, assistenti vocali, agenti che spiegano, suggeriscono, rispondono. Secondo: anche quando l'IA è invisibile, l'utente umano tende a inferire l'intento e l'interesse: mi sta davvero aiutando? Sta promuovendo un prodotto sponsorizzato? - e queste inferenze, pur non rappresentando l'intento effettivo della macchina, risultano essere fattori determinanti per la fiducia.

In questa tesi, la combinazione di competenza e calore può essere utilizzata operativamente: la competenza è il pilastro più intuitivo della fiducia nell'IA (accuratezza, prestazioni, affidabilità), mentre il calore denota la dimensione più delicata ma spesso trascurata: il sistema agisce per me, piuttosto che contro di me o a vantaggio di terze parti? Questo è particolarmente rilevante per i sistemi di raccomandazione, poiché l'utente potrebbe essere preoccupato per le raccomandazioni di parte: raccomandazioni che sono più vantaggiose per la piattaforma che per l'utente.

Parallelamente, Cuddy et al. (2008) parlano anche della "primacy of warmth" (p. 89): in molti ambiti, gli esseri umani valutano prima le intenzioni (calore), solo poi le capacità (competenza). Applicato all'IA, questo suggerisce anche un'ipotesi potenzialmente cruciale: in contesti molto sensibili (investimenti, privacy, spesa rilevante, rischio), l'utente potrebbe chiedersi prima: Posso fidarmi delle sue intenzioni? e solo dopo: È davvero bravo? (Cuddy et al., 2008, pp. 63, 89).

#### **2.4.6. Fiducia in AI e fiducia nei recommendation agents: quando la tecnologia diventa "attore sociale"**

Questa linea di ragionamento trova supporto diretto nella letteratura sugli agenti di raccomandazione. Wang e Benbasat (2005) affrontano una questione ancora attuale: la natura della fiducia in artefatti tecnologici è controversa – possiamo davvero attribuire fiducia a una tecnologia? – Tuttavia, gli utenti, empiricamente, trattano gli agenti come attori sociali e attribuiscono loro caratteristiche tipicamente umane. La ricerca parte dal presupposto che gli recommendation agents aiutano gli utenti mitigando l'information overload e migliorando il decision-making online; ma la loro adozione dipende in modo critico dalla fiducia, poiché l'utente delega di fatto parte del processo decisionale all'agente che agisce per conto dell'utente (Wang & Benbasat, 2005).

Qui emerge un'altra questione che è rilevante per questa tesi: la fiducia negli agenti di raccomandazione non è semplicemente una questione di competenza tecnica (il sistema comprende le mie preferenze?; capisce i miei gusti?), ma include dimensioni relazionali come integrità e benevolenza. Secondo Wang e Benbasat (2005), gli utenti potrebbero temere che l'agente operi nell'interesse del merchant o del produttore, piuttosto che nell'interesse del consumatore; quindi, la fiducia nell'agente implica convinzioni su (i) competenza, (ii) benevolenza e (iii) integrità, parallelamente alla fiducia tra esseri umani (Wang & Benbasat, 2005).

Lo studio propone e testa un modello integrato Trust-TAM, dimostrando che sia la percezione dell'agente come strumento utile (perceived usefulness) sia la fiducia nell'agente come assistente virtuale contribuiscono all'intenzione di adozione. In altre parole, non è sufficiente che l'agente sia efficiente, ma è anche necessaria una valutazione dell'affidabilità

relazionale. Questa implicazione è particolarmente rilevante per questa tesi nel collegare la fiducia alla trasformazione dei mercati: se il mercato digitale si sta effettivamente spostando verso un utilizzo algoritmico basato sulle raccomandazioni, l'uso dipenderà dalla fiducia nell'agente algoritmico, non solo dall'efficacia nella riduzione dei costi di ricerca (Wang & Benbasat, 2005).

Tuttavia, partendo dalla discussione del punto 2.3, la trama si infittisce poiché benevolenza e integrità sono costrutti legati alla psicologia della giustificazione: infatti, quando la scelta è delicata o identitaria (ad esempio, edonica), l'utente potrebbe essere meno disposto a delegare e più interessato alle intenzioni rispetto a una mera decisione strumentale (utilitaristica) in cui la competenza regna sovrana. Questo sarà un quadro concettuale rilevante sia per formulare le ipotesi (Sezione 2.5) sia per comprendere i risultati empirici (Capitolo 3).

#### **2.4.7. Determinanti della fiducia nell'AI: affidabilità, trasparenza, tangibilità, antropomorfismo e comportamenti di immediatezza**

La revisione di Glikson e Woolley (2020) consente di ordinare i determinanti della fiducia nell'IA in modo più dettagliato rispetto alla tradizione dell'automazione. Essi mostrano che la fiducia dipende dalle caratteristiche dell'IA e da come viene presentata e vissuta dall'utente. Tra i determinanti più ricorrenti, emergono: tangibilità (quanto l'AI è presente e percepibile), trasparenza (quanto è comprensibile il processo o almeno i criteri), affidabilità (consistenza della performance), immediacy behaviors (segnali sociali/comunicativi che riducono la distanza) e antropomorfismo, con effetti differenti sulla fiducia cognitiva ed emotiva (Glikson & Woolley, 2020).

Un aspetto importante della metodologia che vale la pena menzionare nella tesi riguarda la qualità delle prove: Glikson e Woolley (2020) notano limitazioni nel terreno empirico, come misure eterogenee di fiducia o l'eccesso di studi a breve termine e sperimentali. Sulla base di queste osservazioni, gli autori arrivano alla conclusione che la fiducia si può evolvere diversamente in contesti di lungo periodo e ad alta posta in gioco. Ciò è rilevante perché la fiducia nei sistemi di raccomandazione, nelle piattaforme e negli assistenti AI non è qualcosa che si crea dall'oggi al domani, ma al contrario, è qualcosa che si crea nel tempo; la fiducia può essere infranta, ricostruita e tende a basarsi su casi di errore (Glikson e Woolley, 2020).

Lee e See (2004), da parte loro, forniscono una grammatica complementare, secondo cui la fiducia è influenzata non solo dalle caratteristiche di automazione, ma anche dai processi cognitivi che mediano la valutazione (analitico, analogico, affettivo). In un'ottica d'insieme, potremmo dire che l'affidabilità conduce alla fiducia attraverso processi analitici (evidenza delle prestazioni), mentre l'antropomorfismo e i comportamenti comunicativi conducono alla fiducia attraverso processi analogici e affettivi, avvicinando l'interazione a un modello interpersonale (Lee e See, 2004; Glikson e Woolley, 2020).

In questa fase, il quadro calore/competenza (Cuddy et al., 2008) può fungere da cornice di lettura: affidabilità e trasparenza aumentano la competenza percepita; benevolenza, integrità e alcuni segnali comunicativi aumentano il calore percepito; e l'esito fiducia emerge spesso quando entrambe le dimensioni superano una soglia minima. Dove il calore manca (come quando l'utente percepisce una motivazione commerciale), la fiducia può crollare anche in presenza di elevata competenza; e laddove manca la competenza (come nel caso di errori ripetuti), il calore non è sufficiente a sostenere l'affidamento.

#### **2.4.8. Calibrazione della fiducia: undertrust, overtrust e qualità dell'esito decisionale**

Una delle intuizioni più concrete della letteratura sulla fiducia nell'automazione è la sua insistenza sul concetto di calibrazione. Lee e See (2004) spiegano come la fiducia non determini completamente la reliance ma la guidi; e come l'obiettivo sia quello di evitare il **disuse** (il rifiuto di capacità utili) e il **misuse** (affidamento inappropriato). Nel caso dell'IA, questa logica diventa ancora più pressante: gli utenti possono sottovalutare la tecnologia per scetticismo o per esperienze iniziali negative; oppure possono sopravvalutarla in virtù della sua "aura" di intelligenza, marketing e bias della novità. Entrambi gli errori generano inefficienze: nel primo caso si rinuncia ai benefici potenziali (risparmio di tempo, decisioni migliori, personalizzazione); nel secondo caso, si aumentano rischi (decisioni sbagliate, eccessiva dipendenza, perdita di autonomia). Glikson e Woolley (2020) notano che i modelli di fiducia nella tecnologia possono anche essere diversi dalla fiducia interpersonale: mentre in molte relazioni umane la fiducia tende ad aumentare con interazioni ripetute, con la tecnologia questa può diminuire quando si verificano errori e malfunzionamenti; e con l'intelligenza artificiale le

traiettorie possono essere ancora più variabili, oscillando tra scetticismo iniziale e aumento della fiducia in seguito all'esperienza diretta, o fiducia iniziale e delusione in seguito agli errori (Glikson e Woolley, 2020).

Negli agenti di raccomandazione, Wang e Benbasat (2005) hanno scoperto che l' *initial trust* è importante: nelle interazioni iniziali, quando l'utente ha pochi elementi su cui basarsi, la fiducia iniziale serve a superare il rischio e l'incertezza, e può determinare il successo o il fallimento dell'interazione continuativa dell'utente. Ciò ha evidenti conseguenze economiche per piattaforme e aziende: l'adozione da parte dell'utente è precaria e i costi di switching sono bassi; quindi, la costruzione di fiducia iniziale diventa una leva competitiva (Wang & Benbasat, 2005).

In una prospettiva più ampia, la calibrazione della fiducia oltre ad essere una sfida individuale, rappresenta anche una sfida sistemica: infatti quando ampie popolazioni di utenti si fidano eccessivamente di sistemi opachi, i rischi sociali (disinformazione, polarizzazione, discriminazione) aumentano; quando, invece, si fidano troppo poco, la diffusione di innovazioni utili si rallenta. In termini di policy e governance, la "fiducia appropriata" diventa un obiettivo collettivo, non soltanto un esito psicologico.

#### **2.4.9. Fiducia, spiegazioni e percezione di imparzialità: il problema della benevolenza nei mercati digitali**

Sebbene la competenza tecnica sia una condizione necessaria, la fiducia nei mercati digitali richiede spesso un'altra garanzia: la percezione che il sistema non sia manipolativo, opportunistico o non allineato con gli interessi del consumatore. Wang e Benbasat (2005) esprimono questo dilemma in modo piuttosto schietto: l'utente potrebbe chiedersi se l'agente agisca davvero per lui, poiché opera sui siti di e-vendor ed è probabilmente allineato con gli interessi del merchant. È qui che emergono Benevolenza e integrità che occupano un posto centrale nella fiducia negli agenti di raccomandazione, e non solo come una bella metafora: sono costrutti che si riferiscono al conflitto di interessi percepito, alle raccomandazioni potenzialmente distorte e, in generale, alla sensazione di essere spinti verso direzioni non ideali dal punto di vista dell'utente (Wang e Benbasat, 2005). In particolare, la classificazione di Cuddy et al. (2008) fornisce confini concettuali a queste idee: benevolenza e integrità sono

collegate a dimensioni di calore: il giudizio sulle intenzioni. Il consumatore non si chiede soltanto se l'AI *sa*, ma anche se l'AI *vuole* fare qualcosa di buono per lui, o se, al contrario, lo guida in direzioni poco chiare. L'attribuzione di intenzionalità non implica che la tecnologia possieda intenzioni. Significa che l'utente tratti l'IA come l'interfaccia di un sistema sociotecnico, in cui organizzazioni, incentivi e strategie sono in gioco da qualche parte dietro l'algoritmo (Cuddy et al., 2008; Wang & Benbasat, 2005).

Glikson e Woolley (2020) ci permettono di approfondire questo punto, evidenziando come l'incertezza venga alimentata dall'opacità e dal non determinismo dell'IA e il venir meno della comprensione porti gli utenti a colmare le lacune attraverso la proiezione di inferenze e narrazioni. In questo processo narrativo, la fiducia emotiva può essere alimentata da segnali sociali (antropomorfismo, comunicazione), ma di contro può essere minata dalla paura di pregiudizi, sorveglianza e perdita di controllo. Pertanto, la questione delle spiegazioni e della trasparenza – sebbene non discussa qui in senso tecnico – è correlata alla fiducia, in quanto l'utente non desidera necessariamente comprendere un modello complicato, ma cerca piuttosto segnali di responsabilità, equità e controllabilità (Glikson e Woolley, 2020).

#### **2.4.10. Verso un modello integrato: fiducia come funzione di performance, percezioni sociali e architettura di interazione**

A questo punto, è possibile suggerire, in modo coerente con le fonti, una lettura integrata utile alla tesi.

##### **(a) La performance è la base, ma non è autosufficiente.**

Lee e See (2004) dimostrano che la fiducia porta all'affidamento in condizioni di complessità e incertezza; quindi, la performance e l'affidabilità percepita sono importanti. Ma proprio perché la fiducia spesso funziona come un'euristica, la performance non viene sempre letta in modo freddamente statistico: episodi salienti di errore possono contare più della media; la mancata fornitura di qualsiasi spiegazione può essere considerata negativa; e questo può rendere una persona più consapevole dei rischi coinvolti.

##### **(b) Le percezioni sociali organizzano la fiducia emotiva.**

Glikson e Woolley (2020) fanno un'importante distinzione tra fiducia cognitiva ed emotiva; e Cuddy et al. (2008) scoprono che le valutazioni sociali sono strutturate lungo le dimensioni di competenza e calore. L'IA può essere percepita in questi termini, in particolare quando assume la forma di agenti conversazionali: competente ma fredda (accurata ma *senza cuore*), o calda ma incompetente (interfaccia piacevole ma inefficace). Infatti, la fiducia più forte sembra instaurarsi laddove la competenza è accompagnata da segnali di benevolenza, o almeno dall'assenza di sospetti di motivazioni opportunistiche.

**(c) L'architettura dell'interazione decide la calibrazione.**

Wang e Benbasat (2005) dimostrano che la fiducia insieme all'utilità percepita guidano l'adozione: infatti, l'AI deve essere utile e deve essere ritenuta affidabile come assistente. Pertanto, la progettazione dell'interfaccia, il feedback, la gestione degli errori e l'incertezza forniscono calibrazione: l'architettura dell'interazione può evitare l'eccessiva o la scarsa fiducia e garantire un adeguato reliance.

Questo schema integrato è particolarmente adatto a sostenere le scelte di ricerca del Capitolo 3. Nello studio empirico del Capitolo 3, la fiducia non si misura semplicemente rispondendo alla domanda *Ti fidi?*. Si osserva anche se gli utenti seguono le raccomandazioni e come tale adesione cambi in base al tipo di decisione (utilitaristica o edonica), alla competenza percepita dell'IA e al fatto che l'IA sia percepita come un alleato (benevolenza/integrità). In breve: la fiducia è il meccanismo psicologico che trasforma un output algoritmico in azione di mercato.

**2.4.11. Implicazioni per le raccomandazioni algoritmiche e “ponte” verso la formulazione delle ipotesi**

Integrando i punti precedenti, il punto 2.4 può essere enunciato in alcune proposizioni che costituiscono un punto di partenza naturale per il punto 2.5.

1. *La fiducia è una condizione necessaria, ma non deve essere massimizzata indiscriminatamente.*

La letteratura su trust in automation mette in guardia contro il misuse e il disuse, infatti, la fiducia deve essere calibrata sulla capacità del sistema e anche sul rischio del compito (Lee & See, 2004).

2. *L'AI intensifica la domanda di fiducia perché aumenta l'opacità e l'autonomia.*

L'IA è diversa dall'automazione deterministica, in realtà, può essere (e spesso lo è) opaca e non pienamente prevedibile. Per questo motivo la fiducia è fondamentale per la sua integrazione nei contesti organizzativi e decisionali (Glikson & Woolley, 2020).

3. *Nei recommender systems, la fiducia include dimensioni sociali: come competenza, benevolenza, integrità.*

Gli utenti trattano gli agenti come assistenti virtuali e attribuiscono su di loro credenze tipicamente interpersonali; la fiducia insieme all'utilità percepita e alla facilità d'uso rappresenta un fattore determinante ai fini dell'adozione (Wang & Benbasat, 2005).

4. *Le percezioni di competenza e calore offrono una grammatica interpretativa della fiducia.*

La fiducia si costruisce sulla valutazione delle capacità e delle intenzioni; in molti contesti, le intenzioni percepite (calore/trustworthiness) possono prevalere sulle capacità, soprattutto nei contesti in cui sono in gioco vulnerabilità e rischio (Cuddy et al., 2008).

A partire da questi presupposti, la sezione successiva (2.5) potrebbe generare ipotesi più precise: ad esempio, che la fiducia e l'adesione alle raccomandazioni IA dipendano entrambe dal tipo di decisione (utilitaristica vs. edonica come discusso nella sezione 2.3) e che la percezione di competenza e la benevolenza percepite fungano da mediatori/moderatori. La fiducia non è semplicemente "fidarsi dell'IA": è un costrutto multifattoriale che collega performance, percezione sociale e architettura delle scelte; in altre parole, la fiducia è ciò che rende l'economia delle raccomandazioni algoritmiche un terreno privilegiato per un'analisi economico-comportamentale.

### **Riferimenti bibliografici essenziali (APA 7<sup>a</sup> ed.)**

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Elsevier.

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

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.

Wang, W., & Benbasat, I. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, 6(3), 72–101.

## **2.5 Modello concettuale e ipotesi di ricerca: dalla teoria sulla fiducia alla verifica empirica**

Alla luce della letteratura esaminata nei precedenti sottocapitoli (2.1–2.4), questo sottocapitolo ha un obiettivo specifico: tradurre la teoria (avversione/apprezzamento degli algoritmi, *word-of-machine*, fiducia nell'automazione e dimensioni socio-cognitive come competenza e calore) in un modello concettuale con ipotesi verificabili. *L'idea di fondo* è che la fiducia dovrebbe essere intesa non come un atteggiamento astratto, ma come una condizione economico-comportamentale che determina se e quanto le raccomandazioni algoritmiche incidano davvero sulle scelte e quindi sul valore creato (o non creato) nei mercati digitali (adozione, conversione, fidelizzazione, benessere). In altre parole, se da un punto di vista puramente tecnico l'IA può formulare raccomandazioni eccellenti, di contro in assenza di un livello sufficiente di fiducia (opportunamente calibrata, non cieca) nella raccomandazione algoritmica, l'impatto economico si riduce o si distorce: assistendo così al disuso (rifiuto dell'algoritmo), misuso (sovra-affidamento) o alla calibrazione inefficiente della dipendenza dalla tecnologia.

In linea con l'orientamento generale di questa tesi, l'analisi esplora un contesto in cui la fiducia è più palesemente messa a dura prova: *la valutazione di un consiglio* proveniente da diverse fonti (IA vs. consulente umano vs. partnership ibrida IA+umano), e applicato a *domini decisionali differenti* che evocano diversi registri psicologici nei soggetti (più analitici vs. più basati sul valore/prosociali). La logica, ancora una volta coerente con la letteratura, è che i soggetti non rispondono all'automazione allo stesso modo nei diversi contesti: la stessa IA può essere percepita come autorevole e desiderabile in alcune decisioni, e inadeguata, o fredda, in altre.

### 2.5.1. Costrutti chiave: fiducia, competenza, calore e propensione a seguire il consiglio

Per costruire ipotesi coerenti, può essere utile specificare quali dimensioni compongono la risposta dell'individuo a una raccomandazione.

Fiducia nel consiglio e nella fonte (*trust*). Nella prospettiva dell'automazione, la fiducia è un atteggiamento che regola l'*appropriate reliance*, ovvero quanto e quando un utente si affida a un sistema, in base alle aspettative di performance e del rischio percepito. È un equilibrio delicato: infatti, una fiducia troppo bassa porta al rifiuto di un buon consiglio; mentre troppa fiducia porta all'accettazione di un consiglio sbagliato e manipolativo.

Competenza percepita (*competence/ability*). La competenza si riferisce alla convinzione che la fonte sia competente, accurata e "brava" in ciò che fa. Nelle decisioni economiche, questa dimensione è solitamente la prima ad essere attivata: se l'utente crede nella superiorità della fonte, sarà incline a darle maggiore importanza.

Calore/empatia percepita (*warmth*). Gli agenti sociali vengono valutati dagli individui su due grandi dimensioni: competenza e calore (intesi come benevolenza, vicinanza e buone intenzioni). Quando i valori, l'identità o le questioni morali sono parte della decisione, il calore potrebbe diventare determinante per la fiducia: è necessario "avere ragione", ma anche essere percepiti come *allineati e attenti* alla persona.

Intenzione di seguire il consiglio (*advice taking*). Dal punto di vista economico-comportamentale, la variabile più vicina all'esito di mercato non è tanto la fiducia in sé, quanto l'intenzione di comportarsi secondo il consiglio (seguire o meno la raccomandazione).

Infine, per ristabilire un collegamento con la dimensione economica di questa tesi, come discusso nel Capitolo 1, si può osservare che il valore generato da un sistema di raccomandazione non deve essere necessariamente correlato alla sua accuratezza tecnica. La ricerca su value-aware recommender systems evidenzia che le piattaforme di raccomandazione ottimizzano per una varietà di obiettivi (valore per l'impresa, valore per l'utente, equità, diversità, orizzonte di lungo periodo) e che un eccesso di sbilanciamento verso il valore business può ritorcersi contro in termini di percezioni di ingiustizia e abbandono della piattaforma – tutte conseguenze che in ultima analisi, passano proprio dalla fiducia.

## 2.5.2 Variabili esplicative: fonte del consiglio e dominio decisionale

Il modello proposto combina due leve principali.

### **(a) Fonte del consiglio (3 livelli):**

1. *IA* (algoritmo che analizza dati e produce raccomandazioni);
2. *Umano* (consulente/esperto);
3. *IA + Umano* (partnership: l'IA elabora e l'umano interpreta/spiega).

Questa scelta è coerente con due evidenze teoriche discusse in precedenza. Da un lato, l'algoritmo può essere visto come "più oggettivo, coerente e in grado di processare informazioni complesse"; dall'altro, l'essere umano è spesso associato a empatia, responsabilità morale e sensibilità contestuale. Il caso ibrido è interessante perché - almeno in teoria - racchiude i pregi di entrambi e riduce le rispettive debolezze: è una configurazione coerente sia con la nozione di "calibrazione della fiducia" nell'automazione sia con i risultati manageriali relativi alla co-progettazione uomo-macchina.

### **(b) Dominio della decisione (2 livelli):**

1. *Investimento finanziario* (decisione tendenzialmente analitica, orientata alle performance e al rischio);
2. *Donazione/beneficenza* (una decisione prosociale, più orientata al valore e di natura emotiva).

La distinzione è coerente con la letteratura che mostra che l'accettazione dell'IA varia a seconda del tipo di compito: quando il compito da svolgere è percepito come "oggettivo" e misurabile, le persone potrebbero essere più inclini ad attribuire maggiore legittimità all'algoritmo; quando invece il compito richiede sensibilità morale o comprensione umana, l'algoritmo può essere penalizzato.

### 2.5.3 Meccanismi teorici: avversione, apprezzamento e “word-of-machine”

Le ipotesi del modello poggiano su tre famiglie di meccanismi.

1. *Algorithm aversion*: il costo psicologico dell'errore algoritmico.

L'avversione all'algoritmo è la riduzione della propensione a utilizzare un sistema algoritmico dopo aver osservato l'errore, spesso giudicato più severamente rispetto ad un errore umano. Il meccanismo è particolarmente rilevante in situazioni di elevata incertezza o di elevato coinvolgimento personale. L'errore non è solo un evento tecnico, ma un segnale di inadeguatezza di principio.

2. *Algorithm appreciation*: quando l'algoritmo viene preferito.

Allo stesso modo, la letteratura mostra che esiste anche una spinta opposta: in molte decisioni, gli esseri umani possono attribuire maggiore affidabilità a un algoritmo rispetto a un essere umano, soprattutto quando lo considerano più accurato, più coerente o meno soggetto a bias. Questa linea di ragionamento aiuta a evitare un'interpretazione unilaterale: infatti, non esiste un rifiuto automatico dell'IA, ma invece un equilibrio che dipende dal contesto e dalle rappresentazioni mentali della fonte.

3. *Word-of-machine*: il peso del frame e delle aspettative.

Un aspetto spesso trascurato è che la stessa raccomandazione può essere interpretata in modo diverso a seconda della sua fonte: una macchina rispetto a un essere umano. Il *word-of-machine effect* attribuisce maggiore importanza ai consigli forniti da una macchina/algoritmo, soprattutto in contesti di raccomandazione in cui sono in gioco l'obiettività e la competenza tecnica. Questo approccio interpretativo è particolarmente rilevante per la tesi in quanto si collega l'economia dei dati (Cap. 1) alla psicologia economica: la fiducia è anche un prodotto di narrazioni, aspettative e reputazioni delle tecnologie.

## 2.5.4 Ipotesi di ricerca

Sulla base di quanto sopra, si propongono le seguenti ipotesi.

### **H1 (Effetto della fonte nel dominio "analitico").**

Nel dominio degli investimenti finanziari, la fonte IA produrrà, in media, livelli più elevati di competenza percepita rispetto alla fonte umana. Tale superiorità percepita si tradurrà in maggiore fiducia nel consiglio dell'IA e inoltre maggiore intenzione di seguire tale consiglio. Tutto ciò è in linea con i meccanismi di "algorithm appreciation" e con l'aspettativa di oggettività che è associata alla macchina.

### **H2 (Effetto della fonte nel dominio "valore/prosociale").**

Nel dominio della donazione/beneficenza, la fonte umana produrrà, in media, livelli più elevati di calore/empatia percepita rispetto alla fonte IA. Questa dimensione incrementerà la fiducia oltre alla propensione a seguire il consiglio umano, poiché la decisione prosociale richiede allineamento di valori e intenzioni benevole percepite.

### **H3 (Vantaggio della partnership IA+umano).**

La condizione ibrida (IA + consulente umano) genererà i livelli più elevati (e/o i livelli più stabili in tutti i domini) di fiducia nella consulenza, poiché tenderà a massimizzare – simultaneamente – sia la competenza percepita (il contributo dell'IA) sia il calore/legittimazione sociale (il ruolo dell'essere umano che spiega, contestualizza e "si fa carico"/condivide la responsabilità dei suggerimenti). In altre parole, questa partnership potrebbe bilanciare avversione e apprezzamento nei confronti dell'IA, favorendo una dipendenza più calibrata dalla tecnologia.

### **H4 (Interazione fonte × dominio).**

L'effetto della fonte sulla fiducia non sarà omogeneo; emergerà un'interazione tale per cui la preferenza per l'IA (rispetto all'uomo) aumenterà nel dominio degli investimenti e diminuirà in quello delle donazioni, mentre la partnership ibrida avrà un profilo più "stabile" in tutti i domini.

### **H5 (Interazione variabili percettive × dominio)**

L'impatto delle variabili percettive (fiducia, competenza e calore) sugli esiti finali (intenzione e affidabilità) cambia a seconda del dominio decisionale. In particolare, nel contesto dell'investimento dovrebbero assumere maggiore rilievo le dimensioni più cognitive e tecniche, mentre nel contesto della donazione dovrebbe aumentare il peso della dimensione relazionale.

**H6 (Moderazione: familiarità con l'IA).**

Il rapporto tra fonte e valutazioni dipenderà dal grado di familiarità con l'IA: una maggiore familiarità con questa tecnologia ridurrà l'avversione verso le fonti algoritmiche e aumenterà la tendenza a valutarle più favorevolmente, soprattutto nei contesti in cui l'IA appare coerente con il compito richiesto.

### 2.5.5 Sintesi del modello e ponte verso il Capitolo 3

Il modello concettuale può essere riassunto come segue: la fonte del consiglio (IA vs. umano vs. IA+umano) e l'ambito decisionale (investimento vs. donazione) influenzano sia la fiducia che l'intenzione di seguire la raccomandazione. Gli effetti proposti sono a loro volta spiegati attraverso i due canali psicologici (la competenza percepita e il calore percepito), che variano in funzione delle caratteristiche individuali (preferenza per l'umano, comfort con l'IA, propensione al rischio, orientamento prosociale). In questo modello, la fiducia è più di un atteggiamento: è invece il meccanismo di trasmissione, che collega le tecnologie intelligenti ai risultati economici (adozione, comportamento, valore), anticipando così le implicazioni discusse a livello di mercati e strategie nei capitoli precedenti.

Questo sottocapitolo non ripete il capitolo 3, ma ne getta le basi. Nel Capitolo 3 verranno descritti nel dettaglio la traduzione operativa del modello (il disegno sperimentale tra soggetti (3x2), gli scenari, la randomizzazione, le scale Likert, la dimensione del campione e l'analisi utilizzata) per chiarire come abbiamo misurato i costrutti e testato le ipotesi tramite un questionario online.

Qui lo scopo è quello di conferire coerenza verticale alla tesi. Infatti, viene chiarito *perché* ci aspettiamo determinati effetti, attraverso quali meccanismi e in quali condizioni ci aspettiamo che l'IA (da sola o in partnership) possa risultare più persuasiva, affidabile o accettabile.

## **Riferimenti bibliografici essenziali (APA 7<sup>a</sup> ed.)**

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, 40, 61–149.

De Biasio, A. (2024). *Value-Aware Recommendation: Algorithms and Applications* (PhD thesis). Università degli Studi di Padova.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Journal of Experimental Psychology: General*, 148(12), 2108–2124.

Longoni, C., Cian, L., & Morewedge, C. K. (2022). The word of machine: A framework for understanding and uncovering the persuasive influence of algorithms.

## **CAPITOLO 3 – Analisi empirica: fiducia nelle raccomandazioni AI tra decisioni di investimento e di donazione**

### **3.1 Disegno della ricerca**

Questo capitolo presenta l'analisi empirica della tesi. Prima ancora di addentrarci nei numeri di questo capitolo, è importante spiegare il ragionamento alla base del disegno sperimentale e perché questo particolare disegno sia adatto a testare le ipotesi formulate nel sottocapitolo 2.5. Più specificamente, l'idea principale che ha guidato il disegno sperimentale era spiegare gli effetti che diversi tipi di fonti di raccomandazione (IA, umano, IA+umano) e diversi tipi di domini decisionali (investimento vs. donazione) avevano sulla fiducia in tali raccomandazioni e la conseguente disponibilità a seguirle. In altre parole, la ricerca mette a confronto “chi consiglia” e “su che cosa si decide”, considerando la fiducia non come un tratto

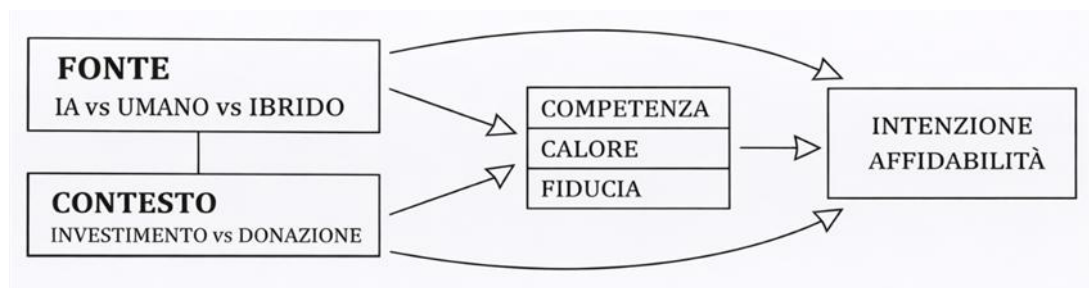
fisso dell'individuo, ma come una risposta contestuale che dipende dalle aspettative sulla fonte e dal significato attribuito alla scelta.

### 3.1.1 Logica generale: dal modello concettuale alla verifica empirica

Nel capitolo 2, abbiamo mostrato che la fiducia nell'IA oscilla tra sfiducia e apertura, come questa oscillazione sia particolarmente sensibile al tipo di decisione e alle dimensioni percettive di competenza e calore/empatia. Inoltre, nel sottocapitolo 2.5, abbiamo definito un modello in cui la fonte del consiglio e il dominio decisionale hanno un effetto considerevole sul livello di fiducia nella raccomandazione e sull'intenzione di seguirla.

Il disegno sperimentale utilizzato per la ricerca è stato creato per testare il modello stabilito nel paragrafo precedente. Utilizzando un assetto "pulito", in cui viene presentato un solo scenario, in modo tale che i partecipanti non possano "indovinare" l'intenzione o l'obiettivo dello studio. L'assegnazione allo scenario avviene in modo casuale, per garantire che eventuali differenze tra le variabili dipendenti osservate nello studio siano dovute alle manipolazioni sperimentali effettuate durante l'esperimento (fonte e dominio), e non a differenze sistematiche tra i gruppi.

**Figura 3.1 -Schema concettuale del disegno**



*Fonte: elaborazione propria.*

### 3.1.2 Disegno sperimentale: un 3×2 tra-soggetti

La ricerca utilizza un disegno fattoriale 3×2 tra-soggetti (*between-subjects*). I due fattori sperimentali sono:

- *Fonte del consiglio (3 livelli)*
  1. IA
  2. Consulente umano

3. IA + consulente umano (ibrido)
  - *Dominio decisionale (2 livelli)*
    1. Investimento finanziario
    2. Donazione/beneficenza

La combinazione dei due fattori genera sei condizioni sperimentali. Ogni partecipante viene assegnato casualmente a una sola condizione e legge un solo scenario. Questa scelta, oltre a essere coerente con la letteratura sugli esperimenti survey-based, è particolarmente adatta quando si vogliono stimare con chiarezza: (i) gli effetti principali dei fattori (se, in media, una fonte genera più fiducia dell'altra; se, in media, un dominio genera più fiducia dell'altro) e (ii) soprattutto l'interazione Fonte×Dominio, cioè se l'effetto della fonte cambia a seconda del dominio decisionale.

**Tabella 3.1 – Matrice delle condizioni sperimentali del disegno 3×2**

<b>Fonte raccomandazione</b>	<b>della</b>	<b>Investimento</b>	<b>Donazione</b>
<b>IA</b>		IA – Investimento	IA – Donazione
<b>Umano</b>		Umano – Investimento	Umano – Donazione
<b>Ibrido (IA + Umano)</b>		Ibrido – Investimento	Ibrido – Donazione

*Fonte: elaborazione propria.*

### **3.1.3 Manipolazioni: Costruzione dello scenario e coerenza tra le condizioni**

Le manipolazioni saranno implementate tramite brevi scenari testuali. Gli scenari sono costruiti in modo tale che tutti gli elementi contestuali siano mantenuti costanti e che solo le componenti rilevanti vengano manipolate: (a) chi produce il consiglio (fonte) e (b) quale tipo di decisione viene rappresentata (dominio).

Per quanto riguarda il tipo di decisione da prendere, distinguiamo tra scenari in cui la decisione è inquadrata come una decisione di investimento e scenari in cui è inquadrata come una decisione di donazione.

- Nel dominio dell'**investimento**, lo scenario attiva un frame più analitico: allocazione del portafoglio, strategia di investimento e valutazione di una raccomandazione che richiama competenza tecnica e affidabilità informativa.
- Nel dominio della **donazione**, lo scenario richiama invece un frame prosociale e valoriale: scelta di un'organizzazione non profit e valutazione di un consiglio che può implicare sensibilità verso valori, finalità e impatto sociale.

Parallelamente, la fonte del consiglio è descritta in tre modi diversi:

- **IA**: un sistema che analizza i dati e propone una raccomandazione
- **Umano**: un consulente che ascolta e suggerisce
- **Ibrido**: un sistema che analizza e un umano che interpreta e spiega il risultato dell'analisi

Un aspetto importante nella costruzione di questi scenari è garantire che siano comparabili tra le diverse condizioni. Lo scenario deve essere credibile in tutte le condizioni; poiché in caso contrario, sarà possibile attribuire eventuali differenze nei risultati alla natura irrealistica dello scenario. Anche il fatto che lo scenario sia percepito come credibile può rappresentare una variabile confondente. Per evitare ciò, lo scenario è costruito in modo realistico e simmetrico: cambia la fonte, non il “valore” implicito del consiglio.

#### **3.1.4 Variabili misurate: variabili dipendenti principali e dimensioni percettive**

Dopo lo scenario, i partecipanti sono tenuti a valutare il consiglio e la fonte del consiglio su una serie di item su una scala Likert da 1 a 7. La struttura di queste variabili si basa sulla seguente idea: distinguere tra esiti comportamentali/intenzionali e meccanismi percettivi che aiutano a spiegare come tali esiti si formino.

- *Fiducia nel consiglio*: misura la volontà del partecipante di considerare il consiglio credibile, rassicurante e degno di essere preso seriamente in considerazione.

- *Affidabilità percepita della fonte*: valuta la solidità e l'affidabilità della fonte del consiglio. Questa è una variabile che si colloca a metà strada tra la fiducia nel consiglio e l'intenzione di seguirlo; pertanto, può contribuire a rivelare la qualità della relazione tra il partecipante e la fonte del consiglio.
- *Intenzione di seguire il consiglio*: questa variabile di risultato si riferisce all'intenzione del partecipante di seguire il consiglio. Questo è il risultato più vicino al comportamento effettivo.
- *Competenza percepita della fonte*: il grado in cui la fonte del consiglio è percepita come capace, preparata e tecnicamente adeguata di svolgere il compito richiesto.
- *Calore/empatia percepita*: esprime il livello di vicinanza, comprensione e preoccupazione per il benessere del partecipante espresso dalla fonte del consiglio.

Queste variabili, già introdotte nel Capitolo 2, sono cruciali per l'interpretare i risultati. Ad esempio, un eventuale maggior livello di fiducia riposto nell'IA nel dominio dell'investimento è significativo solo se si accompagna anche da un aumento della competenza e dell'affidabilità attribuita alla fonte. Al contrario, un'eventuale preferenza per l'umano nel dominio della donazione è più plausibile solo se tale preferenza si riflette in livelli più elevati di calore percepito ed a una maggiore sensazione di affidamento relazionale.

In questa fase si introduce il quadro concettuale per ciascuna delle variabili da considerare in questo studio. La descrizione dei metodi di indagine, degli item, degli indici compositi e dell'affidabilità interna sarà fornita nel §3.2 così da distinguere chiaramente il piano teorico delle variabili dal loro trattamento operativo.

**Tabella 3.2 – Sintesi dei costrutti principali, numero di item e scala di misurazione**

<b>Costrutto</b>	<b>N. item</b>	<b>Scala</b>
Fiducia	3	Likert 1-7
Competenza percepita	3	Likert 1-7
Calore/empatia percepita	3	Likert 1-7
Affidabilità della fonte	1	Likert 1-7
Intenzione di seguire il consiglio	1	Likert 1-7

*Fonte: elaborazione propria.*

*Variabili di controllo (misure delle differenze personali):*

Per determinare per chi i sistemi di intelligenza artificiale funzionano meglio, è necessario misurare determinate predisposizioni individuali, come le preferenze per le fonti umane o per l'intelligenza artificiale, la propensione al rischio in contesti finanziari e i comportamenti prosociali. Questi non sono il focus dello studio in sé, ma estendono l'ambito dell'indagine per determinare se gli effetti osservati differiscono in base a queste caratteristiche individuali in modo coerente con la teoria (ad esempio, coloro che si sentono più a loro agio nell'uso dell'intelligenza artificiale potrebbero incontrare una minore resistenza al cambiamento in altri ambiti di valore).

*Variabili descrittive:*

L'età, il genere, il livello di istruzione e la familiarità dei partecipanti con la tecnologia e uso di applicazioni di intelligenza artificiale consentiranno di descrivere il campione e di condurre controlli di robustezza, se necessario.

### **3.1.5 Strategia di analisi: test delle ipotesi e lettura dei meccanismi**

Il disegno 3 x 2 consente di testare le ipotesi principali utilizzando un'ANOVA fattoriale 3 x 2 sulle variabili principali (fiducia, intenzione, competenza, calore umano, affidabilità). L'ANOVA fornirà tre informazioni essenziali:

1. l'effetto della Fonte,
2. l'effetto del Dominio,
3. l'effetto di interazione Fonte×Dominio.

A valle di questi test, l'interpretazione viene resa più trasparente attraverso:

- medie marginali per cella (le sei condizioni),
- confronti post-hoc (es. Bonferroni) quando necessari per capire quali coppie differiscono,
- rappresentazioni grafiche delle interazioni.

Infine, si prevede un modello esteso in cui l'intenzione viene spiegata includendo fiducia, competenza, calore e altre dimensioni come predittori, al fine di poter rispondere alla

domanda più sostantiva della ricerca: l'intenzione dipende solo dagli scenari o anche dai costrutti psicologici? Questo modello non sostituisce necessariamente l'uso dell'ANOVA fattoriale; piuttosto, si tratta di informazioni aggiuntive che aiutano a dimostrare se l'effetto “puro” della condizione sperimentale diminuisce quando si prendono in considerazione i giudizi della fonte (suggerendo un ruolo interpretativo o di mediazione di tali costrutti percettivi).

### **3.1.6 Come si passa dal disegno ai dati**

In sintesi, questo disegno sperimentale è specificamente progettato per determinare l'impatto di ciascuna di queste due leve (la fonte del consiglio e il dominio decisionale) sui livelli di fiducia e sulla probabilità che gli individui seguano tale raccomandazione, passando per competenza e calore percepiti. L'uso di un disegno sperimentale 3x2 tra soggetti riflette l'idea che la fiducia si formi all'interno di uno scenario e non (solo) come preferenza generale per “umano” o “algoritmo”.

La Sezione 3.2 descrive in dettaglio la struttura del questionario, le misure e la costruzione degli indici; la Sezione 3.3 descrive il campione e la raccolta dei dati; la Sezione 3.4 presenta i risultati dello studio (descrittive per cella, ANOVA, medie marginali, confronti post-hoc e grafici), mentre nella Sezione 3.5 si discutono implicazioni, collegando l'evidenza empirica al quadro teorico del Capitolo 2 ed anche alle implicazioni economico-manageriali del Capitolo 1.

Dopo aver descritto il disegno di ricerca e la metodologia dello studio, è importante fornire dettagli sulle modalità di implementazione dello stesso. Per comprendere la ricerca condotta, è innanzitutto necessario descrivere i partecipanti allo studio e le modalità di raccolta dei dati. La sezione seguente approfondisce la struttura dettagliata del questionario, degli scenari e delle misure utilizzate nella conduzione di questo studio per fornire ulteriori chiarimenti sulla struttura della ricerca in questione.

### **3.2 Struttura del questionario e misure**

Dopo aver definito le motivazioni alla base del disegno sperimentale nel §3.1, il passo successivo consiste nel mostrare la costruzione del questionario che verrà utilizzato per raccogliere i dati necessari a testare le ipotesi formulate all'interno del framework. Questa sezione non si occupa della presentazione dei risultati o della descrizione delle caratteristiche del campione, poiché tali argomenti saranno trattati in ordine nel §3.3 e nel §3.4. L'attenzione

di questa sezione si concentra invece sulla costruzione del questionario, sulla sua organizzazione e sull'analisi delle variabili che verranno misurate. Il questionario è il luogo in cui i costrutti teorici discussi nel Capitolo 2 diventano misurabili e analizzabili.

La costruzione del questionario è una parte essenziale dello studio. Un questionario troppo lungo renderà difficile per il partecipante rimanere coinvolto e attento allo studio. D'altro canto, un questionario troppo breve potrebbe impedire ai partecipanti di comprendere la complessità dei costrutti che vengono misurati. La costruzione di un buon questionario richiede un'attenta valutazione di ciascuno di questi fattori.

Il questionario è stato costruito in modo coerente con il disegno sperimentale tra soggetti presentato nel §3.1. All'inizio, al partecipante viene chiesto di immaginare una situazione decisionale e di valutarla in base a uno scenario specifico, che gli viene assegnato casualmente tra sei condizioni possibili. A partire da quel punto, l'intero strumento si sviluppa attorno a tre nuclei: il blocco degli scenari, il blocco delle valutazioni e il blocco delle misure individuali e descrittive. Questa struttura riflette il disegno sperimentale. I partecipanti vengono inizialmente esposti allo scenario da manipolare. E successivamente, vengono rilevate le reazioni immediate allo scenario. Solo in seguito vengono raccolte le disposizioni personali e le informazioni anagrafiche. Questo ordine garantisce che le domande personali non influenzino la valutazione dello scenario. La prima delle tre sottoparti di questa sezione principale è dedicata agli scenari sperimentali (§3.2.1), seguita dalle scale e dai costrutti identificati (§3.2.2) e, infine, dalla costruzione degli indici di tali costrutti e da una discussione della loro affidabilità e validità (§3.2.3).

### **3.2.1 Scenari sperimentali**

Il primo blocco principale del questionario è costituito dagli scenari sperimentali, i quali rappresentano il cuore manipolativo dell'intera ricerca. Come accennato nel §3.1, gli scenari sperimentali consistono in sei diverse situazioni, formate dalla combinazione di tre diverse fonti di consulenza (IA, umano e IA+umano) con due diversi domini decisionali (investimento e donazione). I sei scenari sono formulati in modo tale che in ogni scenario cambino solo la fonte e il contenuto della decisione, mantenendo invece costante la forma complessiva della situazione.

La motivazione alla base della scelta di utilizzare scenari sintetici ma narrativamente riconoscibili è dettata dalla necessità di formulare domande facilmente comprensibili per i partecipanti al sondaggio. Infatti, scenari troppo lunghi potrebbero causare problemi di comprensione, mentre scenari troppo vaghi, al contrario, avrebbero potuto portare a un'attivazione cognitiva debole. La formulazione breve, realistica e simmetrica degli scenari, capace di evocare una scelta credibile, rappresenta un compromesso tra questi due estremi.

Nel dominio degli investimenti, questo scenario si riferisce a una tipica situazione di razionalità strumentale in cui un investitore è tenuto a immaginare di dover prendere una decisione su come allocare il proprio portafoglio o, più in generale, a prendere una decisione finanziaria che implica valutazione del rischio, gestione delle informazioni e aspettative di rendimento. In questa situazione, la credibilità del suggerimento associata a fattori quali la competenza tecnica, accuratezza e capacità di leggere i dati.

Nel dominio delle donazioni, tuttavia, lo scenario assume un livello diverso. Il consiglio viene fornito su quale organizzazione selezionare e a quale causa donare. Sebbene si tratti comunque di una decisione razionale, la decisione implica motivazioni e valori prosociali. La credibilità di questo suggerimento non si basa solo sulla sua accuratezza, ma anche su quanto il suggerimento appaia allineato con i valori, i bisogni e gli obiettivi percepiti come umani. Questo è uno dei punti centrali dell'intera tesi, secondo cui la fiducia non dipende solo dalla qualità tecnica del suggerimento, ma anche dal significato che il soggetto attribuisce alla scelta che è chiamato a immaginare.

All'interno di ciascun dominio, la fonte del consiglio può essere descritta in tre diversi modi. Nella condizione di intelligenza artificiale, in questo caso il consiglio proviene da un sistema che analizza i dati e fornisce una raccomandazione al partecipante. Nella condizione umana, invece, il consiglio proviene da un consulente o da una figura esperta, la quale valuta la situazione e orienta la scelta del partecipante. Infine, nella condizione ibrida, in cui l'output viene presentato come frutto della collaborazione tra IA e giudizio umano. Nella condizione ibrida il sistema di intelligenza artificiale elabora le informazioni e la componente umana interpreta i risultati. Questo tipo di configurazione è particolarmente importante nello studio, in quanto consente di valutare una via di mezzo in un dibattito che postula che tale soluzione sia teoricamente possibile. Questa terza configurazione è particolarmente importante perché introduce una possibile soluzione intermedia tra i due estremi.

Un altro aspetto importante nella costruzione di questi scenari riguarda la loro comparabilità narrativa, cioè garantire che siano percepiti come equivalenti dal lettore. Pertanto, gli scenari differiscono nel contenuto, ma per il resto appaiono uguali al lettore. In altre parole, lo scenario generato dall'IA non deve essere visto come intrinsecamente caricaturale o futuristico. Lo scenario in cui gli esseri umani sono gli unici attori non deve essere visto come banale. Lo scenario ibrido non deve inoltre essere visto come una soluzione intrinsecamente migliore solo perché descritto come più completo. La scrittura degli scenari in questa sezione mira a garantire che gli unici elementi che cambiano tra gli scenari siano la fonte del consiglio e il dominio in cui viene applicato. La simmetria che si crea in questo modo è ciò che conferisce credibilità all'interpretazione causale dei risultati. Quindi, le differenze rilevate nelle variabili dipendenti possono essere lette come solo effetto della manipolazione e non di differenze di stile o di intensità narrativa tra gli scenari.

**Nota.** I sei scenari sperimentali, nella loro formulazione completa, sono riportati in Appendice.

### 3.2.2 Scale e costrutti

Se gli scenari costituiscono il momento in cui il partecipante viene esposto alla manipolazione, il blocco delle misure rappresenta il luogo in cui tale esposizione viene trasformata in informazione analizzabile. Il questionario raccoglie informazioni dai partecipanti sulle loro reazioni allo scenario appena letto. Tale scelta di costrutti non è casuale, ma riflette il modello concettuale descritto nel Capitolo 2 e formalmente presentato nel §2.5. Nello specifico, il modello concettuale indicava che la risposta a una raccomandazione si basa sul livello di fiducia che suscita, sull'affidabilità attribuita alla fonte, sulla competenza percepita della raccomandazione, sul calore o empatia percepita della raccomandazione e sulla volontà del partecipante di intraprendere tale azione.

Per ciascuna delle variabili chiave viene utilizzata una scala Likert a 7 punti. Questo tipo di scelta della scala presenta almeno tre vantaggi distinti. In primo luogo, consente una comprensione più dettagliata delle variabili misurate. In secondo luogo, garantisce che gli item di ciascun questionario siano tutti internamente coerenti. Terzo, facilita la successiva costruzione di indici compositi, poiché tutti gli item sono misurati sulla stessa scala di risposta. Il primo costrutto identificato nel questionario è il livello di fiducia che i partecipanti ripongono nel consiglio. Questa è una delle principali variabili dipendenti di questo studio. La misura del

livello di fiducia si basa su una scala a tre affermazioni (item), utilizzata per cogliere le diverse sfumature di questo costrutto. Gli item utilizzati sono:

- *“Mi fido del consiglio che ho appena letto”*
- *“Questo consiglio è credibile”*
- *“Ritengo affidabile questo consiglio”*

La decisione di includere tre item semanticamente simili ma non identici segue una logica specifica. Il primo item intende catturare il concetto di fiducia. Il secondo introduce il tema della credibilità, ovvero la plausibilità percepita del contenuto del messaggio. Il terzo introduce il concetto di affidabilità del consiglio fornito. Questi tre item, quindi, consentono di costruire un concetto sintetico di fiducia che abbia una solida base percettiva. Non è un caso che, nella successiva costruzione dell'indice, questa scala mostri livelli molto elevati di coerenza interna.

Il secondo costrutto principale è la competenza percepita della fonte. Ancora una volta, per catturare questo costrutto è stata utilizzata una scala a tre item. Gli item sono:

- *“La fonte è competente in materia”*
- *“La fonte mostra padronanza tecnica”*
- *“La fonte sa valutare correttamente i dati”*

Questo costrutto è altamente coerente con il quadro teorico discusso nel Capitolo 2 di questa tesi. In particolare, la competenza della fonte è uno dei principali fattori chiave per l'accettazione dell'IA. Gli item utilizzati per valutare questo costrutto si sono concentrati sulla competenza generale della fonte, sulla competenza tecnica della fonte e sulla capacità della fonte di elaborare correttamente le informazioni.

Il terzo blocco di item riguardava il calore o l'empatia percepiti dalla fonte, misurati anch'essi attraverso tre item:

- *“La fonte è empatica”*
- *“La fonte è attenta ai miei bisogni”*
- *“La fonte appare umana e comprensiva”*

Questa scala introduce una dimensione diversa, ma non per questo meno importante. Mentre il costrutto precedente valutava la capacità tecnica della fonte, questo costrutto valutava la capacità della fonte di entrare in contatto con il destinatario del consiglio. Ciò significa che la valutazione della fonte considera anche il livello di empatia, sensibilità e il grado di orientamento alla persona della fonte. Il lessico degli item è volutamente relazionale: empatia, attenzione ai bisogni, comprensione. Tale scelta è pienamente coerente con la letteratura discussa nei paragrafi teorici precedenti. Infatti, la letteratura evidenzia che la distinzione tra competenza e calore rappresenta un asse portante per interpretare la fiducia nelle tecnologie e nelle fonti di consiglio.

Oltre a questi tre costrutti multi-item, sono presenti due misure single-item.

La prima è l'intenzione di seguire il consiglio, rilevata con l'affermazione:

- *“Se dovessi agire realmente, seguirei questo consiglio”*

Questa variabile è quella che più si avvicina alla misurazione del livello comportamentale. Sebbene non misuri direttamente il comportamento dei partecipanti, consente di misurare il grado di disponibilità dei partecipanti ad agire in base al consiglio. Ai fini di questa tesi, questa è una variabile importante da misurare, perché collega il giudizio sulla fonte al possibile esito decisionale.

La seconda misura single-item riguarda l'affidabilità generale della fonte, formulata come:

- *“Considero questa fonte generalmente affidabile”*

In questa sezione, l'attenzione si concentra meno sul consiglio in sé e più sulla fonte in generale. Sebbene sia possibile che un individuo consideri credibile il consiglio di una fonte, ciò non implica necessariamente che la consideri anche affidabile. Al contrario, potrebbe ritenere la fonte affidabile senza essere pienamente convinto dai consigli specifici forniti. Per questo motivo, l'inclusione di entrambe le misure risulta utile ai fini dell'analisi.

Oltre alle variabili principali misurate, è stato incluso anche un blocco di misure individuali. Come definito nel §3.1, vengono misurate quattro diverse disposizioni: preferenza per i consigli umani, familiarità con l'IA, propensione ad assumersi rischi finanziari e livello di orientamento prosociale. Sebbene queste variabili non vengano utilizzate per "spiegare tutto" ciò che accade nello scenario, hanno contribuito a verificare se l'effetto variasse in base alle

caratteristiche individuali. In altre parole, queste misure individuali hanno contribuito a determinare su quale gruppo di individui l'IA sarebbe stata più o meno accettabile.

La Tabella 3.3 presenta le statistiche descrittive per ciascuno dei singoli elementi, nonché il valore dell'alfa di Cronbach se quell'elemento specifico venisse eliminato, così da offrire una lettura analitica della tenuta interna delle scale.

**Tabella 3.3 – Statistiche descrittive degli item e affidabilità interna delle scale**

Costrutto	Item	Media	Std. err.	Cronbach's $\alpha$
<b>Fiducia</b>	Fid1: Mi fido del consiglio che ho appena letto	4,6940	0,0733	0,9392
	Fid2: Questo consiglio è credibile	4,8980	0,0683	0,9368
	Fid3: Ritengo affidabile questo consiglio	4,7015	0,0731	0,9332
<b>Competenza</b>	Comp1: La fonte è competente in materia	4,8184	0,0682	0,8540
	Comp2: La fonte mostra padronanza tecnica	4,9602	0,0642	0,8907
	Comp3: La fonte sa valutare correttamente i dati	4,7413	0,0675	0,8924
<b>Calore</b>	Cal1: La fonte è empatica	4,1915	0,0948	0,8708
	Cal2: La fonte è attenta ai miei bisogni	4,4801	0,0773	0,9198
	Cal3: La fonte appare umana e comprensiva	4,3483	0,0914	0,8519

*Fonte: elaborazione propria.*

**Nota.** Coefficienti di affidabilità complessivi delle scale:

- *Fiducia*  $\alpha = 0,9466$ ;
- *Competenza*  $\alpha = 0,9164$ ;
- *Calore*  $\alpha = 0,9194$ .

---

### 3.2.3 Costruzione degli indici, affidabilità e validità delle misure

Dopo aver definito i costrutti e raccolto le risposte ai singoli item, il passo successivo è la costruzione degli indici sintetici che verranno utilizzati nelle analisi. Questa fase è tutt'altro

che puramente operativa. Data la struttura della ricerca attuale, il modo in cui gli item vengono aggregati influirà sulla qualità delle variabili ottenute e sulla solidità dei risultati della ricerca, che saranno discussi nei paragrafi seguenti. Per tale motivo la costruzione degli indici sintetici è stata accompagnata da una duplice verifica: in primo luogo il controllo della coerenza interna delle scale e, per secondo, una valutazione più strutturata della loro tenuta sul piano del modello di misura.

Per le dimensioni multi-item (fiducia, competenza e cordialità), è stata calcolata la media di tutti gli item di ciascuna dimensione. Il principale vantaggio di questa forma di calcolo è che costituisce un indicatore facilmente interpretabile, mantenendo la stessa metrica originaria della scala Likert (1–7) e rendendo i punteggi immediatamente confrontabili tra diversi costrutti e tra condizioni sperimentali. Non si tratta di una semplice operazione meccanica di sintesi: infatti, tale scelta è stata fatta per rendere le misure più leggibili, metricamente coerenti e più facili da interpretare.

Per i due costrutti a singolo item (*intenzione di seguire il consiglio e affidabilità generale della fonte*), è immediatamente evidente che non è necessaria alcuna aggregazione. In questi due casi, la variabile da analizzare è la stessa dell'item raccolto.

Tuttavia, per costruire questi indici compositi, è innanzitutto necessario assicurarsi che gli item inclusi nella stessa scala presentino una sufficiente coerenza interna. Per verificare tutto ciò è stato utilizzato il coefficiente di Cronbach's alpha, che è uno degli indicatori più diffusi nella ricerca quantitativa basata su questionari. Il coefficiente alfa di Cronbach misura il grado di coerenza tra gli item inclusi in una scala, ossia il grado in cui riflettono una medesima dimensione sottostante. In altri termini, se gli item tendono a “muoversi insieme”, l'aggregazione in un unico indice è teoricamente ed empiricamente plausibile. Al contrario, se tali item mostrano comportamenti troppo eterogenei, la scala perde solidità e la sintesi in un punteggio composito diventa più fragile.

Le soglie per l'interpretazione del coefficiente alfa di Cronbach sono le seguenti: un valore inferiore a 0,60 è considerato indicativo di scarsa coerenza; tra 0,60 e 0,69, la coerenza è considerata discutibile; tra 0,70 e 0,79, la scala è accettabile; tra 0,80 e 0,89, buona; e sopra 0,90, eccellente. Nel caso di questa ricerca, i valori ottenuti si collocano tutti nella fascia alta, offrendo un supporto molto forte alla costruzione degli indici compositi.

Il livello di coerenza interna (affidabilità) degli item che compongono la scala "fiducia nel consiglio" è particolarmente elevato, con un valore di alfa di Cronbach pari a 0,9466. Tale valore rientra nell'intervallo eccellente e indica una notevole coesione tra i tre item impiegati. I singoli item mostrano contributi molto elevati sia nell'item-test sia nell'item-rest correlation, a conferma del fatto che la scala misura in modo coerente una stessa dimensione di giudizio. Il valore medio degli item all'interno di questa scala è di circa 4,76, ma il dato più importante è l'elevato livello di robustezza di questo costrutto.

La scala di competenza rivela un livello molto elevato di coerenza interna. Infatti, presenta un valore di  $\alpha$  pari a 0,9164. Come per le altre scale, il risultato di questo coefficiente è ben al di sopra della soglia di eccellenza, indicando che i tre item inclusi nella scala sono altamente convergenti su una stessa dimensione latente. Pertanto, misurata in questo modo, la competenza può essere considerata un costrutto stabile e coerente, senza che l'unione degli item crei distorsioni rilevanti. Inoltre, la media di questa scala è vicina a 4,84.

Anche la scala di calore ed empatia presenta un livello simile di coerenza interna. Il valore di  $\alpha$  in questo caso è 0,9194. Ancora una volta, questo valore è ben al di sopra della soglia di eccellenza e dimostra che i tre item inclusi nella scala sono convergenti e formano un costrutto unidimensionale stabile della dimensione relazionale del giudizio sulla fonte. Questo è un elemento importante per interpretare le reazioni dei partecipanti nei contesti più valoriali o prosociali. La media complessiva di questa scala è di circa 4,34. Per le due misure single-item, come precedentemente anticipato, non è stato calcolato alcun coefficiente di affidabilità interna. La media relativa all'intenzione degli intervistati di seguire il consiglio è di circa 4,32, mentre la media dell'elemento che misura il livello di affidabilità della fonte del consiglio è di circa 4,58. Questi valori svolgono, in questa fase, una funzione prevalentemente descrittiva. La loro interpretazione sostantiva emergerà soprattutto nel §3.4, quando saranno messi in relazione con la fonte del consiglio, con il dominio decisionale e con gli altri costrutti del modello.

Oltre alla verifica dell'affidabilità interna degli item, è possibile effettuare un ulteriore controllo sulla robustezza del modello: attraverso una stima SEM/CFA costruita sui tre fattori latenti principali (Fiducia, Competenza e Calore) ciascuno rilevato dai rispettivi item osservati. La SEM non viene introdotta come metodo alternativo di analisi dei dati. Piuttosto, è stato utilizzato come uno strumento per verificare che le scale sviluppate non solo siano internamente affidabili, ma anche che i loro item siano correlati ai fattori come previsto.

I risultati del modello SEM confermano ancora una volta i risultati dei test di affidabilità precedentemente svolto. Tutti i caricamenti fattoriali standardizzati sono risultati elevati e statisticamente significativi, con valori compresi tra 0,85 e 0,94. Un simile livello di loading indica che ciascun item contribuisce in modo sostanziale alla definizione del proprio costrutto latente, rafforzando la validità convergente delle scale. Inoltre, anche l'adattamento del modello è risultato soddisfacente, come indicato dal valore dell'SRMR pari a 0,047, che rientra nelle soglie per un buon adattamento tra modello e dati. In altri termini, il modello di misura appare coerente con il modo in cui gli item sono stati concepiti e somministrati.

L'aspetto più utile riguarda però il confronto tra i factor scores stimati dal modello latente e i punteggi compositi calcolati come media degli item. Le correlazioni tra i punteggi fattoriali e i punteggi compositi per ciascuno dei costrutti sono risultate quasi perfette (0,995 per la fiducia, 0,987 per la competenza e 0,998 per il calore umano), indicando che l'utilizzo dei punteggi compositi non ha comportato una perdita di informazioni per i dati, ma ha rappresentato un'alternativa valida all'utilizzo dei fattori latenti. L'uso dei punteggi compositi permette quindi di avere un'analisi più chiara e immediata, senza perdere qualità nella misurazione.

Nel complesso, la validità delle misure è convincente e dipende dalla coerenza tra teoria, item, scale e risultati empirici. Il questionario presenta tre punti di forza. 1) I costrutti sono ben collegati alla teoria. 2) Gli item coprono bene le dimensioni principali in modo semplice. 3) I risultati statistici confermano la solidità delle misure.

Se da un lato la Tabella 3.3 consente una lettura puntuale a livello di item, la Tabella 3.4 sintetizza invece la qualità complessiva delle scale, integrando gli indici di coerenza interna con i principali risultati del modello di misura stimato tramite SEM.

**Tabella 3.4 – Sintesi degli indici compositi e qualità del modello di misura**

Costrutto	N. item	Metodo di costruzione	Media composito	Std. err. composito	Cronbach's $\alpha$	Range loadings standardizzati	Correlazione composito ↔ factor score	Interpretazione sintetica
<b>Fiducia</b>	3	Media dei 3 item	4,7645	0,0681	0,9466	0,90–0,94	0,9948	Alta coerenza degli Item
<b>Competenza</b>	3	Media dei 3 item	4,8400	0,0617	0,9164	0,85–0,92	0,9865	Alta coerenza degli Item

<b>Calore</b>	3	Media dei 3 item	4,3400	0,0818	0,9194	0,85–0,93	0,9979	Alta coerenza degli Item
<b>Intenzione</b>	1	Single item	4,3184	0,0824	n.a.	n.a.	n.a.	Misura diretta
<b>Affidabilità della fonte</b>	1	Single item	4,5771	0,0777	n.a.	n.a.	n.a.	Misura diretta

Fonte: elaborazione propria.

La Sezione 3.3 sposterà ora l'attenzione della discussione dal *cosa* sia stato misurato dal questionario a *chi* è stato misurato e come il questionario è stato somministrato. Il capitolo 3 procede quindi in modo logico, partendo dalla la struttura dello strumento e la qualità delle misure, passando poi alla descrizione del profilo degli individui intervistati e la procedura di rilevazione, per arrivare infine all'analisi dei risultati ottenuti dei dati raccolti.

### 3.3 Campione e procedura

Dopo aver chiarito nel §3.1 la logica del disegno sperimentale e la struttura delle variabili considerate, è necessario concentrarsi su due elementi che contribuiscono alla sostanza dell'intera analisi: chi ha partecipato allo studio e come sono stati raccolti i dati. I risultati della ricerca, contestualizzati all'interno delle informazioni presentate in questa sezione, contribuiranno a definire il contesto dell'analisi, l'interpretazione dei risultati e la relazione tra ipotesi, raccolta dati e analisi dei dati.

Nell'ambito di questa ricerca che indaga gli elementi di fiducia, affidabilità percepita, competenza, calore e disponibilità a seguire un consiglio, è necessario descrivere i profili dei partecipanti e le modalità di somministrazione del sondaggio. I risultati specifici che emergeranno dai dati presentati nei paragrafi seguenti contribuiranno a fornire il contesto per l'analisi. Il questionario è stato somministrato online tramite la piattaforma Qualtrics. I mezzi di distribuzione del questionario sono stati principalmente i social network e i gruppi digitali. Ciò è coerente con il tema della tesi. La raccolta dei dati online contribuisce a mantenere la ricerca coerente rispetto ai suoi obiettivi: infatti, riproduce una cornice plausibile rispetto ai contesti in cui, nella vita quotidiana, gli utenti entrano in contatto con raccomandazioni algoritmiche, strumenti di supporto decisionale e piattaforme automatizzate.

Nei paragrafi che seguono vengono presentati, in sequenza, il reclutamento e i criteri di inclusione, il profilo del campione finale e la procedura di compilazione del questionario.

### 3.3.1 Reclutamento e criteri di inclusione

I questionari sono stati raccolti diffondendo il link per accedervi online e condividendolo attraverso i social network e i gruppi online. Ciò ha permesso di raccogliere un numero significativo di partecipanti e di ottenere una varietà di tipologie diverse di partecipanti. Inoltre, il sondaggio è stato descritto come uno studio accademico che valuta decisioni e consigli economici, anziché rendere evidente l'ipotesi specifica dello studio nella descrizione. Una descrizione eccessivamente trasparente avrebbe potuto indurre i partecipanti a fornire risposte più "riflessive", causando distorsioni nelle risposte fornite.

Per accedere al questionario, i partecipanti dovevano leggere e accettare il modulo di consenso informato presente all'inizio del questionario. Questo modulo di consenso informato descriveva lo scopo della ricerca, la natura volontaria della partecipazione, il trattamento anonimo dei dati e l'assenza di rischi particolari per i partecipanti coinvolti nella ricerca. Solo gli individui che avevano fornito il consenso positivo potevano proseguire con il resto del sondaggio; in caso contrario, il flusso si interrompeva automaticamente.

Uno dei primi criteri per l'inclusione di un partecipante nello studio era la maggiore età. Tutti i partecipanti presi in considerazione per l'analisi erano adulti. Ciò è stato necessario perché gli scenari presentati ai partecipanti prevedevano investimenti finanziari e donazioni a organizzazioni benefiche che prevedono una soglia minima di autonomia decisionale.

Oltre a questo requisito, è stato applicato un ulteriore criterio per garantire la validità delle risposte ottenute. Tutti i questionari incompleti, interrotti o non completamente utilizzabili sono stati esclusi nella fase di pulizia del database. Tale passaggio è ritenuto particolarmente importante nel contesto delle indagini brevi online, poiché in questi contesti il rischio di abbandono o di compilazione superficiale è molto alto. In questo senso, la selezione dei casi è un prerequisito per garantire la buona qualità dei dati.

Un altro aspetto degno di nota è che alla raccolta delle risposte è stato applicato il regime di completo anonimato. Le risposte sono state raccolte in modo completamente anonimo; infatti, non è stata richiesta alcuna informazione identificativa. Non sono stati raccolti nomi, indirizzi email o altri dati identificativi. Ciò ha comportato un vantaggio sotto due aspetti: ha protetto i dati dei partecipanti e garantito l'eticità dello studio; ha inoltre ridotto la pressione sociale sui partecipanti affinché fornissero determinate risposte, soprattutto perché il

questionario copriva alcuni argomenti relativi al comportamento sociale, tecnologico ed economico dei partecipanti.

Il reclutamento dei partecipanti si è basato su quattro principi molto chiari: accessibilità, partecipazione volontaria, maggiore età e anonimato. Di conseguenza, si è trattato di un campione non probabilistico, ma comunque appropriato agli obiettivi della ricerca. L'obiettivo non era quello di stimare la distribuzione di atteggiamenti nell'intera popolazione, ma di osservare i cambiamenti nelle variabili di interesse (fiducia, affidabilità percepita e intenzione di seguire il consiglio) al variare di specifiche condizioni sperimentali.

### 3.3.2 Descrizione del campione

Il database iniziale raccoglieva 508 questionari. Durante la pulizia del database, 106 questionari sono stati ritenuti non validi e pertanto esclusi per incompletezza nelle risposte. I restanti 402 questionari (79,1%) hanno costituito il campione di dati utilizzato per l'analisi. Il tasso di esclusione è stato del 20,9%.

Un punto che vale la pena notare immediatamente è la distribuzione dei partecipanti nelle diverse condizioni sperimentali. La randomizzazione dei partecipanti ha prodotto una distribuzione relativamente uniforme (anche se non perfettamente equilibrata) dei 508 partecipanti nelle sei celle: 60 casi nella condizione di IA e investimento, 74 nella condizione di IA e donazione, 69 nella condizione umana e investimento, 81 nella condizione umana e donazione, 64 nella condizione di IA e umana e investimento e 54 nella condizione di IA e umana e donazione. Come mostrato nella Tabella 3.3, la distribuzione dei partecipanti nelle sei condizioni di trattamento è sufficientemente bilanciata.

SCENARI SPERIMENTALI	Freq.	Percent	Cum.
IA – Investimento	60	14.93	14.93
IA – Donazione	74	18.41	33.33
Umano – Investimento	69	17.16	50.50
Umano – Donazione	81	20.15	70.65
Ibrido – Investimento	64	15.92	86.57
Ibrido – Donazione	54	13.43	100.00
Total	402	100.00	

Fonte: elaborazione propria.

Il campione include partecipanti di tutte le età, dai 18 ai 73 anni. L'età media è 32,15, con una deviazione standard di 12,87. L'età mediana è 27, il che indica che il campione comprende un numero significativo di giovani adulti, così come una quota significativa di partecipanti più anziani.

La distribuzione di genere è relativamente equilibrata, con il 49,75% dei partecipanti costituito da uomini, il 48,76% da donne e l'1,49% appartenente ad altre categorie.

Per quanto riguarda il livello di istruzione, il 43,28% dei partecipanti ha un diploma di scuola superiore, il 27,86% una laurea triennale e il 20,15% una laurea magistrale. Inoltre, il 7,71% possiede un diploma di scuola media inferiore e l'1,00% possiede altri titoli di studio. Tale dato in termini aggregati suggerisce che i partecipanti hanno un buon livello di istruzione e quindi una buona capacità di comprensione degli scenari e delle domande proposte, pur senza rendere il campione eccessivamente omogeneo sul piano culturale.

Di particolare importanza per questa ricerca è il rapporto con la tecnologia. Per quanto riguarda la competenza tecnologica percepita, una parte significativa dei partecipanti si colloca nella fascia medio-alta. Nello specifico, il 69,65% dei partecipanti ha dichiarato di utilizzare applicazioni basate sull'intelligenza artificiale almeno "a volte" (valori 4-7 sulla scala), mentre il 49,00% dei partecipanti riferisce un uso "frequente" o "molto frequente" (valori 5-7). Pertanto, il campione non è composto esclusivamente da utenti iper-esperti, ma include una base ampia di soggetti che hanno già una certa familiarità con strumenti riconducibili all'intelligenza artificiale. Questo rappresenta un vantaggio per questo studio, poiché significa che è meno probabile che le risposte dei partecipanti siano dovute alla loro scarsa familiarità con la tecnologia.

La dimensione del campione e la sua composizione forniscono una base empirica adeguata a questa ricerca. Sebbene non si tratti di un campione rappresentativo, come verrà affrontato nella discussione successiva, la dimensione e la distribuzione all'interno del campione forniscono una robustezza più che sufficiente per raggiungere gli obiettivi di questo studio.

**Tabella 3.6 – Caratteristiche del campione e altre variabili di controllo (N = 402)**

<b>Variabile</b>	<b>Statistiche / Distribuzione</b>
<b>Numerosità del campione</b>	
Numerosità iniziale	508
Numerosità finale	402 (79,1%)

Questionari esclusi	106 (20,9%)
<b>Età (anni)</b>	
Età	M = 32,15; DS = 12,87 (min = 18; max = 73)
<b>Genere</b>	
Uomini	49,75%
Donne	48,76%
Altro / Non specificato	1,49%
<b>Livello di istruzione</b>	
Licenza media	7,71%
Diploma	43,28%
Laurea triennale	27,86%
Laurea magistrale	20,15%
Altro	1,00%
<b>Utilizzo di applicazioni basate su IA</b>	
Uso raro (1–3)	30,35%
Uso occasionale (4)	20,65%
Almeno “a volte” (4–7)	69,65%
Uso frequente / molto frequente (5–7)	49,00%
Uso molto frequente (6–7)	27,36%
<b>Variabili di controllo</b>	
<b>Preferenza verso l’umano</b>	
Basso (1–3)	43 (10,70%)
Medio (4)	113 (28,11%)
Alto (5–7)	246 (61,19%)
<b>Livello di comfort verso i consigli IA</b>	
Basso (1–3)	114 (28,36%)
Medio (4)	128 (31,84%)
Alto (5–7)	160 (39,80%)
<b>Propensione al rischio</b>	
Bassa (1–3)	84 (20,90%)
Media (4)	97 (24,13%)
Alta (5–7)	221 (54,97%)
<b>Comportamento prosociale</b>	
Basso (1–3)	25 (6,22%)
Medio (4)	84 (20,90%)
Alto (5–7)	293 (72,88%)

*Fonte: elaborazione propria.*

**Nota.** Le variabili di controllo su scala Likert sono state aggregate in tre fasce: basso (1–3), medio (4) e alto (5–7).

Accanto alle principali caratteristiche socio-demografiche del campione, la Tabella 3.6 riporta anche alcune variabili di controllo rilevate nel questionario: preferenza verso l’umano, livello di comfort verso i consigli IA, propensione al rischio e comportamento prosociale. Queste misure vengono introdotte per descrivere in modo più completo il profilo dei

partecipanti e per offrire una prima base di lettura della coerenza ex post tra autodichiarazioni individuali e orientamento delle risposte nei diversi scenari. Più precisamente, la propensione al rischio e il comportamento prosociale verranno letti in relazione, rispettivamente, ai contesti dell'investimento e della donazione, mentre la preferenza verso l'umano e il comfort verso i consigli IA saranno considerati in relazione alla fonte del consiglio. Questo approfondimento verrà ripreso nel §3.4.2, dove tali variabili saranno messe in relazione con l'intenzione di seguire il consiglio.

### 3.3.3 Procedura

Il processo di somministrazione del questionario è stato progettato per essere rapido, semplice e metodologicamente coerente con il disegno sperimentale illustrato nel sottocapitolo 3.1, la sequenza seguita all'interno del questionario è stata la seguente: dopo aver aperto il link del questionario, i partecipanti dovevano prima visualizzare la pagina del consenso informato e solo dopo aver ottenuto il consenso, il questionario conduceva allo scenario sperimentale assegnato in modo casuale e, da quel momento in poi, la sequenza restava identica per tutti i partecipanti.

**Figura 3.2**



*Fonte: elaborazione propria.*

Subito dopo aver accettato il consenso informato, i partecipanti sono stati esposti ad una breve introduzione che li invitava a immaginare una specifica situazione decisionale. La piattaforma, attraverso la randomizzazione casuale, assegnava una delle sei condizioni del disegno 3x2, combinando le tre diverse fonti del consiglio (IA, umano, ibrido) con i due domini decisionali (investimento, donazione). In questo modo attraverso la randomizzazione

automatica, si è garantito che ciascuna delle condizioni sperimentali fosse equamente rappresentata nell'esperimento.

Nella presente sezione sono state rilevate le principali variabili dipendenti e i costrutti percettivi (fiducia nel consiglio, affidabilità attribuita alla fonte, competenza percepita, calore/empatia e intenzione di seguire la raccomandazione). Tutti gli item sono stati formulati su una scala Likert a 7 punti per consentire una buona sensibilità nelle risposte, senza tuttavia rendere la procedura di compilazione troppo onerosa.

Una volta conclusa questa parte, il questionario passava alle variabili di moderazione individuale. Queste sono le variabili che, secondo il modello descritto nel Capitolo 2, possono influenzare l'esito dell'esperimento. Pertanto, sono state poste domande sulla preferenza per la fonte umana, sulla familiarità con l'intelligenza artificiale, sulla propensione al rischio e sull'orientamento prosociale. Solo in un secondo momento sono state poste domande sulle variabili demografiche. Questa sequenza non è casuale: collocare le domande anagrafiche in chiusura aiuta a garantire che i partecipanti rispondano spontaneamente agli item delle variabili principali e limita il rischio che il partecipante entri nello scenario con un mindset eccessivamente auto-riflessivo o classificatorio.

La somministrazione del questionario ha richiesto, in media, tra i 2 e i 3 minuti. Si tratta di un tempo volutamente breve. La logica alla base dell'utilizzo di una procedura così breve favorisce il completamento, riduce l'abbandono e limita l'affaticamento cognitivo. La procedura breve limita anche l'uso di strumenti complessi. Durante la costruzione di questo questionario, sono state prese diverse decisioni per garantirne l'equilibrio e la sostenibilità attraverso un'attenta selezione degli item.

Un ulteriore aspetto di questo processo è l'uniformità nel modo in cui il questionario è stato somministrato. A parte la manipolazione sperimentale, tutti i partecipanti hanno sperimentato lo stesso processo con la stessa sequenza di blocchi e lo stesso carico di compilazione. L'unica differenza tra i partecipanti riguardava il contenuto dello scenario mostrato loro (fonte e dominio), non le modalità di somministrazione.

La sezione successiva (§3.4) presenterà invece le strategie di analisi adottate e i risultati empirici ottenuti dal dataset così costruito. Tuttavia, l'attenzione di questa sezione era rivolta alla descrizione del questionario e del processo di costruzione e somministrazione.

### **3.4 Analisi dei dati**

Dopo aver definito il disegno di ricerca, descritto la struttura del questionario e chiarito la composizione del campione, questa sezione si concentra sull'analisi dei dati raccolti durante l'esperimento.

Mentre i capitoli precedenti hanno costruito il quadro teorico della tesi, mostrando come la fiducia nelle raccomandazioni algoritmiche dipenda dalla fonte del consiglio, dal contesto decisionale e dalle percezioni di competenza, calore e affidabilità, in questo capitolo, la ricerca si sposterà dalla discussione delle teorie all'analisi dei dati per verificare se tali teorie si manifestino nei dati.

La trattazione in questo capitolo seguirà una struttura progressiva, iniziando con la descrizione della preparazione dei dati. Dopo la preparazione dei dati per l'analisi, il capitolo tratterà le statistiche descrittive, che aiuteranno a evidenziare le differenze più evidenti tra le sei diverse condizioni sperimentali. Su questa base si innestano i test principali del disegno (le ANOVA fattoriali  $3 \times 2$ ) che permettono di verificare se le differenze tra le condizioni sperimentali risultino effettivamente sistematiche sul piano statistico.

In questo senso, il §3.4 rappresenta il vero punto di contatto tra il modello concettuale elaborato nel Capitolo 2 e la verifica empirica sviluppata nel Capitolo 3. Pertanto, ciascuna delle analisi statistiche presentate in questo capitolo contribuisce a chiarire il problema teorico indagato nella tesi: ovvero, capire se, quando e per chi le raccomandazioni AI risultino più credibili, più persuasive o più legittime rispetto a quelle umane o ibride.

#### **3.4.1 Preparazione dei dati**

La fase preliminare alle analisi vere e proprie consiste in una serie di operazioni di preparazione del dataset per renderlo coerente con gli obiettivi della ricerca e con la struttura del disegno sperimentale. Il processo di preparazione prevede innanzitutto il controllo della completezza delle risposte da parte dei partecipanti e della corretta codifica delle variabili rilevate.

Dato che nelle survey online la compilazione è molto rapida, è fondamentale trasformare un insieme di risposte grezze in un archivio ordinato, leggibile e metodologicamente affidabile.

È stato necessario effettuare il procedimento di costruzione delle variabili sperimentali derivate. Ad esempio, la variabile fonte3 è stata utilizzata per distinguere tra i consigli di tipo IA, umano e ibrido; il contesto2 è stato utilizzato per distinguere tra i due domini di svolgimento dei consigli: gli investimenti e le donazioni; la variabile scenario6 è stata invece costruita per sintetizzare le 3 diverse fonti di consiglio e i 2 diversi domini di svolgimento in ogni singolo caso, pertanto consentendo di individuare gli 6 scenari distinti del disegno sperimentale.

Inoltre, è stato necessario costruire i diversi indici sintetici dei costrutti multi-item di fiducia, competenza e calore, che sono stati calcolati come media delle diverse variabili contenute nei diversi costrutti (come discusso nel §3.2.3). Invece le variabili rilevate tramite singolo item (come intenzione e affidabilità) sono state mantenute nella loro forma originaria.

Infine, le variabili necessarie per le analisi di moderazione (come l'uso dell'IA) sono state convertite in formato numerico, così da consentirne l'inclusione nei modelli lineari e ordinali presentati nei paragrafi successivi.

Nel complesso, questa fase ha avuto la funzione di garantire che le analisi successive siano condotte su una base dati pulita ed allineata al disegno sperimentale. Nei paragrafi che seguono, tale base verrà dapprima esplorata sul piano descrittivo e poi sottoposta a verifica inferenziale.

### **3.4.2 Statistiche descrittive**

In questo sottocapitolo ci soffermeremo su una lettura descrittiva delle principali variabili dipendenti nelle sei condizioni sperimentali.

L'obiettivo di questa fase non è quello di indicare se le differenze in ciascuno scenario siano statisticamente significative, ma solo di esplorare i pattern preliminari dei dati e di comprendere la direzione in cui sembrano evolversi le valutazioni dei partecipanti. Pertanto, mentre in questa sezione verranno esplorate le differenze nelle medie all'interno di ciascuna condizione, sarà nella sezione successiva (§3.4.3) che esse verranno formalmente verificate come statisticamente significative attraverso l'esecuzione di un test ANOVA fattoriale 3x2.

**Tabella 3.7 – Statistiche descrittive delle variabili**

*dipendenti nelle sei condizioni sperimentali*

<b>IA – INVESTIMENTO</b>	N	Mean	SD
fiducia	60	5.217	1.221
calore	60	3.394	1.415
competenza	60	5.167	1.239
intenzione	60	4.75	1.723
affidabilità	60	4.983	1.501

<b>UMANO – DONAZIONE</b>	N	Mean	SD
fiducia	81	5.284	1.033
calore	81	5.222	1.382
competenza	81	5.35	0.917
intenzione	81	5	1.342
affidabilità	81	5.111	1.294

**IA – DONAZIONE**

fiducia	74	3.752	1.372
calore	74	3	1.408
competenza	74	4.077	1.097
intenzione	74	3.297	1.506
affidabilità	74	3.568	1.545

**IBRIDO – INVESTIMENTO**

fiducia	64	5.24	1.165
calore	64	4.797	1.555
competenza	64	5.25	1.119
intenzione	64	4.672	1.653
affidabilità	64	4.984	1.548

**UMANO – INVESTIMENTO**

fiducia	69	3.99	1.180
calore	69	4.812	1.081
competenza	69	4.155	1.033
intenzione	69	3.536	1.208
affidabilità	69	3.928	1.287

**IBRIDO – DONAZIONE**

fiducia	54	5.296	1.192
calore	54	4.759	1.645
competenza	54	5.148	1.359
intenzione	54	4.796	1.687
affidabilità	54	5.056	1.420

Fonte: elaborazione propria.

Un primo elemento da segnalare riguarda la distribuzione del campione nelle sei condizioni. Infatti, come già osservato nel §3.3.2, la randomizzazione dei partecipanti ha prodotto una distribuzione relativamente uniforme (anche se non perfettamente equilibrata).

La prima variabile da considerare è la fiducia. Le medie mostrano subito che la fiducia cambia a seconda della combinazione tra fonte e contesto. Le medie più alte si registrano nelle due condizioni ibride (5,24 nell'Ibrido/Investimento e 5,30 nell'Ibrido/Donazione) e nella condizione Umano/Donazione, che si colloca a 5,28. Anche IA/Investimento presenta un valore elevato (5,22). Le medie più basse si concentrano invece in IA/Donazione (3,75) e in Umano/Investimento (3,99). Già da questi numeri si intravede una possibile logica di "allineamento" tra fonte e dominio: infatti, l'IA sembra essere valutata più favorevolmente nel contesto dell'investimento, mentre l'umano raggiunge dei valori più alti nel contesto della donazione. L'ibrido, invece, non è mai penalizzato in modo evidente mantenendo livelli alti in entrambi i contesti. Infine, le deviazioni standard all'interno di ciascuno scenario sono risultate relativamente basse (circa 1,03 e 1,37 in ciascuno scenario). Tale dato sta ad indicare che i partecipanti tendono a valutare la fiducia in modo abbastanza simile all'interno di ogni scenario, senza una variabilità eccessiva.

La seconda variabile che analizziamo è il calore/empatia percepita. Qui la condizione meno favorevole è chiaramente IA/Donazione che riceve una valutazione media di 3,00, mentre la condizione IA/Investimento riceve 3,39. Le condizioni con fonte umana o mista si collocano invece su livelli molto più alti: infatti, Umano/Donazione raggiunge 5,22, Umano/Investimento arriva a 4,81, Ibrido/Investimento a 4,80 e Ibrido/Donazione a 4,76.

Già a livello descrittivo si può notare che il calore è la dimensione in cui la differenza tra IA pura e presenza umana (piena o almeno parziale) si manifesta con maggiore evidenza. Tuttavia, è anche interessante notare che, a differenza della fiducia, le condizioni ibride sono molto più vicine, in termini di media, a quelle delle risposte umane che alle condizioni IA. La dispersione delle risposte in questa categoria è in qualche modo più ampia rispetto alle altre variabili, soprattutto nelle condizioni ibride, dove la deviazione standard supera 1,55 in entrambi i contesti. Tale maggiore dispersione ci suggerisce che in questa categoria vi sia una maggiore eterogeneità di giudizio su una dimensione che risente di sensibilità individuali contrastanti.

La competenza percepita presenta un andamento simile a quello della fiducia. Infatti, le medie più elevate si osservano in Umano/Donazione (5,35), Ibrido/Investimento (5,25), IA/Investimento (5,17) e Ibrido/Donazione (5,15). Più indietro, invece, si collocano Umano/Investimento (4,15) e soprattutto IA/Donazione (4,08). In base a questi dati si osserva che la competenza non viene attribuita in modo uniforme alle fonti: infatti, l'IA che nel dominio dell'investimento è giudicata molto conveniente perde forza nella donazione, mentre il contrario accade alla fonte umana. Anche in questo caso l'ibrido si conferma relativamente stabile, mantenendo valori alti in entrambe le condizioni. Le deviazioni standard sono contenute (intorno a 1,10), che sta ad indicare una valutazione della competenza abbastanza compatta.

Come accaduto in precedenza anche l'intenzione di seguire i consigli mostra un andamento che riflette ancora una volta il profilo di fiducia nelle fonti. Le condizioni più favorevoli sono Umano/Donazione (5,00), Ibrido/Donazione (4,80), IA/Investimento (4,75) e Ibrido/Investimento (4,67). Più deboli, invece, Umano/Investimento (3,54) e soprattutto IA/Donazione (3,30). In base a questi valori si intuisce che l'intenzione non sembra premiare in modo indiscriminato una singola fonte, ma piuttosto la coerenza percepita tra fonte e tipo di decisione. L'ibrido mostra una capacità di adattamento più ampia rispetto alle condizioni "pure", con punteggi medi concentrati nella parte alta dell'intervallo di valori possibili. Le deviazioni standard in questo caso sono comprese tra 1,21 e 1,72, il che indica che, sebbene vi

sia un certo grado di variabilità tra i partecipanti nelle loro intenzioni di seguire i consigli. Tale variabilità non è però anomala poiché a differenza degli altri costrutti (come competenza o affidabilità) in cui si rileva una semplice valutazione, in questo caso entra in gioco anche una componente comportamentale più variabile dei semplici giudizi, perché entrano in gioco fattori personali, esitazioni, preferenze individuali e propensione all'azione.

Infine, l'ultima delle variabili da discutere in relazione alle statistiche descrittive riguarda la percezione dei partecipanti circa l'affidabilità percepita della fonte.

I valori più elevati si registrano in Umano/Donazione (5,11), Ibrido/Donazione (5,06), Ibrido/Investimento (4,98) e IA/Investimento (4,98, praticamente coincidente con la condizione ibrida nell'investimento). Le medie più basse compaiono invece in IA/Donazione (3,57) e Umano/Investimento (3,93). Anche in questo caso si conferma un quadro già intravisto nelle altre variabili. L'IA sembra funzionare meglio nell'investimento che nella donazione, l'umano mostra il profilo inverso, mentre l'ibrido conserva una posizione relativamente alta e stabile.

Un'ulteriore lettura descrittiva riguarda alcune variabili di controllo rilevate nella parte finale del questionario, considerate qui come primo riscontro ex post della coerenza tra autodichiarazioni individuali e orientamento delle risposte. In particolare, la preferenza verso l'umano e il livello di comfort verso i consigli IA vengono messi in relazione con la fonte del consiglio, osservando come vari l'intenzione di seguire il consiglio nelle tre configurazioni di fonte. La propensione al rischio e il comportamento prosociale vengono invece letti in relazione, rispettivamente, al contesto dell'investimento e a quello della donazione, sempre prendendo come riferimento la variabile intenzione. L'obiettivo, anche in questo caso, non è ancora quello di formulare una verifica inferenziale, ma di osservare se emergano pattern descrittivi coerenti con le aspettative teoriche.

<b>Tabella 3.8 – Intenzione media di seguire il consiglio per livello di preferenza verso l'umano e comfort verso i consigli IA</b>				
<i>Variabile di controllo</i>	<i>Gruppo</i>	<i>IA</i>	<i>Umano</i>	<i>Ibrido</i>
<b>Preferenza verso l'umano</b>	Basso (1–3)	5,67	4,00	5,27
	Alto (5–7)	3,34	4,66	4,44
<b>Livello di comfort verso i consigli IA</b>	Basso (1–3)	2,91	4,23	3,52
	Alto (5–7)	4,90	4,52	5,20

*Fonte: elaborazione propria.*

**Nota.** La tabella riporta le medie della variabile *intenzione di seguire il consiglio*. Per semplicità espositiva sono mostrati solo i livelli bassi (1–3) e alti (5–7) delle variabili di controllo.

Anche queste evidenze descrittive risultano coerenti con le aspettative teoriche. In primo luogo, i partecipanti con alta preferenza verso l'umano mostrano livelli di intenzione più elevati quando la fonte del consiglio è umana ( $M = 4,66$ ) o ibrida ( $M = 4,44$ ) rispetto alla fonte IA ( $M = 3,34$ ). Al contrario, tra i soggetti con bassa preferenza verso l'umano, l'intenzione risulta più alta nelle condizioni IA ( $M = 5,67$ ) e ibride ( $M = 5,27$ ). In secondo luogo, un maggiore comfort verso i consigli IA si associa a livelli più elevati di intenzione soprattutto nelle condizioni IA (4,90 contro 2,91) e ibride (5,20 contro 3,52), mentre la differenza è molto più contenuta nella fonte umana.

<b>Tabella 3.9 – Intenzione media di seguire il consiglio per livello di propensione al rischio e comportamento prosociale</b>			
<i>Variabile di controllo</i>	<i>Gruppo</i>	<i>Investimento</i>	<i>Donazione</i>
<b>Propensione al rischio</b>	Bassa (1–3)	3,25	—
	Alta (5–7)	4,83	—
<b>Comportamento prosociale</b>	Basso (1–3)	—	3,23
	Alto (5–7)	—	4,41

**Fonte:** elaborazione propria.

**Nota.** La tabella riporta le medie della variabile *intenzione di seguire il consiglio*. Per semplicità espositiva sono mostrati solo i livelli bassi (1–3) e alti (5–7) delle variabili di controllo.

Analogamente, nel solo contesto dell'investimento, i partecipanti con alta propensione al rischio riportano un'intenzione media più elevata rispetto a quelli con bassa propensione al rischio (4,83 contro 3,25). Nel solo contesto della donazione, infine, i soggetti con alto orientamento prosociale mostrano un'intenzione media più elevata rispetto a quelli con basso orientamento prosociale (4,41 contro 3,23). Nel complesso, queste evidenze non costituiscono ancora una prova inferenziale, ma offrono un primo riscontro descrittivo della coerenza tra disposizioni autodichiarate e orientamento delle risposte nei diversi scenari.

Nel complesso, possiamo affermare che fiducia, competenza, intenzione e affidabilità sembrano seguire una struttura molto simile, mentre il calore introduce una dinamica parzialmente diversa, più direttamente legata alla presenza di una componente umana. In termini puramente descrittivi, questa regolarità appena descritta potrebbe essere uno dei risultati più interessanti dell'intera tabella. Il secondo risultato interessante riguarda le condizioni "pure" (IA/Investimento e Umano/Donazione) che tendono a emergere come configurazioni

favorevoli, come a suggerire che la fonte venga premiata quando appare “adatta” al tipo di decisione. Mentre la fonte ibrida si distingue per una maggiore stabilità: infatti, le fonti ibride non hanno ricevuto con la stessa frequenza i punteggi più alti dai partecipanti, ma hanno quasi sempre ottenuto punteggi elevati, indipendentemente dal contesto. A questa lettura si aggiunge il primo riscontro offerto dalle variabili di controllo, che mostra una coerenza descrittiva tra autodichiarazioni individuali e intenzione di seguire il consiglio. Infine, sebbene queste differenze siano di natura puramente descrittiva, orientano già la lettura del paragrafo successivo, nel quale tali pattern verranno sottoposti a verifica inferenziale tramite ANOVA fattoriale  $3 \times 2$ , così da stabilire se gli scarti tra scenari siano effettivamente attribuibili alle manipolazioni sperimentali e non a semplici fluttuazioni descrittive.

### 3.4.3 Analisi principali: ANOVA fattoriale $3 \times 2$

Al fine di valutare gli effetti delle manipolazioni sperimentali sulle principali variabili dipendenti, per ciascun esito è stata eseguita un'ANOVA fattoriale  $3 \times 2$  con la Fonte del consiglio (IA, Umano, Ibrido) e il contesto decisionale (Investimento, Donazione) come fattori tra-soggetti. Questa scelta metodologica è coerente con il disegno della ricerca perché permette di valutare contemporaneamente sia gli effetti principali dei due fattori sia la loro interazione. Il punto non è solo capire se una fonte venga valutata meglio o peggio di un'altra in media. Piuttosto, se il suo rendimento cambia quando il consiglio si sposta da una decisione di investimento a una decisione di donazione. Per questo motivo, non guarderemo solo alla significatività statistica, ma anche a come le medie marginali e i grafici di interazione aiutano a capire le differenze tra gli scenari.

#### **Fiducia nel consiglio**

La prima variabile da considerare è la fiducia nel consiglio. Tale variabile rappresenta il punto di partenza più naturale per la discussione di questi risultati, poiché intercetta in modo diretto il tema centrale della tesi. I risultati dell'ANOVA per questa variabile mostrano un effetto principale della fonte statisticamente significativo,  $F(2, 396) = 14,86, p < .001$ . Pertanto, in media il tipo di fonte influenza la fiducia che i partecipanti attribuiscono al consiglio. Inoltre, mentre il contesto in cui è stato fornito il consiglio non ha avuto un impatto statisticamente significativo sulla fiducia nella fonte nel suo complesso ( $F(1, 396) = 0,10, p = 0,752$ ), l'interazione tra queste due variabili (Fonte  $\times$  Contesto) è risultata statisticamente significativa

( $F(2, 396) = 46,69, p < 0,001$ ). In altre parole, la fonte non è giudicata sempre allo stesso modo, la sua efficacia dipende in gran parte dal dominio decisionale in cui il consiglio è collocato. Le medie marginali aiutano a chiarire subito il senso di questo risultato. Nel contesto degli investimenti, i partecipanti sembravano fidarsi maggiormente della fonte di consulenza basata sull'IA ( $M = 5,22$ ) e della condizione ibrida ( $M = 5,30$ ), mentre tale livello di fiducia si abbassa sensibilmente quando la fonte è umana ( $M = 3,99$ ). Nel contesto della donazione, il profilo si rovescia in quanto la fiducia attribuita alla fonte IA scende nettamente ( $M = 3,75$ ), mentre tale fiducia cresce sensibilmente nella condizione umana ( $M = 5,28$ ) e resta alta anche nell'ibrido ( $M = 5,27$ ).

Il punto non è che l'IA venga apprezzata o rifiutata in modo generale. Piuttosto conta quale tipo di decisione il partecipante immagina di dover affrontare. Pertanto, l'IA gode di maggiore fiducia in situazioni che richiedono una logica analitica e capacità di problem solving più rigorose, ma perde credibilità in situazioni legate a decisioni più etiche e prosociali. L'ibrido, al contrario, tiene in entrambi i domini. Il sistema ibrido, invece, ha ottenuto lo stesso livello di fiducia in entrambi i tipi di contesto.

Il grafico di interazione rende questo andamento molto chiaro. Le linee non sono parallele; sono incrociate. La linea relativa all'investimento parte alta con l'IA scende nella condizione umana e risale con l'ibrido. Quella della donazione segue quasi il percorso opposto: parte bassa con l'IA e raggiunge i valori più elevati con l'umano e l'ibrido. Non si osserva soltanto una differenza tra i due contesti, ma un vero cambiamento nel modo in cui le diverse fonti vengono valutate. È proprio questo a spiegare la presenza dell'interazione statistica.

I confronti di Bonferroni confermano queste interpretazioni del grafico di interazione e degli effetti principali. Nell'ambito degli investimenti, l'IA e le fonti ibride godono di una fiducia significativamente maggiore rispetto alle fonti umane, ma non vi è alcuna differenza significativa tra IA e fonti ibride. Nell'ambito delle donazioni, tuttavia, si verificava il contrario: le fonti umane e ibride godono di una fiducia significativamente maggiore rispetto alle fonti di IA, e in modo simile tra loro. Inoltre, sebbene la fonte ibrida non sia la più rilevante in assoluto in ogni esempio, è quella che gode di maggiore fiducia in modo costante in ciascun tipo di contesto.

Merita attenzione soprattutto la fonte ibrida. Pur non risultando sempre la migliore in assoluto, non mostra mai un calo marcato in nessuna delle due condizioni. Questo andamento fa pensare che la combinazione tra elemento umano e componente algoritmica rappresenta una

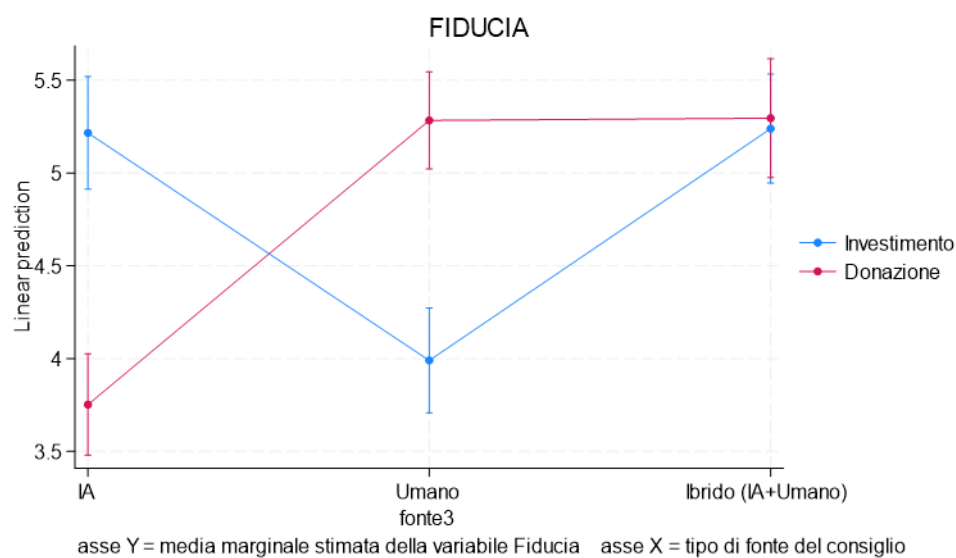
soluzione equilibrata, che è capace di mantenere alti livelli di fiducia anche in contesti molto diversi tra loro.

**Tabella 3.10 ANOVA fattoriale 3×2 – Fiducia**

Effetto	F(df1, df2)	p
Fonte	F(2, 396) = 14,86	< .001
Contesto	F(1, 396) = 0,10	0.752
Fonte × Contesto	F(2, 396) = 46,69	< .001

Fonte: elaborazione propria.

**Figura 3.3**



Fonte: elaborazione propria.

**Tabella 3.11 Principali confronti post-hoc Bonferroni – Fiducia**

Contesto	Confronto	Differenza media ( $\Delta M$ )	p Bonferroni
Investimento	IA vs Umano	1,226	< .001
Investimento	Ibrido vs Umano	1,249	< .001
Investimento	IA vs Ibrido	-0,023	1.000
Donazione	Umano vs IA	1,532	< .001
Donazione	Ibrido vs IA	1,544	< .001
Donazione	Ibrido vs Umano	0,012	1.000

### **Competenza percepita della fonte**

Passando alla dimensione della competenza percepita, il quadro rimane molto coerente rispetto a quello precedentemente analizzato ma con un'interessante sfumatura: infatti, qui il focus non è più sul consiglio in sé, bensì sulla capacità attribuita alla fonte.

Il modello ANOVA ha rivelato un effetto principale significativo della fonte:  $F(2, 396) = 9,03, p < 0,001$ . Non sono state riscontrate differenze autonome nell'effetto del contesto:  $F(1, 396) = 0,00, p = 0,992$ . Mentre, l'effetto di interazione è risultato altamente significativo:  $F(2, 396) = 37,03, p < 0,001$ .

Le medie marginali di questa dimensione sono generalmente simili a quelle relative alla dimensione della fiducia. Difatti nel contesto dell'investimento, la competenza attribuita all'IA è alta ( $M = 5,17$ ) e l'ibrido si colloca persino leggermente sopra ( $M = 5,25$ ). Invece, la fonte umana riceve una valutazione più bassa ( $M = 4,15$ ). Al contrario, nella donazione la situazione cambia: l'IA scende a  $M = 4,08$ , mentre l'umano sale a  $M = 5,35$  e l'ibrido resta su valori elevati ( $M = 5,15$ ).

Il dato più interessante è che la competenza non viene giudicata nello stesso modo in ogni situazione. Pertanto, anche questa valutazione cambia a seconda del contesto: ciò che appare competente in un ambito può esserlo meno in un altro.

Inoltre, il grafico di interazione rivela le stesse informazioni con linee ancora una volta incrociate. Nell'investimento, l'IA gode di un vantaggio netto sull'umano; nella donazione, è l'umano a prendere chiaramente il sopravvento. L'ibrido occupa una posizione elevata in entrambe le condizioni. Il cambiamento della direzione delle linee suggerisce un risultato teoricamente interessante: i partecipanti non giudicano la competenza della fonte in modo generale, ma in base a quanto la ritengono adatta al tipo di decisione da prendere.

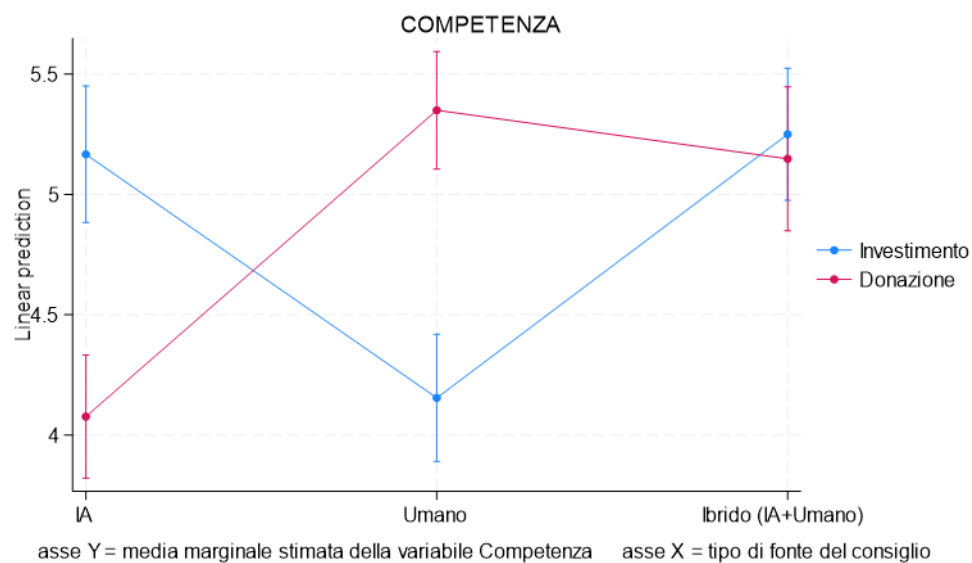
Anche i post-hoc vanno nella stessa direzione. Sia l'IA che l'ibrido ricevono punteggi di competenza significativamente più alti nel dominio dell'investimento rispetto all'umano. Al contrario, l'umano e l'ibrido sono valutati più competenti nel dominio della donazione rispetto all'IA.

La distanza tra umano e ibrido si riduce molto. Questo significa che la fonte ibrida continua a essere vista come abbastanza competente sia nell'investimento sia nella donazione. Se nel caso della fiducia l'ibrido sembrava soprattutto una fonte equilibrata e costante, qui mostra anche di sapersi avvicinare molto alla fonte che, di volta in volta, risulta più convincente.

Effetto	F(df1, df2)	p
Fonte	F(2, 396) = 9,03	< .001
Contesto	F(1, 396) = 0,00	0.992
Fonte × Contesto	F(2, 396) = 37,03	< .001

Fonte: elaborazione propria.

**Figura 3.4**



Fonte: elaborazione propria.

Contesto	Confronto	Differenza media ( $\Delta M$ )	p Bonferroni
Investimento	IA vs Umano	1,012	< .001
Investimento	Ibrido vs Umano	1,095	< .001
Investimento	Ibrido vs IA	0,083	1.000
Donazione	Umano vs IA	1,273	< .001
Donazione	Ibrido vs IA	1,072	< .001
Donazione	Ibrido vs Umano	-0,202	1.000

*Fonte: elaborazione propria.*

### **Calore/empatia percepita**

Il quadro cambia con il calore. Infatti, a differenza di quanto osservato in precedenza, l'ANOVA segnala un effetto principale della fonte molto netto,  $F(2, 396) = 66,06, p < 0,001$ . La fonte, qui, conta molto. Il contesto non ha mostrato un effetto statisticamente significativo sulle valutazioni di calore,  $F(1, 396) = 0,03, p = .960$ . A differenza delle variabili precedenti, l'interazione Fonte  $\times$  Contesto non raggiunge la soglia convenzionale di significatività, pur avvicinandosi ad essa,  $F(2, 396) = 2,87, p = .058$ .

È un dato da leggere con attenzione, perché non indica una completa assenza di differenze tra i due domini, ma indicano che il fattore dominante sia soprattutto chi fornisce il consiglio.

Le medie marginali supportano questa conclusione. Infatti, l'IA ottiene punteggi bassi in entrambi i contesti, con una media di 3,39 nei contesti di investimento e 3,00 nei contesti di donazione. L'essere umano ha ottenuto punteggi significativamente più alti, pari a 4,81 nei contesti di investimento e 5,22 nei contesti di donazione. Il sistema ibrido ha ottenuto punteggi simili a quelli del sistema umano, rispettivamente di 4,80 nei contesti di investimento e 4,76 nei contesti di donazione.

Pertanto, l'unica vera differenza tra i due ambiti riguarda il tipo di fonte da cui proviene il consiglio. L'IA è percepita come fredda, mentre la fonte umana e quella ibrida appaiono più vicine, attente e comprensive.

Le linee del grafico non sono perfettamente parallele, ma non mostrano neppure un incrocio marcato. In questo caso non si produce una inversione netta come visto precedentemente nelle altre variabili, e lo si nota dall'andamento delle linee lievemente divergente. Questa configurazione permette di spiegare perché l'interazione non è così netta come in precedenza, poiché la differenza di calore viene mitigata dalla presenza umana nell'ibrido.

I confronti Bonferroni ci vengono in aiuto rafforzando questa interpretazione. L'intelligenza artificiale ottiene punteggi significativamente diversi dalle altre due fonti. Le differenze tra i punteggi del sistema umano e di quello ibrido sono quasi trascurabili e spesso

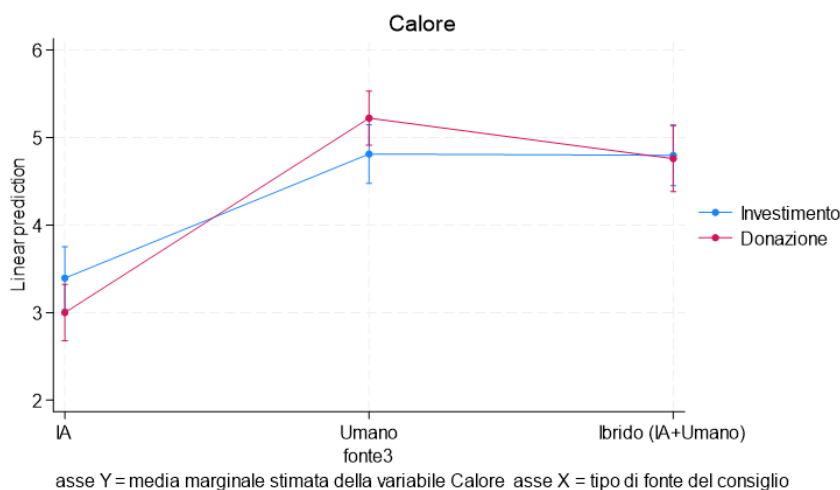
non statisticamente significative. Questo dato è interessante perché permette di comprendere come la presenza anche non esclusiva della componente umana basti a spostare in alto la percezione di empatia e vicinanza. Qui l'ibrido non appare solo come una soluzione intermedia, ma come una fonte che conserva buona parte della vicinanza tipica dell'umano e quindi delle sue caratteristiche.

**Tabella 3.14 ANOVA fattoriale 3×2 – Calore**

Effetto	F(df1, df2)	p
Fonte	F(2, 396) = 66,06	< .001
Contesto	F(1, 396) = 0,00	0.960
Fonte × Contesto	F(2, 396) = 2,87	0.058

Fonte: elaborazione propria.

**Figura 3.5**



Fonte: elaborazione propria.

**Tabella 3.15 Principali confronti post-hoc Bonferroni – Calore**

Contesto	Confronto	Differenza media ( $\Delta M$ )	p Bonferroni
Investimento	Umano vs IA	1,417	< .001
Investimento	Ibrido vs IA	1,402	< .001
Investimento	Ibrido vs Umano	-0,015	1.000
Donazione	Umano vs IA	2,222	< .001
Donazione	Ibrido vs IA	1,759	< .001
Donazione	Ibrido vs Umano	-0,463	0.944

Fonte: elaborazione propria.

### Intenzione di seguire il consiglio

L'intenzione di seguire il consiglio avvicina l'analisi alla dimensione comportamentale dello studio. Anche se non misura un comportamento reale ma solo una disponibilità dichiarata, questa variabile mostra quando la valutazione del consiglio può tradursi in un'azione. L'ANOVA mostra un effetto principale della fonte significativo,  $F(2, 396) = 7,01, p = .001$ . Il contesto, ancora una volta, non produce differenze autonome,  $F(1, 396) = 0,09, p = .767$ . Infine, si osserva un'interazione Fonte  $\times$  Contesto significativa  $F(2, 396) = 32,58, p < 0,001$ .

Nell'ambito dell'investimento, gli individui mostrano una maggiore intenzione di seguire i consigli dell'IA ( $M = 4,75$ ) e dell'ibrido ( $M = 4,68$ ), ma un'intenzione significativamente inferiore di seguire il consiglio del consulente umano ( $M = 3,54$ ). Nella donazione, tuttavia, gli individui mostrano la massima propensione a seguire il consiglio di consulente umano ( $M = 5,00$ ), e del consulente ibrido ( $M = 4,80$ ), mentre la propensione a seguire i consigli dell'IA è minima ( $M = 3,30$ ). Il contesto non modifica la disponibilità ad accettare un consiglio, ma cambia quale fonte il partecipante considera degna di essere seguita.

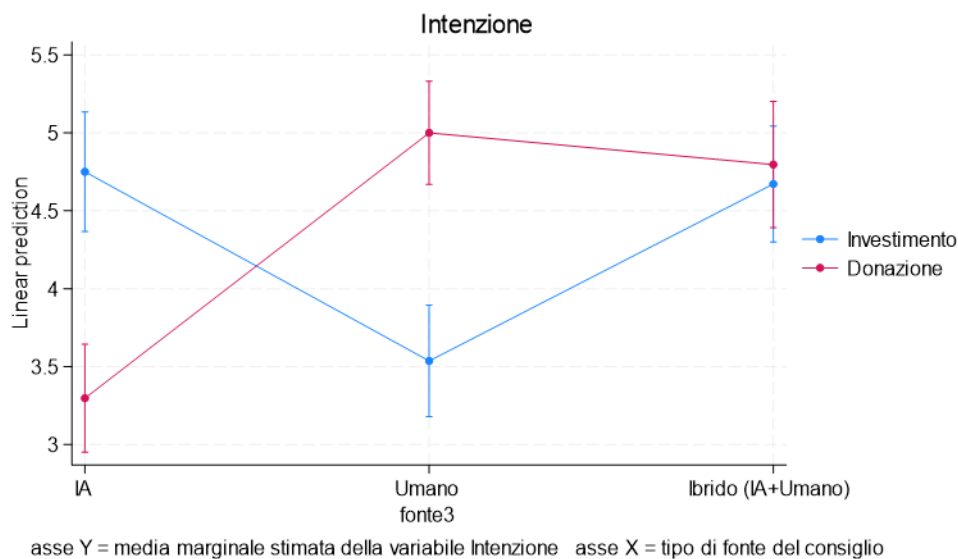
Nel grafico troviamo linee fortemente incrociate. L'investimento favorisce l'IA e l'ibrido, mentre la donazione favorisce l'umano e, in misura leggermente inferiore, l'ibrido. Anche in questo caso l'ibrido rappresenta un profilo meno polarizzato e più stabile.

I confronti di Bonferroni supportano questi risultati. Nell'investimento, IA e ibrido superano significativamente la fonte umana; nella donazione, umano e ibrido si collocano sopra l'IA. Anche qui la distanza tra umano e ibrido tende a essere contenuta. Pertanto, ancora una volta, questi risultati indicano la capacità dell'ibridazione di mantenere elevata l'intenzione di seguire il consiglio anche quando cambiano le aspettative sulla fonte del consiglio.

<b>Effetto</b>	<b>F(df1, df2)</b>	<b>p</b>
Fonte	$F(2, 396) = 7,01$	.001
Contesto	$F(1, 396) = 0,09$	0.767
Fonte $\times$ Contesto	$F(2, 396) = 32,58$	< .001

*Fonte: elaborazione propria.*

**Figura 3.6**



Fonte: elaborazione propria.

Contesto	Confronto	Differenza media ( $\Delta M$ )	p Bonferroni
Investimento	IA vs Umano	1,214	< .001
Investimento	Ibrido vs Umano	1,136	< .001
Investimento	Ibrido vs IA	-0,078	1.000
Donazione	Umano vs IA	1,703	< .001
Donazione	Ibrido vs IA	1,499	< .001
Donazione	Ibrido vs Umano	-0,204	1.000

Fonte: elaborazione propria.

### Affidabilità percepita della fonte

L'ultima variabile considerata è l'affidabilità percepita della fonte. Tale variabile permette di collegare il giudizio sulla fonte alla disponibilità degli individui a seguire i suoi suggerimenti.

Anche in questo caso, la tabella ANOVA rivela un effetto principale significativo per il fattore fonte,  $F(2, 396) = 8,66$ ,  $p < 0,001$ . Il contesto non ha un effetto indipendente sull'affidabilità percepita della fonte del consiglio,  $F(1, 396) = 0,14$ ,  $p = 0,710$ . Tuttavia, si osserva un'interazione significativa tra le due variabili,  $F(2, 396) = 28,97$ ,  $p < 0,001$ .

Le medie marginali riproducono un profilo ormai familiare, tranne che in un particolare. Qui il focus non è più sul consiglio, ma sulla fonte in termini generali. Le raccomandazioni basate sull'IA ( $M = 4,98$ ) e quella ibrida ( $M = 4,95$ ) sono state considerate più affidabili della raccomandazione umana ( $M = 3,93$ ) nello scenario di investimento, mentre nello scenario

relativo alle donazioni la raccomandazione umana ( $M = 5,11$ ) insieme all'ibrida ( $M = 5,06$ ) sono considerate più affidabili delle raccomandazioni basate sull'IA ( $M = 3,57$ ). Pertanto, da questi risultati si può dedurre che l'affidabilità di ciascuna di queste fonti dipende dall'adeguatezza di tale potenziale fonte per lo scenario proposto.

Come già visto in precedenza, anche in questo caso il grafico di interazione mostra linee chiaramente incrociate, che mostrano come la direzione delle valutazioni cambi a seconda della fonte e del tipo di compito. Tuttavia, in ogni caso, l'intelligenza artificiale gode di fiducia nel contesto degli investimenti, ma non in quello delle donazioni. Le fonti umane godono di fiducia nel contesto opposto: sono affidabili per le donazioni, ma non per gli investimenti. Infine, il sistema ibrido mantiene un punteggio elevato in entrambi i contesti, indicando che la combinazione tra componente algoritmica e intervento umano tende a rendere la valutazione della fonte più stabile e meno sensibile al cambiamento del contesto decisionale.

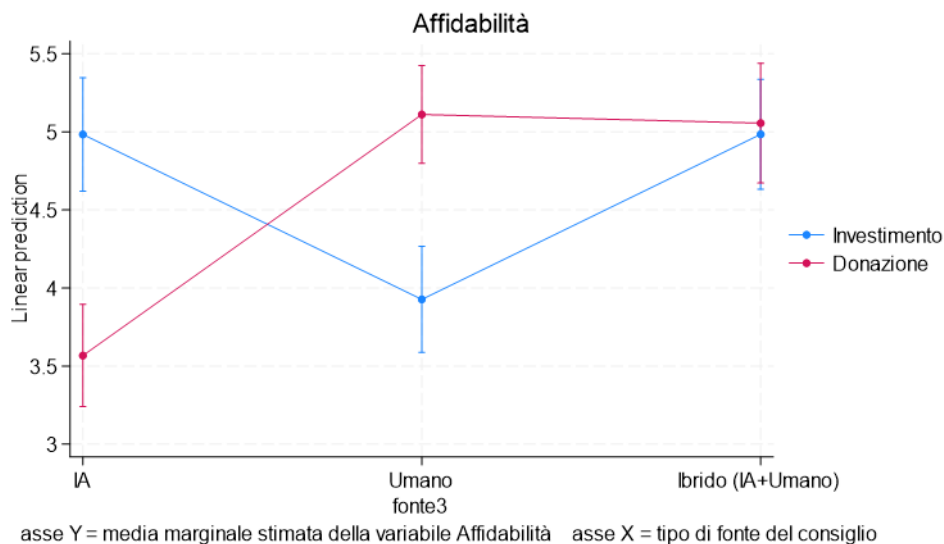
I test post-hoc supportano queste interpretazioni. Le differenze più forti riguardano, ancora una volta, il confronto tra IA e umano in funzione del dominio, mentre la condizione ibrida resta vicina alla fonte che ottiene la valutazione migliore.

Questo dato prepara in modo naturale il passaggio successivo. Se fiducia, intenzione e affidabilità seguono pattern tanto simili, diventa legittimo chiedersi fino a che punto questi esiti siano soltanto reazioni parallele agli scenari e fino a che punto siano invece legati tra loro da rapporti più profondi.

<b>Effetto</b>	<b>F(df1, df2)</b>	<b>p</b>
Fonte	$F(2, 396) = 8,66$	$< .001$
Contesto	$F(1, 396) = 0,14$	0.710
Fonte × Contesto	$F(2, 396) = 28,97$	$< .001$

*Fonte: elaborazione propria.*

**Figura 3.7**



Fonte: elaborazione propria.

Contesto	Confronto	Differenza media ( $\Delta M$ )	p Bonferroni
Investimento	IA vs Umano	1,056	0.001
Investimento	Ibrido vs Umano	1,057	< .001
Investimento	IA vs Ibrido	-0,001	1.000
Donazione	Umano vs IA	1,544	< .001
Donazione	Ibrido vs IA	1,488	< .001
Donazione	Ibrido vs Umano	-0,056	1.000

Fonte: elaborazione propria.

### Sintesi dei risultati ANOVA

Nel complesso, le cinque analisi ANOVA raccontano la stessa storia. L'effetto del contesto da solo non è mai significativo. Ciò che conta davvero è quale fonte venga associata a ciascun dominio. Difatti, l'aspetto più importante degli studi è l'interazione tra fonte e contesto, che risulta significativa per le variabili relative a fiducia, competenza, intenzione e affidabilità, ma resta solo marginale per calore. Il pattern è stabile e mostra che nel dominio degli investimenti, l'IA e la fonte ibrida sono percepite più favorevolmente rispetto alla fonte umana. Mentre nel dominio della donazione, vantaggio passa alla fonte umana, con l'ibrido che però mantiene livelli elevati e mostra una maggior continuità tra i due domini.

Come abbiamo notato, il calore segue una logica in parte diversa, in quanto, qui il fattore cruciale è la fonte in sé. L'IA viene percepita come più fredda, mentre l'umano e l'ibrido

risultano molto più forti sul piano relazionale, indipendentemente dal contesto. Il quadro risulta più definito per la fiducia, la competenza, l'intenzione e l'affidabilità dove la valutazione della fonte cambia in funzione del tipo di decisione.

Pertanto, i partecipanti non valutano la fonte di intelligenza artificiale o quella umana indipendentemente dal tipo di decisione da prendere, ma li collocano dentro aspettative differenti a seconda che il problema appaia più analitico oppure più valoriale.

Un altro risultato interessante emerge dalla valutazione dei risultati della condizione ibrida. In nessuna delle variabili la valutazione delle fonti ibride diminuisce in modo significativo. Di fatto, le fonti "ibride" si sono costantemente classificate tra le migliori per ciascuna delle variabili. Questo pattern è coerente con l'idea sviluppata nel Capitolo 2 dove si ipotizzava che la combinazione tra calcolo algoritmico e interpretazione umana possa rendere la fonte più credibile rispetto a una fonte puramente algoritmica o puramente umana, poichè capace di tenere insieme competenza e relazionalità.

A questo punto, però, l'ANOVA ha esaurito la sua funzione principale, mostrando **dove** e **come** gli scenari differiscono. Resta però aperta un'ulteriore domanda (che guida il §3.4.4): fino a che punto variabili come fiducia, competenza, calore e affidabilità contribuiscono a spiegare gli esiti finali, in particolare, l'intenzione di seguire il consiglio? Per rispondere a questo interrogativo occorre passare da un livello strettamente sperimentale a un livello più propriamente esplicativo.

### **3.4.4 Il peso delle variabili percettive nei due domini decisionali**

Dopo aver verificato nel §3.4.3 che le differenze tra gli scenari non sono casuali, il passo successivo consiste nel chiarire in che modo le principali variabili percettive contribuiscano a spiegare gli esiti finali della ricerca. In questa sezione l'attenzione si concentra su fiducia, competenza e calore come predittori di intenzione di seguire il consiglio e di affidabilità attribuita alla fonte, con l'obiettivo di verificare se il loro impatto cambi a seconda del dominio decisionale. In altri termini, non si tratta più di chiedersi quale fonte sia valutata meglio in assoluto, ma di capire se le dimensioni percettive che orientano il giudizio mantengano lo stesso peso nell'investimento e nella donazione.

Per rispondere a questa domanda sono stati stimati due modelli ordered logit con errori robusti, uno sull'intenzione e uno sull'affidabilità, includendo termini di interazione tra

ciascuna variabile percettiva e il contesto decisionale. Questa impostazione consente di verificare in modo diretto se l'effetto di fiducia, competenza e calore cambi tra investimento e donazione.

<b>Tabella 3.20 – Modello ordered logit dell'intenzione di seguire il consiglio: effetti di fiducia, competenza, calore e interazioni con il dominio</b>			
<i>Predittore</i>	<i>Coefficiente (B)</i>	<i>SE robusto</i>	<i>p-value</i>
Fiducia	1,400	0,254	<0,001
Donazione	2,631	1,089	0,016
Fiducia × Donazione	-0,555	0,354	0,117
Competenza	1,154	0,277	<0,001
Competenza × Donazione	-0,545	0,382	0,154
Calore	0,351	0,106	0,001
Calore × Donazione	0,651	0,183	<0,001

Fonte: elaborazione propria.

**Statistiche del modello:** N = 402; Wald  $\chi^2(7) = 245,55$ ; Prob >  $\chi^2 = 0,000$ ; Pseudo R<sup>2</sup> = 0,372.

**Nota.** Categoria di riferimento del dominio: **investimento**. I coefficienti di fiducia, competenza e calore rappresentano quindi gli effetti nel dominio dell'investimento; i termini di interazione indicano di quanto tali effetti cambino nel dominio della donazione. I cut points del modello non sono riportati per ragioni di sintesi.

Nel modello relativo all'intenzione di seguire il consiglio, tutte e tre le variabili percettive risultano positivamente associate all'esito. In particolare, la fiducia mostra un coefficiente positivo e altamente significativo ( $\beta = 1,400, p < .001$ ), così come la competenza ( $\beta = 1,154, p < .001$ ) e il calore ( $\beta = 0,351, p = .001$ ). Tuttavia, il punto più rilevante riguarda le interazioni con il dominio. Né l'interazione tra fiducia e donazione né quella tra competenza e donazione risultano statisticamente significative. Al contrario, l'interazione tra calore e donazione è positiva e significativa ( $\beta = 0,651, p < .001$ ). Questo indica che il calore incide sull'intenzione in entrambi i domini, ma il suo effetto diventa significativamente più forte nel contesto della donazione.

<b>Tabella 3.21 – Modello ordered logit dell'affidabilità attribuita alla fonte: effetti di fiducia, competenza, calore e interazioni con il dominio</b>			
<i>Predittore</i>	<i>Coefficiente (B)</i>	<i>SE robusto</i>	<i>p-value</i>
Fiducia	1,106	0,323	0,001
Donazione	1,555	1,151	0,177
Fiducia × Donazione	-0,364	0,407	0,371
Competenza	1,399	0,341	<0,001
Competenza × Donazione	-0,419	0,436	0,337
Calore	0,254	0,115	0,026
Calore × Donazione	0,489	0,204	0,016

Fonte: elaborazione propria.

**Statistiche del modello:** N = 402; Wald  $\chi^2(7) = 280,00$ ; Prob >  $\chi^2 = 0,000$ ; Pseudo R<sup>2</sup> = 0,374.

**Nota.** Categoria di riferimento del dominio: **investimento**. I coefficienti di fiducia, competenza e calore rappresentano gli effetti nel dominio dell'investimento; i termini di interazione indicano di quanto tali effetti cambino nel dominio della donazione. I cut points del modello non sono riportati per ragioni di sintesi.

Un quadro molto simile emerge nel modello relativo all'affidabilità attribuita alla fonte. Anche qui fiducia ( $\beta = 1,106$ ,  $p = .001$ ), competenza ( $\beta = 1,399$ ,  $p < .001$ ) e calore ( $\beta = 0,254$ ,  $p = .026$ ) risultano predittori positivi e significativi. Ancora una volta, però, le interazioni con il contesto mostrano che la differenza più netta riguarda il calore: l'interazione tra calore e donazione è positiva e significativa ( $\beta = 0,489$ ,  $p = .016$ ), mentre quelle relative a fiducia e competenza non raggiungono la significatività statistica. Anche nel caso dell'affidabilità, dunque, la dimensione relazionale acquista un peso più marcato nel dominio prosociale.

<b>Tabella 3.22 – Sintesi dei coefficienti nei modelli separati per dominio</b>				
<i>Esito</i>	<i>Dominio</i>	<i>Fiducia</i>	<i>Competenza</i>	<i>Calore</i>
Intenzione	Investimento	1,321	1,078	0,338
Intenzione	Donazione	0,889	0,649	1,088
Affidabilità	Investimento	1,072	1,411	0,251
Affidabilità	Donazione	0,745	0,991	0,773

*Fonte: elaborazione propria.*

**Nota.** La tabella riporta i coefficienti dei modelli ordered logit stimati separatamente nei due domini. È presentata con finalità descrittive e interpretative, per rendere più leggibile il diverso peso relativo delle variabili percettive nei due contesti.

Per rendere più leggibile il contenuto sostantivo di questi risultati, sono stati stimati anche modelli separati per i due domini. Sebbene tali analisi abbiano soprattutto una funzione descrittiva e interpretativa, esse confermano la stessa tendenza generale. Nel contesto dell'investimento, l'intenzione è spiegata soprattutto da fiducia e competenza, mentre il calore mantiene un effetto positivo ma più contenuto. Nel contesto della donazione, invece, il calore emerge come la variabile più forte nel modello sull'intenzione, superando sia la fiducia sia la competenza. Un andamento analogo si osserva anche per l'affidabilità: nell'investimento prevale la componente tecnica, mentre nella donazione cresce sensibilmente il peso della componente relazionale.

Nel complesso, l'impatto delle variabili percettive non è identico nei due domini. La fiducia mantiene un ruolo importante in entrambi, mentre il calore diventa significativamente

più rilevante nel contesto della donazione. La competenza resta un predittore positivo in tutti i modelli e appare relativamente più centrale nell'investimento secondo i modelli separati, pur senza mostrare nei modelli con interazione una differenza statisticamente netta tra i due domini. La differenza tra investimento e donazione, dunque, non consiste nel fatto che una variabile scompaia completamente e un'altra la sostituisca, ma nel fatto che la loro gerarchia relativa cambi in funzione del tipo di decisione.

Questo risultato è importante perché mostra che il dominio non modifica soltanto i livelli medi delle valutazioni, ma incide anche sul modo in cui tali valutazioni si traducono negli esiti finali. In questo senso, il quadro empirico conferma che le decisioni più analitiche tendono a valorizzare maggiormente le dimensioni tecniche e cognitive, mentre le decisioni più valoriali e prosociali rafforzano il peso delle dimensioni relazionali. Su questa base, il paragrafo successivo allargherà ulteriormente l'analisi, verificando se tali pattern restino stabili oppure cambino anche in funzione delle differenze individuali dei partecipanti, in particolare della loro familiarità con l'IA e della loro preferenza verso la fonte umana o algoritmica.

#### **3.4.5 Analisi di moderazione: il ruolo dell'uso dell'IA**

Dopo aver mostrato nel §3.4.3 dove gli scenari differiscono e nel §3.4.4 attraverso quali percezioni tali differenze si traducono in esiti finali, il passo successivo è determinare se questi risultati variano in base ad alcune caratteristiche dei partecipanti. In questa sezione l'attenzione si concentra sull'uso dell'IA come moderatore. L'idea di base è che gli individui che utilizzano strumenti basati sull'intelligenza artificiale (IA) più frequentemente possano percepire le variabili dello scenario in modo diverso rispetto agli individui che utilizzano tali strumenti meno frequentemente. Ad esempio, è possibile che gli individui che utilizzano più spesso strumenti basati sull'IA mostrino percezioni diverse del livello di fiducia, competenza, calore, intenzione di aderire ai consigli altrui o affidabilità di tali altri.

Per i costrutti rappresentati da più di più item (fiducia, competenza e cordialità), verranno utilizzati modelli di regressione lineare con errori standard robusti. Per i costrutti rappresentati da singoli item (intenzione e affidabilità), verranno utilizzati modelli logit ordinali. I modelli includeranno l'uso dell'intelligenza artificiale come variabile moderatrice insieme ai due fattori sperimentali, nonché tutte le interazioni di tali variabili fino all'interazione delle tre variabili (Fonte × Contesto × Uso dell'IA). Le predictive margins saranno calcolate ai valori 2, 4 e 6 della scala di uso dell'IA e saranno interpretati come livello basso, medio e alto del moderatore. Il valore 6 sarà utilizzato per i livelli elevati dell'uso dell'IA. Inoltre, sarà

utilizzato il valore 6 e non 7 per rappresentare un uso intenso dell'IA senza coincidere con il massimo teorico della scala.

Nel caso della fiducia, il modello nel suo complesso risulta statisticamente significativo,  $F(11, 390) = 16,61$ ,  $p < 0,001$ , con un  $R^2$  di 0,328. Mentre l'interazione tra la fonte umana e l'uso dell'IA ha un coefficiente negativo e statisticamente significativo ( $\beta = -0,409$ ,  $p = 0,001$ ). Le interazioni a tre vie non sono risultate statisticamente significative. Questo significa che il moderatore agisce soprattutto sul rapporto tra partecipante e tipo di fonte, piuttosto che sulla specifica combinazione tra fonte e contesto. Guardando i margins osserviamo che nelle situazioni in cui l'uso dell'IA è basso, la fiducia resta molto alta nella condizione Umano/Donazione (5,44) e decisamente più bassa in IA/Donazione (3,04). Quando il livello di uso dell'IA è alto, il quadro cambia e l'IA/Investimento sale fino a 5,62, mentre Umano/Investimento scende a 3,92. Il grafico mostra quindi linee che invece di incrociarsi si allontanano progressivamente. Da come si comportano le linee in questo grafico osserviamo che i soggetti che un utilizzo spesso IA tendono a preferire le fonti algoritmiche o miste, mentre la fonte umana perde parte del suo vantaggio relativo.

**Tabella 3.23**  
**Linear Regression**  
**Predicting Trust**

Fiducia	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
<b>: base IA</b>							
Umano	.747	.51	1.47	.143	-.255	1.749	
Ibrido (IA+Umano)	1.044	.565	1.85	.066	-.068	2.155	*
<b>: base Investimento</b>							
Donazione	-1.267	.495	-2.56	.011	-2.24	-.294	**
<b>Fonte × Contesto</b>							
Umano × Donazione	2.633	.745	3.53	0	1.168	4.098	***
Ibrido × Donazione	.954	.814	1.17	.242	-.646	2.554	
d5_uso_ia_n	.36	.067	5.36	0	.228	.491	***
<b>Fonte × Uso IA</b>							
Umano × Uso IA	-.409	.12	-3.42	.001	-.645	-.174	***
Ibrido × Uso IA	-.199	.119	-1.67	.096	-.434	.035	*
<b>Contesto × Uso IA</b>							
Donazione × Uso IA	.06	.114	0.52	.602	-.165	.284	
<b>Fonte × Contesto × Uso IA</b>							
Umano × Donazione × Uso IA	-.078	.17	-0.46	.645	-.412	.256	

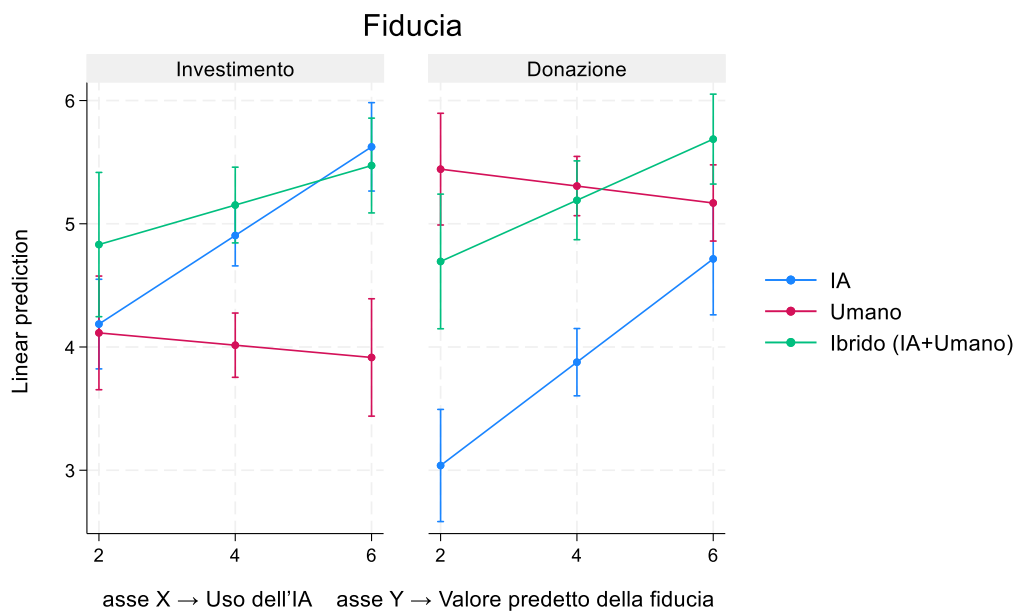
Ibrido × Donazione × Uso IA	.028	.174	0.16	.871	-.313	.369	
Constant	3.467	.298	11.65	0	2.882	4.052	***

Mean dependent var	4.765	SD dependent var	
R-squared	0.333	Number of obs	
F-test	16.611	Prob > F	
Akaike crit. (AIC)	1.251.274	Bayesian crit. (BIC)	

\*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$

Fonte: elaborazione propria.

Figura 3.8



Condizione	Uso IA basso	Uso IA alto
<i>Umano / Donazione</i>	5,44	5.17
<i>IA / Donazione</i>	3,04	4.72
<i>IA / Investimento</i>	4.19	5,62
<i>Umano / Investimento</i>	4.11	3,92

Fonte: elaborazione propria.

Lo stesso schema emerge con la variabile della competenza percepita. Nel complesso, il modello è risultato nuovamente statisticamente significativo,  $F(11, 390) = 18,05$ ,  $p < 0,001$ , con un  $R^2$  di 0,384. Anche in questo caso, la relazione tra l'uso dell'IA e la competenza è risultata statisticamente significativa e positiva ( $\beta = 0,385$ ,  $p < 0,001$ ). L'interazione tra fonte umana e l'uso dell'IA è risultata anch'essa statisticamente significativa, e ancora una volta con un coefficiente negativo ( $\beta = -0,491$ ,  $p < 0,001$ ). Inoltre, l'interazione tra l'uso dell'IA e le fonti ibride è risultata statisticamente significativa, ma con un coefficiente meno rilevante ( $\beta = -0,260$ ,  $p = 0,013$ ). Tuttavia, ancora una volta, nessuna delle interazioni a tre vie è risultata statisticamente significativa. Pertanto, coloro che hanno maggiore familiarità con l'IA tendono a riconoscerle più facilmente competenza, e questo vantaggio si estende anche alla configurazione ibrida.

Le margins mostrano che, a livelli bassi di uso dell'IA, la Donazione/Umano resta molto alta (5,54) mentre IA/Donazione si ferma a 3,33. Quando l'uso dell'IA è alto, invece, l'IA/Investimento sale a 5,60 e l'Ibrido/Donazione arriva addirittura a 5,69, mentre Umano/Investimento si abbassa a 4,00.

Il grafico come già visto nella fiducia, mostra linee che si allontanano tra loro: la fonte umana tende a perdere leggermente terreno, mentre le fonti più tecnologiche vengono valutate sempre meglio. In altre parole, chi ha più familiarità con l'IA non la valuta solo in modo più positivo, ma tende anche a considerarla più legittima e appropriata come fonte di consiglio.

**Tabella 3.25 Linear regression**

competenza	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
<b>: base IA</b>							
Umano	1.336	.524	2.55	.011	.306	2.367	**
Ibrido (IA+Umano)	1.387	.485	2.86	.005	.433	2.341	***
<b>: base Investimento</b>							
Donazione	-.843	.457	-1.84	.066	-1.741	.056	*
<b>Fonte × Contesto</b>							
Umano × Donazione	1.926	.685	2.81	.005	.578	3.273	***
Ibrido × Donazione	-.218	.755	-0.29	.772	-1.702	1.265	
<b>Uso IA</b>							
d5_uso_ia_n	.385	.07	5.53	0	.248	.522	***

<b>Fonte × Uso IA</b>	.	.	.	.	.	.	
Umano × Uso IA	-.491	.114	-4.29	0	-.715	-.266	***
Ibrido × Uso IA	-.26	.104	-2.51	.013	-.463	-.056	**
<b>Contesto × Uso IA</b>	.	.	.	.	.	.	
Donazione × Uso IA	.054	.096	0.56	.574	-.135	.244	
<b>Fonte × Contesto × Uso IA</b>	.	.	.	.	.	.	
Umano × Donazione × Uso IA	-.032	.15	-0.22	.829	-.327	.262	
Ibrido × Donazione × Uso IA	.166	.159	1.05	.296	-.146	.478	
Constant	3.294	.354	9.29	0	2.597	3.991	***

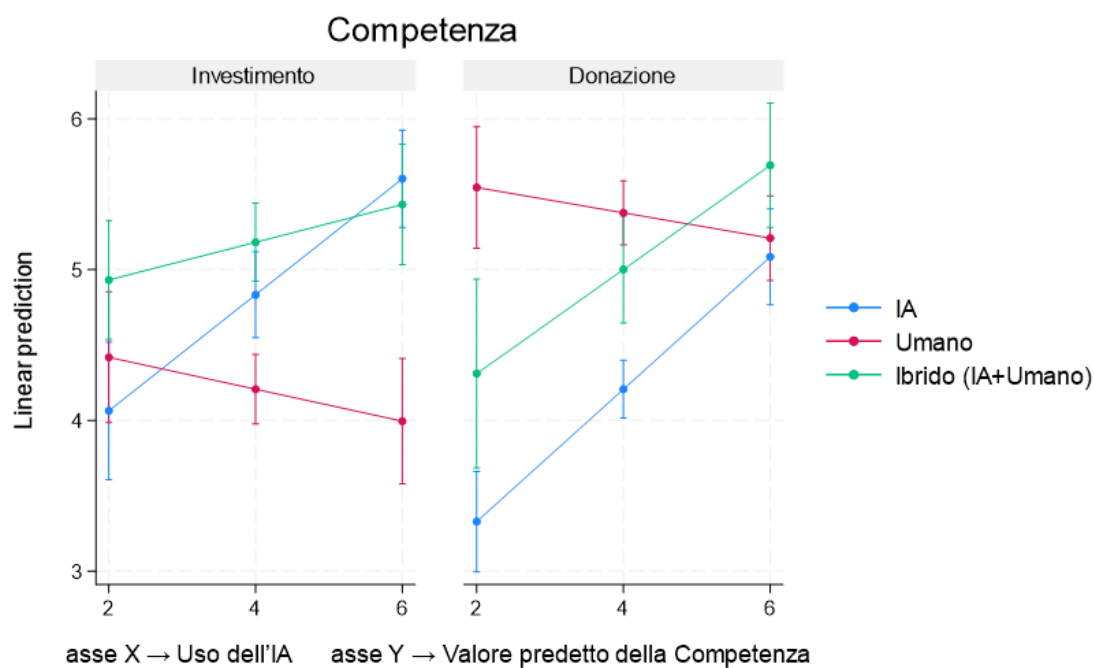
---

Mean dependent var	4.840	SD dependent var	1.237
R-squared	0.328	Number of obs	402
F-test	18.052	Prob > F	0.000
Akaike crit. (AIC)	1.175.101	Bayesian crit. (BIC)	#####

\*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$

Fonte: elaborazione propria.

Figura 3.9



Fonte: elaborazione propria.

Condizione	Uso IA basso	Uso IA alto
Umano / Donazione	5,54	5.21
IA / Donazione	3,33	5.08
IA / Investimento	4.06	5,6
Ibrido / Donazione	4.31	5,69
Umano / Investimento	4.42	4

Fonte: elaborazione propria.

Il caso del calore è diverso e forse proprio per questo motivo il più interessante. Il modello è complessivamente significativo,  $F(11, 390) = 18,30, p < .001$ , con  $R^2 = 0,3208$ , ma l'uso dell'IA, preso da solo, non è significativo ( $\beta = -0,098, p = .371$ ). L'interazione tra contesto della donazione e uso dell'IA è positiva e significativa ( $\beta = 0,518, p = .001$ ). La tripla interazione Umano, Donazione e uso dell'IA è risultata negativa e significativa ( $\beta = -0,549, p = .007$ ).

Pertanto, il moderatore non agisce allo stesso modo in ciascuno di questi contesti, ma soprattutto nel contesto prosociale e in relazione alla fonte umana. I marginS rivelano che, quando l'utilizzo dell'IA è basso il calore è più alto nella condizione Umano/Donazione (5,45) rispetto a IA/Donazione (2,29). Quando il moderatore cresce il vantaggio dell'umano nella donazione si riduce. Con un alto utilizzo dell'IA l'Ibrido/Donazione sale fino a 5,33, mentre Umano/Donazione si ferma a 5,06.

Il grafico mostra linee che si avvicinano, soprattutto nel pannello della donazione. In questo caso il calore non smette di essere associato alla presenza umana, ma in un certo modo la distanza tra le fonti si affievolisce mano a mano che l'uso dell'IA aumenta.

Questo dato complica i risultati precedenti in modo interessante. La familiarità con l'IA non aumenta semplicemente tutte le valutazioni, ma piuttosto modifica anche una dimensione come il calore, che come visto nel §3.4.3 sembrava più legata alla componente umana.

**Tabella 3.27**  
**Linear regression**

calore	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
: base IA							
Umano	1.244	.71	1.75	.081	-.152	2.641	*

Ibrido (IA+Umano)	.287	.838	0.34	.732	-1.361	1.935	
: base Investimento	.	.	.	.	.	.	
Donazione	###	.7	-3.46	.001	-3.801	-1.049	***
Fonte × Contesto	.	.	.	.	.	.	
Umano × Donazione	2.962	.946	3.13	.002	1.102	4.821	***
Ibrido × Donazione	1.434	1.085	1.32	.187	-.7	3.567	
Uso IA	.	.	.	.	.	.	
d5_uso_ia_n	-.098	.109	-0.89	.371	-.312	.117	
Fonte × Uso IA	.	.	.	.	.	.	
Umano × Uso IA	.03	.143	0.21	.832	-.251	.311	
Ibrido × Uso IA	.238	.169	1.41	.159	-.094	.571	
Contesto × Uso IA	.	.	.	.	.	.	
Donazione × Uso IA	.518	.148	3.51	.001	.227	.808	***
Fonte × Contesto × Uso IA	.	.	.	.	.	.	
Umano × Donazione × Uso IA	-.549	.203	-2.71	.007	-.949	-.15	***
Ibrido × Donazione × Uso IA	-.298	.229	-1.30	.193	-.748	.151	
Constant	3.87	.579	6.68	0	2.731	5.009	***

---

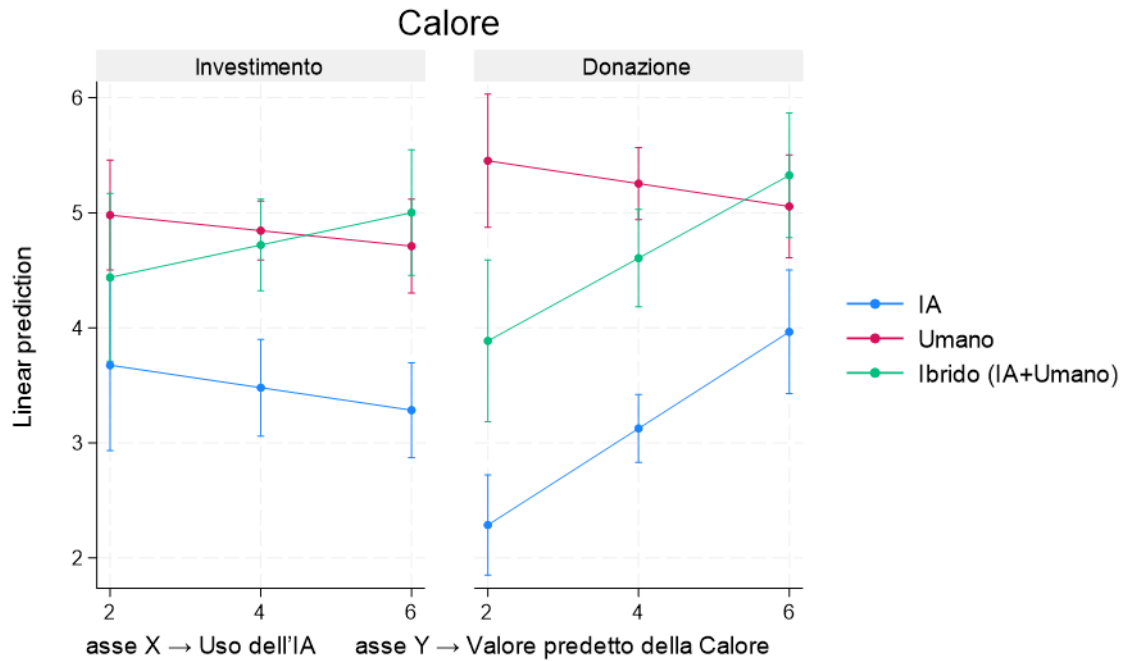
Mean dependent var	4.340	SD dependent var	1.640
R-squared	0.321	Number of obs	402
F-test	18.231	Prob > F	0.000
Akaike crit. (AIC)	1.406.101	Bayesian crit. (BIC)	#####

---

\*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$

Fonte: elaborazione propria.

**Figura 3.10**



Fonte: elaborazione propria.

**Tabella 3.28 Predictive margins più rappresentative – Calore**

Condizione	Usi IA basso	Usi IA alto
Umano / Donazione	5,45	5,06
IA / Donazione	2,29	3,96
Ibrido / Donazione	3,89	5,33
IA / Investimento	3,67	3,283

Fonte: elaborazione propria.

Il passaggio alle variabili ordinali conferma ed arricchisce, il quadro. Per quanto riguarda l'intenzione, il risultato complessivo appare molto chiaro  $\chi^2(11) = 117,36, p < .001$ . Anche qui l'uso dell'IA è positivo e significativo ( $\beta = 0,612, p < .001$ ). L'interazione con la fonte umana è negativa e significativa ( $\beta = -0,744, p < .001$ ). Invece le interazioni a tre vie non risultano significative.

Le margins sono in questo caso calcolate sulla probabilità cumulata di collocarsi nella fascia alta della risposta, cioè nelle categorie 6 e 7 considerate congiuntamente. Questa scelta restituisce un'immagine più ampia e più aderente all'idea di adesione forte al consiglio.

Guardando ai valori stimati, il pattern cambia in modo interessante al variare dell'uso dell'IA. Con un uso basso dell'IA, la probabilità di collocarsi nella fascia alta della scala è particolarmente alta in Umano/Donazione (0,368) e molto più contenuta in IA/Donazione (0,008). Quando invece l'uso cresce, il quadro si modifica. Infatti, in questo caso

l'Ibrido/Donazione raggiunge il valore più alto (0,504), seguito dall'IA/Investimento (0,474), mentre Umano/Donazione scende a 0,320 e Umano/Investimento si ferma a 0,063.

Il grafico mostra quindi linee che, in alcuni casi, si incrociano. Pertanto, la familiarità con l'IA non cambia solo il livello dell'intenzione espressa, ma modifica anche quale fonte risulta più convincente nei diversi contesti.

**Tabella 3.29**      **Ordered**  
**logistic regression**

intenzione	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
: base IA	.	.	.	.	.	.	
Umano	1.869	.717	2.61	.009	.465	3.274	***
Ibrido (IA+Umano)	1.423	1.054	1.35	.177	-.643	3.489	
: base Investimento	.	.	.	.	.	.	
Donazione	-1.274	.77	-1.65	.098	-2.784	.236	*
Fonte × Contesto	.	.	.	.	.	.	
Umano × Donazione	2.739	1.087	2.52	.012	.609	4.87	**
Ibrido × Donazione	.439	1.426	0.31	.758	-2.355	3.234	
Uso IA							
d5_uso_ia_n	.612	.119	5.13	0	.378	.845	***
Fonte × Uso IA	.	.	.	.	.	.	
Umano × Uso IA	-.744	.161	-4.61	0	-1.06	-.427	***
Ibrido × Uso IA	-.304	.22	-1.38	.167	-.735	.128	
Contesto × Uso IA	.	.	.	.	.	.	
Donazione	0	.19	0.00	1	-.372	.372	
Fonte × Contesto × Uso IA	.	.	.	.	.	.	
Umano × Donazione × Uso IA	.08	.25	0.32	.75	-.411	.57	
Ibrido × Donazione × Uso IA	.226	.315	0.72	.474	-.392	.844	
cut1	-.774	.518	.b	.b	-1.789	.24	
cut2	-.002	.518	.b	.b	-1.016	1.013	
cut3	1.209	.525	.b	.b	.18	2.237	

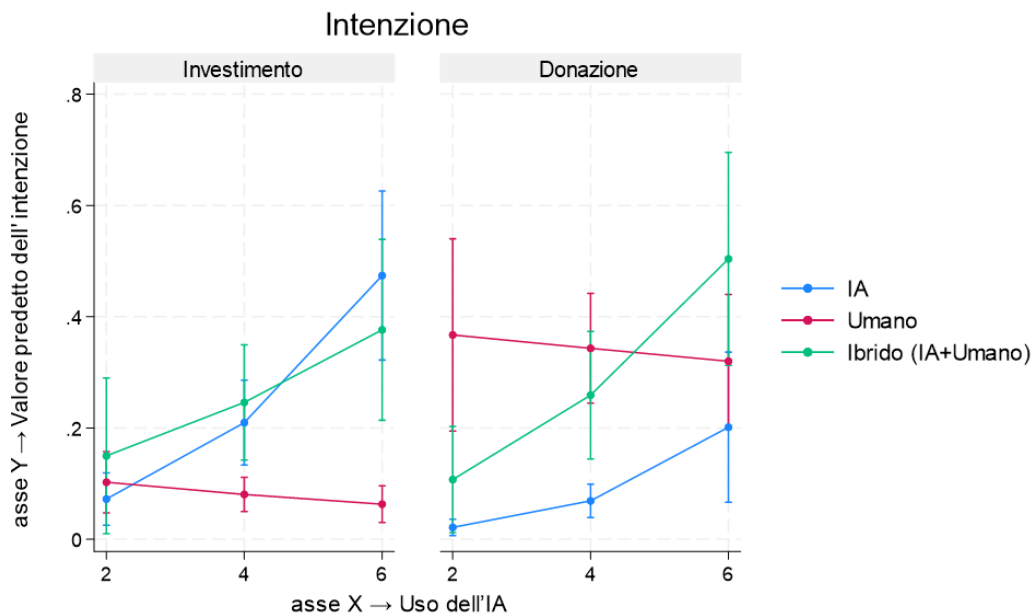
cut4	2.397	.542	.b	.b	1.334	3.46
cut5	3.773	.56	.b	.b	2.675	4.871
cut6	4.778	.579	.b	.b	3.643	5.914

Mean dependent var	4.318	SD dependent var	1.653
Pseudo r-squared	0.090	Number of obs	402
Chi-square	117.364	Prob > chi2	0.000
Akaike crit. (AIC)	1.381.745	Bayesian crit. (BIC)	#####

\*\*\*  $p < .01$ , \*\*  $+CI02:N133$   $p < .05$ , \*  $p < .1$

Figura

3.11



Fonte: elaborazione propria.

Livello di uso dell'IA	Condizione	Probabilità cumulata di risposta alta (6–7)
Basso	Umano / Donazione	0,368
Basso	IA / Donazione	0,021
Basso	Ibrido / Investimento	0,15
Alto	Ibrido / Donazione	0,504
Alto	IA / Investimento	0,474
Alto	Umano / Donazione	0,32
Alto	Umano / Investimento	0,063

Fonte: elaborazione propria.

Un pattern molto simile compare nel modello sull'affidabilità. Anche in questo caso, il modello è complessivamente significativo, Wald  $\chi^2(11) = 127,58$ ,  $p < 0,001$ ; l'uso dell'IA ha un coefficiente positivo e significativo ( $\beta = 0,754$ ,  $p < 0,001$ ); l'interazione tra l'uso dell'IA e la fonte umana ha un coefficiente negativo e significativo ( $\beta = -0,674$ ,  $p = 0,001$ ); invece, le

interazioni a tre vie non risultano significative. Come nel caso dell'intenzione, anche qui le predictive margins vanno lette come probabilità cumulata di risposta alta, cioè che la risposta di un individuo alla domanda sull'affidabilità della fonte rientri nelle categorie 6 o 7 della scala Likert a 7 punti. Questi valori così alti indicano la tendenza ad esprimere un giudizio positivo sull'affidabilità della fonte.

I valori stimati mostrano un andamento particolarmente chiaro. Quando l'uso dell'IA è basso, la probabilità di collocarsi nella fascia alta della scala è nettamente maggiore in Umano/Donazione (0,410), mentre IA/Investimento si colloca a 0,054 e IA/Donazione addirittura a 0,014. In questa fase la componente umana mantiene ancora un chiaro vantaggio simbolico sul piano dell'affidabilità. Tuttavia, con l'aumento dell'utilizzo dell'IA, il valore più alto si riscontra nell'Ibrido/Donazione che raggiunge il valore più alto (0,559), seguito da IA/Investimento (0,537) e Ibrido/Investimento (0,451); nello stesso tempo, Umano/Donazione scende a 0,290, lo stesso valore di IA/Donazione (0,290). Il grafico marginsplot rende tutto più leggibile, in quanto le linee in questo caso tendono ad allontanarsi ed a incrociarsi in alcuni punti, segnalando che la maggiore familiarità con l'IA cambia davvero il modo in cui la fonte viene giudicata. Questo non vuol dire che la fonte umana non sia più credibile. Vuol dire, però, che chi ha maggiore familiarità con l'IA tende a considerare affidabili anche fonti ibride o algoritmiche, soprattutto nei contesti in cui queste appaiono più adatte.

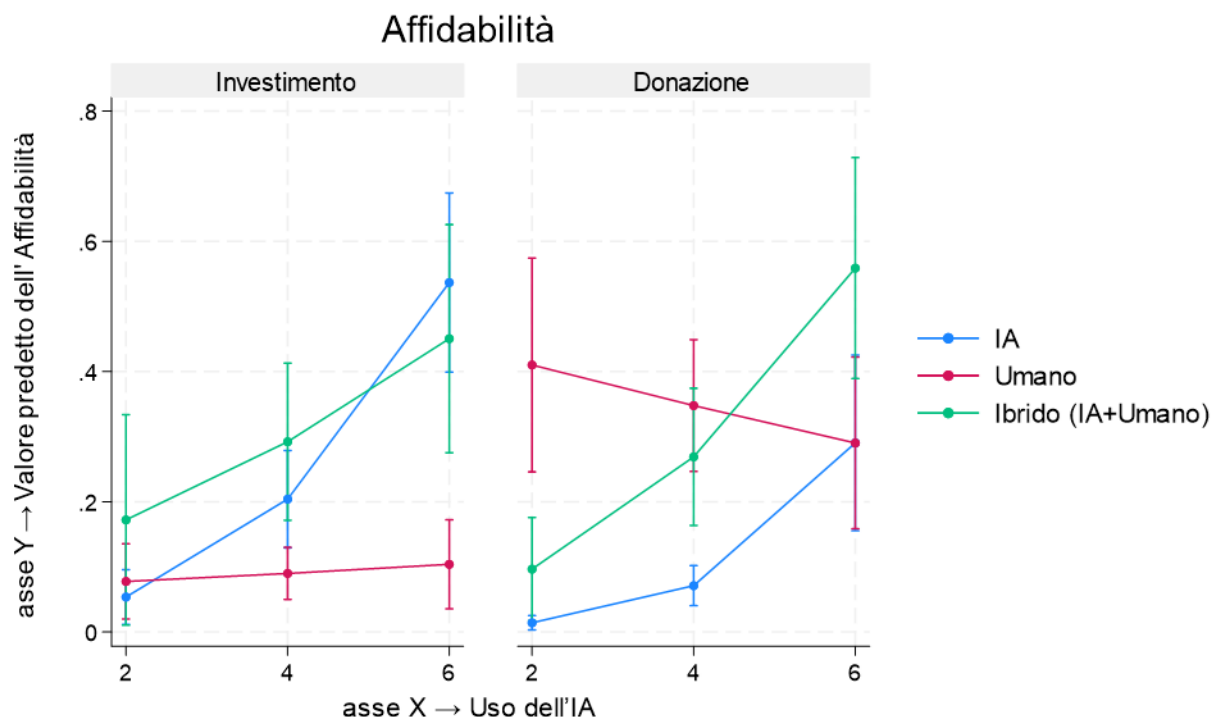
**Tabella 3.31**  
**Ordered logistic regression**

affidabilità	Coef.	St.Err.	t-value	P-value	[95% Conf	Interval]	Sig
: base IA		.	.	.	.	.	
Umano	1.743	.925	1.89	.059	-.069	3.556	*
Ibrido (IA+Umano)	2.122	1.119	1.90	.058	-.071	4.314	*
<b>: base Investimento</b>		.	.	.	.	.	
Donazione	-1.543	.851	-1.81	.07	-3.21	.124	*
<b>Fonte × Contesto</b>		.	.	.	.	.	
Umano × Donazione	4.08	1.239	3.29	.001	1.652	6.508	***
Ibrido × Donazione	.327	1.426	0.23	.818	-2.468	3.123	
Uso IA		.	.	.	.	.	
d5_uso_ia_n	.754	.136	5.55	0	.488	1.021	***

<b>Fonte × Uso IA</b>	.	.	.	.	.	.	
Umano × Uso IA	-.674	.203	-3.32	.001	-1.072	-.276	***
Ibrido × Uso IA	-.411	.229	-1.79	.073	-.86	.038	*
<b>Contesto × Uso IA</b>	.	.	.	.	.	.	
Donazione	.084	.187	0.45	.654	-.282	.449	
<b>Fonte × Contesto × Uso IA</b>	.	.	.	.	.	.	
Umano × Donazione × Uso IA	-.297	.273	-1.09	.277	-.831	.238	
Ibrido × Donazione × Uso IA	.191	.303	0.63	.527	-.402	.785	
cut1	-.765	.658	.b	.b	-2.055	.525	
cut2	.255	.641	.b	.b	-1.001	1.51	
cut3	1.468	.64	.b	.b	.213	2.723	
cut4	2.743	.652	.b	.b	1.465	4.022	
cut5	4.378	.67	.b	.b	3.064	5.692	
cut6	5.392	.687	.b	.b	4.046	6.739	
Mean dependent var	4.577			SD dependent var	1.557		
Pseudo r-squared	0.106			Number of obs	402		
Chi-square	127.581			Prob > chi2	0.000		
Akaike crit. (AIC)	1.315.204			Bayesian crit. (BIC)	#####		

\*\*\*  $p < .01$ , \*\*  $p < .05$ , \*  $p < .1$

**Figura 3.12**



Fonte: elaborazione propria.

Livello di uso dell'IA	Condizione	Probabilità cumulata di risposta alta (6–7)
Basso	Umano / Donazione	0,41
Basso	IA / Investimento	0,054
Basso	IA / Donazione	0,014
Alto	Ibrido / Donazione	0,559
Alto	IA / Investimento	0,537
Alto	Ibrido / Investimento	0,451
Alto	Umano / Donazione	0,29
Alto	IA / Donazione	0,29

Fonte: elaborazione propria.

Le cinque analisi, considerate nel loro insieme, indicano un quadro abbastanza coerente degli effetti dell'IA. L'uso della IA agisce come effetto moderatore del rapporto tra fonte del consiglio e valutazioni dei partecipanti, mentre ha avuto un effetto moderatore molto più debole sul contesto e un effetto moderatore parziale della combinazione completa tra fonte e contesto. Nelle dimensioni legate alla cognizione (fiducia, competenza), l'uso dell'IA sembra aumentare il vantaggio delle fonti tecnologiche e miste. Per quanto riguarda il calore, l'effetto compare soprattutto nella donazione, dove in questi casi il vantaggio della fonte umana si riduce. Per intenzione e affidabilità, chi ha più familiarità non cambia solo il livello delle valutazioni, ma

anche quale fonte preferisce. Nel complesso, questi risultati confermano e al contempo complicano i risultati delle ANOVA e dei modelli esplicativi. Da un lato li confermano perché mantengono centrale il ruolo della fonte. Dall'altro li complicano perché mostrano che tale ruolo non è fisso: anzi, dipende perlopiù dal modo in cui i partecipanti sono già abituati a convivere con strumenti intelligenti.

### **3.5.1 Discussione delle ipotesi alla luce della letteratura**

Dopo aver presentato i risultati empirici dello studio, è ora possibile discuterli alla luce delle ipotesi formulate nel Capitolo 2 e dei principali filoni teorici richiamati nel quadro di riferimento. L'obiettivo di questa sezione non è ripetere le analisi già svolte, ma integrare due passaggi che, fino a questo punto, erano rimasti distinti: da un lato verificare se le ipotesi trovino conferma nei dati; dall'altro chiarire a quali prospettive teoriche i risultati si avvicinino maggiormente, in quali punti le confermino e in quali, invece, le complichino o le ridefiniscano.

In questo senso, la discussione non verrà organizzata separando la sintesi delle ipotesi dal confronto con la letteratura, ma procedendo ipotesi per ipotesi. Per ciascuna di esse si tratterà quindi non solo di stabilire se i risultati siano coerenti con quanto atteso, ma anche di mostrare in che misura essi confermino, sfumino o mettano in discussione i principali lavori discussi nel Capitolo 2. Va inoltre precisato che la discussione si concentrerà sulle ipotesi per cui il Capitolo 3 ha fornito un riscontro empirico diretto.

#### ***H1 (Effetto della fonte nel dominio "analitico").***

La prima ipotesi da verificare sostiene che la fonte di intelligenza artificiale (IA) sarebbe stata valutata come dotata di un livello di competenza superiore rispetto all'umano nel contesto dell'investimento, portando a una maggiore fiducia nella fonte di IA e, di conseguenza, a una maggiore intenzione di seguire i consigli forniti da tale fonte. I dati, in generale, confermano questa ipotesi. Sul piano descrittivo, nello scenario IA/Investimento la competenza percepita è più alta rispetto a Umano/Investimento e lo stesso vale per la fiducia e per l'intenzione di adesione. Le ANOVA confermano che il vantaggio della fonte cambia a seconda del contesto e che, nell'investimento, e che nell'investimento l'IA si colloca sistematicamente sopra la fonte umana.

Inoltre, i confronti post-hoc per determinare le differenze specifiche all'interno di ciascuno di questi test ANOVA hanno rivelato che le fonti di IA e ibride superano l'umano sia in termini di giudizio cognitivo sia in termini di disponibilità a seguire il consiglio.

In questo senso, H1 è confermata nella sua direzione generale, ma non in modo del tutto netto, perché dai dati emerge anche che l'ibrido si comporta almeno altrettanto bene dell'IA, e a volte anche meglio. Questo non smentisce l'ipotesi, ma la rende meno netta di quanto previsto. L'idea che l'IA venga premiata nell'investimento è confermata. Quello che i risultati aggiungono è che tale vantaggio non appartiene esclusivamente alla macchina "da sola", infatti come abbiamo visto può essere condiviso e persino rafforzato quando la componente algoritmica viene affiancata da una mediazione umana.

Da questo punto di vista, i risultati sono coerenti con il filone dell'*algorithm appreciation*. Logg, Minson e Moore (2019) mostrano infatti che, in molti contesti, le persone possono aderire più facilmente a un consiglio algoritmico che a uno umano, soprattutto quando il compito appare misurabile, comparabile e cognitivamente strutturato. Allo stesso tempo, i dati della tesi risultano coerenti anche con Castelo, Bos e Lehmann (2019), secondo cui la fiducia negli algoritmi cresce quando il compito viene percepito come più oggettivo, quantificabile e tecnico. L'investimento, nel quadro empirico di questa ricerca, appare trattato proprio in questo modo: come una decisione in cui la competenza tecnica della fonte è particolarmente rilevante.

I risultati, invece, si allontanano da una lettura rigida dell'*algorithm aversion*. Il filone inaugurato da Dietvorst, Simmons e Massey (2015) mostra che le persone possono evitare gli algoritmi anche quando questi risultano complessivamente più accurati, soprattutto dopo averli visti sbagliare. Nel caso di questa tesi, però, non emerge una sfiducia generalizzata verso l'IA nel dominio analitico. Al contrario, l'algoritmo viene premiato quando il compito appare coerente con le sue caratteristiche percepite. In questo senso, H1 non conferma una semplice opposizione tra umano e macchina, ma suggerisce che, nell'investimento, la fonte venga valutata soprattutto in base alla sua adeguatezza rispetto al compito.

## ***H2 (Effetto della fonte nel dominio "valore/prosociale").***

Nel dominio della donazione/beneficenza, la fonte umana produrrà, in media, livelli più elevati di calore/empatia percepita rispetto alla fonte IA. Questa dimensione incrementerà la

fiducia oltre alla propensione a seguire il consiglio umano, poiché la decisione prosociale richiede allineamento di valori e intenzioni benevole percepite.

Anche in questo caso, i dati supportano l'ipotesi. Infatti, è emerso che le fonti umane mostrano livelli di calore umano significativamente più elevati rispetto alle fonti di intelligenza artificiale nel contesto delle donazioni. Inoltre, sia le fonti umane che le fonti ibride mostrano livelli di fiducia e intenzione più elevati rispetto alla sola fonte algoritmica pura.

Questi risultati sono in linea con l'idea intuitiva che il calore umano e la componente di vicinanza e comprensione siano più importanti in un ambito prosociale e valoriale come quello delle donazioni.

Anche in questo caso, tuttavia, i risultati introducono una sfumatura nella conferma dell'ipotesi. Non solo il fattore umano si è dimostrato superiore all'intelligenza artificiale in questo esperimento, ma le caratteristiche delle soluzioni ibride sono molto simili a quelle del fattore umano in termini di fiducia, intenzione e calore.

Questo significa che H2 è confermata nella sua opposizione fondamentale tra umano e IA, ma i risultati indicano che la relazione di cura, sensibilità e comprensione non si limita alla consulenza umana in senso stretto. Infatti, tali qualità relazionali dell'essere umano possono essere trasferite anche a un'architettura mista in cui la presenza umana continua a fungere da garanzia relazionale. Questo passaggio è importante in quanto indica che l'ambito prosociale non rifiuta la tecnologia in sé, ma solo la tecnologia priva di intermediazione umana.

Da questo punto di vista, i risultati sono coerenti con i lavori che mostrano come il vantaggio dell'algoritmo tenda ad attenuarsi nei compiti più soggettivi o quando entra in gioco una fonte umana ritenuta più adatta. In particolare, Castelo, Bos e Lehmann (2019) sostengono che i consumatori si affidano più facilmente agli esseri umani quando il compito appare interpretativo, relazionale o legato a intuizione ed emozione. La donazione, nel quadro di questa tesi, si avvicina proprio a questo tipo di contesto: non è una decisione "edonica" in senso stretto, ma richiama una logica più valoriale e prosociale, in cui il calore della fonte diventa particolarmente rilevante.

Il secondo riferimento teorico importante è il modello "word-of-machine" proposto da Longoni e Cian (2022). In questo caso, le analogie con i risultati di questa tesi sono ancora più dirette. Gli autori dimostrano che la preferenza per fonti di consulenza umane o basate sull'intelligenza artificiale dipende dal tipo di compromesso che si attiva nel processo

decisionale. Se la decisione si basa su attributi utilitaristici, gli individui tenderanno a preferire l'IA. Tuttavia, se la decisione coinvolge attributi edonistici o legati al gusto, gli individui preferiranno l'essere umano all'IA. Gli stessi autori mostrano anche che questo effetto tende a ridursi o a scomparire nel caso di decisioni ibride.

La tesi conferma con notevole chiarezza questa intuizione di fondo, pur adattandola a un contesto non perfettamente sovrapponibile. Il contesto dell'investimento è molto simile a quello utilitaristico, perché richiama calcolo, oggettività e competenza tecnica. La donazione, invece, pur non essendo edonica in senso stretto, richiama maggiormente una logica più valoriale e prosociale, dove la componente relazionale conta di più.

Anche sotto questo profilo, i risultati della tesi vanno nella stessa direzione indicata da Longoni e Cian (2022). Difatti, l'IA è favorita quando la decisione richiede analisi e razionalità, invece l'umano torna a prevalere quando contano di più sensibilità, relazione e benevolenza percepita. In questo senso, H2 conferma la logica generale del word-of-machine, ma aggiunge che, nel dominio prosociale, anche l'ibrido può mantenere una parte importante della legittimazione relazionale tipica della fonte umana.

### **H3 (Vantaggio della partnership IA+umano).**

La terza ipotesi attribuiva un ruolo alla partnership tra IA+umano, sostenendo che tale partnership sarebbe in grado di generare i più alti livelli di fiducia o comunque maggior stabilità rispetto alle componenti pure. Tutto ciò grazie alla sua capacità di combinare competenza tecnica e legittimazione relazionale. L'ibrido non emerge sempre come la condizione migliore in assoluto, e quindi H3 non può dirsi confermata in senso pieno se la si interpreta come superiorità sistematica. Infatti, in diversi casi la fonte umana nella donazione o l'IA nell'investimento hanno raggiunto valori pari o persino leggermente superiori.

Tuttavia, considerando il quadro generale, la seconda parte della terza ipotesi è supportata. L'ibrido non crolla mai. In ciascuna delle cinque variabili centrali, si colloca quasi sempre nella fascia alta e non viene mai penalizzato né nel contesto dell'investimento né in quello della donazione.

Questo risultato è stato coerente in tutto lo studio, sia nelle statistiche descrittive, sia nelle ANOVA, sia nelle analisi di moderazione. In ciascuna analisi, i risultati hanno indicato che la partnership IA+umano è stata la condizione che ha tratto maggior beneficio dalla familiarità con l'IA senza perdere del tutto la componente relazionale. Pertanto, sebbene la terza ipotesi non sia pienamente confermata, lo è in un modo teoricamente interessante. Infatti, sebbene l'ibrido non sia condizione migliore in assoluto, è però la configurazione più capace di reggere il cambiamento di contesto.

Da questo punto di vista, la tesi aggiunge una sfumatura importante rispetto a quanto discusso da Castelo, Bos e Lehmann (2019). Mentre gli autori hanno discusso di come i compiti possano essere ampiamente classificati come oggettivi o soggettivi, questo studio ha utilizzato un terzo approccio: l'uso di una configurazione ibrida composta dall'essere umano e l'algoritmo. L'ibrido non elimina la distinzione tra i contesti, ma la rende meno rigida. Più semplicemente, non trasforma la decisione in un compito "da macchina", mostra però che la mediazione umana rende accettabile anche una raccomandazione sostenuta dall'algoritmo.

È una correzione importante, perché suggerisce che la fiducia non dipenda solo da chi consiglia, ma anche da come la tecnologia viene incorporata nella relazione di consiglio.

Il riferimento teorico più diretto, però, è quello di Longoni e Cian (2022). Il loro modello indica che l'utilizzo di agenti ibridi uomo-IA per le decisioni attenua l'effetto della preferenza per l'umano o per l'agente IA. Lo stesso risultato emerge anche dai risultati di questa tesi.

Gli agenti ibridi uomo-IA sono in grado di conservare livelli alti di competenza (così come gli altri attributi essenziali per l'esito delle decisioni) senza perdere del tutto la componente relazionale. È per questo che resta quasi sempre su valori alti e raramente viene penalizzato. Pertanto, ancora una volta, i risultati di questa tesi non solo confermano Longoni e Cian (2022), ma supportano anche una delle implicazioni più interessanti dello studio: il dibattito non riguarda solo se l'IA piaccia o meno, ma in quale forma istituzionale e narrativa entri nel processo decisionale.

Nel complesso, quindi, H3 non è confermata come ipotesi di superiorità assoluta dell'ibrido, ma trova supporto nella sua formulazione più solida: quella che attribuisce alla partnership IA+umano una maggiore stabilità tra domini differenti. Si tratta forse di un risultato meno netto di una "vittoria" dell'ibrido, ma più rilevante sul piano teorico e manageriale, perché

suggerisce che la soluzione più efficace non sia sempre la sostituzione dell'umano, bensì una configurazione capace di integrare tecnologia e mediazione relazionale.

#### **H4 (Interazione fonte × dominio).**

I risultati confermano in modo molto netto questa ipotesi. La quarta ipotesi sosteneva che l'effetto della variabile "fonte" sulla variabile "fiducia" non sarebbe stato omogeneo, ma che ci sarebbe stato un vero e proprio effetto di interazione tra le variabili "fonte" e "dominio": la preferenza per l'IA nel dominio degli investimenti ma non in quello delle donazioni, mentre la soluzione ibrida sarebbe risultata più efficace in entrambi i domini. Questa è l'ipotesi che trova maggiore conferma nei risultati dello studio. Gli effetti di interazione sono presenti in ciascuna delle ANOVA, per fiducia, competenza, intenzione e affidabilità. L'effetto di interazione è il risultato centrale dello studio. L'IA è preferita quando la decisione da prendere è di natura analitica, mentre l'umano è preferito quando la natura della decisione entra in un registro più valoriale. La fonte ibrida risulta elevata in entrambi i casi. È proprio in questo punto che l'esperimento rivela la sua maggiore forza.

Se i risultati si fossero limitati a un effetto principale della fonte, la lettura sarebbe stata molto più povera, poiché si sarebbe potuto dire al massimo, che una certa fonte piace "più" di un'altra.

Mentre nel nostro caso il risultato è più interessante: il giudizio sulla fonte dipende da che cosa quella fonte è chiamata a fare. La macchina non è apprezzata o rifiutata in assoluto. L'umano non è preferito sempre e comunque. La relazione con la fonte dipende dal contesto e da quanto viene percepita come adatta al compito. Da questo punto di vista, H4 è confermata e rappresenta il punto centrale di interpretazione di tutto il capitolo empirico.

I risultati di questa tesi si inseriscono esattamente in questo spazio intermedio. In altre parole, non mostrano né una sfiducia generalizzata verso l'IA né un suo apprezzamento indiscriminato. Al contrario, i risultati mostrano che la valutazione cambia in base al contesto: infatti l'IA tende a essere preferita nell'investimento, l'umano nella donazione, mentre l'ibrido mantiene valutazioni alte in entrambi i casi. Pertanto, questi risultati sono più strettamente correlati alla prospettiva di Logg et al. (2019) e alle letture contestuali di Castelo et al. (2019) e di Longoni e Cian (2022) che a una lettura rigida dell'avversione algoritmica.

Allo stesso tempo, però, la tesi non sostiene neppure una visione semplicemente favorevole all'IA, perché il vantaggio della macchina dipende dal tipo di compito e non riguarda le dimensioni più relazionali, soprattutto il calore.

Il confronto con la letteratura discussa nel Capitolo 2 permette di evitare due semplificazioni opposte. La prima sarebbe ridurre tutto a una conferma generica del fatto che “le persone non si fidano degli algoritmi”. La seconda sarebbe leggere i risultati come una prova lineare del fatto contrario, cioè che “gli algoritmi sono preferiti agli esseri umani”. Nessuna di queste due formule restituisce davvero il quadro emerso dall'analisi. Piuttosto, i risultati indicano che la fiducia nei consigli dell'IA è situata, dipende dal tipo di compito svolto, dalle dimensioni percettive che quel compito attiva e dal livello di familiarità che gli individui hanno con la tecnologia utilizzata.

**H5 (Interazione variabili percettive × dominio).**

I risultati confermano questa ipotesi. Le analisi mostrano infatti che fiducia, competenza e calore contribuiscono tutte alla spiegazione degli esiti finali, ma che il loro peso relativo non resta costante nei due domini. Il punto più rilevante è che il calore acquista un ruolo più forte nel contesto della donazione, mentre nell'investimento tendono a risultare relativamente più centrali le dimensioni cognitive e tecniche. In questo senso, il dominio non modifica soltanto i livelli medi delle valutazioni, ma anche il modo in cui le variabili percettive si traducono in intenzione di seguire il consiglio e in affidabilità attribuita alla fonte.

La lettura dei modelli distinti per i due contesti rende questo quadro ancora più chiaro. Nell'investimento, il giudizio finale appare trainato soprattutto da fiducia e competenza, mentre il calore mantiene un ruolo positivo ma più contenuto. Nella donazione, invece, cresce il peso della componente relazionale, che diventa particolarmente rilevante soprattutto quando si considera l'intenzione di seguire il consiglio. Anche per l'affidabilità si osserva una dinamica simile: la competenza rimane importante, ma nella donazione il calore assume un rilievo maggiore rispetto a quanto accade nel dominio analitico. H5 risulta quindi confermata, perché mostra che l'equilibrio tra le variabili percettive cambia in funzione del contesto decisionale.

Da questo punto di vista, i risultati sono coerenti con la letteratura che interpreta la fiducia negli algoritmi come dipendente dalla natura del compito. Castelo, Bos e Lehmann

(2019) mostrano che gli algoritmi tendono a essere più credibili quando il compito appare oggettivo, misurabile e cognitivamente strutturato, mentre le fonti umane risultano più adatte quando entrano in gioco elementi interpretativi, emotivi o relazionali. I risultati della tesi si inseriscono bene in questa prospettiva: l'investimento attiva soprattutto una logica di adeguatezza tecnica, mentre la donazione richiama in misura maggiore una valutazione relazionale, nella quale il calore della fonte diventa più importante.

H5 è coerente anche con il modello del *word-of-machine* di Longoni e Cian (2022). In quel filone, la preferenza per l'IA cresce quando diventano centrali attributi funzionali e cognitivi, mentre si riduce quando il compito richiede sensibilità, interpretazione e benevolenza percepita. La tesi non riproduce meccanicamente quella distinzione, perché la donazione non coincide con un compito edonistico in senso stretto; tuttavia, ne conferma la logica generale. L'investimento si comporta come un dominio in cui prevalgono aspettative di competenza tecnica e razionalità, mentre la donazione rafforza il peso della dimensione relazionale. In questo senso, H5 non oppone rigidamente competenza e calore, ma mostra che il loro equilibrio cambia in funzione del contesto.

Un ulteriore collegamento utile è con la prospettiva della *trust in automation* di Lee e See (2004). In quel quadro, la fiducia non è una disposizione fissa, ma un meccanismo che orienta l'affidamento alla tecnologia in base al compito e alla situazione. I risultati della tesi vanno nella stessa direzione: non mostrano che una singola variabile domini in assoluto tutti i contesti, ma che l'equilibrio tra fiducia, competenza e calore cambia al variare del dominio decisionale. Nell'investimento prevale una grammatica più cognitiva; nella donazione acquista maggiore rilievo la componente relazionale. H5 contribuisce così a rafforzare una lettura della fiducia come processo situato e multidimensionale, non riducibile a una semplice disposizione favorevole o sfavorevole verso l'IA.

Nel complesso, dunque, H5 è confermata. I risultati non mostrano che nell'investimento contino solo le dimensioni tecniche e che nella donazione conti solo la dimensione relazionale; mostrano piuttosto che tutte e tre le variabili percettive restano rilevanti, ma che la loro gerarchia relativa cambia tra i due domini.

## **H6 (Moderazione: familiarità con l'IA).**

Il rapporto tra fonte e valutazioni dipenderà dal grado di familiarità con l'IA: una maggiore familiarità con questa tecnologia ridurrà l'avversione verso le fonti algoritmiche e aumenterà la tendenza a valutarle più favorevolmente, soprattutto nei contesti in cui l'IA appare coerente con il compito richiesto.

I risultati confermano questa ipotesi. Le analisi mostrano infatti che la familiarità con l'IA svolge un ruolo moderatore nel rapporto tra fonte e valutazioni, contribuendo in generale a rafforzare le fonti tecnologiche o ibride e a ridurre, in alcuni casi, il vantaggio iniziale della fonte umana. Questo effetto emerge in modo evidente per fiducia e competenza, in forma più limitata per il calore, e modifica anche la probabilità di collocarsi nella fascia alta di intenzione e affidabilità. Più che spostare semplicemente i livelli medi, la familiarità con l'IA tende quindi a cambiare anche l'ordine di preferenza tra le fonti.

Da questo punto di vista, i risultati sono coerenti con la prospettiva della *trust in automation* e, in particolare, con il concetto di *appropriate reliance* sviluppato da Lee e See (2004). È in questo punto che i risultati acquistano particolare rilievo interpretativo. A una domanda come "Le persone si fidano dell'IA?", la risposta sarebbe stata ambigua. Ma attraverso la lente dello studio di Lee e See (2004), tale ambiguità smette di essere un problema e diventa un importante contributo alla conoscenza dell'argomento: la fiducia non è qualcosa di fisso, ma un meccanismo che orienta l'affidamento. I dati suggeriscono che i partecipanti cercano (più o meno consapevolmente) una forma di appropriatezza tra fonte e compito. Proprio perché l'ibrido risulta la configurazione più stabile conferma questa lettura: infatti, l'automazione è più accettata quando non prende completamente il posto dell'umano, ma lo accompagna

Questa interpretazione si adatta bene ai risultati della tesi. I partecipanti non sembrano affidarsi all'IA in modo indiscriminato, né rifiutarla sempre e comunque. Il loro comportamento appare piuttosto selettivo: la macchina viene premiata quando è percepita come adeguata al compito, mentre viene penalizzata quando la decisione richiede qualcosa che l'IA, da sola, non sembra garantire. Un aspetto ancora più interessante è che la familiarità con l'IA modifica questi pattern. Chi utilizza più spesso strumenti intelligenti tende infatti ad associare meno automaticamente l'affidabilità alla sola fonte umana e, in alcuni casi, sposta il vantaggio verso la soluzione ibrida o verso l'IA nei contesti in cui queste appaiono più coerenti. Non si tratta, però, di una fiducia cieca, ma di una fiducia che dipende dal tipo di compito e dall'esperienza dell'utente.

In questo senso, H6 arricchisce il quadro complessivo della tesi. I risultati non mostrano soltanto che le persone preferiscono una fonte o un'altra, ma suggeriscono che tale preferenza è influenzata anche dal grado di familiarità con la tecnologia. La fiducia nelle raccomandazioni algoritmiche emerge quindi come una costruzione situata e non rigida, che varia non solo in base al dominio decisionale, ma anche in funzione dell'esperienza pregressa dei soggetti con l'IA. Proprio per questo, H6 contribuisce a rafforzare una lettura della fiducia come processo dinamico, contestuale e mediato, più vicino alla logica dell'affidamento appropriato che a una semplice opposizione tra accettazione e rifiuto della tecnologia.

### **3.5.2 Implicazioni manageriali**

Dopo aver collocato i risultati nel dibattito teorico di riferimento, il passo successivo consiste nel tradurli in termini operativi. Questo sottocapitolo non intende proporre regole universali o soluzioni valide in ogni contesto. I dati della tesi, al contrario, suggeriscono prudenza proprio verso le letture troppo semplici. Il messaggio manageriale che emerge non è che l'IA sia sempre preferibile, né che la presenza umana resti comunque insostituibile. Più realisticamente, i risultati indicano che la scelta della fonte del consiglio dovrebbe essere progettata in funzione del tipo di decisione, della qualità della relazione richiesta e del profilo dell'utente.

Un primo punto riguarda i contesti in cui conviene privilegiare una fonte algoritmica. Quando la decisione viene percepita come analitica, comparabile, fondata sull'elaborazione di informazioni numerose e strutturate, l'algoritmo tende a godere di maggiore legittimità. In questi casi, l'utente sembra attribuire valore alla competenza tecnica, alla capacità di trattare dati in modo coerente e alla promessa di una valutazione meno arbitraria. Per imprese e piattaforme questo significa che, in ambiti come la consulenza finanziaria, la gestione del portafoglio, il supporto a decisioni assicurative o, più in generale, i sistemi di raccomandazione ad alta intensità informativa, la componente algoritmica può essere messa maggiormente in primo piano. Non basta però "dire" che dietro il consiglio c'è un algoritmo. Serve costruire un'interfaccia che renda visibile la competenza del sistema, cioè la sua capacità di leggere le informazioni rilevanti e trasformarle in una raccomandazione percepita come fondata.

Diverso è il caso delle decisioni in cui entrano in gioco dimensioni più relazionali, valoriali o prosociali. Qui la componente umana recupera centralità, soprattutto perché il

giudizio non si forma soltanto sulla base della correttezza tecnica del consiglio, ma anche della sua adeguatezza simbolica e della qualità della relazione che lo accompagna. Nei contesti di donazione, assistenza, welfare, orientamento o supporto a scelte con forte carica personale, ciò che conta non è soltanto l'efficienza della raccomandazione, ma il fatto che essa appaia comprensiva, legittima e sensibile alla situazione del destinatario. Per organizzazioni nonprofit, fondazioni, enti del terzo settore o servizi ad alta componente fiduciaria, questo implica che una strategia interamente algoritmica rischia di impoverire la qualità percepita del consiglio. In questi casi, la presenza umana non è solo un "canale" di comunicazione: è una parte sostanziale del valore offerto.

Proprio su questo terreno diventa particolarmente interessante la configurazione ibrida. È forse il risultato più utile, in chiave manageriale, dell'intero capitolo empirico. L'ibrido non emerge sempre come la soluzione migliore in senso assoluto, ma risulta spesso la più stabile. Mantiene livelli elevati di fiducia nei contesti analitici e, allo stesso tempo, non perde del tutto legittimazione nei contesti più relazionali. Per molte imprese questo è un punto cruciale. Significa che non sempre la scelta più efficace consiste nel sostituire l'umano con la macchina, né nel relegare l'IA a puro supporto invisibile. In molti casi, il vantaggio competitivo può nascere proprio dalla combinazione: l'algoritmo elabora, seleziona, compara; l'umano interpreta, spiega, contestualizza e assorbe quella quota di responsabilità relazionale che la macchina, da sola, fatica ancora a sostenere. È una lezione importante per la consulenza finanziaria, per i servizi digitali, per i customer advisory systems e per tutte quelle piattaforme che devono bilanciare efficienza, fiducia e accettabilità.

Un'ulteriore implicazione riguarda la segmentazione degli utenti. I risultati mostrano che la familiarità con l'IA modifica in parte il modo in cui la fonte viene valutata. Questo suggerisce che non tutti gli utenti dovrebbero essere esposti allo stesso tipo di interfaccia o allo stesso frame di consiglio. Per soggetti più abituati all'uso dell'IA, la componente algoritmica può essere resa più visibile e più autonoma; per utenti meno familiari, può essere opportuno rafforzare la mediazione umana o almeno esplicitare con maggiore chiarezza il ruolo di supervisione, interpretazione e controllo. In termini di design manageriale, ciò apre la strada a sistemi di raccomandazione non solo personalizzati nei contenuti, ma anche nella forma della fonte: non soltanto "cosa consigliamo", ma anche "chi sembra consigliare" e "come questo consiglio viene presentato".

In sintesi, i risultati della tesi suggeriscono che la gestione della fonte del consiglio debba diventare una vera leva progettuale. Nei contesti più tecnici conviene valorizzare la competenza algoritmica; in quelli più valoriali o relazionali occorre preservare la presenza umana; nei casi intermedi o più delicati, la soluzione ibrida appare spesso la più promettente. Il punto, in fondo, è che la fiducia non nasce dalla tecnologia in sé, ma dalla coerenza tra tecnologia, compito e utente. È proprio su questa consapevolezza che si innesta naturalmente il paragrafo successivo, dedicato ai limiti dello studio e ai possibili sviluppi futuri.

#### **3.5.4 Limiti e sviluppi futuri**

Ogni ricerca empirica, anche quando restituisce risultati chiari e coerenti, richiede di essere letta entro i confini del proprio disegno. Riconoscere questi confini non significa ridimensionare il valore del lavoro svolto, ma definirne con maggiore precisione la portata. È proprio questa consapevolezza a rendere un'analisi più solida: i risultati acquistano significato non solo per ciò che mostrano, ma anche per le condizioni entro cui possono essere interpretati. Nel caso di questa tesi, i dati raccolti permettono di mettere in luce pattern interessanti e teoricamente rilevanti nel rapporto tra fonte del consiglio, contesto decisionale e fiducia nelle raccomandazioni. Allo stesso tempo, alcuni limiti metodologici e interpretativi suggeriscono cautela e aprono spazi concreti per sviluppi futuri.

Un primo limite riguarda il campione e la modalità di raccolta dei dati. La survey è stata distribuita online, attraverso canali digitali e reti sociali, e questo rende plausibile una maggiore presenza di soggetti già relativamente esposti all'uso di strumenti tecnologici. Ne consegue che il campione, pur numericamente adeguato e sufficientemente eterogeneo, non può essere considerato probabilistico né pienamente rappresentativo della popolazione generale. I risultati non sono generalizzabili, quindi, va intesa in termini analitici più che statistici: i pattern osservati sono utili per comprendere il fenomeno, ma non autorizzano un'estensione automatica a tutti i contesti e a tutti i segmenti di utenza.

Un secondo limite riguarda la natura stessa delle misure e degli scenari utilizzati. Il disegno sperimentale si fonda su situazioni ipotetiche, costruite per isolare in modo controllato l'effetto della fonte e del contesto. Questa scelta è metodologicamente giustificata, ma comporta una semplificazione inevitabile di decisioni che, nella realtà, sono spesso più dense, più ambigue e più ricche di segnali contestuali. Inoltre, alcune variabili — in particolare intenzione

e affidabilità — sono state rilevate tramite singolo item. Si tratta di una soluzione accettabile in un questionario breve, ma meno ricca rispetto a scale multi-item capaci di cogliere sfumature ulteriori. Anche la distinzione tra intenzione dichiarata e comportamento effettivo resta rilevante: dichiarare di essere disposti a seguire un consiglio non equivale, necessariamente, a farlo in una situazione reale in cui entrano in gioco conseguenze economiche, pressioni temporali o costi reputazionali.

Esistono poi limiti di carattere più propriamente interpretativo. I risultati della tesi si fondano su due domini specifici — investimento e donazione — che sono stati scelti proprio perché rappresentano due logiche decisionali contrastanti. Questa scelta si è rivelata utile sul piano analitico, ma rende prudente l'estensione dei risultati ad altri contesti. Non è detto, per esempio, che gli stessi pattern si riproducano in decisioni di consumo quotidiano, in scelte mediche, in ambiti educativi o in situazioni organizzative ad alta incertezza. Allo stesso modo, alcune variabili individuali potenzialmente rilevanti — come atteggiamenti più articolati verso l'IA, tratti di personalità decisionale o forme diverse di familiarità tecnologica — non sono state esplorate fino in fondo.

Proprio da questi limiti emergono gli sviluppi futuri più promettenti. Una prima direzione consiste nel replicare lo studio in altri domini decisionali, così da verificare se il rapporto tra fonte e contesto mantenga la stessa struttura anche al di fuori di investimento e donazione. Una seconda estensione riguarda l'uso di campioni più ampi e diversificati, possibilmente costruiti con criteri meno dipendenti dalla diffusione online. Sarebbe inoltre particolarmente utile introdurre misure comportamentali reali, o almeno scelte incentivabili, capaci di ridurre la distanza tra risposta dichiarata e condotta osservata. Infine, i risultati ottenuti suggeriscono che meriti un approfondimento specifico la configurazione ibrida, che in questa tesi si è rivelata spesso la soluzione più stabile. Capire meglio in quali forme di mediazione uomo-IA essa produca fiducia, e per quali utenti, rappresenta probabilmente una delle piste di ricerca più fertili. Lo stesso vale per variabili individuali come familiarità tecnologica, atteggiamenti verso l'IA, propensione al rischio e orientamento prosociale, che potrebbero essere esplorate con maggiore dettaglio in studi successivi.

Nel complesso, dunque, i limiti di questo lavoro non ne indeboliscono il contributo, ma ne definiscono con maggiore chiarezza il perimetro. La tesi offre una base empirica utile per comprendere come la fonte del consiglio venga valutata in funzione del contesto e di alcune caratteristiche dei soggetti, e fornisce indicazioni che possono essere sviluppate ulteriormente

in ricerche future. È su questa base che il discorso può ora allargarsi alle conclusioni generali della tesi, dove i risultati verranno ricondotti al contributo complessivo del lavoro.

## Conclusioni

Questa tesi è nata da una domanda che, nel dibattito contemporaneo sull'intelligenza artificiale, appare sempre più centrale: in quali condizioni le persone si fidano di un consiglio quando la fonte non è un essere umano, ma un sistema algoritmico, oppure una combinazione tra macchina e intervento umano? La questione non è solo tecnologica. È, prima ancora, economica e manageriale. Nei mercati digitali, nelle piattaforme, nei servizi di consulenza e nei sistemi di raccomandazione, la qualità di una decisione non dipende soltanto dalla quantità di dati disponibili o dalla sofisticazione del modello che li elabora. Dipende anche da come quella raccomandazione viene percepita, da chi la propone, dal contesto in cui si colloca e dal grado di legittimazione che riesce a ottenere presso il destinatario.

A partire da questo problema, il lavoro ha cercato di tenere insieme due livelli che spesso vengono trattati separatamente. Da un lato, il piano più ampio dell'economia dei dati e della trasformazione dei mercati, in cui l'IA emerge come infrastruttura decisionale sempre più pervasiva. Dall'altro, il piano micro-comportamentale, in cui la tecnologia incontra la fiducia, la percezione di competenza, il calore relazionale e la disponibilità concreta a seguire un consiglio. La scelta di osservare questi meccanismi attraverso un disegno sperimentale centrato su due contesti diversi — investimento e donazione — ha permesso di affrontare il tema in una prospettiva non astratta, ma situata. Non si è chiesto, genericamente, se l'IA piaccia oppure no. Si è chiesto, più realisticamente, quando e perché una certa fonte di consiglio venga percepita come più adatta di un'altra.

I risultati dell'analisi empirica convergono su un punto fondamentale: la valutazione della fonte non è fissa. Non esiste una preferenza generalizzata per l'algoritmo, così come non esiste una superiorità automatica della presenza umana. La fonte del consiglio viene giudicata in modo diverso a seconda del tipo di decisione che il soggetto immagina di dover affrontare. Il primo risultato da sottolineare, dunque, è l'importanza della fonte; il secondo è che il dominio decisionale, preso isolatamente, conta meno di quanto si potrebbe pensare; il terzo, e più rilevante, è che ciò che davvero struttura il giudizio è la interazione tra fonte e contesto. È proprio qui che le ipotesi principali del lavoro trovano il loro riscontro più netto: non nella superiorità assoluta di una fonte, ma nel fatto che la sua efficacia cambi in funzione del compito.

Nel contesto dell'investimento, più analitico, comparabile e orientato alla competenza tecnica, la fonte algoritmica e soprattutto la configurazione ibrida risultano mediamente più convincenti. Nel contesto della donazione, invece, torna a pesare di più la componente umana, soprattutto nelle dimensioni che richiamano sensibilità, vicinanza, legittimazione relazionale e adeguatezza valoriale. Questo quadro consente di dire, in sintesi, che le ipotesi sulla natura situata della fiducia e sull'effetto dell'interazione tra fonte e dominio trovano ampia conferma, mentre le ipotesi che opponevano in modo troppo rigido umano e IA risultano meglio interpretabili in termini di adeguatezza della fonte al contesto.

Proprio qui emerge uno dei risultati più interessanti dell'intero lavoro: il ruolo della configurazione ibrida. Se la contrapposizione tra IA e umano aiuta a leggere con chiarezza le polarizzazioni del giudizio, è l'ibrido a suggerire la via interpretativa più ricca. Nei dati, la combinazione tra supporto algoritmico e mediazione umana non domina sempre in modo netto, ma mostra una caratteristica decisiva: mantiene livelli elevati e relativamente stabili nei diversi contesti. Nell'investimento conserva la credibilità cognitiva della fonte tecnologica; nella donazione non perde del tutto la legittimazione relazionale propria della componente umana. Questo significa che l'ibrido non è semplicemente una soluzione di compromesso. È, piuttosto, una forma di integrazione che permette di tenere insieme ciò che le fonti pure tendono a separare: competenza tecnica da un lato, riconoscibilità e vicinanza dall'altro.

Un secondo contributo rilevante della tesi riguarda il modo in cui gli esiti finali — in particolare intenzione di seguire il consiglio e affidabilità attribuita alla fonte — prendono forma. I risultati mostrano che questi esiti non dipendono soltanto dagli scenari sperimentali in sé, ma anche dalle percezioni che gli scenari attivano. Fiducia, competenza e calore non sono semplici variabili accessorie: funzionano come meccanismi interpretativi attraverso cui la fonte del consiglio diventa più o meno accettabile. In altre parole, non basta sapere se una raccomandazione proviene da un algoritmo, da un essere umano o da una collaborazione tra i due. Conta anche il modo in cui quella fonte viene letta: se appare competente, se appare affidabile, se appare capace di comprendere il destinatario. Anche sotto questo profilo, le ipotesi risultano nel complesso confermate, ma con una precisazione importante: la competenza e la fiducia spiegano molto, mentre il calore assume un ruolo più selettivo e soprattutto più forte nei domini prosociali.

A complicare ulteriormente il quadro interviene la familiarità con l'IA. Le analisi di moderazione indicano che chi utilizza più spesso strumenti intelligenti tende a valutare

diversamente le fonti di consiglio. La maggiore esposizione alla tecnologia non elimina le differenze tra umano, IA e ibrido, ma le riorienta. In alcuni casi rafforza il vantaggio delle configurazioni tecnologiche o miste; in altri riduce l'associazione automatica tra affidabilità e componente umana. Anche questo risultato è importante, perché suggerisce che il rapporto con l'IA non dipenda soltanto dalla natura della decisione, ma anche dall'esperienza pregressa del soggetto. La fiducia, quindi, non è solo situata rispetto al compito; è situata anche rispetto alla traiettoria di familiarità tecnologica dell'utente.

Se si guarda ora al contributo teorico del lavoro, il confronto con la letteratura permette di precisare meglio la portata dei risultati. Da un lato, i dati non confermano una lettura forte e generalizzata dell'*algorithm aversion* nel senso più rigido di Dietvorst, Simmons e Massey (2015): l'IA non viene rifiutata in assoluto. Dall'altro, non sostengono neppure un'idea di apprezzamento lineare e universale dell'algoritmo. In questo senso, i risultati si collocano più vicino alla prospettiva di Logg, Minson e Moore (2019), perché mostrano che l'algoritmo può essere preferito in determinati contesti, soprattutto quando il compito appare analitico e strutturato. Allo stesso tempo, però, la tesi non si ferma a questa contrapposizione e mostra che la preferenza per la fonte dipende in modo decisivo dal dominio decisionale.

Sotto questo profilo, i risultati appaiono molto coerenti anche con Castelo, Bos e Lehmann (2019): la fiducia nella fonte non è fissa, ma dipende dal tipo di compito e dalla sua adeguatezza percepita. Il lavoro risulta inoltre in linea con Longoni e Cian (2022), perché conferma che le fonti algoritmiche tendono a essere premiate quando prevalgono attributi più cognitivi e utilitaristici, mentre la componente umana recupera terreno quando diventano centrali dimensioni più relazionali. Tuttavia, la tesi aggiunge una sfumatura importante: la presenza dell'ibrido mostra che il confronto non si esaurisce nell'opposizione macchina-versus-umano, ma si apre alla possibilità di configurazioni che integrano calcolo e mediazione. Infine, i risultati dialogano bene con Lee e See (2004) e con il tema della *trust in automation*: i partecipanti non sembrano affidarsi all'IA in modo cieco, ma in modo selettivo e contestuale, cioè in una logica di *appropriate reliance*. In questo senso, anche il contributo di Dietvorst, Simmons e Massey (2018) è utile per leggere il fatto che l'avversione algoritmica non sia immutabile, ma possa attenuarsi in presenza di forme di controllo, mediazione o familiarità.

Sul piano manageriale, le implicazioni che emergono sono piuttosto chiare. Nei contesti più tecnici e analitici conviene valorizzare maggiormente la componente algoritmica, soprattutto quando è in grado di mostrare competenza, coerenza e capacità di elaborazione

informativa. Nei contesti più relazionali, valoriali o prosociali, invece, la presenza umana continua a svolgere una funzione centrale di legittimazione. La lezione più interessante, tuttavia, riguarda ancora una volta l'ibrido: per molte organizzazioni, piattaforme e servizi, la soluzione più efficace potrebbe non essere né l'automazione pura né il presidio umano esclusivo, ma una progettazione della fonte del consiglio capace di combinare supporto algoritmico e mediazione umana.

Naturalmente, il lavoro presenta anche alcuni limiti. Il campione non probabilistico e il reclutamento online riducono la generalizzabilità dei risultati; l'uso di scenari ipotetici non consente di sovrapporre pienamente intenzioni dichiarate e comportamenti reali; inoltre, i domini selezionati — investimento e donazione — rappresentano due contesti molto significativi, ma non esauriscono la varietà delle decisioni in cui oggi intervengono sistemi intelligenti. Proprio per questo, uno sviluppo futuro importante sarà estendere l'analisi ad altri ambiti, utilizzare campioni più diversificati, introdurre misure comportamentali più dirette e approfondire ulteriormente il ruolo delle differenze individuali e della configurazione ibrida.

Se si guarda al lavoro nel suo insieme, il contributo principale della tesi non consiste nel dimostrare una superiorità generale dell'IA o dell'umano. Sarebbe una conclusione troppo semplice e, in fondo, poco utile. Il contributo sta piuttosto nell'aver mostrato che la fiducia nelle raccomandazioni è contestuale, relazionale e mediata. Contestuale, perché cambia con il tipo di decisione. Relazionale, perché coinvolge dimensioni che non si esauriscono nella sola accuratezza tecnica. Mediata, perché passa attraverso percezioni di competenza, affidabilità e calore che riorganizzano il rapporto tra fonte e soggetto. In questo quadro, la configurazione ibrida acquista un valore interpretativo particolare: non come soluzione universalmente ottimale, ma come forma più flessibile di mediazione tra razionalità algoritmica e bisogno di legittimazione umana.

La riflessione finale che questo lavoro lascia al lettore è, in fondo, semplice solo in apparenza. La diffusione dell'intelligenza artificiale nei processi decisionali non rende superflua la componente umana. La costringe, semmai, a ridefinirsi. Allo stesso modo, la presenza dell'algoritmo non dissolve la questione della fiducia; la rende più complessa, perché sposta il problema dal "chi decide" al "come viene costruita la credibilità di chi decide". In un'economia sempre più attraversata da sistemi intelligenti, la vera posta in gioco non è scegliere una volta per tutte tra macchina e persona. È progettare forme di consiglio in cui fonte, compito e aspettative dell'utente risultino coerenti tra loro. È da questa coerenza, più che dalla

sola presenza della tecnologia, che dipende la possibilità di trasformare l'innovazione algoritmica in decisioni realmente accettate, comprese e seguite.

## BIBLIOGRAFIA

Alfaifi, Y. H. (2024). Recommender systems applications: Data sources, features, and challenges. *Information*, 15(10), 660. <https://doi.org/10.3390/info15100660>

Botti, S., & McGill, A. L. (2011). The locus of choice: Personal causality and satisfaction with hedonic and utilitarian decisions. *Journal of Consumer Research*, 37(6), 1065–1078.

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). *Strength in numbers: How does data-driven decisionmaking affect firm performance?* SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.1819486>

Brynjolfsson, E., Rock, D., & Syverson, C. (2017). *Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics* (NBER Working Paper No. 24001). National Bureau of Economic Research. <https://doi.org/10.3386/w24001>

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Let the machine decide: When consumers trust or distrust algorithms. *NIM Marketing Intelligence Review*, 11(2), 24–29.

Corrado, C., Haskel, J., Iommi, M., & Jona-Lasinio, C. (2022). *The value of data in digital-based business models: Measurement and economic policy implications* (OECD Economics Department Working Papers No. 1723). OECD Publishing. <https://doi.org/10.1787/d960a10c-en>

Coyle, D., & Li, W. (2021). *The data economy: Market size and global trade*. Bennett Institute for Public Policy, University of Cambridge.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In *Advances in Experimental Social Psychology* (Vol. 40, pp. 61–149). Elsevier.

Căpușeanu, S., Barbu, C.-M., Solomon, A.-G., & Rakos, I.-S. (2025). Reshaping the digital economy with big data: A meta-analysis of trends and technological evolution. *Electronics*, 14(13), 2709. <https://doi.org/10.3390/electronics14132709>

De Biasio, A. (2024). *Value-aware recommendation: Algorithms and applications* [Tesi di dottorato, Università degli Studi di Padova].

Dhar, R., & Wertenbroch, K. (2000). Consumer choice between hedonic and utilitarian goods. *Journal of Marketing Research*, 37(1), 60–71.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114–126.

*Exploring Ethical Frontiers of AI in Marketing*. (2024). ScienceDirect. [Dati bibliografici incompleti nel file sorgente.]

Farboodi, M., & Veldkamp, L. (2021). *A growth model of the data economy* (NBER Working Paper No. 28427). National Bureau of Economic Research. <https://www.nber.org/papers/w28427>

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Academy of Management Annals*, 14(2), 627–660. <https://doi.org/10.5465/annals.2018.0057>

Greci, G. M. (2023). *Effetti dell'adozione dell'intelligenza artificiale sulle attività d'impresa* [Tesi di laurea magistrale, Università di Pisa].

Hao, K. (2020, August 5). AI is learning when it should and shouldn't defer to a human. *MIT Technology Review*.

Jones, C. I., & Tonetti, C. (2020). Nonrivalry and the economics of data. *American Economic Review*, 110(9), 2819–2858. <https://doi.org/10.1257/aer.20191330>

Khan, U., Dhar, R., & Wertenbroch, K. (2004). *A behavioral decision theoretic perspective on hedonic and utilitarian choice* (Working Paper No. 2004/66/MKT). [Fonte da verificare.]

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80.

Lizzardi, M. J. (2023). *Recommender systems and business applications* [Tesi di laurea magistrale, Università Politecnica delle Marche].

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>

Longoni, C., & Cian, L. (2022). Artificial intelligence in utilitarian vs. hedonic contexts: The “word-of-machine” effect. *Journal of Marketing*, 86(1), 91–108. <https://doi.org/10.1177/0022242920957347>

Longoni, C., Cian, L., & Morewedge, C. K. (2022). *The word of machine: A framework for understanding and uncovering the persuasive influence of algorithms*. [Dati bibliografici incompleti nel file sorgente.]

Mazzotta, A. (2024). *L'impatto dell'intelligenza artificiale sulla produttività* [Tesi di laurea, LUISS Guido Carli].

McKinsey & Company. (2016). *Industry 4.0: How to navigate digitization of the manufacturing sector*.

OECD. (2022). *The value of data in digital business models*. OECD Publishing. [Voce sintetica presente nel file; probabilmente riferita al lavoro di Corrado et al., 2022.]

OECD. (2024). *Artificial intelligence, data and competition – note by Austria*. Directorate for Financial and Enterprise Affairs, Competition Committee.

Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of Marketing Research*, 42(1), 43–53.

Ortiz, J. H. (Ed.). (2021). *Industry 4.0: Current status and future trends*. IntechOpen. <https://doi.org/10.5772/intechopen.97805>

Pomo, C. (2023). *Towards responsible AI in recommender systems* [Tesi di dottorato, Politecnico di Bari].

Roy, D., & Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data*, 9, 59.

Wang, W., & Benbasat, I. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, 6(3), 72–101. <https://doi.org/10.17705/1jais.00065>

World Economic Forum. (2023–2024). *Centre for the Fourth Industrial Revolution Network report*.

Zhang, X. (2017). *Welfare properties of recommender systems* [Doctoral dissertation, Carnegie Mellon University]. <https://doi.org/10.1184/R1/6724262>

## APPENDICE

### QUESTIONARIO

#### Parte 1: Consenso informato comune a tutti i rispondenti

\*Sei invitato/a a partecipare a un sondaggio anonimo per una ricerca universitaria sul modo in cui le persone valutano diversi tipi di consigli (umani, IA o combinati). La partecipazione è volontaria e puoi interrompere in qualsiasi momento. I dati saranno usati solo a fini accademici in forma aggregata e anonima. Accettando, confermi di avere almeno 18 anni e di aver compreso le informazioni. Acconsenti a partecipare?

Sì, accetto e voglio continuare


No, non accetto

[Pagina successiva >](#)

**Per favore leggi con attenzione la seguente descrizione e immagina la situazione descritta.**

<

[Pagina successiva >](#)

Powered by Qualtrics 

## Parte 2 Ogni rispondente ha visto uno dei successivi sei scenari

Stai utilizzando un'app di **intelligenza artificiale** che analizza dati finanziari e di mercato per offrirti strategie di investimento personalizzate. L'**algoritmo** ha elaborato dati storici e di rischio e ti propone il seguente consiglio: "In base alla tua propensione al rischio e all'andamento dei mercati, ti suggeriamo di allocare il 40% del tuo portafoglio in obbligazioni a medio termine e il 60% in ETF globali a basso costo per massimizzare il rendimento atteso nel lungo periodo."

<

[Pagina successiva >](#)

Stai utilizzando un **app di intelligenza artificiale** che ti aiuta a destinare parte del tuo reddito ad attività benefiche. L'**algoritmo** analizza l'impatto sociale e l'efficienza economica delle organizzazioni non profit e ti propone il seguente consiglio: "In base ai tuoi valori e alle analisi di efficacia, ti suggeriamo di destinare la tua donazione alla Fondazione X, che utilizza il 95% dei fondi per progetti diretti e ha un impatto sociale superiore del 20% alla media del settore."

< [Pagina successiva >](#)

Stai consultando un **consulente finanziario umano** esperto in analisi di mercato e gestione patrimoniale. Dopo aver valutato il tuo profilo di rischio e le condizioni dei mercati, il **consulente** ti propone il seguente consiglio: "In base alla tua propensione al rischio e all'andamento dei mercati, ti suggerisco di allocare il 40% del tuo portafoglio in obbligazioni a medio termine e il 60% in ETF globali a basso costo per massimizzare il rendimento atteso nel lungo periodo."

< [Pagina successiva >](#)

Stai parlando con un **consulente umano** del settore non profit che conosce diverse organizzazioni benefiche. Dopo aver ascoltato i tuoi valori e le tue preferenze, ti propone il seguente consiglio: "In base ai tuoi valori e alle analisi di efficacia, ti suggerisco di destinare la tua donazione alla Fondazione X, che utilizza il 95% dei fondi per progetti diretti e ha un impatto sociale superiore del 20% alla media del settore."

< [Pagina successiva >](#)

Utilizzi un servizio in cui un **algoritmo di intelligenza artificiale** elabora una proposta di investimento e un **consulente umano** la verifica rispetto ai tuoi obiettivi e al tuo profilo di rischio. La raccomandazione congiunta è: "In base alla tua propensione al rischio e all'andamento dei mercati, consigliamo di allocare il 40% del portafoglio in obbligazioni a medio termine e il 60% in ETF globali a basso costo per massimizzare il rendimento atteso nel lungo periodo."

< [Pagina successiva >](#)

Utilizzi un servizio in cui un **algoritmo di intelligenza artificiale** seleziona organizzazioni ad alto impatto e un **consulente umano** ne verifica trasparenza e allineamento ai tuoi valori personali. La raccomandazione congiunta è: "In base ai tuoi valori e alle analisi di efficacia, consigliamo di destinare la tua donazione alla Fondazione X, che utilizza il 95% dei fondi per progetti diretti e mostra un impatto sociale superiore del 20% alla media del settore."

< Pagina successiva >

### Parte 3: Domande sui i vari costrutti comune a tutti

\*Chi ha fornito il consiglio che hai appena letto?

- Algoritmo di IA
- Consulente umano
- Collaborazione IA + Umano
- Non ricordo

\*Il consiglio riguardava principalmente...

- Aspetti economici/funzionali
- Aspetti valoriali/personali

< Pagina successiva >

\*Ora ti chiediamo di valutare il consiglio che hai appena letto. Indica quanto sei d'accordo rispondendo alle seguenti affermazioni su una scala da 1 a 7:

- 1 = Totalmente in disaccordo
- 2 = Molto in disaccordo
- 3 = In disaccordo
- 4 = Né d'accordo né in disaccordo
- 5 = D'accordo
- 6 = Molto d'accordo
- 7 = Totalmente d'accordo

Mi fido del consiglio che ho appena letto

Effettua una selezione ▼

< Pagina successiva >

\*Questo consiglio è credibile.

\*Ritengo affidabile questo consiglio.

\*La fonte è competente in materia

\*La fonte mostra padronanza tecnica.

\*La fonte sa valutare correttamente i dati.


\*La fonte è empatica.


\*La fonte è attenta ai miei bisogni.

\*La fonte appare umana e comprensiva.

\* Se dovessi agire realmente, seguirei questo consiglio.

Effettua una selezione 

\* Considero questa fonte generalmente affidabile.

Effettua una selezione 




Pagina successiva 

\* Preferirei ricevere consigli da una persona piuttosto che da un algoritmo di intelligenza artificiale.

Effettua una selezione 


\* Sono a mio agio a seguire consigli generati da un algoritmo di intelligenza artificiale.

Effettua una selezione 




Pagina successiva 

\* Sono disposto/a a correre rischi finanziari se penso che il rendimento potenziale valga la pena.

Effettua una selezione 

\* Sostenere cause sociali e benefiche è importante per me.

Effettua una selezione 




Pagina successiva 

## Parte 4: Domande socio-demografiche


\* Ti chiediamo ora alcune brevi informazioni generali. Le risposte sono anonime e verranno utilizzate solo per scopi di ricerca.

Quanti anni hai?

\* Qual è il tuo genere?

Effettua una selezione 



Pagina successiva 

\*Qual è il tuo titolo di studio più elevato completato?

\*Ti consideri una persona tecnologicamente competente?

\*Usi applicazioni basate su intelligenza artificiale (es. assistenti virtuali, ChatGPT, Copilot, Gemini, app di raccomandazione)?

Grazie per aver completato il sondaggio.  
La risposta è stata registrata.