UNIVERSITÀ DI GENOVA
SCUOLA DI SCIENZE SOCIALI
DIPARTIMENTO DI ECONOMIA



Tesi di laurea magistrale in

Management

L'adozione dell'intelligenza artificiale nel settore bancario: modelli organizzativi e implicazioni per lo sviluppo delle competenze

Relatore: Angelo Gasparre

Candidato: Giacomo De Bonis

anno accademico 2025-2026

Indice

Introduzione 12

Capitolo I – L'evoluzione dell'intelligenza artificiale e il suo impatto sulla trasformazione organizzativa 17

1.1 Introduzione 17

1.2 L'evoluzione dell'Intelligenza Artificiale: dalle origini teoriche ai sistemi di apprendimento automatico 18

1.3 Dal Machine Learning alla Generative AI: la nuova frontiera dell'intelligenza artificiale 22

1.3.1 Il Deep Learning e la capacità di apprendere rappresentazioni complesse 22

1.3.2 La diffusione dell'Intelligenza Artificiale Generativa 23

1.3.3 Dall'Intelligenza Artificiale Generativa all'Agentic AI 24

1.4 L'intelligenza artificiale come leva della trasformazione organizzativa 25

1.5 L'impatto dell'intelligenza artificiale sul lavoro e sulle competenze 26

1.6 Il settore bancario come contesto privilegiato di osservazione 28

1.7 Metodologia e struttura del lavoro 29

1.8 Verso l'analisi dell'intelligenza artificiale nel settore bancario 30

Capitolo II – L'adozione dell'intelligenza artificiale nel sistema bancario 31

2.1 Introduzione 31

2.2 Dalla banca di prossimità alla banca "aumentata": un approfondimento storico e digitale 32

2.2.1 Le ondate della digitalizzazione del settore bancario 32

2.3 La diffusione dell'intelligenza artificiale nel settore bancario 33

2.4 I principali casi d'uso: dai processi tradizionali all'apporto dell'intelligenza artificiale 36

2.4.1 La valutazione del merito creditizio 37
2.4.2 Il contrasto alle frodi e il presidio antiriciclaggio 38
2.4.3 La relazione con la clientela e la personalizzazione dei servizi 39
2.4.4 L'efficienza operativa e l'automazione dei processi interni 40
2.4.5 La conformità normativa e la sicurezza informatica 41
2.4.6 La gestione del rischio, il monitoraggio andamentale e l'integrazione dei fattori ESG 42
2.5 L'intelligenza artificiale generativa e agentica nel settore bancario 43
2.6 Il quadro regolamentare: il Regolamento europeo sull'intelligenza artificiale dopo l'Omnibus digitale 44
2.6.1 La genesi e le ragioni dell'intervento normativo 45
2.6.2 L'approccio basato sul rischio 46
2.6.3 Il Regolamento (UE) 2026/1744 e il calendario aggiornato 46
2.6.4 Il quadro nazionale 48
2.6.5 Le implicazioni per il settore bancario 49
2.7 Sorveglianza umana e human in the loop: la centralità del presidio umano 49
2.7.1 Che cosa significa "presidio umano": una precisazione terminologica 50
2.7.2 Il fondamento giuridico: gli articoli 14 e 26 dell'AI Act e l'articolo 22 del GDPR 50
2.7.3 Il controllo "a quattro occhi": che cosa prevede l'AI Act e che cosa prevede la disciplina bancaria 52
2.7.4 Quattro esempi tratti da processi bancari 53
2.7.5 I limiti del presidio umano e le condizioni della sua efficacia 55
2.8 Sfide, rischi e fattori critici di adozione 57
2.8.1 La qualità e la governance dei dati 57

2.8.2 La trasparenza e la spiegabilità dei modelli 57
2.8.3 I rischi di distorsione e discriminazione 58
2.8.4 Il rischio di modello e il rischio di terze parti 58
2.8.5 I sistemi legacy e il divario di competenze 58
2.9 Due tipi ideali a confronto: la banca digital native e la banca commerciale tradizionale 59
2.9.1 Premessa metodologica: perché tipi ideali e non casi aziendali 59
2.9.2 Il tipo ideale della banca digital native 60
2.9.3 Il tipo ideale della banca commerciale a radicamento territoriale 61
2.9.4 Il confronto lungo cinque dimensioni 63
2.9.5 I limiti della tipologia 66
2.10 Considerazioni conclusive 67
Capitolo III – I modelli organizzativi per l'adozione dell'intelligenza artificiale 69
3.1 Introduzione 69
3.2 Dalla trasformazione digitale alla trasformazione AI-driven 70
3.2.1 Il quadro concettuale della trasformazione digitale 70
3.2.2 Che cosa l'intelligenza artificiale ha di specifico 71
3.3 Il problema organizzativo posto dall'intelligenza artificiale 72
3.3.1 I problemi fondamentali della progettazione organizzativa 72
3.3.2 Automazione e potenziamento come scelta di progettazione 73
3.4 Le precondizioni: prontezza organizzativa e capacità 74
3.4.1 I fattori di prontezza organizzativa 74
3.4.2 Dalla prontezza alla capacità organizzativa 75
3.4.3 La governance del dato come infrastruttura organizzativa 76

3.5 I modelli organizzativi per l'adozione dell'intelligenza artificiale 77
3.5.1 Il modello centralizzato 78
3.5.2 Il modello decentrato 80
3.5.3 Il modello hub and spoke: la configurazione federata e le sue varianti 81
3.5.4 Il modello esternalizzato e consortile 83
3.5.5 Il modello a piattaforma e prodotto 85
3.5.6 Il modello predittivo o decision-centric 87
3.5.7 Il modello agentico-ibrido 88
3.5.8 Il modello ecosistemico 90
3.5.9 Quattro considerazioni di ordine generale 92
3.6 La dimensione della governance: presidi, ruoli e controlli 93
3.6.1 Il sistema di gestione della qualità come vincolo di progettazione 94
3.6.2 L'articolazione dei controlli e il rapporto con l'assetto esistente 94
3.7 Dai modelli al funzionamento: i meccanismi che ne determinano l'efficacia .. 95
3.7.1 La ridefinizione del controllo 95
3.7.2 La validazione e il problema del riferimento di correttezza 96
3.7.3 Il lavoro invisibile di manutenzione 97
3.7.4 L'identità professionale e le resistenze 97
3.8 La declinazione nel settore bancario 98
3.9 Verso le competenze: le implicazioni dei modelli 99
3.10 Considerazioni conclusive 101
Capitolo IV – Il ruolo dell'uomo nell'organizzazione AI-driven: evoluzione delle competenze 103
4.1 Introduzione 103

4.2 L’evoluzione del lavoro nel settore bancario 104
4.2.1 Il quadro teorico: compiti, sostituzione e reintegrazione 104
4.2.2 Le evidenze empiriche sull’automazione 106
4.2.3 La specificità del lavoro bancario 107
4.2.4 Dalla mansione al compito: la ricomposizione dei ruoli 108
4.3 La collaborazione fra uomo e intelligenza artificiale 109
4.3.1 Il fondamento economico della complementarità 109
4.3.2 Le configurazioni della collaborazione e i loro esiti 110
4.3.3 Fiducia, affidamento e calibrazione 111
4.3.4 Il paradosso dell’apprendimento 112
4.4 Le competenze richieste dall’organizzazione AI-driven 114
4.4.1 Un criterio di classificazione 114
4.4.2 Le competenze di comprensione: l’alfabetizzazione all’intelligenza artificiale 114
4.4.3 Le competenze di valutazione: giudizio critico e validazione 115
4.4.4 Le competenze di relazione 116
4.4.5 Le competenze di governo 117
4.4.6 Il raccordo fra competenze e modelli organizzativi 117
4.5 Ruoli emergenti e trasformazione dei profili esistenti 119
4.5.1 I ruoli emergenti 119
4.5.2 La trasformazione dei profili esistenti 121
4.6 Le politiche di sviluppo delle competenze 122
4.6.1 Dalla formazione all’apprendimento nel lavoro 122
4.6.2 L’obbligo normativo e la sua traduzione organizzativa 123

4.6.3 Le condizioni di efficacia 124
4.7 Rischi e questioni aperte 125
4.8 Risposta alla domanda di ricerca e considerazioni conclusive 126
4.8.1 La risposta complessiva 126
4.8.2 I limiti del lavoro 127
4.8.3 Prospettive di approfondimento 127
4.8.4 Considerazione finale 128
Bibliografia 129

Ringraziamenti

Eccoci qua al termine di un percorso lungo, complesso e ricco di sfide. Il raggiungimento di questo traguardo rappresenta per me molto più di una semplice conclusione accademica: è la testimonianza dell'importanza della determinazione e della capacità di non arrendersi mai. Nei momenti più difficili ho spesso dubitato delle mie capacità e della possibilità di portare a termine questo percorso, ma aver superato questi ostacoli ha reso ancora più significativo il risultato raggiunto.

Desidero esprimere la mia più sincera gratitudine a tutte le persone che, con il loro sostegno, hanno contribuito a rendere possibile questo traguardo.

Un sentito ringraziamento va al professor Angelo Gasparre per la disponibilità, la professionalità e il costante supporto dimostrati durante la stesura dell'elaborato. I suoi suggerimenti e le sue osservazioni hanno contribuito in modo determinante al raggiungimento dell'obiettivo finale.

Alla mia famiglia, che rappresenta da sempre il punto fermo della mia vita. Ai miei genitori, Angelo e Paola, per i valori, l'amore e il sostegno con cui mi hanno accompagnato in ogni passo del mio percorso; a mia sorella Marzia, presenza costante nei momenti belli e in quelli più difficili e ai miei nipoti Leonardo, Vittoria e Alberto, che con la loro crescita, il loro entusiasmo e la loro spontaneità arricchiscono la mia vita ogni giorno. Questo traguardo è anche vostro, perché il vostro affetto, la vostra vicinanza e il vostro esempio mi hanno dato la forza di arrivare fin qui.

Un ringraziamento speciale va a Rebecca, la mia ragazza, che ha avuto un ruolo fondamentale durante questo percorso. È stata costantemente al mio fianco, condividendo con entusiasmo ogni successo e offrendomi sempre il suo sostegno nei momenti più difficili. La sua pazienza, la sua determinazione e la sua fiducia nelle mie capacità mi hanno aiutato a superare gli ostacoli e a non perdere mai di vista l'obiettivo. Se oggi posso celebrare questo

importante traguardo, gran parte del merito va a lei, che ha creduto in me persino quando io stesso faticavo a farlo.

A tutti i miei amici, da quelli che mi accompagnano da sempre a quelli del mare, del Vicolo, della Banca e del calcio. Ognuno di voi ha arricchito il mio percorso a modo proprio, con una parola di incoraggiamento, un consiglio, una risata o semplicemente con la propria presenza. Questo traguardo è anche il risultato dell'energia e della forza che mi avete trasmesso.

Infine, un pensiero speciale va alla nonna Pupo, la persona che più di ogni altra ha rappresentato il cuore e l'anima della nostra famiglia. Con il suo esempio ci ha insegnato il valore dell'amore, del sacrificio e dell'importanza di rimanere uniti, trasmettendoci giorno dopo giorno il vero significato della famiglia. Oggi non è fisicamente qui a condividere questo momento con noi, ma la sua presenza vive costantemente dentro di noi.

Abstract

Il presente elaborato analizza in che modo l'adozione dell'intelligenza artificiale (IA) nel settore bancario richieda una trasformazione dei modelli organizzativi e come tale trasformazione ridefinisca le competenze umane necessarie. La scelta del tema nasce da un'apparente contraddizione: l'IA è ormai ampiamente diffusa nelle banche, ma i benefici che produce non sono automatici, poiché dipendono dalla capacità delle organizzazioni di integrarla nei processi, nella governance e nel lavoro delle persone. Il lavoro adotta un approccio teorico-analitico, fondato sulla revisione critica della letteratura scientifica e delle principali fonti istituzionali e normative, e si avvale del confronto fra due tipi ideali di intermediario, la banca digital native e la banca commerciale a radicamento territoriale, oltre che dell'analisi di otto configurazioni organizzative. Nell'interpretazione proposta e sulla base della tassonomia elaborata dall'autore, l'analisi suggerisce che nessun modello sia superiore in assoluto: l'efficacia dipende dalla coerenza fra tecnologia, assetto organizzativo e competenze.

Abstract (English)

This thesis examines how the adoption of artificial intelligence (AI) in the banking sector requires a transformation of organisational models, and how this transformation redefines the human skills needed. The choice of topic stems from an apparent contradiction: AI is now widespread in banks, yet its benefits are not automatic, since they depend on the ability of organisations to embed it in their processes, governance and people's work. The study follows a theoretical-analytical approach based on a critical review of the scientific literature and of the main institutional and regulatory sources. It compares two ideal types of intermediary, the digital native bank and the territorially rooted commercial bank, and analyses eight organisational configurations. Within the interpretation proposed and the taxonomy developed by the author, the analysis suggests that no model is superior in absolute terms: effectiveness depends on the coherence between technology, organisational design and skills.

Introduzione

Negli ultimi anni l'intelligenza artificiale (IA) si è affermata come una delle innovazioni tecnologiche più rilevanti per il management. A differenza delle precedenti ondate di digitalizzazione, orientate soprattutto alla gestione delle informazioni e all'automazione delle attività operative, l'IA è in grado di supportare e, in alcuni casi, di ridefinire attività cognitive tradizionalmente svolte dalle persone: l'analisi di grandi volumi di dati, la previsione di fenomeni complessi, la produzione di contenuti e il supporto alle decisioni. Non si tratta, quindi, di un semplice strumento da acquistare e installare, ma di una tecnologia a uso generale, il cui valore dipende dalla capacità delle organizzazioni di riprogettare processi, strutture e modalità di collaborazione fra persone e sistemi intelligenti.

Il settore bancario rappresenta un contesto particolarmente adatto a osservare questo fenomeno. L'attività delle banche si fonda sull'acquisizione e sull'elaborazione di grandi quantità di informazioni su clienti, mercati e rischi, si svolge entro un perimetro regolamentare rigoroso ed è esposta a una crescente pressione competitiva. Al tempo stesso, le funzioni svolte dagli intermediari richiedono un'adozione prudente e governata, nella quale l'innovazione deve convivere con i presidi di rischio, di conformità e di vigilanza. Secondo l'Indagine Fintech della Banca d'Italia (2025), nel solo 2024 il sistema finanziario italiano ha investito 593 milioni di euro in progetti di IA e sono stati censiti 798 sistemi, l'11 per cento dei quali classificato come ad alto rischio ai sensi dell'AI Act. Le stesse fonti indicano però che i benefici documentati sono disomogenei e condizionati dal contesto di impiego, dalla qualità dei dati e dalle competenze di chi utilizza la tecnologia. Il quadro normativo europeo, inoltre, attribuisce un ruolo centrale alla sorveglianza umana e all'alfabetizzazione del personale, trasformando in obblighi giuridici questioni che sono, prima di tutto, organizzative. È da questa apparente contraddizione, ossia una diffusione ampia a fronte di risultati non automatici, che nasce la motivazione del presente lavoro.

L'elaborato intende rispondere alla seguente domanda di ricerca: come l'adozione dell'intelligenza artificiale nel settore bancario richieda una trasformazione dei modelli organizzativi e in che modo tale trasformazione ridefinisca le competenze umane necessarie.

La prima parte riguarda il piano organizzativo, ossia quali configurazioni consentano di trasformare la tecnologia in valore; la seconda riguarda il piano delle persone, ossia quali competenze tali configurazioni richiedano.

Dal punto di vista metodologico, il lavoro adotta un approccio teorico-analitico, fondato sulla revisione critica della letteratura scientifica e delle principali fonti istituzionali e normative in materia di intelligenza artificiale, organizzazione e competenze. La scelta risponde alla natura dell'oggetto di studio: un fenomeno recente e in rapida evoluzione, che richiede, prima di una verifica empirica, un lavoro di sistematizzazione dei contributi esistenti. Per confrontare gli intermediari si ricorre inoltre al tipo ideale di matrice weberiana, impiegato non per descrivere casi aziendali, ma come termine di paragone per isolare le condizioni strutturali che influenzano l'adozione della tecnologia.

Il lavoro procede dal generale al particolare ed è articolato in quattro capitoli. Il primo capitolo definisce il quadro teorico di riferimento. Dopo aver ricostruito l'evoluzione dell'IA, dalle riflessioni di Turing e dalla conferenza di Dartmouth fino al machine learning, all'IA generativa e ai più recenti sistemi agentici, il capitolo mostra come essa possa essere interpretata come una General Purpose Technology, il cui impatto sul lavoro combina in modo dinamico automazione e potenziamento delle attività umane.

Il secondo capitolo analizza l'adozione dell'IA nel sistema bancario. Ripercorre le ondate di digitalizzazione del settore, ne misura la diffusione sulla base delle rilevazioni della Banca d'Italia e di ABI Lab ed esamina i principali casi d'uso, confrontando la gestione tradizionale dei processi con l'apporto dell'IA e con le evidenze empiriche disponibili. Il capitolo dedica poi specifici approfondimenti all'IA generativa e agentica, al quadro regolamentare europeo, anche alla luce delle modifiche apportate all'AI Act nel luglio 2026, e al tema della sorveglianza umana, per chiudersi con il confronto fra due tipi ideali: la banca digital native e la banca commerciale a radicamento territoriale.

Il terzo capitolo costituisce il nucleo analitico dell'elaborato. Prende avvio da ciò che l'IA ha di specifico rispetto alle precedenti tecnologie digitali, ossia autonomia, apprendimento e inscrutabilità, ne ricostruisce le implicazioni per la progettazione organizzativa ed esamina

le precondizioni dell'adozione. Confronta quindi otto configurazioni organizzative: quattro modelli di allocazione della competenza (centralizzato, decentrato, hub and spoke ed esternalizzato) e quattro modelli di aggregazione del lavoro (a piattaforma e prodotto, predittivo o decision-centric, agentico-ibrido ed ecosistemico). Esamina infine la governance e i meccanismi che determinano il funzionamento effettivo di ciascun modello. Il quarto capitolo si concentra sul ruolo dell'uomo nell'organizzazione AI-driven: ricostruisce l'evoluzione del lavoro bancario, analizza la collaborazione fra persone e sistemi intelligenti, classifica le competenze richieste e le raccorda alle configurazioni organizzative studiate, descrive i ruoli emergenti e la trasformazione dei profili esistenti e discute le politiche di sviluppo delle competenze, i rischi e le questioni ancora aperte.

Le riflessioni conclusive del lavoro, anziché essere raccolte in un capitolo autonomo, sono presentate in quest'ultima parte dell'introduzione. Il percorso svolto consente di articolare la risposta alla domanda di ricerca in quattro considerazioni. La prima è che la trasformazione organizzativa non è una conseguenza dell'adozione tecnologica, ma la sua condizione. I benefici dell'IA restano disomogenei perché il valore della tecnologia dipende da investimenti complementari di natura organizzativa, quali la governance del dato, l'assetto dei presidi e la definizione dei diritti decisionali. In assenza di tali investimenti, la tecnologia da sola non garantisce alcun vantaggio.

La seconda considerazione è che non esiste un modello organizzativo superiore in assoluto, ma una coerenza da ricercare. Ciascuna delle otto configurazioni analizzate risolve alcuni problemi generandone altri, e la scelta dipende da variabili di contingenza quali la dimensione, la maturità, il numero di sistemi in esercizio e la natura dell'informazione su cui si fonda il modello di business. Le decisioni su dove collocare le competenze e su come aggregare il lavoro sono inoltre indipendenti: una parte rilevante delle difficoltà osservate nei programmi di adozione deriva non dalla scelta di un modello sbagliato, ma dall'incoerenza fra le due. La banca digital native, che dispone di dati strutturati fin dall'origine e non è vincolata da sistemi stratificati, trova coerenza in assetti decentrati o hub and spoke con hub leggero, combinati con un modello a piattaforma e prodotto. La banca commerciale a radicamento territoriale, il cui vantaggio risiede nella conoscenza della clientela e del territorio,

trova invece coerenza in una formula sintetizzabile come «esternalizzazione della piattaforma e internalizzazione del giudizio», associata a un'organizzazione del lavoro costruita attorno alle decisioni.

La terza considerazione riguarda il funzionamento effettivo delle configurazioni. L'analisi mostra che la struttura formale non ne determina, da sola, l'efficacia: due organizzazioni che adottano il medesimo modello possono ottenere risultati opposti a seconda di come sono progettati gli incentivi di chi esercita la sorveglianza, di chi valida i sistemi e con quale indipendenza, di come viene riconosciuto il lavoro continuo di manutenzione e di come vengono ridefiniti i contenuti professionali dei ruoli coinvolti. Lo stesso vale per il presidio umano richiesto dalla normativa, che è efficace solo se sostenuto da competenza, autorità, tempo e tracciabilità. Ne discende che ogni configurazione genera un fabbisogno di competenze specifico e che la domanda pertinente non è quali competenze richieda l'IA in generale, ma quali ne richieda una determinata configurazione organizzativa. In questa prospettiva, la ridefinizione delle competenze non è una conseguenza della trasformazione organizzativa, ma la condizione della sua riuscita.

La quarta considerazione, di carattere conclusivo, è formulata come esito interpretativo del lavoro: sulla base della tassonomia proposta, si ipotizza che la capacità di giudizio possa essere più scarsa della competenza tecnica. Essa comprende la capacità di comprendere che cosa un sistema sia in grado di fare, di calibrare l'affidamento sui suoi esiti, di validarli e di specificare gli obiettivi da raggiungere. Si tratta di capacità che presuppongono la conoscenza del dominio applicativo e che, secondo l'analisi svolta, possono essere meno agevolmente reperibili sul mercato rispetto ai profili tecnici. Per un intermediario che accede alla tecnologia dall'esterno, ciò assume un rilievo strategico: il patrimonio di conoscenza dei processi e della clientela costituisce la principale leva per governare una tecnologia che non produce internamente.

Da queste considerazioni discendono alcune indicazioni operative per il management. Prima di scegliere un modello, le banche dovrebbero valutare la propria prontezza organizzativa in relazione allo specifico obiettivo perseguito, a partire dalla governance del dato, che

costituisce spesso la precondizione più trascurata. Dovrebbero inoltre estendere l'impianto dei controlli esistente anziché costruirne uno parallelo, rendere visibile e attribuire formalmente il lavoro di verifica e manutenzione dei sistemi e accompagnare la formazione con incentivi che non penalizzino chi, motivatamente, si discosta dalle raccomandazioni algoritmiche. In assenza di queste condizioni, il presidio umano rischia di ridursi a un adempimento formale e la struttura organizzativa, quale che sia, di restare priva di effettività.

Il lavoro presenta alcuni limiti. Trattandosi di un lavoro teorico-analitico, non produce evidenza empirica originale e le proposizioni formulate hanno il valore di ipotesi argomentate sulla base della letteratura esistente. La costruzione per tipi ideali non consente inferenze su intermediari determinati, la tassonomia dei modelli organizzativi andrebbe verificata su casi concreti e il quadro normativo è in evoluzione. Ne discendono tre possibili sviluppi: la verifica empirica, attraverso studi di caso comparati, del raccordo fra configurazioni organizzative e fabbisogni di competenze; la misurazione dell'effetto dell'automazione sull'accumulazione di esperienza nei processi bancari; l'approfondimento delle configurazioni agentiche e dell'attribuzione della responsabilità sulle azioni compiute da sistemi autonomi.

In definitiva, la tesi sostenuta è che l'intelligenza artificiale non sia soltanto una sfida tecnologica, ma soprattutto una sfida organizzativa e umana. I sistemi intelligenti non sottraggono alle persone il loro ruolo, ma ne spostano il contenuto dall'esecuzione al giudizio. Le banche in grado di trarne valore saranno quelle che, accanto agli investimenti tecnologici, sapranno costruire assetti coerenti con la propria natura e sviluppare nelle persone la capacità di governare i sistemi intelligenti.

Capitolo I – L'evoluzione dell'intelligenza artificiale e il suo impatto sulla trasformazione organizzativa

1.1 Introduzione

Negli ultimi anni l'intelligenza artificiale (IA) si è affermata come una delle innovazioni tecnologiche più rilevanti in ambito manageriale, cambiando il modo in cui le organizzazioni progettano i processi, prendono decisioni e creano valore.

Nelle fasi precedenti della trasformazione digitale l'innovazione era orientata soprattutto alla digitalizzazione delle informazioni e all'automazione delle attività operative. L'attuale evoluzione dell'IA introduce invece la possibilità di supportare, integrare e, in alcuni casi, ridefinire attività cognitive tradizionalmente svolte dall'essere umano: l'analisi di grandi quantità di dati, la previsione di fenomeni complessi, la produzione di contenuti e il supporto alle decisioni. Emerge così il concetto di processi AI-native, ossia processi ripensati o sviluppati fin dall'origine in un'ottica orientata all'IA.

La diffusione dell'IA è stata resa possibile dalla convergenza di più fattori tecnologici degli ultimi anni. Da un lato, la disponibilità di grandi quantità di dati, strutturati e non, prodotti da imprese e dispositivi connessi, ha fornito la base necessaria per l'addestramento degli algoritmi; dall'altro, l'aumento della capacità di calcolo e il miglioramento delle tecniche di apprendimento automatico hanno permesso di sviluppare modelli sempre più accurati. Come osservano Haenlein e Kaplan (2019), il recente successo dell'IA non nasce da una singola innovazione, ma è il risultato di un lungo percorso in cui progresso algoritmico, disponibilità di dati e potenza di calcolo si sono rafforzati a vicenda.

In questo contesto l'IA non può essere considerata solo una tecnologia per automatizzare attività ripetitive o aumentare l'efficienza. La letteratura sulla trasformazione digitale e sulla gestione dell'intelligenza artificiale evidenzia effetti congiunti sui modelli organizzativi, sui processi aziendali, sulle modalità di creazione del valore e sulle competenze richieste alle persone (Vial, 2019; Berente et al., 2021; Raisch e Krakowski, 2021). L'adozione dell'IA

comporta quindi un cambiamento di tipo sistemico, che coinvolge non solo la dimensione tecnologica, ma anche quella organizzativa, gestionale e culturale.

La letteratura sulla trasformazione digitale e sull'adozione organizzativa dell'IA converge nell'indicare che i benefici non derivano dal solo acquisto della tecnologia, ma dagli investimenti complementari nella riprogettazione dei processi, negli assetti decisionali e nelle competenze (Brynjolfsson, Rock e Syverson, 2017; Vial, 2019; Fountaine, McCarthy e Saleh, 2019; Berente et al., 2021; Raisch e Krakowski, 2021). In questa prospettiva la tecnologia diventa un fattore abilitante di un cambiamento più ampio, che richiede nuove forme di collaborazione tra persone e sistemi intelligenti (Pecis, 2025).

Alla luce di queste considerazioni, il presente elaborato analizza il rapporto tra l'adozione dell'intelligenza artificiale, la trasformazione organizzativa e l'evoluzione delle competenze umane, concentrandosi sul settore bancario, nel quale l'intensità informativa dei processi e la rilevanza della gestione del rischio hanno favorito un ampio sviluppo dei casi d'uso (Financial Stability Board, 2017; OECD, 2021; Banca d'Italia, 2025). Il lavoro si sviluppa attorno alla seguente domanda di ricerca:

"Come l'adozione dell'intelligenza artificiale nel settore bancario richiede una trasformazione dei modelli organizzativi e in che modo tale trasformazione ridefinisce le competenze umane necessarie?"

Per rispondere a questa domanda è necessario, in primo luogo, comprendere il percorso evolutivo dell'intelligenza artificiale, analizzandone lo sviluppo storico, le principali innovazioni tecnologiche e l'ampliamento delle capacità applicative. Solo attraverso questa ricostruzione è possibile comprendere come l'IA sia passata dall'essere una disciplina di ricerca a rappresentare uno dei principali motori della trasformazione organizzativa contemporanea.

1.2 L'evoluzione dell'Intelligenza Artificiale: dalle origini teoriche ai sistemi di apprendimento automatico

L'intelligenza artificiale è oggi considerata una delle tecnologie più strategiche per lo sviluppo economico e organizzativo; le sue origini, tuttavia, risalgono a molti decenni prima

dell'attuale diffusione delle applicazioni basate sul machine learning e sull'IA generativa. Come osservano Haenlein e Kaplan (2019), l'idea di realizzare macchine capaci di simulare il comportamento intelligente dell'uomo non è recente, ma è il risultato di una lunga evoluzione scientifica, segnata da fasi di forte entusiasmo alternate a fasi di rallentamento. Gli autori descrivono questo percorso con la metafora delle stagioni dell'intelligenza artificiale, evidenziando come lo sviluppo della disciplina sia stato influenzato dalle aspettative della comunità scientifica, dalla disponibilità di risorse tecnologiche e dagli investimenti nella ricerca.

Le prime riflessioni teoriche risalgono agli anni Quaranta del Novecento e trovano un punto di riferimento nell'opera del matematico britannico Alan Turing. Durante la Seconda guerra mondiale Turing contribuì allo sviluppo della macchina elettromeccanica The Bombe, progettata per decifrare i codici della macchina Enigma. L'efficacia di tale sistema nel risolvere problemi complessi lo portò a chiedersi se una macchina potesse manifestare forme di comportamento simili all'intelligenza umana (Copeland, 2004; Russell e Norvig, 2021).

Queste riflessioni confluirono nel celebre articolo Computing Machinery and Intelligence, nel quale Turing (1950) affrontò per la prima volta il problema in termini scientifici, proponendo il cosiddetto Turing Test: un esperimento per valutare la capacità di una macchina di produrre risposte indistinguibili da quelle di un essere umano nel corso di una conversazione. Il Test di Turing resta ancora oggi un riferimento teorico importante, perché sposta l'attenzione dalla struttura interna della macchina ai comportamenti che essa è in grado di manifestare.

La nascita ufficiale dell'IA viene convenzionalmente collocata nel 1956, quando, presso il Dartmouth Summer Research Project on Artificial Intelligence, John McCarthy, Marvin Minsky, Claude Shannon, Nathaniel Rochester e altri studiosi diedero vita a un progetto di ricerca interdisciplinare volto a sviluppare macchine capaci di simulare l'intelligenza umana. Fu in questa occasione che venne coniato il termine Artificial Intelligence (McCarthy et al., 1955). Come sottolinea Negri (2026), la conferenza di Dartmouth rappresenta il punto di

partenza simbolico della disciplina, poiché segna il passaggio dalle riflessioni teoriche ai primi programmi di ricerca sistematici.

L'entusiasmo generato dalla conferenza avviò la fase che Haenlein e Kaplan (2019) definiscono AI Spring, caratterizzata da un forte ottimismo sulla possibilità di sviluppare macchine capaci di riprodurre l'intelligenza umana. Tra la fine degli anni Cinquanta e l'inizio degli anni Settanta la comunità scientifica ottenne risultati importanti, sostenuti da consistenti investimenti pubblici e privati. In questi anni furono sviluppati alcuni dei primi programmi di elaborazione del linguaggio naturale e di problem solving automatico, come ELIZA, realizzato da Joseph Weizenbaum presso il MIT, e il General Problem Solver di Allen Newell, Herbert Simon e Cliff Shaw (Weizenbaum, 1966; Newell, Shaw e Simon, 1959). Questi sistemi dimostrarono che un elaboratore poteva simulare alcune forme di ragionamento umano, alimentando aspettative molto elevate.

Le previsioni di quel periodo si rivelarono però troppo ottimistiche. I limiti tecnologici dell'epoca, insieme alla ridotta capacità di calcolo e alla scarsità di dati, impedirono il raggiungimento di quegli obiettivi. A partire dalla metà degli anni Settanta, il ridimensionamento delle aspettative portò a una forte riduzione dei finanziamenti, dando avvio alla fase nota come AI Winter. Come osserva Negri (2026), questi rallentamenti non furono un fallimento della disciplina, ma la conseguenza della distanza tra le aspettative iniziali e le reali possibilità offerte dalle tecnologie disponibili.

Una caratteristica dei sistemi di questa prima fase era l'utilizzo dei cosiddetti Expert Systems. Tali applicazioni si basavano su una logica simbolica, nella quale il comportamento della macchina era determinato da un insieme di regole definite dagli sviluppatori attraverso strutture del tipo if-then. Il sistema poteva quindi prendere decisioni solo sulla base della conoscenza codificata dagli esperti, senza alcuna capacità di apprendimento autonomo. Gli Expert Systems rappresentarono comunque un passo avanti, perché permettevano di automatizzare processi decisionali relativamente complessi in contesti con regole ben definite, trovando applicazione nella consulenza tecnica e nel supporto alle decisioni aziendali.

Proprio la loro architettura ne costituiva, però, il limite principale. Come osservano Haenlein e Kaplan (2019), gli Expert Systems risultavano efficaci solo in contesti molto strutturati, mentre mostravano prestazioni ridotte di fronte a problemi caratterizzati da elevata variabilità, incertezza o complessità, come il riconoscimento di immagini o la comprensione del linguaggio naturale. L'esempio riportato dagli stessi autori è significativo: sistemi come Deep Blue, sviluppato da IBM, erano in grado di sconfiggere il campione mondiale di scacchi Garry Kasparov esplorando milioni di combinazioni, ma non possedevano alcuna reale capacità di apprendimento né sapevano generalizzare la conoscenza acquisita ad altri contesti (Campbell, Hoane e Hsu, 2002).

Il superamento di questi limiti è il vero punto di svolta nella storia dell'IA. Come sottolinea Negri (2026), il cambiamento di paradigma avvenne quando l'obiettivo cessò di essere quello di programmare in dettaglio ogni comportamento della macchina e divenne, invece, quello di permetterle di apprendere direttamente dai dati. Questo passaggio segnò il progressivo abbandono dell'approccio simbolico a favore di metodi statistici e probabilistici, aprendo la strada al Machine Learning.

Il Machine Learning può essere definito come un insieme di tecniche che permettono ai sistemi informatici di individuare in autonomia regolarità, correlazioni e schemi presenti nei dati, migliorando le proprie prestazioni senza ricevere istruzioni esplicite per ogni singolo problema. A differenza degli Expert Systems, gli algoritmi di apprendimento automatico costruiscono modelli matematici analizzando grandi quantità di dati e adattano il proprio comportamento in base all'esperienza acquisita.

Come osservano Jordan e Mitchell (2015), il Machine Learning è una delle innovazioni più significative dell'informatica contemporanea, perché consente di affrontare problemi impossibili da risolvere con regole deterministiche. L'efficacia di questi algoritmi cresce all'aumentare della quantità e della qualità dei dati disponibili, rendendo il dato la principale risorsa strategica per il funzionamento dei sistemi intelligenti. Come osserva Negri (2026), dati, capacità di calcolo e algoritmi sono i tre pilastri della moderna evoluzione dell'IA: solo la loro

convergenza ha reso possibile un'accelerazione senza precedenti nell'addestramento di modelli sempre più sofisticati.

Questa evoluzione segna il passaggio da un'intelligenza artificiale intesa come sistema di automazione basato su regole a un sistema capace di apprendere, adattarsi e migliorare in autonomia attraverso l'esperienza, estendendone l'uso a contesti caratterizzati da elevata complessità informativa e decisionale. Su queste basi si sviluppa la successiva evoluzione rappresentata dal Deep Learning.

1.3 Dal Machine Learning alla Generative AI: la nuova frontiera dell'intelligenza artificiale

L'evoluzione dell'IA non si è fermata con l'introduzione degli algoritmi di apprendimento automatico. Negli ultimi anni l'aumento della capacità di calcolo, la disponibilità di dataset sempre più grandi e la costruzione di datalake hanno favorito la nascita di una nuova generazione di sistemi intelligenti, con capacità di elaborazione molto più avanzate. Da strumenti in grado di eseguire regole predefinite si è passati a sistemi capaci di apprendere dai dati, fino a modelli che generano nuovi contenuti e che, più di recente, sono in grado di pianificare attività e perseguire obiettivi complessi con crescente autonomia.

1.3.1 Il Deep Learning e la capacità di apprendere rappresentazioni complesse

Uno degli sviluppi più significativi del Machine Learning è il Deep Learning, un insieme di algoritmi basati su reti neurali artificiali con più livelli di elaborazione. A differenza dei modelli tradizionali, nei quali le caratteristiche rilevanti devono essere selezionate manualmente dagli sviluppatori, le reti neurali profonde individuano in autonomia le rappresentazioni più significative, costruendo livelli di astrazione sempre più complessi (LeCun, Bengio e Hinton, 2015; Goodfellow, Bengio e Courville, 2016).

Come osservano LeCun, Bengio e Hinton (2015), il vantaggio principale di questa tecnica è la capacità di apprendere direttamente dai dati grezzi, senza una preventiva modellazione manuale delle informazioni. Nel riconoscimento delle immagini, ad esempio, i primi livelli

identificano forme geometriche elementari, mentre quelli successivi ricostruiscono strutture via via più articolate, fino a riconoscere l'oggetto rappresentato.

Secondo Haenlein e Kaplan (2019), il ritorno delle reti neurali artificiali è stato reso possibile da tre fattori: l'ampia disponibilità di dati digitali, l'aumento della potenza di calcolo e il miglioramento delle tecniche di addestramento. Sebbene le basi teoriche fossero note già dagli anni Quaranta, per molti decenni i limiti hardware ne avevano impedito un utilizzo efficace. Il Deep Learning ha prodotto un netto miglioramento delle prestazioni nel riconoscimento delle immagini, nell'elaborazione del linguaggio naturale, nei sistemi di raccomandazione, nella traduzione automatica e nel riconoscimento vocale. In particolare, la capacità di elaborare dati non strutturati ha ampliato il potenziale dell'IA, consentendo alle imprese di estrarre valore da documenti, registrazioni audio e contenuti testuali difficilmente analizzabili con strumenti tradizionali. Questa evoluzione costituisce il fondamento tecnico su cui si sviluppa la successiva rivoluzione dell'intelligenza artificiale generativa.

1.3.2 La diffusione dell'Intelligenza Artificiale Generativa

L'evoluzione più significativa degli ultimi anni è la diffusione dell'Intelligenza Artificiale Generativa (Generative AI), una categoria di sistemi progettati non solo per analizzare o classificare informazioni, ma anche per produrre nuovi contenuti sotto forma di testi, immagini, audio, video e codice informatico.

Secondo Negri (2026), la Generative AI rappresenta una fase distinta rispetto ai paradigmi precedenti. Mentre gli algoritmi tradizionali erano orientati alla classificazione, alla previsione o all'individuazione di correlazioni statistiche, i modelli generativi creano nuovi contenuti coerenti con le conoscenze acquisite durante l'addestramento, estendendo il campo di applicazione dell'IA anche ad attività creative o intellettuali. Alla base di questa evoluzione vi sono i cosiddetti Foundation Models, modelli di grandi dimensioni addestrati su enormi quantità di dati provenienti da molte fonti: una volta completato l'addestramento, possono essere adattati a una vasta gamma di compiti diversi, senza bisogno di sviluppare un algoritmo specifico per ogni singola applicazione (Bommasani et al., 2021; Vaswani et al., 2017).

L’introduzione di questi modelli ha cambiato anche il rapporto tra uomo e tecnologia. L’interazione con l’IA non richiede più competenze strettamente informatiche o la definizione di procedure rigide, ma può avvenire attraverso il linguaggio naturale: gli utenti possono formulare richieste, porre domande o indicare obiettivi, ricevendo risposte articolate, sintesi, analisi o contenuti originali. Come osserva Negri (2026), questa modalità di interazione ha favorito la rapida diffusione dell’intelligenza artificiale nelle organizzazioni, riducendo le barriere di accesso alla tecnologia. Dal punto di vista organizzativo, la Generative AI supporta la redazione di documenti, l’analisi normativa, la gestione della conoscenza aziendale, la produzione di report e l’assistenza alle decisioni: l’IA non interviene più solo nelle attività operative, ma inizia a riguardare anche processi a elevato contenuto cognitivo.

1.3.3 Dall’Intelligenza Artificiale Generativa all’Agentic AI

Nel percorso storico ricostruito da Negri (2026), l’evoluzione dell’IA non si arresta con la Generative AI. L’emergere dei sistemi agentici può essere collocato come uno sviluppo successivo alla diffusione dei modelli generativi e la loro definizione tecnica richiede fonti specificamente dedicate agli agenti autonomi.

A differenza dei modelli generativi impiegati principalmente per produrre un output in risposta a un input, gli agenti basati su grandi modelli linguistici sono configurati per raggiungere obiettivi attraverso cicli di pianificazione e azione. La letteratura scientifica ne descrive architetture che combinano profilo, memoria, pianificazione e azione: il sistema può scomporre un compito in sotto-obiettivi, utilizzare strumenti esterni, acquisire informazioni dall’ambiente e adattare i passaggi successivi in base ai risultati ottenuti (Wang et al., 2024). L’elemento distintivo non è quindi la sola generazione di contenuti, ma l’orchestrazione di una sequenza di azioni entro un perimetro definito.

Questa evoluzione cambia il ruolo dell’IA nelle organizzazioni. Se le prime applicazioni erano pensate come strumenti di supporto per automatizzare singole attività, i sistemi agentici tendono a integrarsi nei processi aziendali, assumendo il ruolo di collaboratori digitali capaci di interagire con persone, banche dati, applicativi gestionali e altri sistemi informatici.

La crescente integrazione dei sistemi intelligenti nei processi rende insufficiente una valutazione fondata sulle sole prestazioni tecniche: la trasformazione riguarda anche la distribuzione delle attività tra persone e sistemi, i diritti decisionali, le responsabilità, i modelli operativi e le competenze (Berente et al., 2021; Raisch e Krakowski, 2021; OECD, 2026). Pecis (2025) offre una lettura divulgativa convergente di questa trasformazione in quanto nell'interpretazione proposta, il passaggio dagli Expert Systems al Machine Learning, dal Deep Learning alla Generative AI e, infine, all'Agentic AI segnala un progressivo ampliamento del ruolo organizzativo dell'IA e la portata effettiva di tale evoluzione dipende dalle scelte di progettazione, governance e supervisione adottate.

1.4 L'intelligenza artificiale come leva della trasformazione organizzativa

L'evoluzione dell'intelligenza artificiale mostra come il progresso tecnologico non sia un fenomeno isolato, ma il motore di una più ampia trasformazione delle organizzazioni. L'attenzione non riguarda più solo le prestazioni degli algoritmi, ma la capacità delle organizzazioni di integrare questi strumenti nei propri modelli operativi e nella gestione del capitale umano.

In questa prospettiva l'IA può essere definita una General Purpose Technology (Bresnahan e Trajtenberg, 1995), cioè una tecnologia pervasiva, in continuo miglioramento e capace di generare innovazioni complementari lungo l'intera struttura economica. Come già avvenuto con l'elettricità o con Internet (Brynjolfsson e McAfee, 2014), il suo valore non deriva dalla semplice disponibilità della tecnologia, ma dalla possibilità di riprogettare processi, strutture organizzative e modalità di creazione del valore. L'IA non sostituisce i precedenti strumenti digitali, ma ne amplia le potenzialità, rendendo possibile una gestione più intelligente dei dati.

La letteratura sulla trasformazione digitale mette in guardia dall'assimilare l'adozione dell'IA a un semplice acquisto di software: il valore dipende da cambiamenti complementari nei processi, nelle strutture, nella governance e nelle competenze (Brynjolfsson, Rock e Syverson, 2017; Vial, 2019; Fountaine, McCarthy e Saleh, 2019; Berente et al., 2021). Pecis (2025) riprende questa impostazione in chiave divulgativa. Un'adozione efficace richiede

pertanto una revisione del modo in cui l'organizzazione progetta i processi, coordina le attività e valorizza le competenze delle persone, in un'ottica di change management.

La letteratura manageriale ha analizzato questa trasformazione anche in termini di equilibrio tra automazione e potenziamento del lavoro umano. Come osservano Raisch e Krakowski (2021), le organizzazioni non devono scegliere in modo netto tra sostituire le persone con le macchine (automation) e affiancarle a esse (augmentation), ma gestire in modo dinamico la tensione tra questi due poli: un uso orientato solo all'automazione rischia di erodere nel tempo le competenze e il vantaggio competitivo dell'impresa.

Da questa prospettiva emerge il concetto di organizzazione AI-native, inteso nel presente elaborato come un assetto nel quale dati, algoritmi e piattaforme sono incorporati fin dall'origine nei processi decisionali e operativi. In tali organizzazioni l'IA non è un elemento aggiuntivo rispetto alle attività tradizionali, ma un'infrastruttura trasversale alle funzioni aziendali.

L'evoluzione verso modelli AI-native richiede il superamento di una logica basata sulla semplice digitalizzazione dei processi. Se la trasformazione digitale ha introdotto strumenti informatici finalizzati soprattutto all'automazione delle attività amministrative e operative, l'intelligenza artificiale consente oggi di intervenire anche sui processi a elevato contenuto cognitivo, cambiando il modo in cui vengono prese le decisioni e distribuite le responsabilità. Di conseguenza, l'adozione dell'IA non può essere affrontata come un semplice progetto tecnologico gestito dalle funzioni informatiche: richiede il coinvolgimento dell'intera organizzazione e, in particolare, del management, chiamato a definire nuove modalità di governance, nuovi criteri di gestione dei dati e nuovi modelli di collaborazione tra persone e sistemi intelligenti.

1.5 L'impatto dell'intelligenza artificiale sul lavoro e sulle competenze

Tra gli aspetti più analizzati dalla letteratura vi è l'impatto dell'intelligenza artificiale sul lavoro umano. Le prime interpretazioni si concentravano soprattutto sul rischio di sostituzione dei lavoratori da parte delle macchine, alimentando il dibattito sugli effetti

occupazionali dell'automazione. La letteratura più recente non ha superato tale prospettiva, ma ne ha evidenziato la complessità: l'introduzione dell'IA tende a produrre simultaneamente effetti di sostituzione e di potenziamento del lavoro, la cui prevalenza dipende dal contesto organizzativo, dalla natura delle attività e dalle concrete scelte di implementazione.

Come osservano Acemoglu e Restrepo (2018, 2019), l'automazione non produce solo un effetto di sostituzione (displacement) delle mansioni svolte dall'uomo, ma genera anche nuovi compiti e nuove funzioni (reinstatement) in cui il lavoro umano mantiene un vantaggio comparato. Anche l'IA contemporanea non si limita ad automatizzare attività ripetitive, ma può supportare processi decisionali, attività creative e funzioni a elevato contenuto cognitivo. Il rapporto tra uomo e tecnologia non si configura quindi come un progressivo passaggio dalla sostituzione alla collaborazione, ma come una combinazione dinamica di automazione e potenziamento (augmentation): come mostrano Raisch e Krakowski (2021), la coesistenza di questi due poli genera una tensione organizzativa che deve essere governata in modo consapevole, poiché la prevalenza dell'uno o dell'altro dipende dalle scelte gestionali e produce effetti differenti sulle competenze e sul vantaggio competitivo dell'impresa.

L'intelligenza artificiale tende quindi a modificare la natura del lavoro, riducendo il peso delle attività standardizzate e accrescendo l'importanza di competenze come il pensiero critico, la capacità interpretativa, la supervisione delle decisioni algoritmiche, la creatività e la gestione delle relazioni. In questo scenario assume particolare rilievo il tema dell'AI Literacy, intesa come l'insieme delle conoscenze e delle competenze necessarie per comprendere il funzionamento dell'IA, valutarne criticamente i risultati e utilizzarne in modo consapevole le potenzialità (Long e Magerko, 2020). L'introduzione dell'IA rende quindi necessario un investimento costante nella formazione, nell'aggiornamento professionale e nello sviluppo delle competenze, affinché le persone possano collaborare in modo efficace con sistemi sempre più autonomi.

La trasformazione organizzativa non riguarda quindi solo la tecnologia, ma coinvolge direttamente il capitale umano. L'evoluzione dei contesti aziendali comporta una ridefinizione dei ruoli professionali, delle responsabilità e delle competenze richieste. Le

organizzazioni sono quindi chiamate a sviluppare politiche di upskilling e reskilling, favorendo l'acquisizione di competenze digitali avanzate senza trascurare le capacità umane difficilmente replicabili dalle macchine, come il giudizio critico, la leadership, l'empatia, la negoziazione e il problem solving complesso. Ne consegue che il vantaggio competitivo non dipenderà solo dalla disponibilità di tecnologie avanzate, ma dalla capacità delle imprese di costruire modelli organizzativi in cui persone e sistemi intelligenti operino in modo complementare.

1.6 Il settore bancario come contesto privilegiato di osservazione

Le trasformazioni analizzate assumono particolare rilievo nel settore bancario, caratterizzato da processi ad alta intensità informativa e da un'ampia presenza di casi d'uso dell'IA documentata da fonti istituzionali e di settore (Financial Stability Board, 2017; OECD, 2021; Banca d'Italia, 2025). L'attività bancaria si fonda infatti sull'acquisizione, sull'elaborazione e sull'interpretazione di grandi volumi di informazioni su clienti, mercati finanziari, operazioni di pagamento, concessione del credito e gestione del rischio. L'IA trova quindi un terreno di applicazione particolarmente favorevole, poiché consente di migliorare l'efficienza operativa, di aumentare la capacità predittiva dei modelli decisionali e di offrire servizi più personalizzati.

L'interesse verso questa tecnologia nel comparto non può però essere ricondotto solo alle opportunità di mercato. Come emerso in questo capitolo, l'introduzione dell'IA comporta una trasformazione più ampia, che riguarda la struttura organizzativa delle banche, le modalità di coordinamento delle attività e le competenze richieste ai professionisti. Le banche sono così uno dei primi settori in cui emerge con chiarezza la necessità di integrare innovazione tecnologica, revisione dei modelli organizzativi e valorizzazione delle competenze, a conferma del fatto che il successo dell'adozione dell'IA dipende dalla capacità di governare il cambiamento in modo sistemico.

1.7 Metodologia e struttura del lavoro

Il presente elaborato adotta un approccio teorico-analitico, fondato sulla revisione e sull'integrazione critica della letteratura scientifica e delle principali fonti istituzionali disponibili in materia di intelligenza artificiale, trasformazione organizzativa e gestione delle competenze. La scelta metodologica risponde alla natura dell'oggetto di studio: un fenomeno recente, in rapida evoluzione e con una produzione scientifica ancora in fase di consolidamento, che richiede un lavoro di sistematizzazione dei contributi esistenti più che una verifica empirica di tipo quantitativo.

L'analisi si sviluppa a partire da fonti di natura diversa, selezionate in base alla loro rilevanza e autorevolezza. Da un lato sono stati utilizzati contributi accademici pubblicati su riviste scientifiche internazionali, che forniscono il fondamento teorico dei principali concetti trattati; dall'altro sono state impiegate fonti primarie di riferimento nella storia della disciplina, utili a ricostruire con precisione l'evoluzione tecnologica dell'intelligenza artificiale. A questi riferimenti si affiancano, nei capitoli successivi, report e documenti istituzionali di settore, utili a inquadrare la concreta diffusione della tecnologia nel comparto bancario.

Sul piano espositivo, il lavoro segue un percorso che si sviluppa dal generale al particolare. Il primo capitolo definisce il quadro teorico di riferimento, ricostruendo l'evoluzione dell'intelligenza artificiale e i suoi effetti sui modelli organizzativi e sulle competenze. Il secondo capitolo analizza l'adozione dell'IA nel sistema bancario, illustrandone gli ambiti applicativi e proponendo un confronto tra due tipi ideali di intermediario, la banca digital native e la banca commerciale a radicamento territoriale. Il terzo capitolo, di carattere più analitico, esamina i modelli organizzativi funzionali all'adozione dell'IA, mettendone a confronto vantaggi e limiti. Il quarto e ultimo capitolo approfondisce il ruolo dell'uomo nell'organizzazione AI-driven, con particolare attenzione all'evoluzione delle mansioni e allo sviluppo delle competenze.

Attraverso questa articolazione, l'elaborato intende offrire una lettura integrata del fenomeno, nella quale la dimensione tecnologica, quella organizzativa e quella delle competenze

umane non vengano considerate isolatamente, ma come componenti interdipendenti di un unico processo di trasformazione digitale.

1.8 Verso l'analisi dell'intelligenza artificiale nel settore bancario

L'intelligenza artificiale può essere definita come il risultato di un lungo percorso evolutivo che, a partire dalle prime riflessioni teoriche di Alan Turing, ha condotto allo sviluppo di sistemi sempre più sofisticati, capaci non solo di elaborare informazioni, ma anche di apprendere, generare nuovi contenuti e operare con un crescente livello di autonomia. In questa prospettiva, la domanda di ricerca che guida l'elaborato trova un primo fondamento teorico: l'adozione dell'IA non richiede soltanto investimenti tecnologici, ma implica una trasformazione dei modelli organizzativi e una ridefinizione delle competenze necessarie per operare in contesti sempre più digitalizzati.

Alla luce di queste considerazioni, il capitolo successivo approfondisce l'evoluzione del settore bancario e analizza il livello di diffusione dell'intelligenza artificiale nelle istituzioni finanziarie, evidenziando i principali ambiti applicativi, le opportunità offerte dalle nuove tecnologie e le trasformazioni organizzative in atto. Tali considerazioni permettono di verificare come i presupposti teorici sviluppati in questo capitolo trovino concreta applicazione in uno dei settori più interessati dalla rivoluzione tecnologica.

Capitolo II – L'adozione dell'intelligenza artificiale nel sistema bancario

2.1 Introduzione

Il capitolo precedente ha ricostruito il percorso evolutivo dell'intelligenza artificiale e ne ha illustrato la portata sul piano organizzativo, evidenziando come l'adozione di queste tecnologie non possa essere ridotta a una semplice implementazione di strumenti informatici, ma rappresenti un processo di cambiamento sistemico. Su queste premesse si sviluppa il presente capitolo, che si concentra su uno specifico contesto applicativo: il settore bancario, individuato come ambito privilegiato di osservazione della trasformazione in atto.

La scelta del comparto bancario non è casuale. Le banche operano in un ambiente caratterizzato da un'elevata intensità informativa, da processi fortemente standardizzati, da un rigoroso perimetro regolamentare e da una crescente pressione competitiva, condizioni che rendono il settore particolarmente adatto all'adozione di soluzioni basate sull'IA. Al tempo stesso, le attività svolte dagli intermediari – dalla gestione del risparmio alla concessione del credito, dalla tutela della stabilità finanziaria alla protezione dei dati della clientela – richiedono un'adozione prudente e governata, nella quale l'innovazione tecnologica deve combinarsi con i presidi di rischio, conformità e vigilanza.

Il capitolo si sviluppa lungo direttrici complementari. In apertura viene proposto un approfondimento storico sull'evoluzione del settore bancario e sulle ondate di digitalizzazione che ne hanno ridisegnato il modello di servizio. Vengono quindi analizzati il grado di diffusione della tecnologia, sulla base delle più recenti rilevazioni istituzionali, e i principali casi d'uso lungo la catena del valore bancaria, esaminati confrontando la modalità tradizionale di gestione del processo, l'apporto dell'intelligenza artificiale e le evidenze empiriche disponibili, comprese quelle di segno critico. Il capitolo dedica poi un approfondimento all'intelligenza artificiale generativa e agentica, ricostruisce il quadro regolamentare di riferimento alla luce delle modifiche introdotte nel luglio 2026, analizza in modo esteso il tema della sorveglianza umana sui sistemi automatizzati e i principali fattori critici di adozione. In ultimo, viene proposto un confronto fra due tipi ideali – la banca digital native e la banca

commerciale tradizionale – per evidenziare come dotazione tecnologica, risorse e cultura organizzativa influenzino le concrete modalità di adozione della tecnologia.

2.2 Dalla banca di prossimità alla banca "aumentata": un approfondimento storico e digitale

Per comprendere la portata dell'attuale transizione verso l'intelligenza artificiale è utile collocarla nel più ampio percorso di evoluzione tecnologica che ha interessato il settore bancario negli ultimi decenni. L'adozione dell'IA non è un evento isolato, ma l'ultima e più avanzata tappa di un processo di digitalizzazione progressiva che ha ridefinito, ondata dopo ondata, il modello di servizio, l'organizzazione del lavoro e il rapporto con la clientela.

2.2.1 Le ondate della digitalizzazione del settore bancario

La prima ondata di innovazione tecnologica, avviata a partire dagli anni Sessanta e Settanta, ha riguardato l'automazione dei processi di back office attraverso l'informatizzazione della contabilità e la centralizzazione dei sistemi informativi (Arner, Barberis e Buckley, 2016; Gomber, Koch e Siering, 2017). In questa fase la tecnologia rispondeva soprattutto a un'esigenza di efficienza interna, riducendo i tempi di elaborazione delle operazioni e i margini di errore legati alla gestione manuale. Il modello di servizio restava però incentrato sulla filiale fisica, luogo esclusivo dell'incontro tra banca e cliente.

La seconda ondata, collocabile tra gli anni Ottanta e Novanta, ha portato la tecnologia verso il cliente finale. La diffusione degli sportelli automatici (ATM), delle carte di pagamento e, in seguito, dei primi servizi di online banking ha esteso la disponibilità dei servizi bancari oltre gli orari e i confini fisici della filiale. È in questa fase che prende avvio la multicanalità, destinata a modificare le abitudini della clientela e l'organizzazione delle reti distributive (Arner, Barberis e Buckley, 2016; Gomber, Koch e Siering, 2017).

La terza ondata coincide con l'avvento di Internet e, in seguito, degli smartphone. L'home banking prima e il mobile banking poi hanno spostato una quota crescente di operazioni sui canali digitali, trasformando l'applicazione mobile nel principale punto di contatto quotidiano tra la banca e il cliente. In parallelo, l'evoluzione del quadro normativo europeo – in

particolare la direttiva sui servizi di pagamento (PSD2) e il paradigma dell'open banking – ha imposto agli istituti l'apertura, previo consenso del cliente, dei propri dati a soggetti terzi, favorendo l'ingresso nel mercato di operatori FinTech e di nuovi modelli di servizio basati sull'integrazione tra piattaforme (Parlamento europeo e Consiglio, 2015). In questa fase la relazione bancaria assume una configurazione sempre più ibrida, in cui i canali fisici e digitali coesistono e si integrano.

L'attuale ondata, oggetto del presente elaborato, è caratterizzata dall'adozione dell'intelligenza artificiale e segna un salto rispetto alle precedenti. Mentre le fasi anteriori avevano soprattutto digitalizzato e reso accessibili a distanza processi già esistenti, l'IA interviene sul contenuto cognitivo delle attività, supportando l'analisi, la previsione e la decisione. Come osservato nel primo capitolo, si passa così da una logica di digitalizzazione dei processi a una logica AI-native, nella quale la tecnologia non si limita ad automatizzare l'esistente, ma consente di ripensare il modo stesso in cui il valore viene prodotto. È in questo senso che si può parlare di una progressiva transizione dalla banca di prossimità, storicamente fondata sulla relazione fisica di sportello, a una banca "aumentata", nella quale il giudizio dei professionisti è potenziato dall'apporto dei sistemi intelligenti.

Questa ricostruzione mette in luce un elemento di fondo: ogni ondata tecnologica ha comportato non solo l'introduzione di nuovi strumenti, ma anche una ridefinizione dei ruoli professionali, delle competenze richieste e dei modelli organizzativi. La transizione verso l'IA, in tal senso, rappresenta l'espressione più matura di una dinamica strutturale, nella quale innovazione tecnologica e trasformazione organizzativa procedono insieme.

2.3 La diffusione dell'intelligenza artificiale nel settore bancario

L'adozione dell'intelligenza artificiale nel settore bancario ha conosciuto negli ultimi anni una decisa accelerazione, sostenuta dalla progressiva maturazione delle tecniche di apprendimento automatico e, più di recente, dalla diffusione dell'IA generativa. Le principali fonti istituzionali che monitorano il fenomeno concordano nel rilevare come l'IA sia passata, in un breve arco di tempo, da tecnologia sperimentale confinata a progetti pilota a leva strategica integrata nei piani industriali delle banche.

Un osservatorio utile in ambito nazionale è l'Indagine Fintech condotta con periodicità biennale dalla Banca d'Italia, che rileva lo stato di adozione delle innovazioni tecnologiche applicate ai servizi finanziari sull'intero sistema bancario e su un campione rappresentativo di intermediari non bancari. La quinta edizione, pubblicata il 23 dicembre 2025 e riferita ai progetti del biennio 2023-2024, consente di quantificare il fenomeno con un grado di precisione superiore rispetto alle rilevazioni precedenti. Nel solo 2024 il sistema finanziario italiano ha investito 593 milioni di euro in progetti di intelligenza artificiale, importo proveniente quasi integralmente dalle banche e relativo per l'89 per cento a iniziative di nuovo avvio, a fronte del 39 per cento dell'anno precedente. L'indagine censisce 798 sistemi di intelligenza artificiale, classificati per il 55 per cento a rischio limitato, per il 34 per cento a rischio minimo e per l'11 per cento ad alto rischio secondo le categorie dell'AI Act. Un terzo degli intermediari, di cui circa la metà banche, ha istituito una struttura dedicata alla governance dell'intelligenza artificiale (Banca d'Italia, 2025).

La composizione tecnologica dei progetti aiuta a precisare dove l'IA sia effettivamente impiegata. Le tecniche più ricorrenti per numero di progetti risultano il machine learning (46 per cento), l'IA generativa (32 per cento) e l'elaborazione del linguaggio naturale (meno del 17 per cento). L'indagine rileva inoltre una netta divaricazione per ambito: la GenAI, che caratterizza oltre la metà delle nuove iniziative di investimento, è impiegata prevalentemente nelle attività operative, mentre nell'intermediazione del credito prevalgono ancora modelli predittivi basati su tecniche di machine learning (Banca d'Italia, 2025). È un dato rilevante ai fini del presente lavoro, perché mostra come le due tecnologie non si sostituiscano, ma presidino aree diverse della catena del valore, con implicazioni organizzative distinte.

A livello di sistema associativo, un contributo di primo piano è offerto da ABI Lab, il centro di ricerca e innovazione per la banca promosso dall'Associazione Bancaria Italiana. Il Rapporto sullo scenario e sui trend del mercato ICT per il settore bancario 2026, condotto su un campione rappresentativo di circa l'89 per cento del settore bancario in termini di dipendenti, rileva che per la quasi totalità degli istituti il budget ICT per il 2026 risulta in crescita o stabile rispetto all'anno precedente. Le principali priorità di investimento indicate sono, nell'ordine, l'intelligenza artificiale – individuata come principale area di intervento per

rendere più efficienti i processi interni e migliorare i servizi alla clientela –, il rafforzamento della sicurezza informatica e l'ottimizzazione dei processi; significativa è anche l'attenzione rivolta alla gestione strutturata dei dati, alla modernizzazione dei sistemi di core banking e all'evoluzione dei sistemi di gestione della relazione con la clientela (ABI Lab, 2026; ABI, 2026). Guardando alla ricerca e allo sviluppo, il rapporto conferma l'intelligenza artificiale come principale driver di innovazione anche per gli anni successivi.

Sul piano internazionale, l'analisi del fenomeno si arricchisce del contributo dell'OCSE, che ha dedicato specifici approfondimenti all'impiego dell'intelligenza artificiale, del machine learning e dei big data nel settore finanziario (OECD, 2021). Le analisi dell'ente mettono in luce, accanto alle opportunità in termini di efficienza, personalizzazione dei servizi e migliore gestione del rischio, una serie di sfide di natura regolamentare e di tutela: la trasparenza degli algoritmi, la qualità e la governance dei dati, i rischi di distorsione (bias) e discriminazione, nonché le implicazioni per la stabilità finanziaria derivanti da un impiego diffuso di modelli analoghi da parte di più operatori. Analoghe considerazioni erano state anticipate dal Financial Stability Board (2017), che segnalava già come la concentrazione dell'offerta tecnologica e l'adozione di modelli simili da parte di più intermediari possano generare comportamenti correlati e nuove forme di rischio sistemico.

A questi contributi si affianca una crescente produzione accademica. Come osservano Königstorfer e Thalmann (2020), l'adozione della tecnologia si concentra soprattutto nelle aree della gestione del rischio, dell'individuazione delle frodi e del miglioramento dell'interazione con la clientela, pur permanendo sfide significative in termini di disponibilità di competenze, qualità dei dati e interpretabilità dei modelli. Il quadro che emerge dall'insieme di queste fonti è quello di un settore in rapida trasformazione, nel quale l'intelligenza artificiale si sta affermando come componente strutturale del modello di business.

A questo quadro si aggiunge la prospettiva delle autorità di supervisione. L'Autorità bancaria europea ha analizzato l'impiego delle tecniche di machine learning nei modelli interni per il calcolo dei requisiti patrimoniali, richiamando la necessità di adeguati presidi in materia di governance dei modelli, qualità dei dati e comprensibilità dei risultati da parte delle

funzioni aziendali e dell'autorità di vigilanza (European Banking Authority, 2021a). La convergenza fra le indicazioni delle autorità di vigilanza e il quadro delineato dall'AI Act configura, per le banche, un sistema di aspettative articolato, nel quale l'innovazione tecnologica deve procedere di pari passo con il rafforzamento dei controlli interni.

Se, da un lato, i grandi gruppi bancari si collocano tipicamente all'avanguardia, disponendo delle risorse necessarie per investire in competenze, infrastrutture e progettualità dedicate, dall'altro anche gli intermediari di minori dimensioni sono progressivamente coinvolti nel processo, spesso attraverso soluzioni offerte da fornitori esterni o mediante la partecipazione a iniziative di sistema.

2.4 I principali casi d'uso: dai processi tradizionali all'apporto dell'intelligenza artificiale

L'IA trova applicazione lungo l'intera catena del valore bancaria, dalle attività rivolte al cliente finale (front office) ai processi interni di supporto (back office), fino alle funzioni di controllo e di gestione del rischio. Per apprezzarne il concreto valore aggiunto, i principali casi d'uso sono di seguito esaminati con una prospettiva uniforme: per ciascun ambito viene prima descritta la modalità tradizionale di gestione del processo, poi illustrato l'apporto dell'intelligenza artificiale e, infine, riportate le evidenze empiriche disponibili insieme alle criticità che ne condizionano la portata.

Quest'ultimo passaggio richiede una premessa. La letteratura economica ha da tempo segnalato che fra l'introduzione di una tecnologia a uso generale e la misurazione dei suoi effetti sulla produttività intercorre un intervallo anche lungo, dovuto alla necessità di investimenti complementari in capitale organizzativo e in competenze: è il cosiddetto paradosso della produttività, formulato per l'intelligenza artificiale da Brynjolfsson, Rock e Syverson (2017). Ne consegue che le stime di beneficio riportate nella pubblicistica di settore, spesso di fonte aziendale e riferite a singole implementazioni, non possono essere assunte come misura generale dell'impatto della tecnologia. Nei paragrafi che seguono si è quindi scelto di riportare, per ciascun caso d'uso, evidenze provenienti da studi sottoposti a revisione tra pari o da rilevazioni istituzionali, dando conto anche dei risultati di segno critico.

2.4.1 La valutazione del merito creditizio

Il processo tradizionale: la valutazione del merito creditizio è una delle attività fondamentali della banca. Nel modello tradizionale, l'istruttoria di affidamento si fonda su un insieme definito di informazioni – bilanci, documentazione reddituale, dati andamentali, centrale rischi – elaborate attraverso modelli statistici di scoring e integrate dal giudizio dell'operatore competente. Il processo, per quanto strutturato, tende a essere caratterizzato da tempi di istruttoria non trascurabili, da una capacità predittiva legata al numero e alla natura delle variabili considerate e da un margine di soggettività nella valutazione delle situazioni non standard (European Banking Authority, 2020).

L'apporto dell'intelligenza artificiale: i modelli di credit scoring basati su tecniche di machine learning consentono di analizzare un'ampia gamma di variabili – anche non tradizionali e non strutturate – per stimare la probabilità di insolvenza del richiedente. Khandani, Kim e Lo (2010) documentano, su dati di una banca commerciale statunitense, un miglioramento significativo della capacità predittiva rispetto ai modelli lineari tradizionali. La capacità degli algoritmi di individuare relazioni non lineari fra i dati consente inoltre una maggiore granularità nella definizione delle condizioni di offerta.

Evidenze e criticità: il guadagno di accuratezza non si distribuisce in modo uniforme. Fuster, Goldsmith-Pinkham, Ramadorai e Walther (2022) mostrano, sul mercato statunitense dei mutui, che il passaggio a modelli di machine learning migliora la previsione complessiva ma produce effetti distributivi diseguali: alcune categorie di richiedenti ottengono condizioni migliori, altre risultano penalizzate, senza che ciò dipenda da un intento discriminatorio. Bartlett, Morse, Stanton e Wallace (2022) documentano come, pur in presenza di processi algoritmici, permangano differenziali di prezzo a sfavore di mutuatari appartenenti a minoranze. Si tratta di risultati che smentiscono l'idea di una neutralità intrinseca dei modelli e che contribuiscono a spiegare la scelta del legislatore europeo di collocare i sistemi di valutazione dell'affidabilità creditizia delle persone fisiche fra quelli ad alto rischio (Regolamento UE 2024/1689, allegato III, punto 5, lettera b), con obblighi rafforzati in materia di governance dei dati, documentazione e sorveglianza umana. A ciò si aggiunge un quadro di trasparenza e garanzie desumibile dalla lettura congiunta degli articoli 13, 14, 15 e 22 del

Regolamento (UE) 2016/679: i primi richiedono, nei casi previsti, informazioni significative sulla logica utilizzata e sull'importanza e le conseguenze del trattamento; l'articolo 22 disciplina il diritto a non essere sottoposti a decisioni basate unicamente sul trattamento automatizzato e, quando esse sono ammesse, prevede salvaguardie quali l'intervento umano, la possibilità di esprimere la propria opinione e di contestare la decisione. Gli orientamenti applicativi richiedono che tali informazioni siano comprensibili e che l'intervento umano sia sostanziale (Article 29 Working Party, 2018). Ne deriva un'esigenza organizzativa di trasparenza, comprensibilità e contestabilità della decisione, non un autonomo e generalizzato obbligo di motivazione fondato sul solo articolo 22.

2.4.2 Il contrasto alle frodi e il presidio antiriciclaggio

Il processo tradizionale: il contrasto alle frodi e gli adempimenti in materia di antiriciclaggio (Anti-Money Laundering, AML) si fondano, nel modello tradizionale, su sistemi di controllo basati su regole predefinite (rule-based), che segnalano le operazioni al superamento di determinate soglie o al verificarsi di condizioni prestabilite. Questa impostazione, per quanto necessaria, genera un elevato numero di segnalazioni, molte delle quali si rivelano falsi positivi, con conseguente impiego di ingenti risorse nelle verifiche manuali e con il rischio di non intercettare schemi fraudolenti nuovi o non previsti dalle regole (FATF, 2021).

L'apporto dell'intelligenza artificiale: gli algoritmi di apprendimento automatico si prestano all'analisi di grandi volumi di transazioni, consentendo di individuare schemi anomali che sfuggirebbero ai controlli statici e di adattarsi con maggiore rapidità all'evoluzione delle tecniche fraudolente. Il Gruppo di azione finanziaria internazionale ha dedicato al tema un rapporto specifico, nel quale riconosce alle nuove tecnologie la possibilità di rendere i presidi antiriciclaggio più rapidi, più accurati e meno costosi, in particolare nella profilatura della clientela e nel monitoraggio continuo delle operazioni (FATF, 2021).

Evidenze e criticità: lo stesso rapporto del FATF (2021) individua con chiarezza i limiti dell'approccio: la difficoltà di spiegare gli esiti dei modelli alle autorità, la scarsità di dati etichettati – le frodi accertate sono eventi rari, il che rende l'addestramento strutturalmente sbilanciato –, i vincoli alla condivisione delle informazioni fra intermediari e il rischio che

l'automazione si traduca in un adempimento formale privo di reale capacità di intercettazione. Va inoltre osservato che la riduzione dei falsi positivi non è un risultato automatico: opera su un compromesso esplicito con i falsi negativi, la cui accettabilità è una decisione di rischio che resta in capo agli organi aziendali e non al modello. Sul piano delle responsabilità, infine, la valutazione conclusiva e l'inoltro della segnalazione di operazione sospetta restano attribuiti a una funzione aziendale individuata secondo la disciplina antiriciclaggio: l'automazione interviene sulla fase di selezione, non sulla titolarità dell'obbligo.

2.4.3 La relazione con la clientela e la personalizzazione dei servizi

Il processo tradizionale: nel modello tradizionale, la relazione con la clientela si fonda soprattutto sul contatto diretto in filiale e sui canali di assistenza telefonica, con orari e capacità di risposta necessariamente limitati. La segmentazione della clientela e la formulazione delle offerte commerciali si basano su categorie ampie e su analisi condotte ex post, con una capacità di personalizzazione contenuta e tempi di reazione non sempre allineati alle esigenze del cliente (European Central Bank Banking Supervision, 2024).

L'apporto dell'intelligenza artificiale: la diffusione di assistenti virtuali e chatbot, progressivamente potenziati dalle capacità conversazionali dell'IA generativa, ha consentito di offrire un servizio di assistenza continuativo e di gestire in autonomia una quota crescente di richieste; l'Indagine Fintech rileva proprio nell'arricchimento dell'offerta di servizi digitali tramite chatbot, mediante cui raccomandare prodotti e offrire consulenza e assistenza, uno degli ambiti in cui l'adozione è più visibile (Banca d'Italia, 2025). In ambito consulenziale, le soluzioni di robo-advisory rendono accessibili a una platea più ampia servizi di consulenza finanziaria automatizzata, tradizionalmente riservati a segmenti di clientela elevati.

Evidenze e criticità: D'Acunto, Prabhala e Rossi (2019), analizzando l'introduzione di un servizio di robo-advising su un campione di investitori, rilevano benefici concentrati sui clienti in partenza meno diversificati, mentre per gli investitori già diversificati il vantaggio si riduce e i costi aumentano: l'effetto del servizio dipende quindi in misura decisiva dal punto di partenza dell'utente. La ricerca sull'adozione dei consulenti algoritmici documenta una diffusa resistenza degli utenti ad affidarsi a raccomandazioni automatizzate, fenomeno

noto come avversione all'algoritmo (Dietvorst, Simmons e Massey, 2015). Ne deriva che il ruolo del professionista non viene eliminato ma ridefinito, spostandosi verso le componenti relazionali e la gestione delle situazioni complesse, coerentemente con la logica di complementarità delineata nel capitolo precedente. Resta infine il vincolo di trasparenza: i sistemi destinati a interagire direttamente con le persone devono essere progettati in modo che l'utente sia informato di stare interagendo con un sistema di IA, salvo che ciò risulti evidente dal contesto (Regolamento UE 2024/1689, art. 50).

2.4.4 L'efficienza operativa e l'automazione dei processi interni

Il processo tradizionale: numerose attività di back office – l'elaborazione documentale, la riconciliazione dei dati, la gestione delle pratiche, l'estrazione di informazioni da documenti – sono state storicamente svolte in modo prevalentemente manuale, con un elevato dispendio di tempo, un rischio di errore non trascurabile e una scarsa scalabilità. La gestione di documenti non strutturati, in particolare, ha tradizionalmente richiesto l'intervento diretto di personale dedicato (Banca d'Italia, 2025).

L'apporto dell'intelligenza artificiale: l'integrazione fra tecniche di intelligenza artificiale e strumenti di automazione dei processi consente di gestire in modo più rapido queste attività, grazie alla capacità di interpretare il linguaggio naturale e di trattare informazioni contenute in documenti testuali, contratti e corrispondenza. È l'ambito nel quale l'adozione risulta oggi più concentrata: secondo l'Indagine Fintech, poco meno della metà dei progetti censiti insiste sulle attività operative, in aumento rispetto alla precedente rilevazione, e l'IA generativa è impiegata in via prevalente proprio in quest'area, facilitando l'automazione dei processi interni e la semplificazione dei flussi operativi (Banca d'Italia, 2025).

Evidenze e criticità: gli studi sperimentali disponibili confermano l'esistenza di guadagni di produttività, ma ne precisano l'entità e la distribuzione. Noy e Zhang (2023) rilevano, su compiti di scrittura professionale, una riduzione sensibile dei tempi di esecuzione e un miglioramento della qualità, con benefici maggiori per i partecipanti inizialmente meno performanti. Brynjolfsson, Li e Raymond (2025), analizzando l'introduzione di un assistente conversazionale in un contesto di assistenza clienti, misurano un aumento medio della

produttività concentrato sugli operatori meno esperti, a fronte di effetti modesti o nulli sui più esperti. Dell'Acqua e altri (2026) segnalano un profilo di rischio speculare: su compiti collocati appena oltre la "frontiera irregolare" delle capacità del modello, l'uso dell'assistente ha peggiorato le prestazioni dei partecipanti, che hanno accolto acriticamente output plausibili ma errati. Il beneficio dipende quindi dalla corretta selezione dei compiti e dalla capacità dell'utilizzatore di riconoscere i limiti dello strumento: un risultato che riporta direttamente al tema della sorveglianza umana trattato al § 2.7 e a quello delle competenze, oggetto del capitolo conclusivo.

2.4.5 La conformità normativa e la sicurezza informatica

Il processo tradizionale: il presidio della conformità normativa e della sicurezza informatica ha tradizionalmente richiesto un'attività di monitoraggio prevalentemente manuale, affidata a personale specializzato chiamato a interpretare un corpus normativo esteso e mutevole e a sorvegliare i sistemi informativi a fronte di minacce in costante evoluzione (European Banking Authority, 2021b; Parlamento europeo e Consiglio dell'Unione europea, 2022).

L'apporto dell'intelligenza artificiale: nel primo ambito, le soluzioni basate sull'IA supportano l'interpretazione e il monitoraggio del quadro regolamentare, agevolando l'adeguamento degli intermediari a una normativa in continua evoluzione, secondo la prospettiva che la letteratura riconduce al filone della RegTech (Arner, Barberis e Buckley, 2017). Nel secondo, gli algoritmi di apprendimento automatico rafforzano i presidi di cybersecurity, consentendo di individuare tempestivamente attività anomale e potenziali minacce ai sistemi informativi; si tratta, non a caso, della seconda priorità di investimento indicata dalle banche italiane dopo l'intelligenza artificiale (ABI Lab, 2026).

Evidenze e criticità: l'impiego dell'IA in questi ambiti si colloca all'interno di un perimetro regolamentare già denso. Il Regolamento (UE) 2022/2554 sulla resilienza operativa digitale del settore finanziario (DORA) impone agli intermediari obblighi puntuali in materia di gestione dei rischi informatici, segnalazione degli incidenti e presidio dei fornitori terzi critici, che si applicano anche alle componenti basate sull'intelligenza artificiale. L'adozione

di strumenti di IA in ambito compliance introduce inoltre un rischio specifico: la delega a un sistema generativo dell'interpretazione di una norma comporta la possibilità di risposte plausibili ma inesatte, che in questo dominio si traducono direttamente in rischio di non conformità. Il ricorso a queste soluzioni richiede quindi la verifica documentata dell'output da parte della funzione competente, secondo quanto illustrato al § 2.7.

2.4.6 La gestione del rischio, il monitoraggio andamentale e l'integrazione dei fattori ESG

Il processo tradizionale: la misurazione e il monitoraggio dei rischi – di credito, di mercato, operativi e di liquidità – si fondano tradizionalmente su modelli quantitativi alimentati da dati storici e su rilevazioni condotte con cadenza periodica. Questa impostazione, per quanto rigorosa, ha un'intrinseca natura retrospettiva: i segnali di deterioramento vengono spesso colti con ritardo, quando il peggioramento della posizione si è già manifestato. Allo stesso modo, l'integrazione dei fattori ambientali, sociali e di governance (ESG) nelle valutazioni ha risentito della scarsa disponibilità e della difficile comparabilità dei dati non finanziari (Basel Committee on Banking Supervision, 2013; European Banking Authority, 2020).

L'apporto dell'intelligenza artificiale: gli algoritmi di apprendimento automatico consentono di affinare i modelli di rischio, incorporando un numero più ampio di variabili e cogliendo relazioni non lineari fra i fenomeni. Particolarmente rilevante è il contributo al monitoraggio andamentale del credito: l'analisi continua di dati transazionali e comportamentali permette di individuare precocemente i segnali di allerta (early warning), in coerenza con l'impostazione degli orientamenti dell'Autorità bancaria europea sulla concessione e sul monitoraggio del credito (European Banking Authority, 2020). Nell'ambito ESG, le tecniche di elaborazione del linguaggio naturale consentono di estrarre e strutturare informazioni da fonti eterogenee a supporto della valutazione dei profili di sostenibilità della clientela.

Evidenze e criticità: l'impiego di queste tecniche nei modelli regolamentari incontra un vincolo esplicito: l'Autorità bancaria europea, nel documento di consultazione dedicato all'uso del machine learning nei modelli interni, ha chiarito che l'adozione di tecniche più complesse è ammissibile soltanto se l'intermediario è in grado di comprendere e spiegare il

funzionamento del modello, di giustificarne le scelte metodologiche e di garantirne la stabilità nel tempo, evitando l'eccessivo adattamento ai dati storici (European Banking Authority, 2021a). Nell'ambito ESG, la letteratura ha inoltre documentato la scarsa convergenza fra i giudizi di sostenibilità prodotti da fornitori diversi a partire da dati analoghi (Berg, Kölbel e Rigobon, 2022): un limite che riguarda i dati sottostanti e che nessuna tecnica di elaborazione può risolvere autonomamente. Il rischio, in questo caso, è che l'automazione conferisca un'apparenza di oggettività a valutazioni che restano largamente convenzionali.

La molteplicità dei casi d'uso illustrati evidenzia la natura trasversale dello strumento nel settore bancario, coerentemente con la sua qualificazione come tecnologia ad applicazione generale. L'adozione di queste soluzioni non produce, però, benefici automatici: presuppone adeguati investimenti in infrastrutture e competenze, oltre a una revisione dei processi e dei modelli organizzativi, aspetti che saranno approfonditi nei capitoli successivi.

2.5 L'intelligenza artificiale generativa e agentica nel settore bancario

Tra le tecnologie che stanno ridisegnando più in profondità l'attività bancaria, l'intelligenza artificiale generativa merita un approfondimento specifico. Come illustrato nel primo capitolo, i modelli generativi si distinguono per la capacità di produrre contenuti originali a partire da istruzioni formulate in linguaggio naturale; è proprio questa caratteristica a renderli particolarmente adatti a un settore ad alta intensità documentale e normativa come quello bancario, nel quale una quota rilevante del lavoro consiste nella lettura, nella sintesi, nella redazione e nella verifica di testi.

Sul piano delle applicazioni, l'IA generativa interviene principalmente su tre fronti. In primo luogo, la produttività individuale dei professionisti: strumenti di assistenza integrati negli applicativi di uso quotidiano supportano la redazione di documenti, la sintesi di normative e di corrispondenza, l'analisi di testi complessi e la preparazione di reportistica. In secondo luogo, la gestione della conoscenza aziendale: gli assistenti conversazionali fondati su tecniche di recupero delle informazioni consentono di interrogare in linguaggio naturale l'ampio patrimonio documentale interno – normative, disposizioni operative, manuali di prodotto – restituendo risposte contestualizzate e rendendo fruibile una conoscenza in molti casi

dispersa. In terzo luogo, il supporto alla relazione commerciale e all'assistenza alla clientela, attraverso interfacce conversazionali evolute in grado di comprendere richieste articolate.

L'evoluzione più recente, già delineata sul piano teorico nel primo capitolo, conduce dai sistemi generativi ai sistemi agentici. Se i primi operano in risposta a un input dell'utente, i secondi sono progettati per perseguire un obiettivo attraverso una sequenza di azioni, interagendo con applicativi, banche dati e strumenti esterni (Wang et al., 2024). In ambito bancario, la presente analisi formula l'ipotesi applicativa di collaboratori digitali capaci di svolgere compiti articolati – come la raccolta e la riconciliazione di informazioni provenienti da fonti diverse, la predisposizione di pratiche o l'esecuzione di controlli – sotto la supervisione di un operatore. Si tratta di una trasposizione prospettica delle capacità e dei casi organizzativi descritti nella letteratura sugli agenti, non dell'affermazione che tali applicazioni siano già diffuse nel settore (OECD, 2026).

Occorre tuttavia evidenziare come l'introduzione dell'IA generativa e agentica sollevi questioni in parte diverse da quelle poste dai tradizionali modelli di machine learning. Tra queste, la possibilità che i modelli producano risposte errate ma plausibili, le cosiddette allucinazioni (Ji et al., 2023), la necessità di tutelare la riservatezza e la sicurezza dei dati trattati, il rischio connesso all'utilizzo di soluzioni offerte da fornitori terzi e la questione della localizzazione dei dati, particolarmente rilevante in un contesto regolamentato. Nell'interpretazione qui proposta, per i sistemi agentici il baricentro del rischio si estende dall'accuratezza dell'output alle conseguenze delle azioni autorizzate: diventano centrali l'identità propria dell'agente, la delimitazione dei permessi secondo il principio del privilegio minimo, la tracciabilità delle operazioni, la supervisione proporzionata e la possibilità di sospensione (NIST, 2026; OECD, 2026). Queste misure non eliminano il rischio di azione, ma rendono controllabile e ricostruibile il perimetro entro cui l'agente opera.

2.6 Il quadro regolamentare: il Regolamento europeo sull'intelligenza artificiale dopo l'Omnibus digitale

L'adozione dell'intelligenza artificiale nel settore bancario si colloca all'interno di una cornice regolamentare di recente definizione e tuttora in movimento. Il principale riferimento

normativo è costituito, a livello europeo, dal Regolamento (UE) 2024/1689, noto come AI Act, primo quadro giuridico organico al mondo in materia di intelligenza artificiale (Parlamento europeo e Consiglio, 2024). Tale quadro è stato modificato, in modo rilevante per la tempistica di applicazione, dal Regolamento (UE) 2026/1744 dell'8 luglio 2026, pubblicato nella Gazzetta ufficiale dell'Unione europea il 24 luglio 2026 ed entrato in vigore, in via d'urgenza, il terzo giorno successivo alla pubblicazione, ossia il 27 luglio 2026: si tratta della prima modifica sostanziale all'AI Act dalla sua adozione, denominata Omnibus digitale sull'IA (Parlamento europeo e Consiglio, 2026).

2.6.1 La genesi e le ragioni dell'intervento normativo

Il percorso che ha condotto all'approvazione dell'AI Act affonda le proprie radici nella strategia europea avviata nella seconda metà dello scorso decennio, nell'ambito della quale le istituzioni dell'Unione hanno progressivamente maturato la consapevolezza della necessità di un quadro normativo armonizzato. La proposta di regolamento è stata presentata dalla Commissione europea il 21 aprile 2021 (Commissione europea, 2021) e, nel corso dell'iter legislativo, il testo è stato oggetto di una significativa revisione per tenere conto della rapida diffusione dell'intelligenza artificiale generativa, le cui caratteristiche di applicazione generale non erano pienamente riconducibili all'impianto originario. Il Regolamento, adottato il 13 giugno 2024, è stato pubblicato nella Gazzetta ufficiale dell'Unione europea il 12 luglio 2024 ed è entrato in vigore il 1° agosto 2024.

Le ragioni dell'intervento normativo sono sintetizzate dalla stessa Commissione europea, che individua nell'opacità dei sistemi di intelligenza artificiale uno dei principali fattori di rischio: poiché spesso non è possibile ricostruire le ragioni per le quali un sistema ha assunto una determinata decisione, può risultare difficile accertare se una persona sia stata ingiustamente svantaggiata, ad esempio nell'accesso a un finanziamento. L'obiettivo dichiarato del Regolamento è, quindi, quello di promuovere un'intelligenza artificiale antropocentrica e affidabile, capace di garantire un elevato livello di protezione della salute, della sicurezza e dei diritti fondamentali, sostenendo al contempo l'innovazione.

2.6.2 L'approccio basato sul rischio

L'impianto dell'AI Act si fonda su un approccio basato sul rischio (risk-based approach), che classifica i sistemi di intelligenza artificiale in funzione del livello di rischio che comportano per la sicurezza e per i diritti fondamentali delle persone. Il Regolamento individua quattro livelli di rischio, ai quali si affianca una specifica disciplina per i modelli di intelligenza artificiale per finalità generali (General Purpose AI). Al vertice si collocano i sistemi a rischio inaccettabile, considerati una minaccia evidente per la sicurezza, i diritti e la libertà delle persone e, come tali, vietati. Seguono i sistemi ad alto rischio, ammessi ma soggetti a stringenti obblighi in materia di sicurezza, trasparenza, qualità dei dati e valutazione di conformità. I sistemi a rischio limitato sono soggetti a soli obblighi di trasparenza, mentre i sistemi a rischio minimo non sono sottoposti a specifici obblighi.

Questa architettura assume particolare rilievo per il settore bancario. I sistemi destinati a valutare l'affidabilità creditizia delle persone fisiche o a stabilirne il merito di credito rientrano fra le applicazioni ad alto rischio elencate nell'allegato III, punto 5, lettera b), del Regolamento, con la sola eccezione di quelli impiegati per individuare frodi finanziarie; vi rientrano altresì i sistemi utilizzati per la valutazione del rischio e la determinazione dei prezzi nelle assicurazioni sulla vita e sulla salute e, in quanto datore di lavoro, quelli impiegati nella selezione e nella gestione del personale (allegato III, punto 4). Ne consegue che l'adozione di questi sistemi è subordinata al rispetto di obblighi rafforzati, fra i quali assumono particolare rilievo quelli in materia di trasparenza, governance dei dati e sorveglianza umana. È utile osservare, sulla base dei dati richiamati al § 2.3, che tale perimetro riguarda comunque una minoranza dei sistemi in uso: l'Indagine Fintech classifica come ad alto rischio l'11 per cento dei sistemi censiti presso gli intermediari italiani (Banca d'Italia, 2025).

2.6.3 Il Regolamento (UE) 2026/1744 e il calendario aggiornato

La modifica introdotta nel luglio 2026 non incide sull'architettura del Regolamento – non riapre le categorie di rischio né la logica di fondo – ma interviene sulle date di applicazione della parte più onerosa della disciplina. Gli obblighi del Capo III, sezioni 1, 2 e 3, relativi ai sistemi ad alto rischio, sono stati rinviati: per i sistemi indipendenti classificati ai sensi

dell'articolo 6, paragrafo 2, ed elencati nell'allegato III la scadenza passa dal 2 agosto 2026 al 2 dicembre 2027, mentre per i sistemi che costituiscono componenti di sicurezza di prodotti disciplinati dalla normativa di armonizzazione dell'Unione, ai sensi dell'articolo 6, paragrafo 1, la nuova data è il 2 agosto 2028. La ragione dichiarata del differimento è di natura operativa: il ritardo nella predisposizione delle norme tecniche armonizzate, che forniscono le soluzioni tecniche necessarie ai fornitori per dimostrare la conformità, e nella costituzione dei quadri nazionali di governance e di valutazione della conformità ha determinato un onere di adeguamento più gravoso del previsto.

Il rinvio non ha invece toccato le altre scadenze, che restano operative. Il calendario applicabile può essere così riassunto:

- 2 febbraio 2025 – divieti relativi ai sistemi a rischio inaccettabile (art. 5) e obblighi in materia di alfabetizzazione all'intelligenza artificiale (art. 4);
- 2 agosto 2025 – obblighi per i fornitori di modelli di IA per finalità generali (Capo V), disposizioni sulla governance e regime sanzionatorio;
- 2 agosto 2026 – applicazione della generalità delle restanti disposizioni del Regolamento, fra cui gli obblighi di trasparenza di cui all'articolo 50;
- 2 dicembre 2026 – nuovo termine per la marcatura dei contenuti generati o manipolati artificialmente, con riduzione da sei a tre mesi del periodo di tolleranza concesso ai fornitori, e applicazione dei nuovi divieti introdotti nell'articolo 5 in materia di generazione di immagini intime non consensuali e di materiale pedopornografico;
- 2 agosto 2027 – termine, differito, entro il quale le autorità nazionali competenti devono istituire gli spazi di sperimentazione normativa (regulatory sandbox);
- 2 dicembre 2027 – obblighi relativi ai sistemi ad alto rischio dell'allegato III, fra cui la sorveglianza umana di cui all'articolo 14 e gli obblighi dei deployer di cui all'articolo 26;
- 2 agosto 2028 – obblighi relativi ai sistemi ad alto rischio che costituiscono componenti di sicurezza di prodotti.

Il Regolamento del 2026 interviene inoltre estendendo alle piccole imprese a media capitalizzazione le semplificazioni già previste per le piccole e medie imprese e precisando il ruolo dell'Ufficio per l'IA nella supervisione dei sistemi costruiti su modelli per finalità generali quando il modello e il sistema siano sviluppati dal medesimo fornitore; sono contestualmente individuate le ipotesi nelle quali permane la competenza delle autorità nazionali, fra le quali è espressamente richiamato l'ambito degli istituti finanziari.

Per il settore bancario il differimento ha un effetto pratico non trascurabile: l'ambito applicativo più rilevante – la valutazione dell'affidabilità creditizia – rientra nell'allegato III e vede quindi spostarsi al 2 dicembre 2027 il termine per l'adeguamento. Sono invece già applicabili gli obblighi di alfabetizzazione dell'articolo 4, in vigore dal febbraio 2025 e rilevanti per il tema delle competenze affrontato nel capitolo conclusivo, e, dall'agosto 2026, gli obblighi di trasparenza dell'articolo 50, che riguardano una platea di soggetti assai più ampia e che incidono direttamente sui casi d'uso conversazionali descritti al § 2.4.3. Il rinvio, va sottolineato, riguarda la data e non il contenuto degli obblighi: la costruzione di un sistema di gestione del rischio, della documentazione tecnica e della registrazione degli eventi richiede tempi lunghi e non può essere ricostruita a posteriori.

2.6.4 Il quadro nazionale

Sul piano interno, l'Italia si è dotata di una disciplina di raccordo con la legge 23 settembre 2025, n. 132, recante disposizioni e deleghe al Governo in materia di intelligenza artificiale, pubblicata in Gazzetta Ufficiale il 25 settembre 2025 ed entrata in vigore il 10 ottobre 2025. La legge non duplica il regolamento europeo: l'articolo 1, comma 2, stabilisce infatti che le sue disposizioni si interpretano e si applicano conformemente al Regolamento (UE) 2024/1689. L'articolo 20 designa l'Agenzia per l'Italia digitale e l'Agenzia per la cybersicurezza nazionale quali autorità nazionali per l'intelligenza artificiale, con ruoli distinti rispettivamente di promozione e di vigilanza, mentre l'articolo 24 delega il Governo ad adeguare l'ordinamento nazionale al regolamento europeo in materia di vigilanza, sanzioni e percorsi di alfabetizzazione. Il quadro nazionale si affianca, per gli intermediari, alle competenze delle

autorità di settore, che restano titolari della vigilanza prudenziale e di quella sulla trasparenza e correttezza dei comportamenti.

2.6.5 Le implicazioni per il settore bancario

Il quadro regolamentare delineato dall'AI Act non deve essere interpretato come un semplice vincolo allo sviluppo tecnologico. Alle disposizioni del Regolamento si affiancano, del resto, le indicazioni delle autorità di vigilanza del settore finanziario, orientate a garantire che l'impiego di modelli di intelligenza artificiale avvenga nel rispetto dei principi di sana e prudente gestione, di adeguato presidio dei rischi e di tutela della clientela. La definizione di regole chiare e di presidi affidabili costituisce una condizione abilitante per un'adozione sostenibile dell'IA: in un settore fondato sulla fiducia, garantire trasparenza, correttezza e responsabilità nell'impiego degli algoritmi rappresenta non solo un obbligo di conformità, ma anche un fattore di competitività. Va peraltro osservato che per gli intermediari l'AI Act non opera in isolamento, ma si sovrappone a discipline preesistenti – la protezione dei dati personali, la resilienza operativa digitale, la disciplina dell'esternalizzazione e dei controlli interni – la cui composizione rappresenta essa stessa un onere organizzativo, tanto più significativo per gli operatori di dimensione minore. La compliance normativa si configura, in questa prospettiva, come parte integrante della strategia di adozione dell'intelligenza artificiale, con rilevanti implicazioni sui modelli organizzativi e sulle competenze richieste ai professionisti del settore.

2.7 Sorveglianza umana e human in the loop: la centralità del presidio umano

Fra i temi di maggiore rilievo nel dibattito sull'adozione dell'intelligenza artificiale in ambito bancario vi è quello del presidio umano sui sistemi automatizzati. Si tratta di un tema che attraversa l'intero elaborato: costituisce il punto di congiunzione fra la dimensione regolamentare trattata nel paragrafo precedente, i modelli organizzativi esaminati nel terzo capitolo e l'evoluzione delle competenze di cui si occupa il capitolo conclusivo. Per questa ragione il paragrafo ne propone una trattazione distesa, che parte dalle definizioni, ne ricostruisce il fondamento giuridico, ne chiarisce la portata effettiva e la cala in processi bancari concreti.

2.7.1 Che cosa significa "presidio umano": una precisazione terminologica

L'espressione human in the loop viene spesso impiegata in senso generico per indicare qualunque forma di coinvolgimento umano in un processo automatizzato. In realtà essa si rifà a una configurazione precisa, che è utile distinguere da altre due.

- Human in the loop in senso proprio: il sistema produce un esito che non diventa operativo finché una persona non lo convalida. L'intervento umano è un passaggio necessario del flusso; se manca, il processo si arresta. È il caso della proposta di delibera generata da un motore di scoring che deve essere deliberata da un organo competente.
- Human on the loop: il sistema opera in autonomia e la persona ne sorveglia il funzionamento, potendo intervenire per correggerlo o interromperlo. L'intervento è possibile ma non necessario. È il caso del monitoraggio di un sistema antifrode che blocca automaticamente le operazioni sospette, con l'analista che ne verifica l'operato e gestisce le eccezioni.
- Human in command: la persona non interviene sul singolo esito, ma decide se, dove e con quali limiti il sistema debba essere impiegato, e conserva la responsabilità complessiva del suo utilizzo. È il livello proprio degli organi aziendali e delle funzioni di controllo.

La distinzione non è nominalistica: i tre livelli comportano requisiti diversi di competenza, di tempo disponibile e di progettazione dell'interfaccia, e generano garanzie di intensità diversa per il cliente. Confonderli è la premessa più comune di un presidio che appare robusto sulla carta ma non lo è nei fatti.

2.7.2 Il fondamento giuridico: gli articoli 14 e 26 dell'AI Act e l'articolo 22 del GDPR

Sul piano normativo il principio ha trovato riconoscimento nell'articolo 14 dell'AI Act, dedicato alla sorveglianza umana (human oversight), che si applica ai sistemi ad alto rischio e le cui disposizioni, per effetto del rinvio illustrato al § 2.6.3, troveranno applicazione dal 2 dicembre 2027 per i sistemi dell'allegato III. La norma stabilisce che tali sistemi debbano

essere progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui sono in uso; la sorveglianza umana mira a prevenire o ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali che possono emergere quando il sistema è utilizzato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile. Il paragrafo 3 dell'articolo precisa, inoltre, che le misure di sorveglianza sono commisurate ai rischi, al livello di autonomia e al contesto di utilizzo del sistema.

Il paragrafo 4 dell'articolo declina il principio in cinque capacità che devono essere messe a disposizione di chi esercita la sorveglianza: comprendere correttamente le capacità e i limiti pertinenti del sistema ed essere in grado di monitorarne debitamente il funzionamento, anche al fine di individuare e affrontare anomalie, disfunzioni e prestazioni inattese; restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'output, tendenza che la norma stessa definisce «distorsione dell'automazione», in particolare in relazione ai sistemi utilizzati per fornire informazioni o raccomandazioni per decisioni che devono essere prese da persone fisiche; interpretare correttamente l'output del sistema; decidere, in qualsiasi situazione particolare, di non usare il sistema o di ignorarne, annullarne o ribaltarne l'output; intervenire sul funzionamento del sistema o interromperlo mediante un pulsante di arresto o una procedura analoga che consenta al sistema di arrestarsi in condizioni di sicurezza. Si tratta, si noti, di requisiti che gravano in primo luogo sul fornitore, chiamato a progettare il sistema in modo che tali capacità siano effettivamente esercitabili dal deployer.

A questa disposizione se ne affianca una seconda, spesso trascurata ma più direttamente rilevante per le banche che acquistano soluzioni da terzi. L'articolo 26 pone infatti a carico del deployer – l'utilizzatore professionale del sistema – l'obbligo di impiegarlo conformemente alle istruzioni per l'uso e di affidarne la sorveglianza umana a persone fisiche che dispongano della necessaria competenza, formazione e autorità, oltre che del necessario supporto.

Il quadro si completa con l'articolo 22 del Regolamento (UE) 2016/679, che precede l'AI Act e continua ad applicarsi autonomamente. La norma riconosce all'interessato il diritto di non essere sottoposto a una decisione basata unicamente su un trattamento automatizzato che produca effetti giuridici che lo riguardano o che incida in modo analogamente significativo sulla sua persona; ove la decisione automatizzata sia ammessa, devono essere garantiti il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione. Le linee guida del Gruppo di lavoro Articolo 29, poi fatte proprie dal Comitato europeo per la protezione dei dati, hanno chiarito che l'intervento umano, per essere tale, deve essere effettuato da una persona dotata dell'autorità e della competenza necessarie a modificare la decisione e deve consistere in un esame sostanziale di tutti gli elementi rilevanti: un passaggio meramente formale non è sufficiente a sottrarre il trattamento all'ambito applicativo della norma (Article 29 Working Party, 2018). È da questa precisazione che discende, sul piano pratico, l'intera differenza fra una supervisione reale e una supervisione simbolica.

2.7.3 Il controllo "a quattro occhi": che cosa prevede l'AI Act e che cosa prevede la disciplina bancaria

L'articolo 14, paragrafo 5, dell'AI Act prevede una misura rafforzata di sorveglianza – la verifica e la conferma separata dell'azione o della decisione da parte di almeno due persone fisiche dotate della necessaria competenza, formazione e autorità – esclusivamente per i sistemi ad alto rischio di cui all'allegato III, punto 1, lettera a), ossia i sistemi di identificazione biometrica remota. Non si tratta quindi di una regola generale applicabile a tutti i sistemi ad alto rischio, né di una regola che riguardi i casi d'uso bancari descritti al § 2.4: né i sistemi di valutazione del merito creditizio, né quelli antifrode, né quelli impiegati nella gestione del personale ricadono in quella specifica previsione. Le banche possono trovarsi nel perimetro solo in ipotesi marginali e comunque estranee al nucleo dell'attività di intermediazione, quali l'eventuale impiego di sistemi di identificazione biometrica remota per finalità di sicurezza fisica; peraltro, lo stesso allegato III esclude espressamente dalla categoria i sistemi destinati alla verifica biometrica la cui unica finalità sia confermare che una determinata persona fisica

è la persona che dice di essere, ipotesi nella quale rientrano le procedure di riconoscimento impiegate nell'acquisizione a distanza della clientela.

Ciò non significa che il principio dei quattro occhi sia estraneo alla banca: significa che nel settore bancario esso ha una fonte diversa e assai più risalente. Il principio della separazione fra chi propone e chi delibera, e più in generale della doppia verifica delle operazioni rilevanti, appartiene alla tradizione dei controlli interni ed è codificato negli orientamenti dell'Autorità bancaria europea sulla governance interna, che ne fanno un elemento costitutivo dell'assetto organizzativo (European Banking Authority, 2021b), oltre che nelle disposizioni di vigilanza nazionali in materia di sistema dei controlli interni (Banca d'Italia, Circolare n. 285/2013). Nell'ordinamento italiano il medesimo principio trova espressione nell'articolazione dei poteri di delibera, nella distinzione fra proponente e deliberante nel processo del credito e nell'organizzazione dei controlli su più livelli.

La conseguenza pratica è rilevante e vale la pena esplicitarla: quando una banca introduce un modello di scoring basato su machine learning nel processo di affidamento, il doppio controllo non nasce con l'algoritmo, ma è già presente nell'architettura dei poteri deliberativi. Il problema organizzativo non è quindi introdurre un secondo paio di occhi, ma evitare che i due paia di occhi esistenti finiscano per guardare entrambi lo stesso output algoritmico, convergendo su di esso anziché verificarlo in modo indipendente. È un rischio concreto e documentato dalla letteratura sull'automazione, che i paragrafi seguenti illustrano con riferimento a processi determinati.

2.7.4 Quattro esempi tratti da processi bancari

Per rendere operativo quanto precede, si considerano quattro processi rappresentativi delle diverse configurazioni di presidio, con l'indicazione, per ciascuno, di chi sorveglia, che cosa può fare e quali condizioni ne rendono effettiva l'azione.

Delibera di affidamento con scoring algoritmico: il modello elabora la pratica e restituisce una classe di rischio e una raccomandazione. La configurazione è di tipo human in the loop: la raccomandazione non produce effetti finché l'organo competente non delibera. Perché la sorveglianza sia sostanziale occorre che l'analista disponga delle informazioni sui

principali fattori che hanno determinato l'esito, che abbia la facoltà di discostarsene e che tale facoltà sia esercitabile senza oneri sproporzionati; che lo scostamento sia motivato e tracciato; e che i tempi assegnati all'istruttoria siano compatibili con un esame effettivo. Un indicatore organizzativo utile è il tasso di scostamento dalla raccomandazione, monitorato nel tempo: un valore prossimo allo zero non segnala necessariamente la bontà del modello, ma può indicare che la supervisione si è ridotta a una ratifica. Sul piano della tutela del cliente, questo assetto è anche ciò che consente di soddisfare i requisiti dell'articolo 22 del GDPR in caso di esito negativo.

Alert antiriciclaggio generato da un modello comportamentale: il sistema seleziona le posizioni anomale e le assegna all'analista. La configurazione è ibrida: il sistema opera in autonomia nella fase di selezione (human on the loop), mentre la decisione conclusiva sulla segnalazione resta necessariamente umana. Il presidio effettivo dipende qui da due condizioni spesso sottovalutate: il carico di lavoro assegnato al singolo analista, poiché un numero di alert eccedente la capacità di esame trasforma la verifica in un adempimento formale; e la possibilità di segnalare all'unità che governa il modello i casi in cui la selezione appare sistematicamente inefficace, alimentando un circuito di miglioramento. Il presidio umano, in questo caso, non è soltanto un controllo a valle, ma la principale fonte di informazione per la manutenzione del sistema.

Assistente generativo per la consultazione della normativa interna: il collaboratore interroga in linguaggio naturale il corpus documentale e ottiene una risposta sintetica con i riferimenti. Non si tratta di un sistema ad alto rischio ai sensi dell'AI Act, ma il presidio resta essenziale per ragioni di rischio operativo e di conformità: la risposta va verificata sulla fonte citata prima di essere utilizzata in un atto o comunicata al cliente, e la responsabilità dell'atto resta di chi lo compie. Il richiamo all'articolo 50 dell'AI Act deve essere circoscritto: la disposizione disciplina una categoria specifica, ossia i testi generati o manipolati dall'IA che siano pubblicati allo scopo di informare il pubblico su questioni di interesse pubblico. Soltanto per tali testi l'obbligo informativo non si applica quando il contenuto è stato sottoposto a revisione umana o controllo editoriale e una persona fisica o giuridica assume la responsabilità editoriale della pubblicazione. Non si tratta, dunque, di un'esenzione generale dagli

obblighi di trasparenza, né di una regola applicabile indistintamente agli altri casi d'uso; il riferimento vale qui soltanto come analogia sull'importanza della verifica umana accompagnata da responsabilità. Il caso illustra inoltre il rischio segnalato da Dell'Acqua e altri (2026): la plausibilità formale dell'output riduce la propensione a verificarlo proprio nei casi in cui la verifica sarebbe più necessaria.

Agente digitale che predispone una pratica operando su più applicativi: è la configurazione più recente e più delicata, poiché il sistema non produce un contenuto da valutare ma compie una sequenza di azioni. Il presidio si sposta dal risultato al perimetro: quali applicativi l'agente può utilizzare, con quali profili autorizzativi, entro quali soglie di importo, con quale registrazione delle operazioni compiute e con quale possibilità di sospensione. La logica applicabile è quella già nota alla banca in materia di deleghe operative e di separazione dei compiti, estesa a un soggetto non umano; l'elemento nuovo è la necessità di una registrazione degli eventi sufficientemente granulare da consentire, ex post, la ricostruzione della catena di azioni e l'attribuzione della responsabilità.

2.7.5 I limiti del presidio umano e le condizioni della sua efficacia

La letteratura sui fattori umani nell'automazione mostra che la semplice presenza di una persona nel processo non garantisce un controllo effettivo. Parasuraman e Manzey (2010), in una rassegna sistematica, documentano due fenomeni distinti e complementari: la compiacenza verso l'automazione, che porta a ridurre l'attenzione dedicata al monitoraggio quando il sistema si è dimostrato affidabile, e la distorsione dell'automazione, che induce ad accettare acriticamente l'indicazione del sistema (errori di commissione) o a non rilevare ciò che il sistema non segnala (errori di omissione). Skitka, Mosier e Burdick (1999) mostrano sperimentalmente come tali errori aumentino in condizioni di carico cognitivo elevato e come la presenza di un secondo operatore non sia di per sé sufficiente a eliminarli, poiché entrambi tendono a convergere sulla stessa fonte automatizzata: è la ragione per cui, come osservato al § 2.7.3, il doppio controllo perde efficacia se non è indipendente.

Un secondo ordine di critiche riguarda il rischio che la sorveglianza umana si riduca a un espediente di allocazione della responsabilità. Wagner (2019) descrive la condizione

dell'operatore chiamato a rispondere di decisioni sulle quali non esercita un controllo effettivo, mentre Green e Chen (2019), analizzando l'uso di strumenti di valutazione del rischio a supporto di decisioni umane, rilevano che i supervisori non solo non correggono in modo affidabile gli errori del modello, ma tendono ad adottarne le indicazioni in modo selettivo, con esiti che non migliorano rispetto a quelli conseguibili dal solo modello. Elish (2019) ha coniato per questa configurazione l'espressione moral crumple zone, per indicare la posizione dell'operatore che assorbe la responsabilità delle disfunzioni di un sistema complesso senza avere avuto la possibilità di prevenirle. Sono osservazioni che invitano a valutare i dispositivi di supervisione non per la loro esistenza formale, ma per la loro capacità di modificare effettivamente gli esiti.

Ne discendono alcune condizioni organizzative, che il terzo capitolo riprenderà in chiave di progettazione dei modelli. La prima è la competenza: chi sorveglia deve comprendere che cosa il sistema misura, su quali dati è stato addestrato e in quali situazioni tende a sbagliare, il che rimanda direttamente al tema dell'AI literacy e all'obbligo dell'articolo 4 dell'AI Act, applicabile dal febbraio 2025. La seconda è l'autorità: la facoltà di discostarsi dall'indicazione del sistema deve essere formalmente riconosciuta e non deve esporre chi la esercita a un onere motivazionale sproporzionato rispetto a quello richiesto per conformarsi. La terza è il tempo: un presidio realistico presuppone che il dimensionamento delle attività tenga conto del tempo necessario alla verifica, e non lo consideri un residuo. La quarta è la tracciabilità: la registrazione degli scostamenti e delle motivazioni è al tempo stesso una garanzia per il cliente, uno strumento di miglioramento del modello e la base della prova in caso di contestazione.

Il tema si collega, in conclusione, al paradigma della complementarità fra automazione e potenziamento del lavoro umano. Come osservano Raisch e Krakowski (2021), un impiego dell'intelligenza artificiale orientato unicamente all'automazione, che escluda progressivamente il contributo umano, rischia di erodere nel tempo le competenze dell'organizzazione e la sua stessa capacità di controllo: paradossalmente, più la supervisione è formale, più si degradano proprio le competenze che la renderebbero possibile. Il presidio umano non è quindi un adempimento normativo aggiuntivo, ma un elemento costitutivo di un'adozione

sostenibile dell'intelligenza artificiale, nella quale la tecnologia potenzia il giudizio umano senza sostituirlo. È su questa premessa che si innesta l'analisi delle competenze richieste dai modelli AI-driven, oggetto del capitolo conclusivo.

2.8 Sfide, rischi e fattori critici di adozione

L'analisi condotta finora ha evidenziato l'ampiezza delle opportunità offerte dall'intelligenza artificiale al settore bancario. Una trattazione equilibrata del fenomeno richiede, tuttavia, di dedicare attenzione anche ai fattori critici che ne condizionano l'adozione. Come rilevato dalle fonti istituzionali e dalla letteratura, il divario tra le potenzialità della tecnologia e i benefici effettivamente conseguiti dipende in larga misura dalla capacità degli intermediari di presidiare una serie di sfide di natura tecnica, organizzativa e culturale.

2.8.1 La qualità e la governance dei dati

Il primo fattore critico riguarda i dati, base di ogni sistema intelligente. L'efficacia dei modelli dipende dalla disponibilità di dati completi, accurati, aggiornati e ben organizzati. Molti intermediari, in particolare quelli che operano su sistemi informativi consolidati nel tempo, si trovano a gestire un patrimonio informativo frammentato, distribuito su archivi eterogenei e non sempre facilmente integrabile. La costruzione di un'adeguata governance del dato – che ne presidi la qualità, la tracciabilità, la sicurezza e la conformità alle norme sulla protezione dei dati personali – costituisce, quindi, una precondizione imprescindibile e, non di rado, il principale ostacolo all'avvio dei progetti.

2.8.2 La trasparenza e la spiegabilità dei modelli

Un secondo profilo riguarda l'interpretabilità dei modelli. I sistemi più avanzati, in particolare quelli basati su reti neurali profonde, operano spesso come "scatole nere" (black box), fornendo risultati accurati ma difficilmente riconducibili a una spiegazione chiara. In un settore nel quale le decisioni devono poter essere motivate nei confronti del cliente e delle autorità, l'explainability non è un requisito accessorio, ma una condizione di legittimità (Arrieta et al., 2020). Da qui l'importanza crescente delle tecniche di explainable AI e la necessità di bilanciare la capacità predittiva dei modelli con la loro trasparenza.

2.8.3 I rischi di distorsione e discriminazione

Strettamente connesso è il rischio di bias. Poiché i modelli apprendono dai dati storici, possono ereditare e amplificare distorsioni preesistenti, con il rischio di produrre esiti discriminatori, ad esempio nell'accesso al credito (Mehrabi et al., 2021). Le evidenze richiamate al § 2.4.1 mostrano peraltro come effetti distributivi indesiderati possano prodursi anche in assenza di distorsioni nei dati, come conseguenza della maggiore flessibilità funzionale dei modelli (Fuster et al., 2022). Il presidio di questo rischio richiede un'attenta selezione e verifica dei dati di addestramento, l'adozione di metriche di equità e un monitoraggio continuo dei risultati, in coerenza con le finalità di tutela dei diritti fondamentali poste alla base dell'AI Act.

2.8.4 Il rischio di modello e il rischio di terze parti

L'impiego di modelli complessi introduce, inoltre, uno specifico rischio di modello (model risk), connesso alla possibilità che il sistema produca risultati errati a causa di errori di progettazione, di degrado delle prestazioni nel tempo o di un utilizzo improprio (Board of Governors of the Federal Reserve System, 2011). A ciò si aggiunge, soprattutto per gli intermediari che ricorrono a soluzioni esterne, un rilevante rischio di terze parti: la dipendenza da fornitori tecnologici per componenti critiche solleva questioni di continuità operativa, di controllo e di localizzazione dei dati, tanto più delicate quanto maggiore è la concentrazione del mercato dell'offerta. Il profilo è oggi presidiato in via diretta dagli orientamenti dell'Autorità bancaria europea in materia di esternalizzazione (European Banking Authority, 2019) e dal Regolamento (UE) 2022/2554 sulla resilienza operativa digitale, che introduce un regime di sorveglianza sui fornitori terzi critici di servizi informatici.

2.8.5 I sistemi legacy e il divario di competenze

Sul piano infrastrutturale, la presenza di sistemi informativi stratificati (legacy) può rendere onerosa l'integrazione delle nuove soluzioni nei processi esistenti; non a caso la necessità di fare fronte all'obsolescenza delle infrastrutture informatiche è indicata dalla Banca d'Italia (2025) fra i fattori che orientano le strategie digitali degli intermediari, e la modernizzazione dei sistemi di core banking figura fra le aree di intervento rilevate da ABI Lab

(2026). Sul piano del capitale umano, la scarsità di competenze specialistiche – dai data scientist agli esperti di governance dell'IA – rappresenta uno dei principali fattori limitanti, comune sia ai grandi gruppi sia, in misura più accentuata, agli intermediari di minori dimensioni. Questo profilo, di natura organizzativa e formativa, rimanda direttamente al cuore della domanda di ricerca del presente elaborato e sarà oggetto di specifico approfondimento nei capitoli conclusivi.

L'insieme di questi fattori critici conferma una conclusione ricorrente: l'adozione dell'IA non è un processo neutro né automatico, ma richiede investimenti mirati, presidi robusti e, soprattutto, un'evoluzione dei modelli organizzativi e delle competenze a seconda del contesto preso in analisi. Le sfide descritte non costituiscono, in ultima analisi, un argomento contro l'innovazione, ma la misura della sua complessità e la ragione per cui il governo del cambiamento assume un rilievo decisivo.

2.9 Due tipi ideali a confronto: la banca digital native e la banca commerciale tradizionale

2.9.1 Premessa metodologica: perché tipi ideali e non casi aziendali

Il confronto proposto in questo paragrafo è costruito fra due tipi ideali e non fra due intermediari determinati. La scelta merita di essere esplicitata, perché condiziona il valore delle affermazioni che seguono.

La nozione di tipo ideale è quella elaborata da Max Weber nel saggio metodologico del 1904 e ripresa in Economia e società: un costrutto concettuale ottenuto accentuando unilateralmente uno o più punti di vista e componendo una molteplicità di fenomeni singoli, diffusi e discreti in un quadro concettuale unitario e coerente. Weber precisa che tale quadro, nella sua purezza concettuale, non è mai empiricamente riscontrabile nella realtà: esso non descrive la realtà, ma fornisce alla descrizione della realtà mezzi espressivi univoci e un termine di paragone rispetto al quale misurare la distanza dei casi concreti (Weber, 1904; 1922). Il tipo ideale non è quindi un'ipotesi da verificare né una media empirica: è uno strumento euristico e la sua utilità si misura sulla capacità di rendere intelligibili le differenze osservate.

L'alternativa sarebbe stata costruire un confronto fra due banche reali e identificate e tale scelta è stata scartata per tre ragioni. La prima è di ordine metodologico: un confronto fra due soli casi, per quanto documentato, non consente generalizzazioni e rischia di far dipendere le conclusioni da circostanze idiosincratiche, mentre l'obiettivo del capitolo è isolare le condizioni strutturali che influenzano l'adozione della tecnologia. La seconda è di ordine documentale: le informazioni pubblicamente verificabili sulle scelte di adozione dell'IA dei singoli intermediari sono in larga parte di fonte aziendale e a finalità comunicativa, quindi non comparabili in modo simmetrico. La terza discende dalla domanda di ricerca, che riguarda la relazione fra assetto organizzativo e capacità di valorizzare la tecnologia, e non la valutazione della performance di operatori determinati.

Di conseguenza, ogni attributo dei due modelli è di seguito ricondotto a una fonte documentale, e la trattazione si chiude con l'esplicitazione dei limiti della tipologia adottata. Si sottolinea inoltre che i due tipi non esauriscono lo spazio delle configurazioni possibili: fra di essi si collocano i grandi gruppi bancari nazionali ed europei, che presentano tratti di entrambi, e forme ulteriori quali le banche promosse da operatori non finanziari.

2.9.2 Il tipo ideale della banca digital native

La banca digital native ha come caratteristica principale l'essere nata in un contesto interamente digitale, privo di infrastrutture tecnologiche stratificate, e l'essere costruita attorno a una piattaforma proprietaria e a un modello di servizio incentrato sull'applicazione mobile. La letteratura sulla disruption digitale nel settore bancario ne ha analizzato la logica economica: Vives (2019) e Thakor (2020) individuano nei minori costi di struttura, nell'assenza di rete fisica e nella scalabilità della piattaforma i tratti distintivi del modello, mentre Boot, Hoffmann, Laeven e Ratnovski (2021) sottolineano come l'innovazione tecnologica agisca sulla struttura stessa dell'intermediazione, riducendo il valore dei costi di ricerca e di transazione su cui si fondava una parte della rendita bancaria tradizionale.

Sul piano della tecnologia informativa, il modello si fonda in via prevalente sulla hard information: dati transazionali, comportamentali e digitali, raccolti in forma strutturata fin dall'origine e immediatamente utilizzabili per l'addestramento di modelli statistici. È la

condizione che Berg, Burg, Gombović e Puri (2020) descrivono documentando il potere predittivo dell'impronta digitale, e che Berg, Fuster e Puri (2022) collocano al centro della loro rassegna sul credito fintech: la sostituzione della relazione con il dato come fonte primaria di informazione sul cliente. Buchak, Matvos, Piskorski e Seru (2018) mostrano peraltro come la crescita di questi operatori sia stata favorita anche da fattori regolamentari, oltre che tecnologici.

Ne discendono, quanto all'adozione dell'intelligenza artificiale, alcuni vantaggi strutturali: assenza di sistemi legacy e conseguente possibilità di integrare l'IA nativamente nei processi, secondo la logica AI-native descritta nel primo capitolo; disponibilità di dati digitali omogenei e di elevata qualità; cultura organizzativa orientata alla sperimentazione. Gli ambiti applicativi tipici sono l'individuazione delle frodi, la personalizzazione dell'offerta e l'automazione dell'assistenza alla clientela.

Il modello presenta però limiti speculari ai propri punti di forza. Petralia, Philippon, Rice e Véron (2019) osservano come i nuovi operatori abbiano finora inciso soprattutto sui segmenti a minore intensità di capitale e di relazione, incontrando maggiori difficoltà nel credito alle imprese, dove la valutazione richiede informazioni non codificabili. Frost, Gambacorta, Huang, Shin e Zbinden (2019) segnalano, sul piano sistemico, le questioni poste dalla concentrazione dei dati e dal rapporto fra scala tecnologica e perimetro regolamentare. Il Financial Stability Board (2017) richiama infine il rischio che modelli fondati su tecnologie analoghe generino comportamenti correlati.

2.9.3 Il tipo ideale della banca commerciale a radicamento territoriale

Il primo tratto è di natura istituzionale e storica. Una quota rilevante delle banche commerciali italiane a base territoriale discende dalle casse di risparmio e dalle banche popolari, sorte fra Otto e Novecento con finalità mutualistiche o di previdenza e cresciute per progressiva estensione della rete di sportelli. Il passaggio decisivo per il comparto delle casse di risparmio si colloca nel 1990, quando, in attuazione della cosiddetta legge Amato, l'attività bancaria è stata conferita a società per azioni secondo il modello di separazione tra ente conferente – la fondazione di origine bancaria – e impresa bancaria (legge 30 luglio 1990, n.

218; d.lgs. 20 novembre 1990, n. 356). Negli anni successivi molti di questi intermediari hanno intrapreso percorsi di crescita per linee esterne, dando vita a gruppi di dimensioni intermedie che hanno però conservato la connotazione territoriale originaria.

Il secondo tratto riguarda la tecnologia di produzione dell'informazione creditizia. La letteratura sul relationship banking descrive un modello di intermediazione fondato sulla raccolta ripetuta e riservata di informazioni sul cliente attraverso la relazione di lungo periodo, informazioni in larga parte qualitative e difficilmente codificabili, definite soft information in contrapposizione alla hard information di natura quantitativa e trasferibile (Boot, 2000). Stein (2002) mostra come la produzione di soft information sia favorita da strutture organizzative decentrate e da gerarchie corte, nelle quali chi raccoglie l'informazione dispone anche di autonomia decisionale; Berger e Udell (2002) collegano esplicitamente la struttura organizzativa della banca alla sua capacità di finanziare le imprese minori, e Alessandrini, Presbitero e Zazzaro (2009) documentano per l'Italia come l'aumento della distanza funzionale fra il luogo della decisione e il luogo della relazione si associ a maggiori vincoli finanziari per le imprese affidate. Il tipo ideale qui costruito assume, quindi, come tratto caratterizzante non la "prossimità" intesa in senso retorico, ma una specifica tecnologia informativa e la struttura organizzativa che la sostiene.

Il terzo tratto è di ordine tecnologico e dimensionale. Secondo l'Indagine Fintech della Banca d'Italia (2025), gli investimenti in tecnologie innovative realizzati nel biennio 2023-2024 da banche e intermediari non bancari sono stati pari a circa 1 miliardo di euro, con una stima di importo analogo per il biennio successivo; la stessa fonte precisa che la spesa risulta molto concentrata fra pochi intermediari e che le banche costituiscono i principali investitori in tecnologie innovative. Ne discende, come tratto tipico e non come giudizio di merito, che gli intermediari di dimensione intermedia dispongono di una capacità di investimento autonomo limitata e tendono ad accedere alla tecnologia attraverso fornitori esterni, piattaforme standardizzate e iniziative di sistema promosse dagli organismi associativi di categoria. La medesima indagine segnala inoltre, fra i fattori che orientano le strategie digitali, la necessità di fare fronte a fenomeni di obsolescenza delle infrastrutture informatiche.

Il quarto tratto riguarda il rapporto fra canale fisico e canale digitale. Sempre secondo la Banca d'Italia (2025), nel 2024 i clienti che hanno operato in prevalenza attraverso canali digitali sono stati 43,6 milioni, pari al 58,2 per cento della clientela complessiva, mentre i depositi detenuti in conti aperti digitalmente ammontavano a 73,7 miliardi di euro, pari al 5,1 per cento del totale dei depositi a fine 2024; il rapporto fra prestiti digitali e prestiti totali risultava pari al 10,6 per cento per le famiglie consumatrici e all'1,2 per cento per le imprese. Il dato descrive con precisione la fisionomia del tipo ideale: canali digitali ormai diffusi come strumento di operatività corrente, ma raccolta, erogazione e relazione ancora largamente ancorate al canale tradizionale.

Composti insieme, questi quattro tratti definiscono un modello nel quale il vantaggio informativo deriva dalla relazione e dalla conoscenza del tessuto economico locale, mentre il vincolo principale è di scala: capacità di investimento, disponibilità di competenze specialistiche e flessibilità delle infrastrutture. È precisamente questa combinazione – patrimonio relazionale elevato, dotazione tecnologica mediata da terzi – a rendere il modello interessante per la domanda di ricerca dell'elaborato, poiché vi si osserva in forma nitida la distanza fra disponibilità della tecnologia e capacità organizzativa di tradurla in valore.

2.9.4 Il confronto lungo cinque dimensioni

Il confronto è organizzato lungo cinque dimensioni, scelte perché rilevanti rispetto alla domanda di ricerca: la tecnologia informativa prevalente, l'architettura tecnologica, la modalità di accesso all'innovazione, la struttura organizzativa e il rapporto con la clientela. Per ciascuna dimensione la tabella seguente indica la configurazione tipica dei due modelli e la fonte da cui l'attributo è desunto.

Dimensione	Banca digital native (tipo ideale)	Banca commerciale a radicamento territoriale (tipo ideale)	Fonti
Tecnologia informativa prevalente	Hard information: dati transazionali, comportamentali e digitali, strutturati fin dall'origine	Soft information prodotta nella relazione di lungo periodo, integrata da hard information	Boot (2000); Stein (2002); Berger e Udell (2002); Berg, Burg, Gombović e Puri (2020); Berg, Fuster e Puri (2022)
Architettura tecnologica	Piattaforma unica, costruita senza vincoli di stratificazione; integrazione nativa dei modelli	Sistemi informativi consolidati nel tempo, affidabili ma con vincoli di flessibilità e integrazione	Vives (2019); Thakor (2020); Banca d'Italia (2025); ABI Lab (2026)
Accesso all'innovazione	Sviluppo interno, controllo diretto della piattaforma e dei dati	Fornitori esterni, piattaforme standardizzate, iniziative di sistema promosse dagli organismi di categoria	Banca d'Italia (2025): investimenti pari a circa 1 miliardo nel biennio 2023-2024, spesa molto concentrata fra pochi intermediari

Dimensione	Banca digital native (tipo ideale)	Banca commerciale a radicamento territoriale (tipo ideale)	Fonti
Struttura organizzativa	Gerarchia corta, decisione centralizzata sulla piattaforma, orientamento alla sperimentazione	Rete distributiva articolata, autonomia decisionale periferica funzionale alla produzione di soft information	Stein (2002); Berger e Udell (2002); Alessandrini, Presbitero e Zazzaro (2009)
Rapporto con la clientela	Mediato dai canali digitali; ampiezza della base di clientela, minore profondità della relazione	Presenza fisica e continuità della relazione; raccolta ed erogazione ancora prevalentemente ancorate al canale tradizionale	Petralia, Philippon, Rice e Véron (2019); Banca d'Italia (2025): 58,2% di clienti digitali a fronte del 5,1% di depositi in conti aperti digitalmente

Tabella 2.1 – Confronto fra i due tipi ideali lungo cinque dimensioni. Elaborazione propria sulle fonti indicate.

La lettura congiunta delle cinque dimensioni consente una prima considerazione. I due modelli non differiscono soltanto per dotazione tecnologica, ma per la natura dell'informazione su cui fondano l'intermediazione, e questa differenza si riflette sulla struttura organizzativa. Il modello digital native dispone di un vantaggio nell'applicazione di tecniche statistiche perché la sua materia prima è già codificata; il modello territoriale dispone di un

patrimonio informativo che le tecniche di elaborazione del linguaggio naturale rendono oggi, per la prima volta, parzialmente trattabile, ma che richiede un investimento preliminare in governance del dato. La discriminante non è quindi la quantità di tecnologia acquistata, ma la coerenza fra tecnologia adottata e tecnologia informativa propria del modello di business.

Una seconda considerazione riguarda la modalità di accesso all'innovazione. Il ricorso prevalente a fornitori esterni, tipico del secondo modello, consente di contenere investimenti e rischi, ma comporta una minore capacità di personalizzazione, una maggiore dipendenza da soggetti terzi e uno spostamento del baricentro delle competenze: le capacità critiche diventano la selezione dei fornitori, la definizione dei requisiti, la verifica degli esiti e il presidio contrattuale, più che lo sviluppo dei modelli. È un punto che sarà ripreso nel capitolo conclusivo, poiché ridefinisce il contenuto delle competenze richieste e non soltanto la loro intensità.

2.9.5 I limiti della tipologia

La costruzione proposta presenta limiti che è corretto esplicitare. In primo luogo, come ricordato, nessun intermediario reale coincide con uno dei due tipi: i grandi gruppi bancari europei, che rappresentano la quota prevalente degli investimenti in intelligenza artificiale, si collocano in posizione intermedia, disponendo sia di risorse per lo sviluppo interno sia di reti fisiche e di sistemi stratificati. In secondo luogo, i due tipi sono costruiti su una fotografia recente e sono soggetti a convergenza: gli operatori digitali stanno acquisendo licenze bancarie complete e presidi di conformità sempre più articolati, mentre gli operatori tradizionali stanno modernizzando le proprie infrastrutture, come mostra la centralità della modernizzazione dei sistemi di core banking fra le priorità rilevate da ABI Lab (2026). In terzo luogo, la tipologia non incorpora la dimensione regolamentare come variabile discriminante, poiché entrambi i modelli, in quanto banche autorizzate, sono soggetti al medesimo quadro prudenziale.

Con questi limiti, il confronto conserva la funzione che gli è propria: non stabilire quale modello sia superiore, ma rendere visibile che la capacità di tradurre l'intelligenza artificiale in valore dipende dalla coerenza fra tecnologia, assetto organizzativo e competenze delle

persone. La tecnologia, di per sé, non garantisce alcun vantaggio. È questa considerazione a introdurre il capitolo successivo, dedicato ai modelli organizzativi funzionali all'adozione dell'IA.

2.10 Considerazioni conclusive

L'analisi sviluppata nel presente capitolo ha consentito di ricostruire il quadro dell'adozione dell'intelligenza artificiale nel settore bancario, collocandola all'interno del più ampio percorso storico di digitalizzazione del comparto ed evidenziandone il livello di diffusione, i principali casi d'uso, il ruolo crescente dell'IA generativa e agentica, la cornice regolamentare di riferimento, la centralità del presidio umano, i fattori critici di adozione e le differenti modalità di implementazione in funzione delle caratteristiche degli intermediari.

L'excursus storico ha mostrato come l'attuale transizione non costituisca una rottura improvvisa, ma l'ultima e più matura tappa di un lungo processo di adattamento tecnologico e organizzativo. L'esame dei singoli casi d'uso, condotto ponendo a confronto la modalità tradizionale di gestione del processo, l'apporto atteso dall'intelligenza artificiale e le evidenze empiriche disponibili, ha reso evidente come il valore della tecnologia non risieda nella semplice automazione delle attività e come i benefici documentati siano condizionati dal contesto di impiego, dalla qualità dei dati e dalle competenze di chi la utilizza.

Sul piano regolamentare, l'aggiornamento del quadro europeo intervenuto nel luglio 2026 ha spostato in avanti i termini di adeguamento per i sistemi ad alto rischio, senza modificarne il contenuto: per le banche il differimento amplia il tempo disponibile, non riduce il lavoro necessario. In parallelo, l'approfondimento dedicato alla sorveglianza umana ha mostrato come il presidio della persona sui sistemi automatizzati non possa essere ridotto a un adempimento formale: la sua efficacia dipende da competenza, autorità, tempo e tracciabilità, ossia da variabili organizzative. È in questo passaggio che il tema regolamentare si converte in tema di progettazione dei modelli organizzativi e di sviluppo delle competenze.

Il confronto fra i due tipi ideali ha infine reso evidente come non esista un'unica via all'adozione dell'IA, ma una pluralità di percorsi, ciascuno coerente con la tecnologia

informativa e con l'assetto organizzativo dell'intermediario. La reale discriminante, si è osservato, non risiede nella dotazione tecnologica, ma nella capacità di governare il cambiamento sul piano organizzativo e delle competenze. Su queste premesse si innesta il capitolo successivo, che approfondirà i modelli organizzativi attraverso i quali le banche possono governare in modo efficace l'integrazione dell'intelligenza artificiale nei propri processi, ponendo le basi per l'analisi, condotta nel capitolo conclusivo, del ruolo dell'uomo e dell'evoluzione delle competenze.

Capitolo III – I modelli organizzativi per l'adozione dell'intelligenza artificiale

3.1 Introduzione

I due capitoli precedenti hanno stabilito una premessa e ne hanno verificato la portata in un contesto applicativo. Il primo ha ricostruito l'evoluzione dell'intelligenza artificiale, mostrando come il suo valore non risieda nelle prestazioni degli algoritmi ma nella capacità delle organizzazioni di integrarli nei propri processi. Il secondo ha documentato, per il settore bancario, che l'adozione è ormai diffusa ma che i benefici non sono automatici: dipendono dalla qualità dei dati, dal contesto di impiego, dall'effettività del presidio umano e dalle competenze di chi utilizza la tecnologia. In entrambi i capitoli la conclusione è la stessa e resta: la tecnologia da sola non basta.

Il presente capitolo costituisce il nucleo analitico dell'elaborato: se la dotazione tecnologica non è sufficiente, occorre stabilire quali elementi la rendano produttiva e, poiché la risposta rimanda all'organizzazione, è necessario esaminare quali configurazioni organizzative rendano possibile trasformare l'IA in valore, quali siano i loro vantaggi e i loro limiti, e a quali condizioni ciascuna funzioni. È la parte della domanda di ricerca che racchiude in che modo l'adozione dell'IA richieda una trasformazione dei modelli organizzativi.

Si procede in cinque passaggi: l'analisi inizia collocando l'intelligenza artificiale nel più ampio fenomeno della trasformazione digitale, per stabilire che cosa l'IA abbia di specifico rispetto alle precedenti ondate tecnologiche; si prosegue con l'inquadramento teorico del problema, mostrando che l'adozione dell'IA riapre i problemi fondamentali della progettazione organizzativa: la divisione del lavoro, l'allocazione dei compiti, il coordinamento e il controllo. Il § 3.4 esamina le precondizioni dell'adozione, ossia i fattori di prontezza organizzativa e le capacità che l'impresa deve possedere prima ancora di scegliere quale modello adottare. Il § 3.5 valuta in modo sistematico otto configurazioni organizzative, raggruppate in due categorie: i modelli di allocazione della competenza – centralizzato, decentrato, *hub and spoke* ed esternalizzato – e i modelli di aggregazione del lavoro – a piattaforma e prodotto, predittivo o *decision-centric*, agentico-ibrido ed ecosistemico – dei quali si propone il confronto. La parte finale del capitolo si sviluppa attraverso l'esame degli assetti di

governance e controllo, delle condizioni che ne influenzano l'efficacia operativa, della loro declinazione nel settore bancario e delle competenze richieste all'operatore per assicurarne una corretta applicazione nel tempo.

3.2 Dalla trasformazione digitale alla trasformazione AI-driven

3.2.1 Il quadro concettuale della trasformazione digitale

L'adozione dell'intelligenza artificiale non nasce in un vuoto teorico: si inserisce su un filone di studi che ha analizzato, nell'ultimo decennio, gli effetti organizzativi delle tecnologie digitali. Il contributo di riferimento è la rassegna di Vial (2019), che attraverso l'esame di 282 lavori costruisce un quadro concettuale della trasformazione digitale articolato in otto blocchi, nei quali viene rappresentata la trasformazione digitale come un processo in cui le tecnologie digitali producono discontinuità che innescano risposte strategiche da parte delle organizzazioni, le quali cercano di modificare i propri percorsi di creazione del valore, in quanto devono gestire anche i cambiamenti strutturali e le barriere organizzative che condizionano gli esiti, positivi e negativi, del processo.

Tre elementi di questo quadro sono rilevanti. Il primo è che la trasformazione digitale è definita come processo e non come stato: non esiste un punto di arrivo, ma una successione di fasi; il secondo è che le tecnologie non producono effetti diretti sulle prestazioni, ma agiscono attraverso la mediazione di cambiamenti strutturali – nella struttura organizzativa, nei processi, nella cultura e nei ruoli – e possono essere ostacolate da barriere organizzative quali l'inerzia e la resistenza. Il terzo e ultimo elemento è che gli esiti non sono univocamente positivi: il quadro contempla espressamente conseguenze negative, il che offre un antidoto alla lettura celebrativa che caratterizza parte della pubblicistica sull'IA. Su queste basi Vial (2019) propone, tra le direttrici di ricerca, l'approfondimento del ruolo delle capacità dinamiche e delle questioni etiche.

Il riferimento alle capacità dinamiche collega la trasformazione digitale a un filone consolidato degli studi strategici. Secondo Teece (2007) i fondamenti microeconomici legati alle performance in ambienti caratterizzati da rapido mutamento tecnologico sono la capacità

di percepire le opportunità (sensing), di coglierle attraverso scelte di investimento e di modello di business (seizing) e di progettare la base di riconfigurazione delle risorse all'interno dell'impresa (reconfiguring). Warner e Wäger (2019) hanno declinato questo schema al caso della trasformazione digitale, descrivendola come un processo continuo di rinnovamento strategico nel quale l'impresa costruisce e ricostruisce le proprie capacità dinamiche a seconda delle necessità e degli obiettivi prefissati. È una prospettiva che sarà utile nel prosieguo, perché suggerisce di valutare i modelli organizzativi non solo per la loro efficienza in un dato momento, ma per la loro capacità di evolvere nel tempo.

3.2.2 Che cosa l'intelligenza artificiale ha di specifico

Inserire l'IA nell'onda della trasformazione digitale è corretto ma insufficiente: occorre stabilire quali problemi organizzativi specifici essa ponga. La risposta più netta e più utile ai fini di questo capitolo è quella di Berente, Gu, Recker e Santhanam (2021), i quali sostengono che gestire l'intelligenza artificiale non equivalga a gestire la tecnologia dell'informazione come si è fatto in passato. L'IA, osservano gli autori, non è una tecnologia o un insieme di tecnologie, ma una frontiera in continua evoluzione di capacità computazionali emergenti; governarla significa assumere decisioni su tre aspetti interdipendenti – l'autonomia, l'apprendimento e l'inscrutabilità – nella ricerca continua di ampliare le prestazioni e l'ambito di applicazione dei sistemi.

Le tre dimensioni meritano di essere lette come altrettanti problemi organizzativi. L'autonomia riguarda il grado in cui il sistema opera senza intervento umano e pone quindi il problema di come ripartire i diritti decisionali fra persone e sistemi: è la questione affrontata, sul piano del presidio, al § 2.7 e ripresa qui sul piano della progettazione. L'apprendimento riguarda il fatto che il comportamento del sistema muta nel tempo in funzione dei dati: ne consegue che un modello non è un artefatto stabile da rilasciare una volta, ma richiede manutenzione continua, monitoraggio delle prestazioni e aggiornamento periodico; la conseguenza organizzativa è che la responsabilità sul sistema non può esaurirsi con la messa in esercizio. L'inscrutabilità riguarda la difficoltà di ricostruire le ragioni di un esito e incide direttamente sui meccanismi di controllo: quando l'output non è pienamente spiegabile, la

verifica non può basarsi sull'ispezione del processo ma deve spostare l'attenzione sui risultati, sui dati di addestramento e sulle procedure di validazione.

A queste tre dimensioni se ne aggiunge una quarta, di natura economica, già richiamata al § 2.4: l'intervallo che separa l'investimento tecnologico dalla manifestazione dei benefici. Brynjolfsson, Rock e Syverson (2017) attribuiscono tale intervallo alla necessità di accumulare capitale in larga parte intangibile – nuovi processi, nuove competenze, nuovi assetti organizzativi – la cui accumulazione richiede tempo e comporta costi che precedono i ricavi. L'implicazione per questo capitolo è diretta: il modello organizzativo non è un accessorio dell'investimento tecnologico, ma è esso stesso l'investimento complementare che ne determina la redditività.

3.3 Il problema organizzativo posto dall'intelligenza artificiale

3.3.1 I problemi fondamentali della progettazione organizzativa

Per analizzare in modo ordinato i modelli organizzativi conviene individuare alcune categorie generali, che consentano di confrontarli lungo le medesime dimensioni anziché descriverli uno per uno. Puranam, Alexy e Reitzig (2014), interrogandosi su che cosa vi sia di realmente nuovo nelle forme organizzative emergenti, riconducono ogni organizzazione alla soluzione di quattro problemi universali: la divisione del lavoro in compiti, l'allocazione dei compiti agli attori, la predisposizione degli incentivi e la circolazione delle informazioni necessarie al coordinamento. Qualunque configurazione, tradizionale o innovativa, si distingue per il modo in cui risolve questi quattro problemi.

La divisione del lavoro viene rimessa in discussione perché i compiti possono essere scomposti diversamente quando una parte del contenuto cognitivo è eseguibile da un sistema. L'allocazione cambia perché fra gli attori compare un soggetto non umano, al quale possono essere assegnati compiti ma non responsabilità. Gli incentivi sono toccati perché il contributo individuale diventa più difficile da isolare quando la prestazione è congiunta, poiché un sistema di misurazione mal impostato può indurre la supervisione formale descritta al § 2.7.5. Il coordinamento incide in modo netto: le informazioni necessarie a coordinare l'azione non

circolano più soltanto fra persone, ma passano attraverso dati, modelli e interfacce, con la conseguenza che la qualità dell'informazione diventa una variabile di progettazione organizzativa e non solo un tema tecnico.

A questa quadripartizione si può affiancare la prospettiva classica del trattamento delle informazioni, secondo la quale la struttura organizzativa è funzione della quantità di informazione che l'organizzazione deve elaborare per fronteggiare l'incertezza dei propri compiti (Galbraith, 1974). La prospettiva è utile perché rende immediatamente intelligibile una tensione che percorrerà l'intero paragrafo dedicato ai modelli: le tecnologie basate sull'IA aumentano la capacità di elaborazione informativa dell'organizzazione, ma la loro introduzione genera nuove esigenze di coordinamento e collaborazione – fra chi possiede il dominio applicativo, chi possiede la competenza tecnica e chi presidia il rischio – che devono essere risolte a livello di struttura.

3.3.2 Automazione e potenziamento come scelta di progettazione

Il secondo elemento da chiarire riguarda la natura della scelta che l'organizzazione compie quando introduce l'intelligenza artificiale in un processo. Raisch e Krakowski (2021) mostrano che automazione e potenziamento non costituiscono due opzioni alternative fra le quali scegliere una volta per tutte, ma gli estremi di un paradosso che deve essere gestito nel tempo. La loro tesi è che l'automazione e il potenziamento siano intrinsecamente collegati: perseguire l'uno senza l'altro genera conseguenze indesiderate, poiché un'organizzazione che automatizzi in via esclusiva erode progressivamente le competenze umane che sono necessarie per governare la tecnologia, mentre un potenziamento privo di automazione rinuncia ai guadagni di efficienza che rendono sostenibile l'investimento.

La rilevanza di questa impostazione è duplice: in primo luogo, qualifica la scelta fra automazione e potenziamento come una decisione di progettazione organizzativa, non come una conseguenza tecnica: due imprese che adottano la medesima tecnologia possono ripartire diversamente i compiti fra persone e sistemi, e tale ripartizione va quindi stabilita in funzione del modello organizzativo prescelto. In secondo luogo, la natura paradossale della relazione

implica che la scelta non sia definitiva ma vada rivista periodicamente, il che richiede assetti capaci di apprendere: torna qui la prospettiva delle capacità dinamiche richiamata al § 3.2.1.

Sul piano della configurazione operativa del lavoro, lo studio di Grønsund e Aanestad (2020) offre l'evidenza empirica più pertinente: attraverso uno studio condotto in un'organizzazione del commercio marittimo internazionale che aveva introdotto un supporto algoritmico automatizzato per attività di analisi e previsione, gli autori osservano che le configurazioni fra persone e algoritmo si sono ridefinite più volte nel tempo. Anziché sostituire il lavoro umano, le configurazioni emerse hanno richiesto nuovi ruoli e una redistribuzione delle competenze esistenti al fine di potenziare e migliorare l'accuratezza dell'algoritmo. Gli autori descrivono l'assetto risultante come uno schema di tipo *human in the loop*, composto sia dal lavoro di verifica – la costruzione di un riferimento di correttezza e la valutazione dell'output algoritmico rispetto a esso – sia dal lavoro di modifica dell'algoritmo e dell'architettura di acquisizione dei dati e ne sottolineano l'importanza strategica, ossia la capacità dell'organizzazione di assicurare che le prestazioni dell'algoritmo continuino a corrispondere ai propri requisiti e agli eventuali mutamenti del contesto.

Il risultato ha una conseguenza che vale la pena rendere esplicita, perché rovescia un'aspettativa diffusa: l'introduzione dell'IA non riduce il lavoro umano necessario, ma ne modifica la natura, generando attività prima inesistenti. Il presidio umano descritto al § 2.7 non è dunque soltanto una garanzia per il cliente e un obbligo normativo; è anche il meccanismo attraverso il quale l'organizzazione mantiene il sistema allineato ai propri obiettivi. Ne segue che un modello organizzativo che non preveda risorse dedicate a queste attività non è un modello più efficiente, ma un modello che non funziona.

3.4 Le precondizioni: prontezza organizzativa e capacità

3.4.1 I fattori di prontezza organizzativa

Prima di scegliere una configurazione, l'organizzazione deve valutare se possieda i presupposti per adottare l'IA. Il contributo di riferimento su questo punto è quello di Jöhnk, Weißert e Wyrtki (2021), che affrontano il tema della prontezza organizzativa legata al

mondo dell'intelligenza artificiale a partire da uno studio basato su interviste con venticinque esperti, i cui risultati sono stati intrecciati con la letteratura scientifica e professionale. Gli autori partono da una constatazione: pur consapevoli della portata dello strumento, le imprese incontrano insidie nell'adozione dell'IA a causa della molteplicità degli ambiti applicativi, della complessità intrinseca della tecnologia, delle nuove esigenze organizzative e delle competenze degli operatori da sviluppare che essa determina. Su questa base essi individuano cinque categorie di fattori di prontezza, corredate di indicatori concretamente rilevabili e ne derivano implicazioni per l'intero processo di adozione.

Le cinque categorie coprono l'allineamento strategico dell'iniziativa agli obiettivi aziendali, la disponibilità di risorse finanziarie, tecnologiche e umane, la dotazione di conoscenza specialistica sull'IA e sui suoi ambiti di applicazione, la cultura organizzativa e infine il patrimonio di dati. Si possono effettuare due considerazioni in merito: la prima è che quattro categorie su cinque non riguardano la tecnologia in senso stretto e ne consegue che la prontezza è un attributo prevalentemente organizzativo mentre la seconda, sottolineata dagli stessi autori, è che la prontezza va valutata in relazione allo specifico scopo di adozione perseguito: un'impresa può essere pronta per un caso d'uso e non per un altro. Ne discende un'indicazione operativa rilevante per il seguito, ossia che la scelta del modello organizzativo dovrebbe seguire e non precedere la valutazione dei presupposti.

L'ordine dei fattori non è indifferente. Le rilevazioni istituzionali richiamate nel secondo capitolo indicano che il vincolo più frequente si colloca a monte sul patrimonio informativo: la costruzione di una governance del dato che ne presidi qualità, tracciabilità e conformità costituisce, come osservato al § 2.8, la precondizione più disattesa e il principale ostacolo all'avvio dei progetti.

3.4.2 Dalla prontezza alla capacità organizzativa

La prontezza descrive la condizione iniziale dell'organizzazione, mentre la capacità identifica l'insieme delle competenze e delle risorse che essa sviluppa e consolida nel tempo. La distinzione tra i due concetti assume particolare rilevanza ai fini dell'analisi, poiché i modelli organizzativi esaminati nei paragrafi successivi si differenziano proprio per le

modalità con cui le capacità vengono accumulate e per la loro collocazione all'interno della struttura aziendale.

Il contributo di Tarafdar, Beath e Ross (2019) esamina l'impiego dell'intelligenza artificiale a supporto delle operazioni aziendali, inteso come installazione di algoritmi nelle applicazioni che sostengono i processi organizzativi e individua cinque capacità e quattro pratiche necessarie per generare valore. Il dato più significativo riguarda il divario fra previsioni e risultati: in un'indagine condotta dagli autori nel 2017 su dirigenti di 106 imprese, metà dei rispondenti dichiarava che la propria azienda non aveva in essere alcuna applicazione di questo tipo e, fra coloro le cui aziende ne disponevano, soltanto la metà riteneva che avessero prodotto risultati aziendali misurabili. Un'evidenza convergente proviene da Fountaine, McCarthy e Saleh (2019), i quali riportano che, sulla base di indagini condotte su un ampio numero di dirigenti, soltanto l'otto per cento delle imprese risultava adottare le pratiche fondamentali che sostengono una diffusione ampia dell'IA, mentre la maggioranza aveva realizzato soltanto sperimentazioni sporadiche o applicazioni circoscritte a un singolo processo. Gli autori attribuiscono la lentezza dei progressi non alla tecnologia ma alla mancata riconfigurazione dell'organizzazione, e individuano nella cultura, più che nella tecnologia, l'ostacolo principale.

Il concetto di capacità organizzativa per l'intelligenza artificiale è stato successivamente formalizzato nella letteratura sui sistemi informativi. Mikalef e Gupta (2021) propongono una concettualizzazione della capacità di intelligenza artificiale come strumento multidimensionale, che integra risorse tangibili quali dati e infrastrutture, risorse umane quali competenze tecniche e manageriali e risorse intangibili quali la cultura orientata al dato e la capacità di apprendimento organizzativo. L'impostazione è utile perché rende esplicito che le componenti sono complementari: la carenza di una di esse riduce il rendimento delle altre, il che spiega perché investimenti tecnologici anche cospicui possano non tradursi in risultati.

3.4.3 La governance del dato come infrastruttura organizzativa

Fra le condizioni preliminari, il patrimonio informativo merita una trattazione autonoma, perché la sua gestione non è una questione tecnica ma un assetto organizzativo a tutti

gli effetti. La governance del dato consiste nella definizione dei diritti decisionali e delle responsabilità relative alla produzione, alla qualità, alla custodia e all'uso dei dati: stabilire chi possa decidere che cosa in materia di dati equivale a risolvere, per una risorsa specifica, i medesimi problemi di allocazione e coordinamento richiamati al § 3.3.1.

Nel contesto bancario questa dimensione ha una particolarità che ne accresce il rilievo: l'aggregazione dei dati di rischio e la relativa reportistica sono già oggetto di principi di vigilanza consolidati, elaborati dal Comitato di Basilea, i quali richiedono agli intermediari assetti di governance, architetture dei dati e capacità di aggregazione adeguate (Basel Committee on Banking Supervision, 2013). Ne consegue che una banca che affronti oggi la governance del dato per finalità di intelligenza artificiale non parte da zero, ma dispone di un impianto al quale ricondurre le nuove esigenze: si tratta, in prospettiva organizzativa, di un vantaggio di percorso e non di un vincolo, purché l'impianto esistente non venga duplicato ma esteso.

3.5 I modelli organizzativi per l'adozione dell'intelligenza artificiale

I modelli esaminati nei paragrafi seguenti si distinguono per il modo in cui rispondono a una domanda di progettazione precisa: dove collocare, all'interno dell'organizzazione, le competenze specialistiche necessarie a sviluppare e governare i sistemi di intelligenza artificiale, e con quali meccanismi collegarle alle strutture che possiedono la conoscenza dei processi e alle funzioni che presidiano i rischi.

Sono stati presi in esame otto modelli di adozione, che possono essere suddivisi in due macrocategorie: la prima comprende i quattro modelli di allocazione della competenza, che rispondono alla domanda su *dove* collocare le capacità specialistiche: centralizzato, decentrato, *hub and spoke* ed esternalizzato. La seconda comprende quattro modelli di aggregazione del lavoro, che rispondono alla domanda su *attorno a che cosa* aggregare le persone: il modello a piattaforma e prodotto, il modello predittivo o *decision-centric*, il modello agentico-ibrido e il modello ecosistemico. Le due famiglie non si pongono in termini di alternatività, bensì di complementarità. Un istituto di credito può adottare una configurazione appartenente alla prima classificazione e, simultaneamente, una riconducibile alla seconda e ne

consegue che il grado di coerenza tra le due scelte rappresenta esso stesso una variabile di progettazione organizzativa.

L'osservazione ha una conseguenza pratica rilevante, ripresa al § 3.5.9: gran parte delle difficoltà osservate nei programmi di adozione non deriva dalla scelta di un modello sbagliato, ma dall'incoerenza fra il luogo in cui la competenza è collocata e il criterio secondo cui il lavoro è aggregato.

Per ciascun modello si esaminano la logica sottostante, i meccanismi di coordinamento su cui si fonda, i vantaggi, i limiti e le condizioni alle quali risulta appropriato. Va precisato fin d'ora che i modelli non sono esclusivi né rappresentano stadi di una progressione necessaria: costituiscono alternative, fra le quali la scelta dipende da variabili di contingenza, secondo l'impostazione classica per cui non esiste una forma organizzativa universalmente superiore ma una coerenza da ricercare fra struttura, tecnologia e ambiente.

3.5.1 Il modello centralizzato

Logica: nel modello centralizzato le competenze specialistiche sono concentrate in un'unica unità organizzativa dedicata, alla quale sono attribuite la definizione degli standard metodologici, lo sviluppo delle soluzioni, il presidio della piattaforma tecnologica e il governo dei rischi connessi. Le unità di business formulano le richieste, mentre l'unità centrale le valuta, le ordina secondo priorità e le realizza. La logica sottostante è quella dell'economia di specializzazione: quando una competenza è scarsa, costosa da acquisire e soggetta a rendimenti crescenti nell'apprendimento, concentrarla consente di raggiungere un livello adeguato, di evitare duplicazioni e di uniformare la qualità.

Meccanismi di coordinamento: il coordinamento è gestito in via gerarchica e per standardizzazione: l'unità centrale definisce metodologie, strumenti e criteri di accettazione e il collegamento con le funzioni di business avviene attraverso processi formalizzati di richiesta e di rilascio.

Vantaggi: la concentrazione delle competenze genera quattro benefici principali, a partire dalla maggiore capacità di attrarre e trattenere profili specialistici che, in

un'organizzazione frammentata, avrebbero maggiori difficoltà a individuare risorse qualificate e percorsi di sviluppo professionale adeguati. Assicura uniformità metodologica e qualitativa, poiché tutte le soluzioni nascono dal medesimo insieme di regole. Facilita il presidio dei rischi e la conformità, perché il perimetro da monitorare è circoscritto e identificabile. Evita la duplicazione degli investimenti infrastrutturali, che possono essere condivisi. In un contesto altamente regolamentato come quello bancario, l'argomento della tracciabilità è particolarmente rilevante: un modello centralizzato rende disponibile l'inventario dei sistemi in uso in tempi brevi, condizione preliminare per qualunque adempimento documentale.

Limiti: il principale limite strutturale è di natura informativa, in quanto concentrando la competenza tecnica, il modello la separa dalla conoscenza del dominio applicativo, che risiede nelle unità operative. Di conseguenza la qualità della soluzione dipenderà dalla capacità di trasferire una conoscenza in larga parte tacita attraverso l'interfaccia fra le due strutture e tale trasferimento è il punto in cui il processo si degrada più facilmente. Si tratta della stessa difficoltà che, in un contesto diverso, Stein (2002) ha analizzato mostrando come l'informazione qualitativa perda valore quando deve percorrere una gerarchia. Un secondo limite è la congestione: se la domanda supera la capacità dell'unità centrale, i tempi di risposta si allungano e le unità di business tendono a dotarsi autonomamente di strumenti, generando quel fenomeno di adozione non censita che vanifica proprio la finalità di controllo per cui il modello era stato predisposto. Un terzo limite riguarda l'appropriazione: soluzioni progettate centralmente incontrano una resistenza all'adozione maggiore di quelle costruite con il coinvolgimento di chi dovrà utilizzarle.

Condizioni di appropriatezza: il modello risulta coerente nelle fasi iniziali di adozione, quando i sistemi in uso sono pochi e le competenze non sono ancora evolute; nelle organizzazioni di dimensione contenuta, in quanto la distanza fra centro e periferia è ridotta e il problema del trasferimento informativo risulta essere meno difficoltoso e nei settori fortemente regolamentati, nei quali il valore attribuito all'uniformità dei presidi è elevato. Sono le condizioni in cui si trovano oggi la maggior parte degli intermediari bancari, e lo conferma il dato richiamato al § 2.3, secondo cui un terzo degli intermediari italiani, di cui circa la metà

banche, ha istituito una struttura dedicata alla governance dell'intelligenza artificiale (Banca d'Italia, 2025).

3.5.2 Il modello decentrato

Logica: nel modello decentrato le competenze sono distribuite all'interno delle singole unità organizzative, ciascuna delle quali sviluppa e gestisce le proprie soluzioni. La logica è speculare alla precedente: privilegia la prossimità fra competenza tecnica e conoscenza del dominio, riducendo il costo di trasferimento dell'informazione e accorciando i tempi di realizzazione.

Meccanismi di coordinamento: il coordinamento avviene per adattamento reciproco all'interno di ciascuna unità, mentre il coordinamento fra unità è debole ed è una caratteristica delle strutture organizzate a silos. È questa asimmetria a generare sia vantaggi, sia limiti del modello.

Vantaggi: la vicinanza al processo consente di individuare con maggiore precisione i casi d'uso realmente rilevanti e di evitare l'errore di sviluppare soluzioni tecnicamente valide ma poco adeguate rispetto ai problemi operativi. La responsabilità sui risultati è chiara, poiché coincide con la responsabilità sul processo, i tempi di realizzazione si riducono e l'appropriazione da parte degli utilizzatori è più elevata, perché la soluzione nasce all'interno della struttura che la impiega.

Limiti: i limiti sono evidenti perché la frammentazione degli standard produce soluzioni non comparabili né riutilizzabili, con duplicazione degli sforzi e dei costi. La qualità diviene disomogenea, poiché dipende dalle competenze disponibili in ciascuna unità e il presidio dei rischi è il profilo più critico: in assenza di un punto di consolidamento, l'organizzazione può perdere la visione d'insieme dei sistemi in uso, il che in un contesto regolamentato non è soltanto inefficiente ma può risultare non conforme alla normativa vigente. Inoltre, il modello è esposto in misura maggiore al rischio segnalato da Lebovitz, Levina e Lifshitz-Assaf (2021): quando manca una funzione indipendente che presidi la costruzione del riferimento di correttezza rispetto al quale i sistemi sono addestrati e valutati, la validazione tende

a essere condotta dagli stessi soggetti che hanno definito il problema, con evidenti conseguenze sull'affidabilità del giudizio.

Una variante contemporanea: la diffusione di strumenti generativi accessibili senza competenze di programmazione ha fatto emergere una forma di decentramento che non è oggetto di scelta ma di constatazione, talvolta descritta come sviluppo da parte di utenti non specialisti. Il fenomeno amplifica i vantaggi del modello – prossimità al processo, rapidità, appropriazione – e ne amplifica simmetricamente i limiti, poiché le soluzioni prodotte sfuggono all'inventario dei sistemi che il § 3.6.2 indica come requisito minimo di governance. Non si tratta quindi di un modello nuovo, ma del modello decentrato realizzato in assenza delle precondizioni di maturità che ne giustificherebbero l'adozione deliberata.

Condizioni di appropriatezza: il modello richiede una maturità elevata, una diffusione ampia delle competenze e una cultura tecnica solida e presuppone che i requisiti di uniformità e di controllo siano stabili e ben consolidati. Sono condizioni difficilmente riscontrabili nel settore bancario, in cui il modello puramente decentrato si presenta di regola non come scelta deliberata ma come esito non governato di iniziative locali.

3.5.3 Il modello hub and spoke: la configurazione federata e le sue varianti

Logica: il modello *hub and spoke* distingue fra ciò che viene concentrato e ciò che viene distribuito. Un nucleo centrale, l'*hub*, mantiene la piattaforma tecnologica, gli standard metodologici, il presidio dei rischi e le competenze più specialistiche, mentre le unità di business dispongono di risorse dedicate, gli *spoke*, che sviluppano e adattano soluzioni all'interno del perimetro attraverso gli strumenti definiti dal centro. La logica è quella di separare ciò che beneficia di economie di scala e richiede uniformità da ciò che richiede prossimità al dominio. La denominazione e la descrizione più diffusa della configurazione provengono dalla pubblicistica manageriale, in particolare da Fountaine, McCarthy e Saleh (2019), ma la logica sottostante è pienamente riconducibile alla teoria organizzativa: si tratta di una forma matriciale, nella quale l'aumento della differenziazione è compensato da un rafforzamento dei dispositivi di integrazione (Lawrence e Lorsch, 1967) e la struttura è progettata per accrescere la capacità di trattamento delle informazioni dell'organizzazione (Galbraith, 1974).

Meccanismi di coordinamento: il modello poggia su meccanismi di integrazione più articolati dei precedenti e ne dipende in misura maggiore: standardizzazione degli strumenti e delle metodologie da parte dell'*hub*, ruoli di collegamento fra centro e unità, organi di coordinamento per la definizione delle priorità, comunità professionali che assicurino la circolazione delle conoscenze fra gli *spoke*. Quest'ultimo meccanismo richiede particolare attenzione, poiché la comunità professionale non costituisce un elemento accessorio dell'assetto organizzativo, ma rappresenta il principale canale attraverso cui si sviluppa e si diffonde la conoscenza pratica. La letteratura sull'apprendimento organizzativo ha infatti evidenziato come le comunità di pratica siano i luoghi nei quali il sapere effettivamente impiegato nelle attività lavorative viene prodotto, condiviso e consolidato (Brown e Duguid, 1991).

Le due varianti: sono ammesse due declinazioni, che si distinguono per la collocazione dei poteri decisionali e che conviene tenere distinte perché generano fabbisogni di competenze diversi.

- *Hub forte:* il centro mantiene la responsabilità sullo sviluppo e sulla validazione; gli *spoke* presidiano la definizione dei requisiti, la sperimentazione e l'adozione. È la variante più prossima al modello centralizzato, dal quale si distingue per la presenza di risorse stabilmente assegnate alle unità di business anziché di semplici referenti occasionali.
- *Hub leggero:* il centro mantiene la piattaforma, gli standard, la validazione indipendente e l'inventario dei sistemi, mentre lo sviluppo è realizzato dagli *spoke*. La variante presuppone una maturità superiore e una diffusione più ampia delle competenze tecniche, ma consente una scalabilità che l'*hub* forte non raggiunge.

La scelta fra le due varianti non è definitiva: la transizione dall'*hub* forte all'*hub* leggero costituisce la traiettoria evolutiva più frequentemente osservabile e rappresenta, in termini di capacità dinamiche, un'operazione di riconfigurazione della base di risorse nel senso di Teece (2007).

Vantaggi: il modello consente in linea di principio di combinare i vantaggi delle due configurazioni precedenti: uniformità dei presidi e delle piattaforme, prossimità al dominio

applicativo, scalabilità superiore a quella del modello centralizzato. Consente una certa gradualità poiché il perimetro affidato agli *spoke* può essere ampliato progressivamente con il crescere della maturità e delle competenze, il che lo rende compatibile con la prospettiva evolutiva delle capacità dinamiche. Un vantaggio ulteriore riguarda la validazione: la separazione fra chi sviluppa e chi valida è nel modello *hub and spoke* strutturalmente fattibile, poiché le due attività sono collocate in strutture distinte.

Limiti: il modello è il più esigente sul piano gestionale e i suoi limiti derivano da questa esigenza, partendo dall'ambiguità dei poteri decisionali: se la ripartizione delle responsabilità fra centro e unità non è definita con precisione, il modello degenera e l'*hub* finisce per esercitare un ruolo di semplice appoggio privo di effettività. Il secondo limite è il costo di coordinamento, che non è trascurabile e che può eccedere i benefici in organizzazioni di dimensione contenuta. Il terzo è la disomogeneità qualitativa fra gli *spoke*, che permane in misura minore rispetto al modello decentrato. Il quarto, di natura professionale, riguarda la posizione degli specialisti collocati negli *spoke*, costretti a un doppio riporto e a un'identità professionale ambivalente, secondo la dinamica che Vaast e Pinsonneault (2021) hanno analizzato con riferimento ai profili tecnici emergenti.

Condizioni di appropriatezza: il modello risulta appropriato in contesti caratterizzati da una dimensione organizzativa tale da assorbire i costi di coordinamento, da un numero di sistemi in esercizio sufficientemente elevato da rendere inefficiente la centralizzazione e da un'allocazione dei poteri decisionali formalizzata e stabile; nel settore bancario, tale configurazione è tipicamente adottata dai gruppi di maggiori dimensioni.

3.5.4 Il modello esternalizzato e consortile

Logica: in questo modello la capacità di sviluppo non risiede all'interno dell'organizzazione: l'impresa accede alla tecnologia attraverso fornitori esterni, piattaforme standardizzate e iniziative promosse da organismi di sistema, mantenendo internamente le funzioni di definizione dei requisiti, selezione, verifica e presidio. È il modello che il secondo capitolo ha individuato come prevalente per gli intermediari di dimensione intermedia e la sua inclusione nella tassonomia non è una concessione descrittiva: si tratta di una configurazione

organizzativa con una propria logica, riconducibile alla scelta fra produzione interna e acquisizione esterna di una capacità.

Meccanismi di coordinamento: il coordinamento avviene per contratto e per specifica: il rapporto con il fornitore è regolato da accordi che definiscono livelli di servizio, obblighi informativi e diritti di verifica.

Vantaggi: il modello consente di accedere a soluzioni mature senza sostenere i costi fissi di costruzione della capacità interna, riducendo il rischio di investimento e i tempi di attivazione. Nel caso delle iniziative di sistema, consente di ripartire tra più intermediari i costi di sperimentazione e di sviluppo degli standard e di accedere a conoscenze che il singolo operatore non potrebbe produrre autonomamente.

Limiti: i limiti sono di tre ordini, il primo è la ridotta capacità di differenziazione: soluzioni standardizzate difficilmente incorporano le specificità che costituiscono il vantaggio competitivo dell'intermediario, il che, per un modello di business fondato sulla relazione, non è un dettaglio. Il secondo è la dipendenza dal fornitore, che comporta rischi di continuità operativa, costi di transizione elevati e una limitata capacità di negoziazione. Il terzo, e più rilevante ai fini della domanda di ricerca, è il rischio di erosione della capacità interna: l'organizzazione che acquisti sistematicamente all'esterno può perdere la capacità di valutare ciò che acquista, con la conseguenza paradossale che le competenze necessarie a governare l'esternalizzazione si degradano proprio per effetto dell'esternalizzazione. È la trasposizione, sul piano delle capacità organizzative, del meccanismo che Raisch e Krakowski (2021) descrivono a proposito dell'automazione, e trova un antecedente teorico nella nozione di capacità di assorbimento, secondo cui la possibilità di riconoscere il valore di una conoscenza esterna e di utilizzarla dipende da uno stock di conoscenza già posseduto internamente (Cohen e Levinthal, 1990).

Condizioni di appropriatezza: il modello è coerente per organizzazioni di dimensione contenuta, per casi d'uso non differenzianti e in fasi iniziali di adozione. La sua sostenibilità nel tempo dipende però da una condizione precisa e spesso trascurata: il mantenimento all'interno di una capacità di valutazione autonoma, sufficiente a definire i requisiti, a verificare

gli esiti e a esercitare la sorveglianza che la normativa pone a carico dell'utilizzatore professionale ai sensi dell'articolo 26 dell'AI Act, come illustrato al § 2.7.2. Il modello esternalizzato non esonera dalla competenza: ne sposta il contenuto dalla costruzione alla valutazione.

3.5.5 Il modello a piattaforma e prodotto

Logica: i quattro modelli finora esaminati rispondono alla domanda su dove collocare la competenza, mentre il modello a piattaforma e prodotto risponde a una domanda diversa: attorno a che cosa aggregare il lavoro quotidiano. La sua logica consiste nel sostituire il progetto con il prodotto come unità organizzativa elementare: anziché costituire gruppi temporanei che si sciolgono al rilascio, l'organizzazione istituisce squadre interfunzionali stabili – che comprendono competenze di dominio, tecniche, di dato e, ove necessario, di conformità – responsabili di un prodotto lungo l'intero ciclo di vita: concezione, realizzazione, esercizio, monitoraggio e ritiro. A queste squadre si affianca una piattaforma interna che fornisce in modo standardizzato le capacità comuni – accesso ai dati, ambienti di addestramento, registri dei modelli, strumenti di monitoraggio, presidi di sicurezza – riducendo il carico cognitivo e tecnico che ciascuna squadra dovrebbe altrimenti sostenere.

Il fondamento teorico della configurazione è duplice. Sul piano della progettazione, essa costituisce una risposta al problema del coordinamento per riduzione della necessità di comunicazione: inserendo in una squadra tutte le competenze necessarie a un prodotto, si minimizza il fabbisogno di scambio informativo fra unità, secondo la strategia che Galbraith (1974) definisce di creazione di compiti autosufficienti. Sul piano della gestione dell'informazione, la configurazione riprende il principio per cui la responsabilità sul dato deve essere attribuita al dominio che lo produce anziché a un'unità centrale che lo raccoglie (Dehghani, 2022).

Meccanismi di coordinamento: il coordinamento avviene per standardizzazione dei risultati e delle interfacce anziché dei processi: la piattaforma stabilisce come i prodotti si connettono e quali requisiti devono soddisfare, lasciando alle squadre autonomia su come impostare l'operatività. Si aggiungono meccanismi di allineamento per obiettivi e una governance

federata, nella quale le regole comuni sono definite con la partecipazione dei responsabili di prodotto.

Vantaggi: il modello affronta direttamente il problema che il § 3.2.2 ha individuato come specifico dell'intelligenza artificiale: poiché il sistema apprende e muta, la responsabilità non può esaurirsi con la messa in esercizio. Attribuendo a una squadra stabile la titolarità dell'intero ciclo di vita, la configurazione rende presidiato il lavoro continuo di monitoraggio, correzione e manutenzione che il § 3.7.3 descrive come sottovalutato e lasciato in secondo piano. Il modello riduce i tempi di rilascio, favorisce l'accumulo di conoscenza sul prodotto e attenua il problema del trasferimento informativo, poiché competenza tecnica e competenza di dominio coesistono nella stessa squadra.

Limiti: il primo limite è il costo in quanto squadre stabili e interfunzionali richiedono una dotazione di competenze che organizzazioni di dimensione intermedia difficilmente possiedono e la piattaforma interna comporta a sua volta un investimento rilevante. Il secondo è il rischio di frammentazione, che riproduce a livello di prodotto il limite del modello decentrato: senza una governance effettiva, le squadre divergono negli standard e nelle scelte tecniche. Il terzo, specifico dei contesti regolamentati, riguarda l'indipendenza della validazione: se la squadra che costruisce il prodotto è la stessa che ne verifica gli esiti, viene meno il requisito di indipendenza argomentato al § 3.7.2, sicché la configurazione richiede necessariamente una funzione di validazione esterna alle squadre. Il quarto riguarda il rapporto con l'assetto dei controlli interni descritto al § 3.6.2: l'autonomia delle squadre va conciliata con l'articolazione dei controlli su più livelli, il che richiede di collocare i controlli di secondo livello al di fuori del perimetro del prodotto.

Condizioni di appropriatezza: la configurazione è coerente per organizzazioni con un numero elevato di sistemi in esercizio, con una dotazione interna di competenze sufficiente a comporre squadre stabili e con un portafoglio di casi d'uso destinati a permanere. È tipicamente adottata dagli operatori digitali e dai gruppi bancari di maggiori dimensioni, e può essere combinata con un assetto *hub and spoke*, nel quale l'*hub* assume il ruolo di gestore della piattaforma anziché di realizzatore delle soluzioni. Questa combinazione – *hub* come

piattaforma, *spoke* come squadre di prodotto – costituisce, allo stato, la configurazione più avanzata e matura osservabile nel settore.

3.5.6 Il modello predittivo o decision-centric

Logica: un'altra impostazione parte da una domanda diversa, da che cosa l'intelligenza artificiale effettivamente produca. La risposta più utile alla progettazione organizzativa è quella proposta dall'economia dell'innovazione: l'intelligenza artificiale riduce il costo della previsione, ossia dell'operazione con cui, a partire dall'informazione disponibile, si colma l'informazione mancante. Poiché la previsione è una componente della decisione, ma non l'unica – vi concorrono il giudizio sulle conseguenze, i dati e l'azione – la riduzione del costo di una componente modifica il valore relativo delle altre e accresce il valore del giudizio umano, che resta complementare (Agrawal, Gans e Goldfarb, 2019).

Da questa premessa discende un criterio di progettazione: l'unità attorno a cui organizzare il lavoro non è né il progetto né il prodotto, ma la decisione. L'organizzazione individua le decisioni ricorrenti che governano il proprio funzionamento, le scompone nelle loro componenti – quali informazioni le alimentano, quale previsione richiedono, quale giudizio presuppongono, quale azione ne consegue – e riprogetta ciascuna componente stabilendo quale sia eseguita da un sistema e quale da una persona. È l'applicazione più diretta e più rigorosa del principio secondo cui la ripartizione dei compiti fra persone e sistemi è una scelta di progettazione organizzativa e non una conseguenza tecnica.

Meccanismi di coordinamento: il coordinamento avviene per esplicitazione dei diritti decisionali: per ciascuna decisione sono definiti chi la assume, su quali informazioni, con quale supporto algoritmico, entro quali soglie e con quali possibilità di scostamento. Ne risulta una mappatura formale che costituisce il documento di progettazione organizzativa e la base dell'inventario dei sistemi richiesto ai fini di conformità normativa.

Vantaggi: il principale è la selettività. Le rilevazioni richiamate al § 3.4.2 documentano un ampio divario fra adozione della tecnologia e conseguimento di benefici misurabili; una parte rilevante di tale divario è attribuibile all'applicazione della tecnologia a compiti che

non costituiscono il vincolo effettivo del processo. Organizzare attorno alle decisioni impone di individuare preliminarmente quale previsione, se migliorata, produca un effetto positivo sull'esito, il che riduce la dispersione degli investimenti. Un secondo vantaggio, particolarmente rilevante in ambito bancario, è la coerenza con l'impianto dei controlli: poiché il settore dispone già di un'architettura formalizzata di deleghe e poteri deliberativi, come osservato al § 2.7.3, la mappatura delle decisioni si innesta su una struttura esistente. Un terzo vantaggio è che la configurazione rende esplicito, decisione per decisione, il grado di intervento umano richiesto, evitando che il livello di sorveglianza sia stabilito implicitamente dalla soluzione tecnica.

Limiti: il primo limite è la complessità dell'esercizio di mappatura, che richiede un investimento analitico rilevante e competenze di analisi organizzativa non sempre disponibili in azienda. Il secondo è il rischio di formalismo in quanto una mappatura delle decisioni costruita a tavolino può non corrispondere al modo in cui le decisioni sono effettivamente assunte e in tal caso produce un documento di analisi ulteriore anziché un cambiamento. Il terzo è che l'impostazione si presta meglio alle decisioni ricorrenti e strutturate che a quelle impreviste e ambigue, ossia proprio a quelle in cui il contributo del giudizio umano è maggiore: il rischio è di riprogettare soltanto la parte più semplice del lavoro.

Condizioni di appropriatezza: sulla base della tassonomia elaborata, si formula l'ipotesi che la configurazione possa risultare coerente nei processi ad alta frequenza decisionale e a elevata intensità informativa – credito, monitoraggio andamentale, antifrode, pricing – e quindi meritevole di particolare considerazione nel settore bancario. In termini analitici è compatibile con tutte le configurazioni di allocazione della competenza, poiché non prescrive dove collocare gli specialisti ma come impostare l'analisi che ne influenza il lavoro; l'appropriatezza effettiva richiede comunque una verifica sul caso concreto.

3.5.7 Il modello agentico-ibrido

Logica: l'evoluzione descritta al § 2.5 – dal sistema che produce un contenuto al sistema che compie una sequenza di azioni – pone un problema organizzativo che nessuna delle configurazioni precedenti affronta direttamente: l'inserimento, all'interno del flusso di

lavoro, di entità non umane dotate di un grado di autonomia operativa. Il modello agentico-ibrido consiste nel progettare unità di lavoro composte da persone e agenti digitali, nelle quali la ripartizione dei compiti è definita esplicitamente e la responsabilità resta integralmente umana.

Il riferimento scientifico più pertinente è la linea di ricerca che analizza i sistemi intelligenti non come strumenti ma come partecipanti al lavoro di gruppo, esaminandone le condizioni di efficacia e i problemi di progettazione (Seeber et al., 2020). Ne discende un'avvertenza metodologica di rilievo: la metafora del collaboratore digitale è utile a descrivere l'interazione, ma non deve indurre a trattare l'agente come un soggetto organizzativo. L'agente esegue compiti e non assume responsabilità e la sua introduzione modifica la divisione del lavoro, non l'attribuzione dei doveri.

Meccanismi di coordinamento: il coordinamento si fonda sulla delimitazione del perimetro di azione: quali applicativi l'agente può utilizzare, con quali profili autorizzativi, entro quali soglie, con quale possibilità di sospensione. Il secondo elemento è la registrazione granulare delle operazioni, utile a ricostruire ex post la catena delle azioni e ad attribuire la responsabilità. Infine, occorre considerare la designazione formale di un responsabile umano per ciascun agente in esercizio, al quale competano la sorveglianza, l'autorizzazione delle modifiche e la decisione di sospensione. Si tratta di scelte di progettazione coerenti con le indicazioni istituzionali su identità autonoma dell'agente, autorizzazioni granulari e controllo umano proporzionato (NIST, 2026; OECD, 2026).

Vantaggi: il modello consente di estendere l'automazione a sequenze di attività che attraversano più applicativi e che l'automazione tradizionale, fondata su regole rigide, non riusciva a coprire; consente di liberare capacità nelle attività di raccolta e riconciliazione delle informazioni. Un vantaggio ulteriore, di natura organizzativa, è che la configurazione impone di esplicitare la divisione del lavoro entro il processo, rendendo visibili passaggi che restano impliciti.

Limiti: i limiti sono i più significativi fra quelli esaminati e ne consigliano un'adozione prudente. Il primo è lo spostamento del profilo di rischio dall'accuratezza dell'output

all'ampiezza delle azioni consentite, con la conseguenza che un errore non produce un contenuto da correggere ma un'operazione da rettificare. Il secondo è l'attenuazione della sorveglianza: poiché l'agente opera su una sequenza, la persona non osserva i singoli passaggi ma l'esito complessivo, condizione nella quale i fenomeni di compiacenza e di distorsione dell'automazione documentati al § 2.7.5 operano con particolare intensità. Il terzo è la difficoltà di attribuzione della responsabilità quando più agenti interagiscono tra di loro, che richiede un disegno esplicito delle dipendenze. Il quarto, di natura professionale, è il rischio che l'assegnazione agli agenti delle attività elementari sottragga alle persone proprio le occasioni di apprendimento attraverso le quali si forma la competenza necessaria a sorvegliarli.

Condizioni di appropriatezza: nell'interpretazione proposta, la configurazione può risultare appropriata per processi interni, ripetitivi, a basso impatto sul cliente e con esiti verificabili, nei quali il perimetro di azione possa essere delimitato in modo stretto. Si formula l'ipotesi che essa sia inappropriata per i processi che incidono direttamente su diritti o posizioni della clientela, per i quali una configurazione *human in the loop* descritta al § 2.7.1 offre, in linea generale, un presidio più coerente. La configurazione presuppone un livello elevato di maturità nella governance del dato e nella registrazione degli eventi; in assenza di tali presupposti la capacità di controllo risulta sensibilmente ridotta. Queste valutazioni costituiscono un'applicazione della tassonomia proposta e richiedono verifica sul processo concreto.

3.5.8 Il modello ecosistemico

Logica: il modello ecosistemico consiste nell'accedere alle capacità di intelligenza artificiale attraverso un insieme coordinato di attori interdipendenti – fornitori tecnologici, operatori specializzati, centri di ricerca, imprese innovative, altre banche – legati da rapporti di complementarità e da standard condivisi, senza che esista fra essi un rapporto gerarchico. Si distingue dal modello esternalizzato, che è una relazione bilaterale regolata da contratto, per la pluralità degli attori e per la natura non gerarchica del coordinamento, fondato su complementarità e su standard anziché su specifiche contrattuali (Jacobides, Cennamo e Gawer, 2018).

Meccanismi di coordinamento: il coordinamento avviene per standard tecnici e per interfacce: la compatibilità tra i diversi attori è assicurata da regole di interoperabilità anziché da un'unità centrale. Nel settore bancario europeo questa modalità trova una premessa nel quadro normativo che ha imposto l'apertura dei dati a soggetti terzi previo consenso del cliente.

Vantaggi: il modello consente di accedere a competenze che nessun singolo operatore potrebbe produrre internamente e di distribuire il rischio di innovazione su una pluralità di attori, in aggiunta consente una velocità di sperimentazione superiore a quella ottenibile con lo sviluppo interno. Per gli intermediari di dimensione intermedia rappresenta la via più praticabile per accedere a capacità avanzate senza sostenerne interamente i costi fissi.

Limiti: il primo limite riguarda la governance: in assenza di gerarchia, la definizione delle regole comuni dipende da equilibri tra gli attori con interessi divergenti e l'intermediario che vi partecipa dispone di una capacità di influenza direttamente proporzionata al proprio peso. Il secondo è l'appropriazione del valore: la partecipazione a un ecosistema comporta la condivisione di informazioni e di apprendimenti, con il rischio che il vantaggio si trasferisca agli attori che controllano gli snodi. Il terzo è regolamentare: il ricorso a una pluralità di soggetti moltiplica il rischio di terze parti descritto al § 2.8 e le esigenze di presidio che la disciplina sulla resilienza operativa digitale pone a carico dell'intermediario. Il quarto riguarda la stabilità: gli ecosistemi sono configurazioni mutevoli e la dipendenza da attori la cui permanenza non è garantita introduce un rischio di continuità operativa che il contratto non copre.

Condizioni di appropriatezza: nell'interpretazione proposta, il modello può risultare coerente per attività non distintive, per la sperimentazione precompetitiva e per lo sviluppo di standard di settore, mentre appare meno appropriato per le componenti che incorporano il vantaggio competitivo dell'intermediario. La partecipazione efficace a un ecosistema di questa portata richiede comunque competenze valutative analoghe a quelle del modello esternalizzato, accresciute dalla necessità di compiere scelte collettive. Anche questa valutazione ha

natura analitica e deve essere verificata rispetto alle caratteristiche dell'intermediario e dell'ecosistema considerati.

3.5.9 Quattro considerazioni di ordine generale

Dall'analisi dettagliata dei modelli emergono quattro considerazioni.

La prima è che nessuna configurazione domina le altre lungo tutte le dimensioni: ciascuna risolve alcuni problemi generandone altri, il che conferma l'impostazione contingente adottata. Il modello centralizzato acquista uniformità cedendo prossimità; quello decentrato compie lo scambio opposto; quello *hub and spoke* tenta di conservare entrambe assumendosi un costo di coordinamento; quello esternalizzato riduce l'investimento cedendo differenziazione e autonomia. Fra i modelli di aggregazione, quello a piattaforma e prodotto risolve il problema del ciclo di vita al prezzo di una dotazione di competenze elevata; quello predittivo risolve il problema della selettività ma presuppone un investimento analitico; quello agentico estende il perimetro dell'automazione accrescendo l'esposizione al rischio; quello ecosistemico attenua il vincolo di scala cedendo controllo.

La seconda riguarda la variabile discriminante: si osserva che la dimensione lungo la quale i modelli si differenziano in modo più netto è l'accesso alla conoscenza del dominio applicativo. È un risultato coerente con l'impianto teorico del § 3.3.1: se il problema organizzativo posto dall'IA è in larga parte un problema di circolazione dell'informazione, la configurazione appropriata è quella che minimizza la distanza fra il luogo in cui il problema è conosciuto e il luogo in cui la soluzione è costruita, compatibilmente con i requisiti di uniformità e di controllo.

La terza riguarda la coerenza fra le due categorie: le scelte di allocazione e di aggregazione sono indipendenti e sono possibili combinazioni coerenti e/o contraddittorie. Un modello a piattaforma e prodotto adottato in un assetto puramente centralizzato è contraddittorio, poiché le squadre di prodotto non dispongono delle competenze necessarie e la piattaforma degrada a strumento di controllo, mentre un modello predittivo adottato in un assetto esternalizzato è coerente soltanto se la mappatura delle decisioni resta interna, perché

altrimenti l'intermediario delega al fornitore la scelta di quali decisioni automatizzare. Un modello agentico adottato in un assetto decentrato è, allo stato attuale, la combinazione più esposta, poiché unisce la massima ampiezza di azione alla minima capacità di consolidamento dei controlli.

La quarta considerazione è di natura dinamica. I modelli sono stati presentati come alternative, ma nella realtà le organizzazioni si spostano fra essi al variare della propria maturità e del numero di sistemi in esercizio. La traiettoria più ricorrente muove dal modello esternalizzato o centralizzato verso configurazioni *hub and spoke* e da queste verso assetti a piattaforma e prodotto, con l'impostazione predittiva che può innestarsi in qualunque stadio e con quella agentica che presuppone lo stadio più avanzato. La transizione non è però priva di costi: modificare la localizzazione delle competenze significa ridefinire ruoli, responsabilità e rapporti di influenza, con resistenze prevedibili. Ne consegue che la scelta iniziale ha effetti persistenti e che la capacità di far evolvere il proprio assetto costituisce essa stessa una capacità dinamica nel senso di Teece (2007) e di Warner e Wäger (2019). È opportuno sottolineare che la presenza, all'interno del medesimo ente, di configurazioni organizzative differenti non rappresenta necessariamente un elemento di incoerenza. Un assetto più strutturato e formalizzato può risultare appropriato per i sistemi integrati nei processi regolamentati e ad elevata criticità, mentre soluzioni più flessibili e sperimentali possono essere adottate per le applicazioni generative a supporto delle attività operative. Tale coesistenza riflette una logica di ambidestrismo organizzativo, intesa come la capacità di combinare lo sfruttamento delle competenze e delle risorse esistenti con l'esplorazione di nuove opportunità, attraverso assetti differenziati ma coordinati a livello di governo centrale (O'Reilly e Tushman, 2013).

3.6 La dimensione della governance: presidi, ruoli e controlli

La scelta della configurazione non esaurisce la progettazione organizzativa. Occorre stabilire, in parallelo, come siano presidiati i rischi e come si articolino le responsabilità: si tratta di una dimensione ortogonale alla precedente, nel senso che ciascuna delle configurazioni esaminate deve dotarsi di un impianto di governance che assume forme diverse a seconda della configurazione.

3.6.1 Il sistema di gestione della qualità come vincolo di progettazione

Il quadro regolamentare esaminato nel secondo capitolo non si limita a imporre obblighi documentali: contiene prescrizioni che incidono direttamente sull'assetto organizzativo. L'articolo 17 dell'AI Act richiede ai fornitori di sistemi ad alto rischio di istituire un sistema di gestione della qualità, documentato in modo sistematico e ordinato, comprensivo della strategia adottata per la conformità normativa, delle procedure di progettazione e di verifica, delle procedure di gestione dei dati, del sistema di gestione del rischio e di un quadro di responsabilità che definisca le prerogative della direzione e del personale. Si tratta della richiesta di un assetto organizzativo formalizzato: la norma non prescrive quale configurazione impostare ma esige che, qualunque essa sia, le responsabilità siano definite e documentate.

3.6.2 L'articolazione dei controlli e il rapporto con l'assetto esistente

Nel settore bancario l'impianto dei controlli interni è già definito e non deve essere costruito ex novo. Gli orientamenti dell'Autorità bancaria europea sulla governance interna e le disposizioni di vigilanza nazionali articolano il sistema dei controlli su più livelli, distinguendo i controlli di linea, incorporati nelle procedure operative, dai controlli sui rischi e sulla conformità, affidati a funzioni indipendenti, e dalla revisione interna (European Banking Authority, 2021b; Banca d'Italia, Circolare n. 285/2013).

La questione organizzativa rilevante è comprendere se l'intelligenza artificiale richieda un impianto di governance separato o un'estensione di quello esistente. L'analisi condotta suggerisce la seconda risposta, per tre ragioni. In primo luogo, i rischi introdotti dall'IA – il rischio di modello, il rischio di terze parti, il rischio operativo e di non conformità – sono in larga parte categorie già presidiate, per le quali esistono metodologie consolidate quali quelle sviluppate in materia di gestione del rischio di modello (Board of Governors of the Federal Reserve System, 2011). In secondo luogo, la creazione di un presidio parallelo genera duplicazioni e zone di incerta attribuzione, che sono la causa più frequente di lacune nei controlli. In terzo luogo, l'esistenza di un impianto consolidato costituisce, come osservato al § 3.4.3 a proposito dei dati di rischio, un vantaggio di percorso per gli intermediari bancari rispetto a organizzazioni di altri settori.

Ciò che l'IA aggiunge non è un nuovo sistema di controlli, ma alcune esigenze specifiche da inserire in quello esistente: la tenuta di un inventario dei sistemi in uso e della loro classificazione di rischio, la definizione di criteri di validazione preventiva e di monitoraggio continuo delle prestazioni, la tracciabilità degli scostamenti fra indicazione del sistema e decisione umana e l'attribuzione formale della responsabilità di sorveglianza a soggetti individuati e dotati di competenza e autorità.

3.7 Dai modelli al funzionamento: i meccanismi che ne determinano l'efficacia

L'analisi condotta ha riguardato la struttura formale, mentre ora è necessario affrontare una questione che la sola struttura non risolve e che costituisce il punto più rilevante del capitolo: due organizzazioni possono adottare la medesima configurazione e ottenere risultati opposti. La ragione è che il funzionamento dipende da meccanismi che operano al di sotto dell'assetto formale.

3.7.1 La ridefinizione del controllo

Il primo meccanismo riguarda il tema del controllo, partendo dall'elaborato di Kellogg, Valentine e Christin (2020) che sistematizza la ricerca interdisciplinare sull'impiego degli algoritmi nei luoghi di lavoro e mostra che il controllo algoritmico opera tramite i sei meccanismi principali, che le autrici designano come le sei R: i datori di lavoro possono impiegare gli algoritmi per indirizzare i lavoratori limitandone le azioni, per valutarli registrandone e classificandone le prestazioni e per sanzionarli sostituendoli o remunerandoli. Le autrici osservano che le capacità tecniche dei sistemi algoritmici rendono possibile una forma di controllo razionale distinta da quelle tecniche e burocratiche impiegate nel secolo precedente e l'uso di algoritmi sta favorendo la nascita di nuove occupazioni con i lavoratori che sviluppano forme di resistenza individuale e collettiva.

La rilevanza di tale questione per la progettazione organizzativa è immediata: un sistema introdotto come supporto al processo decisionale può trasformarsi in uno strumento di direzione e controllo. Quando gli scostamenti rispetto alla raccomandazione algoritmica vengono registrati e utilizzati come elemento di valutazione della performance individuale, il

presidio umano descritto al § 2.7 come garanzia perde la propria funzione di tutela e si configura come un potenziale fattore di penalizzazione per chi lo esercita. In tali condizioni, l'esito più probabile è una progressiva riduzione della supervisione umana a mera ratifica delle indicazioni prodotte dal sistema. Ne deriva che chi sceglie l'assetto organizzativo non può limitarsi a individuare il soggetto responsabile della sorveglianza, ma deve valutare attentamente gli effetti che i meccanismi di misurazione e incentivazione generano su chi è chiamato a esercitarla.

3.7.2 La validazione e il problema del riferimento di correttezza

Il tema della validazione introduce una questione elementare, ma spesso trascurata: come osservano Lebovitz, Levina e Lifshitz-Assaf (2021), i sistemi di IA vengono addestrati e valutati sulla base di parametri di accuratezza considerati affidabili e oggettivi. Tuttavia, tali parametri non costituiscono un riferimento neutrale poiché sono definiti e costruiti da esseri umani e di conseguenza incorporano inevitabilmente limiti cognitivi, assunzioni interpretative e possibili distorsioni.

La valutazione delle prestazioni di un sistema di IA non può essere considerata un processo esclusivamente tecnico, ma richiede una riflessione critica sulla qualità, sull'affidabilità e sulla stessa legittimità dei criteri utilizzati come termine di confronto. Nei domini in cui il giudizio esperto è soggetto a variabilità, il riferimento non è un dato oggettivo ma una convenzione, e le misure di accuratezza costruite su di esso possono risultare fuorvianti.

In ambito organizzativo, la validazione richiede l'intervento di soggetti che possiedano il dominio applicativo e siano indipendenti da chi ha costruito il sistema. Questo requisito discrimina fra i modelli esaminati: il modello centralizzato assicura più agevolmente l'indipendenza ma difficilmente garantisce la competenza di dominio; il modello decentrato presenta la difficoltà opposta; il modello *hub and spoke* può soddisfare entrambe le condizioni, purché la funzione di validazione sia attribuita esplicitamente all'*hub* e non lasciata implicita; il modello esternalizzato deve assicurarsi che la validazione non sia interamente delegata al fornitore, che si troverebbe a giudicare il proprio operato. Il modello a piattaforma e prodotto presenta la criticità più netta, poiché la squadra che costruisce il sistema è anche quella che

ne osserva gli esiti: la sua adozione richiede necessariamente una funzione di validazione collocata all'esterno del perimetro del prodotto.

3.7.3 Il lavoro invisibile di manutenzione

Il terzo meccanismo è quello già anticipato al § 3.3.2 sulla scorta di Grønsund e Aanestad (2020): il funzionamento di un sistema di intelligenza artificiale richiede un lavoro continuo di verifica e correzione che raramente compare nei documenti di progetto. La tendenza a non riconoscerlo produce due effetti negativi: il primo è di dimensionamento, poiché le risorse assegnate risultano insufficienti e l'attività sottrae tempo, che risulta necessario allo svolgimento di altro con un conseguente abbassamento del livello di qualità. Il secondo è di riconoscimento: un'attività non formalizzata non è valorizzata nei percorsi professionali, il che ne compromette la sostenibilità nel tempo e alimenta la resistenza.

Ne discende un criterio di valutazione dei modelli che raramente viene esplicitato: la configurazione appropriata non è quella che minimizza le risorse impiegate, ma quella che rende visibile e attribuisce esplicitamente il lavoro necessario a mantenere i sistemi allineati agli obiettivi.

3.7.4 L'identità professionale e le resistenze

Il quarto meccanismo riguarda le persone coinvolte. Vaast e Pinsonneault (2021) analizzano come le tecnologie digitali possano al tempo stesso abilitare e minacciare l'identità professionale, mostrando l'equilibrio delicato che i profili tecnici emergenti devono mantenere. Il fenomeno non riguarda soltanto i nuovi profili, ma interessa in misura maggiore le figure professionali consolidate, il cui sapere specialistico costituisce la base della loro identità e del loro riconoscimento. Quando un sistema automatizza o mette in discussione la componente distintiva di quel sapere, la resistenza che ne deriva non è irrazionale e non si supera con la formazione tecnica, ma richiede una ridefinizione del contenuto professionale del ruolo.

Questo punto costituisce un raccordo diretto con il capitolo conclusivo e consente di chiarire il significato attribuito alla barriera culturale nel contesto dell'adozione

dell'intelligenza artificiale. Essa non deve essere interpretata come una semplice avversione al cambiamento, bensì come la risposta prevedibile di professionisti che vedono modificarsi il proprio ruolo, i propri margini di autonomia o le proprie responsabilità senza che tali trasformazioni siano accompagnate da una chiara ridefinizione delle funzioni che sono chiamate a svolgere.

3.8 La declinazione nel settore bancario

Le considerazioni svolte trovano nel settore bancario alcune specificazioni che è opportuno rendere esplicite, riprendendo i due tipi ideali costruiti nel secondo capitolo.

Per il tipo ideale della banca digital native, la configurazione coerente è quella decentrata o di tipo *hub and spoke* con *hub* leggero, tipicamente combinata, sul piano dell'aggregazione del lavoro, con un modello a piattaforma e prodotto. La ragione non è dimensionale ma informativa, perché il modello fonda l'intermediazione sulla hard information e dispone di dati strutturati fin dall'origine, mentre la conoscenza del dominio è in larga misura già codificata nei dati e il problema del trasferimento informativo fra centro e periferia, che costituisce il limite principale del modello centralizzato, si presenta in forma attenuata. L'assenza di sistemi stratificati riduce il vincolo infrastrutturale che spinge alla concentrazione delle competenze e rende sostenibile la costruzione di una piattaforma interna attorno alla quale organizzare squadre di prodotto stabili.

Per il tipo ideale della banca commerciale a radicamento territoriale il quadro è opposto e più complesso: la configurazione osservabile è prevalentemente quella esternalizzata, con un nucleo centralizzato di dimensioni contenute. Il patrimonio informativo distintivo di questo modello è costituito da soft information prodotta nella relazione, ossia dalla risorsa meno codificata e più difficilmente trasferibile; l'accesso alla tecnologia, invece, avviene attraverso soluzioni standardizzate, progettate per essere indipendenti dal contesto. Ne deriva il rischio che l'adozione dell'IA rafforzi le componenti dell'attività meno distintive, lasciando inutilizzato proprio ciò che costituisce il vantaggio competitivo dell'intermediario.

La soluzione presentata ha natura organizzativa più che tecnologica: il vincolo è di scala sullo sviluppo, ma il vantaggio risiede nella conoscenza del dominio e la configurazione appropriata è quella che esternalizza la componente infrastrutturale mantenendo internamente le attività a maggiore contenuto informativo come la definizione dei casi d'uso, la specificazione dei requisiti, la validazione degli esiti e la sorveglianza. È una configurazione che si può descrivere come esternalizzazione della piattaforma e internalizzazione del giudizio, che sposta il fabbisogno di competenze dalla capacità di costruire alla capacità di valutare. In aggiunta, le iniziative di sistema promosse dagli organismi associativi possono attenuare il vincolo di scala anche sulla componente distintiva, poiché consentono di sviluppare in forma collettiva soluzioni adattate alle specificità di una categoria di intermediari.

Una terza indicazione riguarda il criterio di aggregazione del lavoro per un intermediario di dimensione intermedia. Sulla base della tassonomia elaborata, si formula l'ipotesi che un modello completo a piattaforma e prodotto sia difficilmente sostenibile quando il vincolo di scala non consente una dotazione interna adeguata; il modello predittivo o *decision-centric* del § 3.5.6 può invece risultare praticabile e, nell'interpretazione proposta, particolarmente indicato, per due ragioni. La prima è che esso non richiede di costruire capacità di sviluppo, ma di condurre un'analisi delle decisioni ricorrenti, attività per la quale la conoscenza del processo è il fattore critico. La seconda è che la mappatura delle decisioni si innesta sull'architettura dei poteri deliberativi già esistente, secondo quanto osservato al § 2.7.3. Si delinea così, come ipotesi da verificare empiricamente, una combinazione basata su esternalizzazione della piattaforma, internalizzazione del giudizio e organizzazione del lavoro attorno alle decisioni. Quanto al modello agentico-ibrido, la presente analisi ne circoscrive la possibile appropriatezza a processi interni delimitati e successivi alla costruzione di adeguati assetti di governance del dato, autorizzazione e registrazione degli eventi.

3.9 Verso le competenze: le implicazioni dei modelli

L'analisi condotta consente di anticipare, in chiusura, il tema del capitolo conclusivo. Ciascuna delle configurazioni esaminate genera un fabbisogno di competenze diverso non solo per intensità ma per contenuto ed è l'implicazione più rilevante dell'intero capitolo.

Il modello centralizzato concentra la domanda di competenze tecniche specialistiche in un'unità ristretta, ma genera un fabbisogno diffuso di competenze di interfaccia: la capacità, nelle unità operative, di formulare un problema in termini trattabili e di valutare criticamente una soluzione prodotta altrove. Il modello decentrato richiede una diffusione ampia delle competenze tecniche, con i costi formativi e di reperimento che ne conseguono. Il modello *hub and spoke* richiede competenze di coordinamento e di governo delle relazioni fra le strutture, che sono di natura manageriale e non tecnica, e pone ai professionisti collocati negli *spoke* il problema identitario segnalato al § 3.5.3. Il modello esternalizzato, che a un esame superficiale parrebbe il meno esigente, richiede in realtà competenze particolarmente sofisticate di natura valutativa: la capacità di specificare requisiti, di selezionare fornitori, di verificare esiti che non si è contribuito a produrre e di esercitare una sorveglianza informata su sistemi dei quali non si conosce il funzionamento interno.

I modelli di aggregazione del lavoro aggiungono a questo quadro fabbisogni ulteriori e di natura diversa: il modello a piattaforma e prodotto che richiede la figura del responsabile di prodotto, il quale combina conoscenza del dominio, capacità di definizione delle priorità e comprensione dei limiti tecnici del sistema; il modello predittivo richiede competenze di analisi e di scomposizione delle decisioni, più vicine all'analisi organizzativa che all'informatica; il modello agentico-ibrido richiede una competenza nuova, consistente nel saper delimitare il perimetro di azione di un'entità non umana e nel sorvegliarne l'operato su sequenze anziché su singoli esiti; il modello ecosistemico richiede capacità di gestione delle relazioni interorganizzative e di orientamento di scelte collettive. Si tratta, in tutti i casi, di competenze ibride, che non coincidono né con il profilo tecnico né con quello di business tradizionale: è un punto sul quale il capitolo conclusivo tornerà in modo sistematico.

Attraversa tutte le configurazioni un nucleo comune, che il secondo capitolo ha già individuato sul piano normativo con l'obbligo di alfabetizzazione all'intelligenza artificiale e che il § 3.7 ha mostrato essere anche una condizione di funzionamento: la capacità di comprendere che cosa un sistema misuri, su quali dati sia stato costruito e in quali circostanze tenda a errare. Senza tale capacità, il presidio umano si riduce a un adempimento e la struttura formale, quale che sia, resta priva di effettività. Colbert, Yee e George (2016) avevano già

osservato come le organizzazioni siano chiamate a riprogettare il lavoro per valorizzare le competenze digitali attenuandone al contempo gli effetti indesiderati: la questione non nasce con l'adozione dell'intelligenza artificiale ma da essa viene resa più stringente.

3.10 Considerazioni conclusive

Il capitolo ha affrontato la parte della domanda di ricerca relativa alla trasformazione dei modelli organizzativi, muovendo dalla constatazione, emersa nei capitoli precedenti, che la dotazione tecnologica non basta a generare valore.

L'analisi ha stabilito, in primo luogo, che l'IA pone problemi organizzativi in parte nuovi, riconducibili all'autonomia, all'apprendimento e all'inscrutabilità dei sistemi e che tali problemi riaprono simultaneamente le questioni fondamentali della progettazione organizzativa. Ha mostrato, in secondo luogo, che l'adozione presuppone condizioni di prontezza in larga parte non tecnologiche e che la capacità organizzativa si costruisce nel tempo attraverso investimenti complementari di natura intangibile.

Il confronto sistematico fra le otto configurazioni esaminate ha evidenziato che nessuna di esse è superiore in assoluto e che, fra i modelli di allocazione della competenza, la variabile discriminante è l'accesso alla conoscenza del dominio applicativo, da bilanciare con i requisiti di uniformità e di controllo. L'ampliamento ai modelli di aggregazione del lavoro – a piattaforma e prodotto, predittivo, agentico-ibrido ed ecosistemico – ha consentito di mostrare che le due scelte sono indipendenti e che una parte rilevante delle difficoltà osservate nei programmi di adozione non deriva dalla scelta di un modello sbagliato, ma dall'incoerenza tra il luogo in cui la competenza è collocata e il criterio secondo cui il lavoro è aggregato. L'esame della dimensione di governance ha mostrato che il quadro regolamentare incide sulla progettazione organizzativa richiedendo un assetto formalizzato e documentato e che per gli intermediari bancari tale assetto costituisce un'estensione dell'impianto dei controlli esistente più che una costruzione nuova.

Il passaggio più rilevante è quello del § 3.7: la struttura formale non determina il funzionamento. L'efficacia di una configurazione dipende da come sono progettati gli incentivi

di chi esercita la sorveglianza, da chi valida i sistemi e con quale indipendenza, dal riconoscimento del lavoro continuo di manutenzione e dalla ridefinizione esplicita dei contenuti professionali dei ruoli incisi. Sono tutti aspetti che rimandano alle persone e che spiegano perché la trasformazione organizzativa non possa essere separata dall'evoluzione delle competenze.

Su questa constatazione si innesta il capitolo conclusivo, dedicato al ruolo dell'uomo nell'organizzazione AI-driven. Se ogni modello genera un fabbisogno di competenze specifico e nessun modello funziona senza un presidio umano effettivo, la ridefinizione delle competenze non è una conseguenza della trasformazione organizzativa, ma la condizione della sua riuscita.

Capitolo IV – Il ruolo dell'uomo nell'organizzazione AI-driven: evoluzione delle competenze

4.1 Introduzione

Le analisi condotte finora convergono verso una medesima conclusione: il valore dell'intelligenza artificiale dipende meno dalla tecnologia in sé che dalle condizioni organizzative che ne rendono possibile l'utilizzo efficace. Il primo capitolo ha ricostruito l'evoluzione dell'IA e ne ha qualificato la natura di tecnologia a uso generale, il cui valore non risiede nelle prestazioni degli algoritmi ma nella capacità delle organizzazioni di integrarli. Il secondo ha verificato tale premessa nel settore bancario, mostrando che l'adozione è ormai diffusa ma che i benefici non sono automatici e che il presidio umano, lungi dall'essere un adempimento formale, dipende da variabili organizzative quali competenza, autorità, tempo e tracciabilità. Il terzo ha analizzato le configurazioni organizzative attraverso le quali l'adozione può essere governata, concludendo con una constatazione che costituisce il punto di partenza del presente capitolo: la struttura formale non determina il funzionamento e ciascun modello genera un fabbisogno di competenze diverso.

La domanda di ricerca che guida l'elaborato chiedeva come l'adozione dell'intelligenza artificiale nel settore bancario richieda una trasformazione dei modelli organizzativi e in che modo tale trasformazione ridefinisca le competenze umane necessarie. La prima parte della domanda ha trovato risposta nel terzo capitolo mentre la seconda parte è l'oggetto di quest'ultimo. L'impostazione che se ne è data merita di essere esplicitata: la ridefinizione delle competenze non è una conseguenza della trasformazione organizzativa ma la condizione della sua riuscita, e la sua direzione non è univoca, poiché varia a seconda della configurazione organizzativa adottata.

Il capitolo si articola in sette passaggi. Il § 4.2 ricostruisce l'evoluzione del lavoro nel settore bancario, a partire dal quadro teorico dell'economia dell'automazione e dagli effetti di sostituzione, di reintegrazione e di ricomposizione dei ruoli. Il § 4.3 affronta la collaborazione fra persone e sistemi intelligenti, esaminandone i fondamenti economici, le configurazioni concrete e le evidenze empiriche disponibili, comprese quelle di segno critico. Il § 4.4

individua e classifica le competenze richieste dall'organizzazione *AI-driven*, raccordandole alle otto configurazioni analizzate nel terzo capitolo, mentre il § 4.5 esamina i ruoli professionali emergenti e la trasformazione di quelli esistenti. Nella parte finale, il § 4.6 tratta le politiche di sviluppo delle competenze e le condizioni della loro efficacia, il § 4.7 spiega i rischi e le questioni aperte e il § 4.8 formula la risposta complessiva alla domanda di ricerca, ne esplicita i limiti e indica alcune prospettive di approfondimento.

Una precisazione metodologica riguarda il trattamento delle previsioni. La pubblicistica sull'impatto dell'intelligenza artificiale sul lavoro abbonda di stime quantitative relative al numero di posti di lavoro che sarebbero destinati a scomparire o a trasformarsi. Tali stime hanno un valore informativo limitato e un elevato potenziale fuorviante. Nell'analisi che segue si è pertanto scelto di fondare l'argomentazione sulla letteratura economica e organizzativa sottoposta a revisione tra pari e sulle rilevazioni istituzionali, riportando i dati di fonte aziendale o professionale soltanto ove non esistano equivalenti e dichiarandone espressamente la natura.

4.2 L'evoluzione del lavoro nel settore bancario

4.2.1 Il quadro teorico: compiti, sostituzione e reintegrazione

Per analizzare in modo ordinato gli effetti dell'IA sul lavoro occorre disporre di un quadro teorico che eviti tanto la previsione catastrofica quanto la rassicurazione generica. La letteratura economica ne offre uno consolidato, fondato su un'intuizione apparentemente semplice: l'unità rilevante di analisi non è la professione ma il compito. Autor, Levy e Murnane (2003) hanno mostrato che le tecnologie informatiche non sostituiscono le professioni in quanto tali, ma le attività che possono essere descritte mediante regole esplicite; ne consegue che l'effetto su una professione dipende dalla composizione dei compiti che essa comprende e che la stessa tecnologia può ridurre la domanda di alcune attività e accrescere quella di altre all'interno del medesimo ruolo.

Su questa base Acemoglu e Restrepo (2018; 2019) hanno costruito un quadro analitico più articolato, nel quale la tecnologia produce simultaneamente tre effetti. Il primo è l'effetto

di sostituzione (*displacement*): le macchine subentrano all'uomo in compiti precedentemente svolti da lavoratori, riducendo la domanda di lavoro. Il secondo è l'effetto di produttività: la riduzione dei costi accresce la domanda di beni e servizi e, indirettamente, quella di lavoro nei compiti non automatizzati. Il terzo, e il più rilevante ai fini del presente capitolo, è l'effetto di reintegrazione (*reinstatement*): la tecnologia genera compiti nuovi, nei quali il lavoro umano conserva un vantaggio comparato e che prima non esistevano. Gli autori sottolineano che l'esito netto dipende dall'equilibrio fra questi effetti e non è determinato dalla tecnologia in sé: periodi storici caratterizzati da forte automazione hanno conosciuto sia crescita sia stagnazione della domanda di lavoro, in funzione dell'intensità con cui si generavano compiti nuovi.

Autor (2015) aggiunge un'osservazione che ha valore di avvertenza metodologica: le previsioni sulla scomparsa del lavoro hanno accompagnato ogni ondata di automazione senza realizzarsi, e la ragione risiede in due fattori sottovalutati: la complementarità fra compiti, per cui l'automazione di una parte di un processo accresce il valore delle fasi che restano umane e la difficoltà di codificare il sapere tacito, ossia il know-how che le persone utilizzano senza poterlo enunciare in forma di regola.

Tuttavia, l'IA generativa introduce un elemento di discontinuità rispetto a questo quadro, che è stato costruito osservando tecnologie precedenti. Se l'automazione tradizionale incideva sui compiti routinari e di media difficoltà, i sistemi generativi incidono su compiti cognitivi non routinari e di qualificazione elevata – la redazione di testi, la sintesi documentale, l'analisi – che il quadro classico collocava fra quelli protetti. Tschang e Almirall (2021) hanno analizzato questa discontinuità, osservando che l'IA opera come automazione che potenzia (*augmenting automation*) e che i suoi effetti occupazionali si distribuiscono in modo diverso da quelli delle tecnologie precedenti, ovvero interessano segmenti professionali finora non esposti. La conseguenza non è che il quadro teorico vada abbandonato, ma che la sua applicazione richieda cautela: l'esposizione all'automazione non coincide più con la ripetitività del compito.

4.2.2 Le evidenze empiriche sull'automazione

Sul piano empirico, l'evidenza più importante riguarda le tecnologie di automazione fisica: Acemoglu e Restrepo (2020) hanno analizzato l'introduzione dei robot industriali nei mercati del lavoro locali statunitensi, documentando effetti negativi significativi su occupazione e salari nelle aree maggiormente esposte, senza che tali effetti risultino compensati da guadagni in altri settori nel periodo osservato. Il risultato è rilevante perché mostra che gli effetti di sostituzione e di reintegrazione non si compensano necessariamente nel breve periodo né uniformemente nello spazio: anche quando l'esito aggregato è neutro, la distribuzione dei costi e dei benefici può essere fortemente diseguale.

Per le tecnologie cognitive l'evidenza è più recente e di natura diversa, poiché deriva in prevalenza da studi sperimentali o quasi sperimentali sulla produttività individuale più che da analisi di mercato del lavoro. Il secondo capitolo ne ha già dato conto al § 2.4.4 e conviene qui richiamarne il contenuto: Brynjolfsson, Li e Raymond (2025) e Noy e Zhang (2023) rilevano entrambi che i guadagni di produttività associati all'impiego di assistenti generativi si concentrano sui lavoratori meno esperti, mentre risultano modesti o nulli per i più esperti. Se il risultato si confermasse su scala più ampia, avrebbe un'implicazione importante per il tema delle competenze: la tecnologia agirebbe da livellatore delle prestazioni, comprimendo il differenziale fra chi possiede e chi non possiede l'esperienza. Occorre effettuare due considerazioni in merito: da un lato, essa abbassa la soglia di accesso a compiti complessi e accresce la produttività dei profili meno esperti; dall'altro, riduce il rendimento dell'esperienza accumulata e, come si vedrà al § 4.3.4, può compromettere il meccanismo attraverso il quale l'esperienza si forma.

A queste evidenze si affianca la prospettiva delle organizzazioni internazionali. L'OCSE, in una rassegna dedicata all'impatto dell'intelligenza artificiale sul mercato del lavoro, osserva che l'evidenza disponibile non mostra, allo stato, effetti negativi generalizzati sull'occupazione aggregata, ma segnala una ricomposizione della domanda di competenze e una crescente esposizione dei lavoratori altamente qualificati, con implicazioni rilevanti per le politiche formative (Lane e Saint-Martin, 2021). L'ente ha dedicato al tema delle

competenze necessarie alla transizione digitale un'analisi specifica, fondata sull'esame di dati di grandi dimensioni relativi agli annunci di lavoro (OECD, 2023).

4.2.3 La specificità del lavoro bancario

Applicare questa analisi al settore bancario richiede di considerarne alcune specificità, che emergono dai capitoli precedenti.

La prima è che il lavoro bancario è storicamente esposto all'automazione e che tale esposizione non costituisce una novità. Come ricostruito al § 2.2.1, il settore ha attraversato più ondate di digitalizzazione: l'informatizzazione della contabilità, gli sportelli automatici e i canali digitali hanno ridotto il peso di specifici compiti esecutivi. Le evidenze disponibili mostrano anche effetti strutturali: per le banche italiane la maggiore digitalizzazione è associata a una minore articolazione della rete fisica, mentre la vigilanza europea rileva cambiamenti nel front office, nel back office, nei modelli operativi e nei profili di competenza (Arnaudo et al., 2022; European Central Bank Banking Supervision, 2024). Tali risultati non autorizzano a sostenere che l'occupazione sia rimasta invariata; suggeriscono però che l'effetto non sia descrivibile come semplice eliminazione del lavoro. Sulla base del quadro per compiti di Acemoglu e Restrepo (2019), la presente analisi interpreta l'evoluzione come una ricomposizione del contenuto professionale, con riduzione di attività esecutive e crescita relativa di compiti consulenziali, commerciali e di controllo.

La seconda specificità è che la composizione attuale dell'adozione indica con precisione dove il lavoro si stia modificando. L'Indagine Fintech della Banca d'Italia (2025) rileva che poco meno della metà dei progetti censiti insiste sulle attività operative e che l'intelligenza artificiale generativa vi è impiegata in via prevalente, mentre nell'intermediazione del credito prevalgono modelli predittivi basati su tecniche di apprendimento automatico. Il risultato è una mappa dell'esposizione: le attività di raccolta, riconciliazione ed elaborazione documentale sono quelle su cui l'effetto di sostituzione agisce più intensamente; le attività di valutazione e decisione sono quelle in cui prevale la trasformazione del contenuto, poiché il sistema fornisce una previsione che la persona deve interpretare, validare e, ove necessario, disattendere.

La terza specificità è di natura regolamentare e costituisce un tratto distintivo del settore rispetto ad altri. Come illustrato ai §§ 2.6 e 2.7, l'impiego di sistemi ad alto rischio – categoria nella quale rientra la valutazione dell'affidabilità creditizia delle persone fisiche – è subordinato a obblighi di sorveglianza umana e l'articolo 26 dell'AI Act pone a carico dell'utilizzatore professionale l'obbligo di affidarne l'esercizio a persone fisiche dotate della necessaria competenza, formazione e autorità. Ne discende una conseguenza che merita di essere resa esplicita, perché differenzia il settore bancario da contesti meno regolamentati: la normativa istituisce una domanda di lavoro umano che non deriva da ragioni di efficienza. Anche laddove fosse tecnicamente possibile ed economicamente conveniente automatizzare integralmente una decisione, il quadro normativo ne impedisce l'esito e richiede il mantenimento di un presidio umano qualificato. Il presidio, in altri termini, non è un residuo destinato a essere eroso dal progresso tecnico, ma una funzione stabile, la cui qualità dipende dalle competenze di chi la esercita.

4.2.4 Dalla mansione al compito: la ricomposizione dei ruoli

La conseguenza organizzativa di questa analisi è che l'effetto dell'intelligenza artificiale sul lavoro bancario non è descrivibile in termini di professioni che scompaiono e professioni che nascono, ma in termini di ricomposizione del contenuto dei ruoli esistenti. Il fenomeno può essere descritto lungo tre direttrici.

La prima è la sottrazione della componente esecutiva nel quale le attività di raccolta, inserimento, riconciliazione e trasposizione di informazioni – che occupano una quota rilevante del tempo di lavoro in molte funzioni di supporto – sono quelle sulle quali l'automazione incide più direttamente. La loro riduzione libera capacità ma pone un problema di riallocazione che non si risolve automaticamente: il tempo liberato produce valore soltanto se viene indirizzato ad attività a maggior valore aggiunto, il che presuppone che le persone possiedano le competenze per svolgerle e che l'organizzazione ne abbia ridefinito il ruolo.

La seconda è l'aggiunta della componente di supervisione, come mostrato al § 3.3.2 sulla scorta di Grønsund e Aanestad (2020), l'introduzione di sistemi intelligenti non riduce il lavoro umano necessario ma ne genera di nuovo: la verifica degli esiti, la costruzione e la

manutenzione del riferimento di correttezza, la segnalazione delle anomalie, l'alimentazione del circuito di miglioramento. Si tratta di attività che nella maggior parte dei casi non compaiono nei documenti di progetto e che, come osservato al § 3.7.3, tendono a restare invisibili: la loro esplicitazione è una condizione sia di efficacia del presidio sia di riconoscimento professionale di chi lo esercita.

La terza è lo spostamento verso il giudizio, in quanto nei compiti in cui il sistema fornisce una previsione, il contributo umano si concentra su ciò che la previsione non contiene: la valutazione delle conseguenze, la considerazione degli elementi non codificati, la gestione delle situazioni che si discostano dai casi su cui il sistema è stato costruito. È la logica che il § 3.5.6 ha descritto come fondamento del modello *decision-centric* e che qui si traduce in un'indicazione sul contenuto del lavoro: la componente di giudizio, che nel modello tradizionale era distribuita su tutte le pratiche, tende a concentrarsi sulle situazioni non standard, con la conseguenza che il lavoro diviene mediamente più complesso e meno ripetitivo, ma anche più esigente in termini di competenza.

Mettendo insieme le tre direttrici si delinea un profilo di lavoro bancario nel quale la quota di attività esecutive si riduce, quella di supervisione e di giudizio si accresce e la relazione con il cliente – componente che i sistemi possono supportare ma non sostituire nella sua dimensione fiduciaria – conserva un rilievo centrale. È un profilo coerente con l'esito delle precedenti ondate di digitalizzazione, ma con una differenza rilevante: le ondate precedenti hanno sottratto compiti manuali e amministrativi, questa sottrae compiti cognitivi e la ricomposizione che ne deriva incide su segmenti professionali più qualificati.

4.3 La collaborazione fra uomo e intelligenza artificiale

4.3.1 Il fondamento economico della complementarità

La nozione di collaborazione fra uomo e macchina è spesso invocata in termini retorici e per renderla analiticamente utile è necessario chiarire in che cosa consista la complementarità. La formulazione più rigorosa è quella già richiamata al § 3.5.6: l'IA riduce il costo della previsione, ossia dell'operazione con cui, a partire dall'informazione disponibile, si

colma il dato mancante; poiché la previsione è soltanto una delle componenti della decisione, alla quale concorrono anche il giudizio sulle conseguenze, i dati e l'azione, la riduzione del costo di una componente accresce il valore relativo delle altre (Agrawal, Gans e Goldfarb, 2019).

La complementarità è una proposizione economica da cui discendono conseguenze precise. La prima è che il valore del giudizio umano aumenta, anziché diminuire, nella misura in cui la previsione diviene abbondante e a basso costo. La seconda è che tale aumento riguarda una capacità specifica: la determinazione di ciò che conta, ossia della funzione obiettivo rispetto alla quale la previsione è utile. La terza, e più rilevante per il presente capitolo, è che la complementarità non si realizza automaticamente: richiede che l'organizzazione ridisegni il compito in modo da assegnare alla persona la componente in cui il suo contributo è effettivamente superiore, il che riporta alla scelta di progettazione descritta al § 3.3.2.

L'evidenza empirica conferma il carattere condizionato della complementarità: Choudhury, Starr e Agarwal (2020), attraverso un disegno sperimentale, mostrano che l'apporto dell'apprendimento automatico alla qualità delle valutazioni dipende dalla combinazione con il capitale umano specifico degli operatori e che la tecnologia risulta particolarmente efficace nell'attenuare distorsioni sistematiche del giudizio individuale. Il risultato indica che la relazione tra sistema e persona dipende da quali distorsioni ciascuno dei due porterà nel processo e da come il compito è pensato.

4.3.2 Le configurazioni della collaborazione e i loro esiti

Il secondo capitolo ha distinto, al § 2.7.1, tre configurazioni del presidio umano: *human in the loop*, nella quale l'esito non diviene operativo senza convalida; *human on the loop*, nella quale il sistema opera in autonomia e la persona ne sorveglia il funzionamento; *human in command*, nella quale la persona decide se e con quali limiti il sistema sia impiegato. Il terzo capitolo ha mostrato, al § 3.5.7, che l'evoluzione verso i sistemi agentici introduce una quarta configurazione, nella quale persone e agenti digitali compongono un'unità di lavoro con una divisione dei compiti esplicita e una responsabilità integralmente umana. La ricerca sui sistemi intelligenti come partecipanti al lavoro di gruppo (Seeber et al., 2020) ha

individuato in questa direzione un'agenda di ricerca specifica, sottolineando che la progettazione dell'interazione – tra chi propone, chi decide, chi verifica, come le informazioni circolano fra i partecipanti – costituisce una variabile determinante dell'esito.

Una rassegna sistematica degli studi che confrontano le prestazioni di combinazioni uomo-macchina con quelle dell'uomo da solo e del sistema da solo rileva che, in media, le combinazioni non superano il migliore dei due componenti presi singolarmente; la sinergia emerge in condizioni determinate – in particolare nei compiti di creazione di contenuti e quando la persona, da sola, supera le prestazioni del sistema – mentre nei compiti di decisione la combinazione tende a produrre risultati inferiori a quelli del componente migliore (Vaccaro, Almaatouq e Malone, 2024). Il risultato non nega la complementarità teorica, ma mostra che essa non si realizza per il solo fatto di affiancare una persona a un sistema: dipende dalla progettazione del compito e dalla capacità della persona di riconoscere quando delegare e quando mantenere completamente il presidio sull'attività.

È esattamente il punto sul quale il secondo capitolo aveva già raccolto evidenze convergenti: Green e Chen (2019) rilevano che i supervisori non correggono in modo affidabile gli errori del modello e ne adottano le indicazioni in modo selettivo; Dell'Acqua e altri (2026) documentano un peggioramento delle prestazioni su compiti collocati appena oltre la frontiera delle capacità del sistema. Le tre evidenze, lette insieme, conducono alla medesima conclusione: la variabile determinante non risiede nella mera presenza dell'essere umano all'interno del processo decisionale, bensì nella sua capacità di valutare criticamente e calibrare il grado di affidamento riposto nel sistema.

4.3.3 Fiducia, affidamento e calibrazione

La letteratura sui fattori umani nell'automazione ha analizzato in profondità questa capacità, definendola in termini di affidamento appropriato. Lee e See (2004) mostrano che la fiducia nei sistemi automatizzati è un costrutto dinamico, influenzato dalle prestazioni osservate, dalla comprensione del funzionamento e dal contesto e che l'obiettivo della progettazione non è massimizzare la fiducia ma allinearla alle effettive capacità del sistema. Un affidamento eccessivo produce i fenomeni di compiacenza e di distorsione dell'automazione già

descritti al § 2.7.5 sulla scorta di Parasuraman e Manzey (2010); un affidamento insufficiente produce il fenomeno speculare dell'avversione all'algoritmo, documentato da Dietvorst, Simmons e Massey (2015), per cui gli utenti abbandonano un sistema dopo averlo visto sbagliare anche quando esso resta superiore al giudizio umano e non ne vengono comprese le reali funzionalità.

La calibrazione dell'affidamento è la competenza più innovativa richiesta dall'organizzazione *AI-driven* ed è utile approfondirne il contenuto. Essa non consiste nel conoscere il funzionamento interno del modello, conoscenza che nella maggior parte dei casi è preclusa e che non sarebbe utilizzabile nel tempo della decisione. Consiste piuttosto nel possedere una rappresentazione operativa dei casi in cui il sistema tende a sbagliare: quali situazioni sono sottorappresentate nei dati su cui è stato costruito, quali categorie di clientela o di operazione ricadono ai margini della sua esperienza, quali segnali nell'output indicano una situazione non standard. È una conoscenza di natura pratica, che si forma attraverso l'uso e il confronto sistematico fra previsione ed esito e che non può essere trasmessa esclusivamente attraverso la formazione d'aula.

Ne discende un'indicazione organizzativa precisa, che raccorda questo paragrafo al terzo capitolo: la calibrazione dell'affidamento presuppone la restituzione dell'esito: se chi esercita la sorveglianza non riceve informazione sistematica su come si sono evoluti i casi che ha convalidato o disatteso, non dispone del materiale su cui costruire la propria rappresentazione degli errori del sistema e il suo affidamento resta necessariamente non calibrato.

4.3.4 Il paradosso dell'apprendimento

Resta da affrontare la questione più delicata del capitolo, che costituisce il punto di massima tensione tra efficienza di breve periodo e sostenibilità di lungo periodo. La letteratura sull'automazione si interroga da tempo su questo tema: Bainbridge (1983), in un contributo divenuto classico, ha osservato che l'automazione dei compiti più semplici lascia all'operatore i soli compiti difficili, sottraendogli le occasioni di pratica attraverso le quali si acquisisce e si mantiene la capacità di affrontarli. L'autrice definì tale circostanza una delle

«ironie» dell'automazione: si chiede alla persona di intervenire proprio nelle situazioni eccezionali, dopo averle tolto il modo di allenarsi su quelle ordinarie.

Il meccanismo si applica con particolare evidenza al lavoro bancario. La competenza dell'analista si forma esaminando un numero elevato di pratiche, la maggior parte delle quali ordinarie: è dall'accumulazione di casi che si costruisce la capacità di riconoscere l'anomalia. Se il sistema istruisce autonomamente le pratiche ordinarie e sottopone alla persona soltanto le eccezioni, la persona si trova a dover giudicare i casi difficili senza aver costruito il riferimento implicito rispetto al quale essi sono difficili. Lo stesso vale per l'addetto ai controlli antiriciclaggio che esamini soltanto gli allarmi selezionati dal modello o per il gestore che riceva proposte commerciali già formulate e dettagliate.

Il fenomeno è la manifestazione, sul piano delle competenze individuali, del paradosso che Raisch e Krakowski (2021) descrivono sul piano organizzativo: l'automazione perseguita in via esclusiva erode le competenze che sarebbero necessarie a governarla. La sua rilevanza è duplice. Sul piano dell'efficacia, esso mostra che il presidio umano richiesto dalla normativa può divenire progressivamente inconsistente non per negligenza ma per effetto del disegno del processo; sul piano della sostenibilità, esso pone un problema di ricambio generazionale: se i percorsi attraverso i quali si formavano i professionisti esperti passavano per le attività ora automatizzate, occorre costruire percorsi alternativi, poiché la competenza necessaria alla supervisione non si genera più come sottoprodotto del lavoro ordinario.

Le soluzioni organizzative possibili sono di tre ordini: la prima consiste nel mantenere deliberatamente una quota di casi ordinari trattati in modo non assistito con finalità formative, una scelta che comporta un costo di efficienza esplicito. La seconda consiste nel costruire percorsi di formazione basati su casi, che ricostruiscano in ambiente protetto l'esperienza che il processo non genera più. La terza consiste nell'organizzare in modo sistematico la restituzione degli esiti, trasformando la supervisione stessa in occasione di apprendimento. Nessuna delle tre è priva di costi e questa è precisamente la ragione per cui la questione va posta come scelta strategica e non come problema formativo.

4.4 Le competenze richieste dall'organizzazione AI-driven

4.4.1 Un criterio di classificazione

L'analisi condotta consente ora di individuare le competenze richieste. Per evitare l'elencazione generica che caratterizza gran parte della pubblicistica sul tema, si adotta un criterio di classificazione fondato sull'analisi dei capitoli precedenti: le competenze sono raggruppate in base alla funzione che svolgono nel processo di adozione, distinguendo fra competenze di comprensione, che consentono di sapere che cosa il sistema faccia; competenze di valutazione, che consentono di giudicarne gli esiti; competenze di relazione, che presidiano la componente dell'attività che il sistema non può svolgere e competenze di governo, che consentono di decidere se, dove e con quali limiti impiegarlo. A queste si affiancano le competenze tecniche specialistiche, la cui intensità dipende dalla configurazione organizzativa adottata e che, per le ragioni esposte nel terzo capitolo, non costituiscono il fabbisogno prevalente nella maggior parte degli intermediari.

4.4.2 Le competenze di comprensione: l'alfabetizzazione all'intelligenza artificiale

Il primo gruppo corrisponde a ciò che la letteratura definisce *AI literacy* e che il quadro normativo europeo ha reso oggetto d'obbligo. Long e Magerko (2020) la definiscono come l'insieme delle competenze che consentono a una persona di valutare criticamente le tecnologie di IA, di comunicare e collaborare efficacemente con esse e di utilizzarle come strumento in contesti diversi; ne articolano il contenuto in competenze relative al riconoscimento dei sistemi, alla comprensione del loro funzionamento generale, alla consapevolezza dei loro limiti e alle implicazioni etiche del loro impiego.

Sul piano normativo, l'articolo 4 dell'AI Act – applicabile, come illustrato al § 2.6.3, già dal 2 febbraio 2025 e tra le disposizioni non interessate dal differimento – impone a fornitori e deployer di adottare misure per sostenere lo sviluppo dell'alfabetizzazione in materia di intelligenza artificiale del proprio personale e delle altre persone che si occupano del funzionamento e dell'utilizzo dei sistemi per loro conto. Le misure devono tenere conto delle conoscenze tecniche, dell'esperienza, dell'istruzione e della formazione di tali soggetti, del

contesto d'uso e delle persone o dei gruppi sui quali i sistemi sono destinati a essere utilizzati. La formulazione vigente precisa espressamente che l'obbligo non richiede di garantire uno specifico livello di alfabetizzazione per ciascun individuo. Ne consegue che la norma vincola alla predisposizione di misure contestualizzate, ma non prescrive un esito uniforme né impone testualmente un determinato modello formativo; la differenziazione dei percorsi costituisce una modalità organizzativa ragionevole per dare attuazione ai fattori che l'articolo richiede di considerare.

Il contenuto minimo di questa competenza, nel contesto bancario, può essere articolato in quattro elementi: sapere riconoscere quando un esito è prodotto o assistito da un sistema di IA; comprendere in termini generali su quali dati e con quale logica il sistema sia costruito; conoscerne i limiti tipici, inclusa la possibilità di risposte plausibili ma errate nei sistemi generativi e il rischio di distorsione nei sistemi predittivi; conoscere le regole interne d'uso e i confini entro cui lo strumento è ammesso. Va sottolineato che si tratta di un requisito di base e non di una competenza specialistica: la sua funzione è rendere possibile l'uso consapevole.

4.4.3 Le competenze di valutazione: giudizio critico e validazione

Nella prospettiva del presente elaborato, il nucleo qualificante del fabbisogno è costituito da un secondo gruppo di competenze, articolato in tre capacità distinte.

La prima è la calibrazione dell'affidamento, descritta al § 4.3.3, che consente di stabilire quando accettare e quando mettere in discussione l'esito di un sistema. La seconda è la capacità di validazione, che il § 3.7.2 ha individuato come condizione di affidabilità dell'intero impianto: la capacità di verificare se un sistema produca risultati corretti rispetto a un riferimento a sua volta costruito, e quindi di interrogarsi sulla qualità di tale riferimento, secondo l'avvertenza di Lebovitz, Levina e Lifshitz-Assaf (2021). La terza è la capacità di individuazione dei requisiti, ossia di tradurre un problema operativo in termini che ne consentano il trattamento e di stabilire in anticipo quali esiti siano accettabili: competenza che assume rilievo decisivo nel modello esternalizzato, dove costituisce l'unico strumento attraverso cui l'intermediario orienta una soluzione che non costruisce.

Vale la pena osservare che queste tre capacità hanno una caratteristica comune: presuppongono la conoscenza del dominio applicativo. Non è possibile calibrare l'affidamento su un sistema di scoring senza conoscere il credito, né validare un esito senza sapere che cosa si sarebbe dovuto ottenere. Ne discende un'indicazione controcorrente rispetto a una diffusa enfasi sulle competenze tecniche: la competenza più scarsa nell'organizzazione *AI-driven* non è quella di chi sa costruire modelli, ma quella di chi conosce il processo abbastanza bene da giudicare se il modello stia funzionando. È, del resto, l'esatta traduzione sul piano delle persone del risultato emerso dal confronto fra i modelli al § 3.5.9, secondo cui la variabile discriminante è l'accesso alla conoscenza del dominio applicativo.

4.4.4 Le competenze di relazione

Il terzo gruppo riguarda la componente dell'attività bancaria che i sistemi supportano ma non sostituiscono. Il secondo capitolo ha documentato, al § 2.4.3, che i benefici dei servizi di consulenza automatizzata dipendono in misura decisiva dal profilo di partenza del cliente (D'Acunto, Prabhala e Rossi, 2019) e che permane una diffusa resistenza ad affidarsi a raccomandazioni algoritmiche. Ne consegue che il ruolo del professionista è spostato verso la gestione delle situazioni complesse, l'interpretazione dei bisogni non espressi e la costruzione della fiducia.

La rilevanza economica di questa componente non è congetturale in quanto Deming (2017) documenta la crescente importanza delle competenze sociali nel mercato del lavoro, mostrando che la loro domanda è cresciuta in modo importante e che esse risultano complementari e non alternative rispetto alle competenze cognitive: le professioni che richiedono entrambe hanno registrato la crescita più marcata. Applicata al settore bancario, l'evidenza suggerisce che la combinazione fra conoscenza tecnica del prodotto e capacità relazionale costituisca il profilo a maggiore valore e che la riduzione della componente esecutiva del lavoro renda tale combinazione più accessibile in termini di tempo disponibile.

Va aggiunto un elemento che discende direttamente dal quadro normativo. L'articolo 50 dell'AI Act impone che l'interlocutore sia informato di star interagendo con un sistema di intelligenza artificiale e l'articolo 22 del GDPR riconosce all'interessato il diritto di ottenere

l'intervento umano e di contestare la decisione automatizzata. Ne deriva una competenza specifica, oggi poco presidiata: la capacità di spiegare al cliente il ruolo che il sistema ha avuto nella decisione che lo riguarda, in termini comprensibili e senza attribuire alla macchina una responsabilità che resta dell'intermediario. È una competenza a metà fra la conoscenza tecnica, la comunicazione e la conformità e la sua assenza costituisce un rischio reputazionale e di contenzioso.

4.4.5 Le competenze di governo

Il quarto gruppo riguarda i ruoli di responsabilità e comprende la capacità di valutare la prontezza organizzativa secondo i fattori individuati al § 3.4.1; quella di selezionare i casi d'uso in funzione del contributo atteso e non della disponibilità tecnologica; quella di scegliere e far evolvere la configurazione organizzativa, che il § 3.5.9 ha qualificato come capacità dinamica; quella di presidiare i rischi, incluso il rischio di terze parti, e di comporre gli obblighi derivanti da discipline sovrapposte; e quella di progettare i sistemi di misurazione delle prestazioni tenendo conto degli effetti descritti al § 3.7.1, ossia evitando che la sorveglianza umana si converta in un costo per chi la esercita.

Su questo gruppo di competenze è utile un'osservazione relativa alla diffusione osservata. Il dato istituzionale richiamato al § 2.3 – secondo cui un terzo degli intermediari, di cui circa la metà banche, ha istituito una struttura dedicata alla governance dell'intelligenza artificiale (Banca d'Italia, 2025) – indica che la funzione di governo si sta formalizzando, ma il dato complementare, ossia che due terzi non l'hanno fatto, indica che nella maggioranza dei casi tali competenze restano distribuite fra ruoli esistenti senza un'attribuzione esplicita. È una condizione che il terzo capitolo ha individuato come fonte tipica di lacune nei controlli.

4.4.6 Il raccordo fra competenze e modelli organizzativi

Il punto centrale di questo paragrafo, e la ragione per cui esso segue anziché precedere l'analisi organizzativa, è che il fabbisogno di competenze non è indipendente dalla configurazione adottata. La tabella seguente raccorda i due piani, riprendendo le otto configurazioni analizzate al § 3.5.

Configurazione	Competenza prevalente	Dove è collocata	Rischio di competenza
Centralizzato	Tecnica specialistica concentrata; di interfaccia diffusa	Unità dedicata; competenze di interfaccia diffuse nelle unità operative	Isolamento dell'unità centrale rispetto al processo
Decentrato	Tecnica diffusa; di validazione	Nelle singole unità di business	Disomogeneità qualitativa; validazione non indipendente
Hub and spoke	Di coordinamento e di traduzione fra centro e unità	Ripartita fra *hub* e *spoke*	Ambiguità di ruolo e identità professionale ambivalente negli *spoke*
Esternalizzato	Di valutazione: requisiti, selezione, verifica, sorveglianza	Nucleo interno ristretto; funzioni committenti	Erosione della capacità di valutare ciò che si acquista
Piattaforma e prodotto	Responsabilità di prodotto: dominio, priorità, limiti tecnici	Squadre interfunzionali stabili	Validazione interna alla squadra; costo di dotazione
Predittivo / decision-centric	Analisi e scomposizione delle decisioni; giudizio	Funzioni di organizzazione e titolari di processo	Formalismo della mappatura; perdita del contatto con la prassi

Configurazione	Competenza prevalente	Dove è collocata	Rischio di competenza
Agentico-ibrido	Delimitazione dei permessi; sorveglianza su sequenze	Responsabile umano designato per ciascun agente	Attenuazione della sorveglianza; perdita di tracciabilità
Ecosistemico	Gestione delle relazioni interorganizzative; orientamento di scelte collettive	Funzioni di sviluppo e organismi associativi	Dipendenza; appropriazione del valore da parte di terzi

Tabella 4.1 – Raccordo fra configurazioni organizzative e fabbisogni di competenza. Elaborazione propria sull'analisi dei §§ 3.5 e 4.4.

Dalla lettura della tabella emergono due considerazioni: la prima è che le competenze tecniche specialistiche costituiscono il fabbisogno prevalente in due sole configurazioni su otto, mentre in tutte le altre prevalgono competenze di valutazione, di coordinamento o di governo: è un risultato che ridimensiona l'idea, diffusa, secondo cui la trasformazione *AI-driven* si risolverebbe in un problema di reperimento di profili tecnici. La seconda è che la caratteristica in comune – l'alfabetizzazione di base e la capacità di calibrare l'affidamento – attraversa tutte le configurazioni: è il requisito minimo senza il quale nessun assetto formale produce gli effetti attesi, secondo quanto già anticipato al § 3.9.

4.5 Ruoli emergenti e trasformazione dei profili esistenti

4.5.1 I ruoli emergenti

Sulla base della tassonomia elaborata nel presente lavoro si propone un elenco di cinque ruoli funzionali. Il primo, relativo alla governance dell'IA, trova un riscontro diretto nella

rilevazione istituzionale richiamata di seguito; gli altri quattro – validatore indipendente, responsabile di prodotto, referente locale e responsabile dell'agente – sono deduzioni organizzative dell'autore ricavate dal raccordo fra modelli, meccanismi di controllo e competenze. Non vanno pertanto intesi come un repertorio di ruoli già uniformemente formalizzati nelle banche. Le denominazioni restano indicative, poiché variano sensibilmente fra organizzazioni.

Il primo è il responsabile della governance dell'IA, cui competono la tenuta dell'inventario dei sistemi e della loro classificazione di rischio, il presidio del sistema documentale richiesto dalla normativa, il coordinamento fra le funzioni di controllo e la verifica della coerenza fra impieghi previsti ed effettivi. È il ruolo che corrisponde alla struttura dedicata rilevata da Banca d'Italia (2025) presso un terzo degli intermediari e che, nelle organizzazioni che non l'hanno istituita, resta distribuito fra conformità, rischi e sistemi informativi.

Il secondo è il validatore indipendente, che il § 3.7.2 ha individuato come requisito di affidabilità dell'intero impianto. La sua caratteristica distintiva è la combinazione fra conoscenza del dominio e indipendenza da chi ha costruito il sistema; nel settore bancario tale figura può innestarsi sull'assetto esistente di validazione dei modelli, il che costituisce un vantaggio di percorso rispetto ad altri settori.

Il terzo è il responsabile di prodotto, proprio delle configurazioni a piattaforma e prodotto, che presidia il ciclo di vita di un sistema, ne definisce le priorità evolutive e ne risponde nel tempo. La sua rilevanza deriva direttamente dalla dimensione dell'apprendimento individuata al § 3.2.2: poiché il sistema muta, occorre un titolare stabile della sua evoluzione.

Il quarto è il referente locale, corrispondente allo *spoke* del modello *hub and spoke*, la cui funzione è duplice: tradurre le esigenze del processo in requisiti trattabili e diffondere nell'unità di appartenenza le regole d'uso e le competenze di base per una corretta adozione. È la figura sulla quale si concentra la tensione identitaria segnalata al § 3.5.3, poiché opera fra due appartenenze.

Il quinto, proprio delle configurazioni agentiche, è il responsabile dell'agente: la persona fisica cui è formalmente attribuita la sorveglianza di un agente digitale, l'autorizzazione delle modifiche al suo perimetro di azione e la decisione di sospensione. La sua istituzione non è una raffinatezza organizzativa ma la condizione perché la responsabilità resti attribuibile, secondo quanto argomentato al § 3.5.7.

4.5.2 La trasformazione dei profili esistenti

Più rilevante dei ruoli nuovi, sul piano quantitativo, è la trasformazione di quelli esistenti, che riguarda la generalità del personale e non una minoranza specialistica. Se ne analizzano quattro, riprendendo i processi già indicati.

L'analista del credito vede ridursi la componente istruttoria e accrescersi quella di valutazione delle situazioni non standard e di motivazione degli scostamenti dalla raccomandazione algoritmica. La competenza che diviene critica è la capacità di argomentare una decisione difforme, che presuppone la comprensione dei fattori che hanno determinato l'esito del sistema e il riconoscimento formale della facoltà di discostarsene, secondo le condizioni indicate al § 2.7.5.

L'addetto ai controlli antiriciclaggio vede la propria attività spostarsi dalla selezione delle posizioni alla valutazione conclusiva alla segnalazione delle inefficienze sistematiche della selezione: una funzione, quest'ultima, che il § 2.7.4 ha descritto come principale fonte informativa per la manutenzione del sistema e che raramente è formalizzata fra i compiti del ruolo.

Il gestore della relazione vede ridursi le attività di preparazione e di rendicontazione e accrescersi la componente consulenziale, con l'aggiunta della competenza esplicativa descritta al § 4.4.4. È il profilo nel quale il rapporto fra tempo liberato e valore generato dipende più direttamente dalla ridefinizione del ruolo: in assenza di questa, il tempo liberato tende a essere assorbito dalle attività preesistenti anziché riallocato.

Le funzioni di conformità e di controllo vedono ampliarsi il perimetro di competenza a una materia nuova e in evoluzione, con la necessità di comporre obblighi derivanti da

discipline sovrapposte – protezione dei dati, resilienza operativa, esternalizzazione, intelligenza artificiale – e di sviluppare una capacità di lettura tecnica dei sistemi che tradizionalmente non apparteneva al profilo.

Il tratto comune alle quattro trasformazioni è che nessuna di esse consiste nella sostituzione del ruolo, e tutte richiedono un intervento organizzativo esplicito: la ridefinizione formale del contenuto professionale, senza la quale, come osservato al § 3.7.4, la reazione prevedibile è la resistenza di gruppi professionali la cui posizione è modificata senza che ne sia stata ridefinita la funzione.

4.6 Le politiche di sviluppo delle competenze

4.6.1 Dalla formazione all'apprendimento nel lavoro

L'individuazione dei fabbisogni pone il problema di come costruire le competenze corrispondenti. Il dibattito su questo punto ha assunto, negli ultimi anni, un rilievo crescente: il World Economic Forum (2023), sulla base di una rilevazione condotta su imprese di dimensione rilevante a livello globale, stimava che una quota consistente delle competenze dei lavoratori sarebbe stata oggetto di trasformazione nell'arco del quinquennio successivo e che la maggioranza dei lavoratori avrebbe avuto necessità di percorsi formativi; l'OCSE, nel già citato lavoro sulle competenze per la transizione digitale, ha documentato l'evoluzione della domanda di competenze attraverso l'analisi degli annunci di lavoro (OECD, 2023). Si tratta di rilevazioni di natura diversa – la prima fondata su dichiarazioni di intenzione delle imprese, la seconda su dati osservati – e la differenza va tenuta presente nel valutarne l'attendibilità relativa.

Sul piano delle prassi d'uso, un'indicazione di ordine di grandezza proviene da rilevazioni di fonte aziendale, che vanno assunte con la cautela che tale provenienza impone. Il rapporto annuale sulle tendenze del lavoro elaborato congiuntamente da Microsoft e LinkedIn (Microsoft e LinkedIn, 2024) rileva, su un campione internazionale di lavoratori della conoscenza, una diffusione dell'utilizzo di strumenti di IA generativa superiore a quella dichiarata dalle organizzazioni di appartenenza, con una quota significativa di utilizzatori che

ricorre a strumenti personali non forniti dall'azienda. Il dato, pur non verificabile con gli standard della ricerca accademica, è coerente con il fenomeno dell'adozione non censita descritto ai §§ 3.5.1 e 3.5.2 e ha una conseguenza pratica rilevante: la formazione non interviene su una platea inesperta, ma su una platea che già utilizza gli strumenti, spesso al di fuori delle regole aziendali. Il primo obiettivo di una politica formativa non è quindi l'introduzione allo strumento, ma la correzione dell'uso e la definizione dei confini. Tale attività assume un'importanza cruciale nel settore bancario, poiché gli istituti di credito gestiscono quotidianamente ingenti quantità di dati sensibili appartenenti alla clientela. In questo contesto, è essenziale saper utilizzare in modo corretto e consapevole gli strumenti di intelligenza artificiale, al fine di evitare rischi connessi alla divulgazione non autorizzata o alla perdita di informazioni riservate.

4.6.2 L'obbligo normativo e la sua traduzione organizzativa

Le politiche formative in materia di intelligenza artificiale non sono, nel settore bancario, una scelta interamente discrezionale. L'articolo 4 dell'AI Act, come illustrato al § 4.4.2, impone a fornitori e deployer di adottare misure per sostenere lo sviluppo dell'alfabetizzazione del personale e delle altre persone che operano per loro conto, tenendo conto delle conoscenze, dell'esperienza, dell'istruzione, della formazione e del contesto d'uso; non richiede di garantire uno specifico livello per ciascun individuo. Sul piano organizzativo, la mappatura dei ruoli esposti ai sistemi, la definizione dei contenuti di competenza, l'erogazione di percorsi differenziati e la conservazione di evidenze sulle attività svolte costituiscono buone misure di attuazione. In particolare, la documentazione è utile per verificare l'efficacia delle iniziative e per fornire elementi probatori sulla ragionevolezza delle misure adottate, ma non è formulata dall'articolo 4 come prescrizione autonoma.

Merita di essere sottolineato che l'obbligo riguarda, per espressa previsione, anche le persone che operano per conto dell'organizzazione. Per un intermediario che acceda alla tecnologia attraverso fornitori esterni – configurazione prevalente per il tipo ideale descritto al § 2.9.3 – può risultare opportuno disciplinare contrattualmente le misure di alfabetizzazione

riferite al personale esterno che opera per suo conto, in raccordo con i presidi già previsti dalla disciplina sull'esternalizzazione.

4.6.3 Le condizioni di efficacia

L'analisi condotta nei paragrafi precedenti consente di individuare quattro condizioni alle quali una politica di sviluppo delle competenze produce effetti e che discendono direttamente dai meccanismi esaminati.

La prima è la differenziazione: il fabbisogno varia in funzione del ruolo e della configurazione organizzativa adottata, secondo quanto mostrato nella tabella 4.1. Un programma indifferenziato è quindi, nell'interpretazione proposta, una misura organizzativamente debole e difficilmente idonea a dimostrare che siano stati considerati tutti i fattori indicati dall'articolo 4; la disposizione, tuttavia, non impone testualmente percorsi differenziati né un livello prestabilito per ciascun individuo.

La seconda è la contestualizzazione: le competenze di valutazione si costruiscono sui casi propri del processo, non su esempi generici. Ne discende che la formazione più efficace è quella che impiega materiale dell'organizzazione e che coinvolge chi il processo lo conosce.

La terza è la continuità del circuito di retroazione: come argomentato al § 4.3.3, senza restituzione sistematica degli esiti la calibrazione dell'affidamento non si forma. La retroazione è una componente della politica formativa e non soltanto del sistema di controllo.

La quarta, e più esigente, è la coerenza degli incentivi. Il § 3.7.1 ha mostrato che un sistema di misurazione che penalizzi lo scostamento dalla raccomandazione algoritmica converte la sorveglianza in un costo; per analogia, una politica formativa che non sia accompagnata dal riconoscimento del tempo dedicato all'apprendimento e delle competenze acquisite nei percorsi professionali produce adesione formale e non cambiamento. È il punto in cui la questione delle competenze cessa di essere un tema di gestione delle risorse umane e diviene una questione di progettazione organizzativa.

4.7 Rischi e questioni aperte

L'analisi svolta non deve indurre a una rappresentazione lineare. Quattro questioni restano aperte e vanno esplicitate.

La prima è la distribuzione diseguale degli effetti: le evidenze richiamate al § 4.2.2 indicano che i guadagni di produttività associati agli assistenti generativi si concentrano sui profili meno esperti, mentre l'esposizione all'automazione, come osservato al § 4.2.1, interessa oggi anche segmenti qualificati. La combinazione dei due fenomeni può produrre una compressione dei differenziali professionali il cui effetto sulle carriere e sulle motivazioni non è determinabile a priori e merita attenzione.

La seconda è il rischio di depauperamento delle competenze, descritto al § 4.3.4. È il rischio più sottovalutato, perché si manifesta con ritardo e non è rilevabile dagli indicatori di breve periodo: un presidio umano che si degrada continua ad apparire in essere finché non si verifica il caso che avrebbe dovuto intercettare.

La terza riguarda l'attribuzione della responsabilità. Il secondo capitolo ha richiamato, al § 2.7.5, la critica secondo cui la sorveglianza umana può risolversi in un dispositivo di allocazione della responsabilità su operatori che non esercitano un controllo effettivo (Wagner, 2019; Elish, 2019). La questione si acuisce nelle configurazioni agentiche, nelle quali la persona risponde di sequenze di azioni che non ha osservato singolarmente. Se ne ricava un criterio: l'attribuzione formale della responsabilità è legittima soltanto se accompagnata dalle condizioni – competenza, autorità, tempo, informazione – che ne rendono possibile l'esercizio.

La quarta riguarda i limiti dell'evidenza disponibile: la ricerca sugli effetti dell'intelligenza artificiale generativa sul lavoro è recente, condotta in prevalenza su compiti circoscritti e in contesti sperimentali, e non consente ancora conclusioni robuste sugli effetti di medio periodo a livello di organizzazione. Le indicazioni formulate in questo capitolo hanno natura di ipotesi argomentate sulla letteratura esistente, non di risultati dimostrati.

4.8 Risposta alla domanda di ricerca e considerazioni conclusive

4.8.1 La risposta complessiva

L'elaborato era guidato dalla domanda relativa al modo in cui l'adozione dell'intelligenza artificiale nel settore bancario richieda una trasformazione dei modelli organizzativi e a come tale trasformazione ridefinisca le competenze umane necessarie. Il percorso svolto consente di articolare la risposta in quattro proposizioni.

La prima è che la trasformazione organizzativa non è una conseguenza dell'adozione tecnologica ma la sua condizione. Il secondo capitolo ha documentato una diffusione ormai ampia della tecnologia a fronte di benefici disomogenei; il terzo ha mostrato che l'investimento complementare di natura organizzativa – governance del dato, assetto dei presidi, definizione dei diritti decisionali – costituisce ciò che ne determina la redditività. La tecnologia, di per sé, non produce vantaggio.

La seconda è che non esiste un modello organizzativo superiore, ma una coerenza da ricercare. L'analisi delle otto configurazioni ha evidenziato che ciascuna risolve alcuni problemi generandone altri e che la scelta dipende da variabili di contingenza: dimensione, maturità, numero di sistemi in esercizio, natura dell'informazione su cui si fonda il modello di business. L'ampliamento della tassonomia alle configurazioni più recenti ha mostrato che le scelte di allocazione della competenza e di aggregazione del lavoro sono indipendenti e che una parte rilevante delle difficoltà osservate deriva dall'incoerenza fra le due.

La terza è che ciascuna configurazione genera un fabbisogno di competenze specifico per contenuto, secondo il raccordo proposto nella tabella 4.1. Ne discende che la questione delle competenze non può essere affrontata in astratto: la domanda pertinente non è quali competenze richieda l'IA, ma quali competenze richieda l'IA in una determinata configurazione organizzativa.

La quarta, e conclusiva, è formulata come ipotesi interpretativa: sulla base della tassonomia elaborata, la capacità di giudizio può risultare più scarsa della competenza tecnica. Il fabbisogno prevalente nella maggioranza delle configurazioni esaminate riguarda la capacità

di comprendere che cosa un sistema è in grado di portare al termine, di calibrare l'affidamento sugli esiti, di validarli e di specificare qual è l'obiettivo finale: capacità che presuppongono la conoscenza del dominio applicativo e che, nell'analisi proposta, possono essere meno agevolmente acquisibili sul mercato del lavoro rispetto ai profili tecnici. Per il tipo ideale della banca commerciale a radicamento territoriale, che accede alla tecnologia prevalentemente dall'esterno, questa ipotesi ha un corollario di rilievo strategico: il patrimonio di conoscenza del processo e della clientela, che costituisce il tratto distintivo del modello, è precisamente la risorsa su cui poggia la capacità di governare una tecnologia che non si produce internamente. La sua valorizzazione non è un'operazione difensiva, ma, nell'interpretazione proposta, la principale leva di cui questo tipo di intermediario dispone.

4.8.2 I limiti del lavoro

L'elaborato presenta limiti che è corretto esplicitare. È un lavoro di natura teorico analitica, fondato sulla revisione critica della letteratura e delle fonti istituzionali e non produce evidenza empirica originale: le proposizioni formulate hanno valore di ipotesi argomentate. La costruzione per tipi ideali adottata nel secondo capitolo, per quanto metodologicamente esplicitata, non consente inferenze su intermediari determinati. La tassonomia dei modelli organizzativi proposta nel terzo capitolo è un'elaborazione dell'autore fondata su categorie della teoria organizzativa e la sua validità andrebbe verificata su casi concreti. Il quadro normativo, infine, è in evoluzione: le modifiche intervenute nel luglio 2026 e gli schemi di decreto attuativi nazionali indicano che alcune delle indicazioni fornite potranno richiedere aggiornamento.

4.8.3 Prospettive di approfondimento

Da questi limiti discendono tre direttrici di approfondimento. La prima è la verifica empirica del raccordo fra configurazione organizzativa e fabbisogno di competenze proposto nella tabella 4.1, che potrebbe essere condotta attraverso studi di caso comparati su intermediari di dimensione e assetto diversi. La seconda riguarda il fenomeno descritto al § 4.3.4: la misurazione dell'effetto dell'automazione sull'accumulazione di esperienza nei processi bancari costituisce, allo stato, una lacuna della letteratura e un tema di evidente rilievo pratico.

La terza riguarda le configurazioni agentiche, la cui diffusione è recente e per le quali l'evidenza organizzativa è ancora limitata: la questione di come si attribuisca e si eserciti la responsabilità su sequenze di azioni compiute da sistemi autonomi è destinata a divenire centrale nei prossimi anni.

4.8.4 Considerazione finale

La ricostruzione condotta restituisce, in conclusione, un'immagine dell'intelligenza artificiale distante tanto dalla prospettiva sostitutiva quanto da quella meramente strumentale. I sistemi intelligenti non sottraggono all'uomo il proprio ruolo nell'organizzazione bancaria: ne modificano il contenuto, spostandolo dall'esecuzione al giudizio, dalla produzione dell'informazione alla sua valutazione, dalla conoscenza delle procedure alla comprensione dei limiti degli strumenti che le eseguono. È una trasformazione che accresce, anziché ridurre, l'esigenza di competenza, e che per questa ragione non può essere governata come un progetto tecnologico. Il vantaggio competitivo, in un settore fondato sulla fiducia, continuerà a dipendere dalla qualità del giudizio delle persone; ciò che l'intelligenza artificiale modifica è il contesto entro cui quel giudizio si esercita e la misura in cui l'organizzazione è capace di renderlo possibile.

Bibliografia

ABI (2026), Banche: al Forum ABI Lab 2026 le priorità degli investimenti tra sicurezza e intelligenza artificiale, comunicato stampa, Associazione Bancaria Italiana, Roma, 25 marzo 2026.

ABI Lab (2026), Scenario e trend del mercato ICT per il settore bancario 2026, Centro di Ricerca e Innovazione per la Banca, Roma, marzo 2026.

Acemoglu, D., Restrepo, P. (2018), The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment, American Economic Review, 108(6), 1488–1542.

Acemoglu, D., Restrepo, P. (2019), Automation and New Tasks: How Technology Displaces and Reinstates Labor, Journal of Economic Perspectives, 33(2), 3–30.

Acemoglu, D., Restrepo, P. (2020), Robots and Jobs: Evidence from US Labor Markets, Journal of Political Economy, 128(6), 2188–2244.

Agrawal, A., Gans, J. S., Goldfarb, A. (2019), Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction, Journal of Economic Perspectives, 33(2), 31–50.

Alessandrini, P., Presbitero, A. F., Zazzaro, A. (2009), Banks, Distances and Firms' Financing Constraints, Review of Finance, 13(2), 261–307.

Arnaudo, D., Del Prete, S., Demma, C., Manile, M., Orame, A., Pagnini, M., Rossi, C., Rossi, P., Soggia, G. (2022), La trasformazione digitale nell'industria bancaria italiana, Questioni di Economia e Finanza (Occasional Papers), n. 682, Banca d'Italia, Roma.

Arner, D. W., Barberis, J., Buckley, R. P. (2016), The Evolution of Fintech: A New Post-Crisis Paradigm?, Georgetown Journal of International Law, 47(4), 1271–1319.

Arner, D. W., Barberis, J., Buckley, R. P. (2017), FinTech, RegTech, and the Reconceptualization of Financial Regulation, Northwestern Journal of International Law & Business, 37(3), 371–413.

Arrieta, A. B. et al. (2020), Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI, Information Fusion, 58, 82–115.

Article 29 Data Protection Working Party (2018), Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679, WP251rev.01, adottate il 3 ottobre 2017 e riviste il 6 febbraio 2018.

Autor, D. H. (2015), Why Are There Still So Many Jobs? The History and Future of Workplace Automation, Journal of Economic Perspectives, 29(3), 3–30.

Autor, D. H., Levy, F., Murnane, R. J. (2003), The Skill Content of Recent Technological Change: An Empirical Exploration, The Quarterly Journal of Economics, 118(4), 1279–1333.

Bainbridge, L. (1983), Ironies of Automation, Automatica, 19(6), 775–779.

Banca d'Italia (2025), Indagine Fintech nel sistema finanziario italiano – Indagine Fintech sugli intermediari bancari e finanziari, Roma, 23 dicembre 2025.

Bartlett, R., Morse, A., Stanton, R., Wallace, N. (2022), Consumer-Lending Discrimination in the FinTech Era, Journal of Financial Economics, 143(1), 30–56.

Basel Committee on Banking Supervision (2013), Principles for Effective Risk Data Aggregation and Risk Reporting, BCBS 239, Bank for International Settlements, Basilea.

Berente, N., Gu, B., Recker, J., Santhanam, R. (2021), Managing Artificial Intelligence, MIS Quarterly, 45(3), 1433–1450.

Berg, F., Kölbel, J. F., Rigobon, R. (2022), Aggregate Confusion: The Divergence of ESG Ratings, Review of Finance, 26(6), 1315–1344.

Berg, T., Burg, V., Gombović, A., Puri, M. (2020), On the Rise of FinTechs: Credit Scoring Using Digital Footprints, The Review of Financial Studies, 33(7), 2845–2897.

Berg, T., Fuster, A., Puri, M. (2022), FinTech Lending, Annual Review of Financial Economics, 14, 187–207.

Berger, A. N., Udell, G. F. (2002), Small Business Credit Availability and Relationship Lending: The Importance of Bank Organisational Structure, The Economic Journal, 112(477), F32–F53.

Board of Governors of the Federal Reserve System (2011), Supervisory Guidance on Model Risk Management, SR Letter 11-7, Washington D.C.

Bommasani, R. et al. (2021), On the Opportunities and Risks of Foundation Models, Stanford Center for Research on Foundation Models (CRFM).

Boot, A. W. A. (2000), Relationship Banking: What Do We Know?, Journal of Financial Intermediation, 9(1), 7–25.

Boot, A., Hoffmann, P., Laeven, L., Ratnovski, L. (2021), Fintech: What's Old, What's New?, Journal of Financial Stability, 53, 100836.

Bresnahan, T. F., Trajtenberg, M. (1995), General Purpose Technologies: "Engines of Growth"?, Journal of Econometrics, 65(1), 83–108.

Brown, J. S., Duguid, P. (1991), Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation, Organization Science, 2(1), 40–57.

Brynjolfsson, E., Li, D., Raymond, L. (2025), Generative AI at Work, The Quarterly Journal of Economics, 140(2), 889–942.

Brynjolfsson, E., McAfee, A. (2014), The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, W. W. Norton & Company, New York.

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, Cambridge (MA).

Buchak, G., Matvos, G., Piskorski, T., Seru, A. (2018), Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks, Journal of Financial Economics, 130(3), 453–483.

Campbell, M., Hoane, A. J., Hsu, F. (2002), Deep Blue, Artificial Intelligence, 134(1–2), 57–83.

Choudhury, P., Starr, E., Agarwal, R. (2020), Machine Learning and Human Capital Complementarities: Experimental Evidence on Bias Mitigation, Strategic Management Journal, 41(8), 1381–1411.

Cohen, W. M., Levinthal, D. A. (1990), Absorptive Capacity: A New Perspective on Learning and Innovation, Administrative Science Quarterly, 35(1), 128–152.

Colbert, A., Yee, N., George, G. (2016), The Digital Workforce and the Workplace of the Future, Academy of Management Journal, 59(3), 731–739.

Copeland, B. J. (a cura di) (2004), The Essential Turing, Oxford University Press, Oxford.

D'Acunto, F., Prabhala, N., Rossi, A. G. (2019), The Promises and Pitfalls of Robo-Advising, The Review of Financial Studies, 32(5), 1983–2020.

Dehghani, Z. (2022), Data Mesh: Delivering Data-Driven Value at Scale, O'Reilly Media, Sebastopol (CA).

Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., Lakhani, K. R. (2026), Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality, Organization Science, 37(2), 403–423,

Deming, D. J. (2017), The Growing Importance of Social Skills in the Labor Market, The Quarterly Journal of Economics, 132(4), 1593–1640.

Dietvorst, B. J., Simmons, J. P., Massey, C. (2015), Algorithm Aversion: People Erroneously Avoid Algorithms Seeing Them Err, Journal of Experimental Psychology: General, 144(1), 114–126.

Elish, M. C. (2019), Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction, Engaging Science, Technology, and Society, 5, 40–60.

European Banking Authority (2019), Guidelines on Outsourcing Arrangements, EBA/GL/2019/02.

European Banking Authority (2020), Guidelines on Loan Origination and Monitoring, EBA/GL/2020/06.

European Banking Authority (2021a), Discussion Paper on Machine Learning for IRB Models, EBA/DP/2021/04.

European Banking Authority (2021b), Guidelines on Internal Governance under Directive 2013/36/EU, EBA/GL/2021/05.

European Central Bank Banking Supervision (2024), Digitalisation: key assessment criteria and collection of sound practices, Frankfurt am Main, July 2024.

Financial Action Task Force (2021), Opportunities and Challenges of New Technologies for AML/CFT, FATF, Parigi.

Financial Stability Board (2017), Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications, FSB, Basilea.

Fountaine, T., McCarthy, B., Saleh, T. (2019), Building the AI-Powered Organization, Harvard Business Review, 97(4), luglio–agosto 2019, 62–73.

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., Zbinden, P. (2019), BigTech and the Changing Structure of Financial Intermediation, Economic Policy, 34(100), 761–799.

Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., Walther, A. (2022), Predictably Unequal? The Effects of Machine Learning on Credit Markets, The Journal of Finance, 77(1), 5–47.

Galbraith, J. R. (1974), Organization Design: An Information Processing View, Interfaces, 4(3), 28–36.

Gomber, P., Koch, J.-A., Siering, M. (2017), Digital Finance and FinTech: Current Research and Future Research Directions, Journal of Business Economics, 87(5), 537–580.

Goodfellow, I., Bengio, Y., Courville, A. (2016), Deep Learning, MIT Press, Cambridge (MA).

Green, B., Chen, Y. (2019), The Principles and Limits of Algorithm-in-the-Loop Decision Making, Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), art. 50.

Grønsund, T., Aanestad, M. (2020), Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations, The Journal of Strategic Information Systems, 29(2), 101614.

Haenlein, M., Kaplan, A. (2019), A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence, California Management Review, 61(4), 5–14.

Jacobides, M. G., Cennamo, C., Gawer, A. (2018), Towards a Theory of Ecosystems, Strategic Management Journal, 39(8), 2255–2276.

Ji, Z. et al. (2023), Survey of Hallucination in Natural Language Generation, ACM Computing Surveys, 55(12), art. 248.

Jöhnk, J., Weißert, M., Wyrtki, K. (2021), Ready or Not, AI Comes – An Interview Study of Organizational AI Readiness Factors, Business & Information Systems Engineering, 63(1), 5–20.

Jordan, M. I., Mitchell, T. M. (2015), Machine Learning: Trends, Perspectives, and Prospects, Science, 349(6245), 255–260.

Kellogg, K. C., Valentine, M. A., Christin, A. (2020), Algorithms at Work: The New Contested Terrain of Control, Academy of Management Annals, 14(1), 366–410.

Khandani, A. E., Kim, A. J., Lo, A. W. (2010), Consumer Credit-Risk Models via Machine-Learning Algorithms, Journal of Banking & Finance, 34(11), 2767–2787.

Königstorfer, F., Thalmann, S. (2020), Applications of Artificial Intelligence in Commercial Banks – A Research Agenda for Behavioral Finance, Journal of Behavioral and Experimental Finance, 27, 100352.

Lane, M., Saint-Martin, A. (2021), The Impact of Artificial Intelligence on the Labour Market: What Do We Know So Far?, OECD Social, Employment and Migration Working Papers No. 256, OECD Publishing, Parigi.

Lawrence, P. R., Lorsch, J. W. (1967), Differentiation and Integration in Complex Organizations, Administrative Science Quarterly, 12(1), 1–47.

Lebovitz, S., Levina, N., Lifshitz-Assaf, H. (2021), Is AI Ground Truth Really True? The Dangers of Training and Evaluating AI Tools Based on Experts' Know-What, MIS Quarterly, 45(3), 1501–1526.

LeCun, Y., Bengio, Y., Hinton, G. (2015), Deep Learning, Nature, 521(7553), 436–444.

Lee, J. D., See, K. A. (2004), Trust in Automation: Designing for Appropriate Reliance, Human Factors, 46(1), 50–80.

Long, D., Magerko, B. (2020), What is AI Literacy? Competencies and Design Considerations, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ACM, 1–16.

McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E. (1955), A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, ripubbl. in AI Magazine, 27(4), 2006, 12–14.

Mehrabi, N. et al. (2021), A Survey on Bias and Fairness in Machine Learning, ACM Computing Surveys, 54(6), art. 115.

Microsoft, LinkedIn (2024), Work Trend Index Annual Report – AI at Work Is Here. Now Comes the Hard Part, Microsoft Corporation. (Rilevazione di fonte aziendale; la sua natura è dichiarata nel testo, § 4.6.1.)

Mikalef, P., Gupta, M. (2021), Artificial Intelligence Capability: Conceptualization, Measurement Calibration, and Empirical Study on Its Impact on Organizational Creativity and Firm Performance, Information & Management, 58(3), 103434.

National Institute of Standards and Technology (2026), Accelerating the Adoption of Software and Artificial Intelligence Agent Identity and Authorization, NCCoE Concept Paper, Initial Public Draft, U.S. Department of Commerce.

Negri, C. (2026), Storia dell'Intelligenza Artificiale: da Turing ai giorni nostri, Osservatorio Artificial Intelligence, Politecnico di Milano, osservatori.net (pubblicato il 16 maggio 2019, ultimo aggiornamento 29 gennaio 2026).

Newell, A., Shaw, J. C., Simon, H. A. (1959), Report on a General Problem-Solving Program, Proceedings of the International Conference on Information Processing, UNESCO, Paris, 256–264.

Noy, S., Zhang, W. (2023), Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence, Science, 381(6654), 187–192.

OECD (2021), Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges and Implications for Policy Makers, OECD Publishing, Parigi.

OECD (2023), Skills for the Digital Transition: Assessing Recent Trends Using Big Data, OECD Publishing, Parigi.

OECD (2026), Agentic AI in organisations, OECD Publishing, Parigi, settembre 2026.

O'Reilly, C. A., Tushman, M. L. (2013), Organizational Ambidexterity: Past, Present, and Future, Academy of Management Perspectives, 27(4), 324–338.

Parasuraman, R., Manzey, D. H. (2010), Complacency and Bias in Human Use of Automation: An Attentional Integration, Human Factors, 52(3), 381–410.

Pecis, E. (2025), L'impatto trasformativo dell'intelligenza artificiale sulle organizzazioni aziendali, Il Sole 24 Ore, 9 ottobre 2025.

Petralia, K., Philippon, T., Rice, T., Véron, N. (2019), Banking Disrupted? Financial Intermediation in an Era of Transformational Technology, Geneva Reports on the World Economy 22, ICMB e CEPR, Ginevra.

Puranam, P., Alexy, O., Reitzig, M. (2014), What's "New" About New Forms of Organizing?, Academy of Management Review, 39(2), 162–180.

Raisch, S., Krakowski, S. (2021), Artificial Intelligence and Management: The Automation–Augmentation Paradox, Academy of Management Review, 46(1), 192–210.

Russell, S. J., Norvig, P. (2021), Artificial Intelligence: A Modern Approach, 4ª ed., Pearson, Hoboken (NJ).

Seeber, I., Bittner, E., Briggs, R. O. et al. (2020), Machines as Teammates: A Research Agenda on AI in Team Collaboration, Information & Management, 57(2), 103174.

Skitka, L. J., Mosier, K. L., Burdick, M. (1999), Does Automation Bias Decision-Making?, International Journal of Human-Computer Studies, 51(5), 991–1006.

Stein, J. C. (2002), Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, The Journal of Finance, 57(5), 1891–1921.

Tarafdar, M., Beath, C. M., Ross, J. W. (2019), Using AI to Enhance Business Operations, MIT Sloan Management Review, 60(4), estate 2019, 37–44.

Teece, D. J. (2007), Explicating Dynamic Capabilities: The Nature and Microfoundations of (Sustainable) Enterprise Performance, Strategic Management Journal, 28(13), 1319–1350.

Thakor, A. V. (2020), Fintech and Banking: What Do We Know?, Journal of Financial Intermediation, 41, 100833.

Tschang, F. T., Almirall, E. (2021), Artificial Intelligence as Augmenting Automation: Implications for Employment, Academy of Management Perspectives, 35(4), 642–659.

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

Vaast, E., Pinsonneault, A. (2021), When Digital Technologies Enable and Threaten Occupational Identity: The Delicate Balancing Act of Data Scientists, MIS Quarterly, 45(3), 1087–1112.

Vaccaro, M., Almaatouq, A., Malone, T. (2024), When Combinations of Humans and AI Are Useful: A Systematic Review and Meta-Analysis, Nature Human Behaviour, 8, 2293–2303.

Vaswani, A. et al. (2017), Attention Is All You Need, Advances in Neural Information Processing Systems (NeurIPS), 30, 5998–6008.

Vial, G. (2019), Understanding Digital Transformation: A Review and a Research Agenda, The Journal of Strategic Information Systems, 28(2), 118–144.

Vives, X. (2019), Digital Disruption in Banking, Annual Review of Financial Economics, 11, 243–272.

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J. et al. (2024), A survey on large language model based autonomous agents, Frontiers of Computer Science, 18(6), 186345.

Wagner, B. (2019), Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems, Policy & Internet, 11(1), 104–122.

Warner, K. S. R., Wäger, M. (2019), Building Dynamic Capabilities for Digital Transformation: An Ongoing Process of Strategic Renewal, Long Range Planning, 52(3), 326–349.

Weber, M. (1904), Die «Objektivität» sozialwissenschaftlicher und sozialpolitischer Erkenntnis, Archiv für Sozialwissenschaft und Sozialpolitik, 19(1), 22–87; trad. it. L'«oggettività» conoscitiva della scienza sociale e della politica sociale, in Il metodo delle scienze storico-sociali, Einaudi, Torino.

Weber, M. (1922), Wirtschaft und Gesellschaft, Mohr, Tübingen; trad. it. Economia e società, Edizioni di Comunità, Milano.

Weizenbaum, J. (1966), ELIZA – A Computer Program For the Study of Natural Language Communication Between Man and Machine, Communications of the ACM, 9(1), 36–45.

World Economic Forum (2023), The Future of Jobs Report 2023, World Economic Forum, Ginevra.

Fonti normative e regolamentari

Parlamento europeo e Consiglio dell'Unione europea (2015), Direttiva (UE) 2015/2366 relativa ai servizi di pagamento nel mercato interno (PSD2), GUUE L 337 del 23 dicembre 2015.

Parlamento europeo e Consiglio dell'Unione europea (2016), Regolamento (UE) 2016/679 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali e alla libera circolazione di tali dati (GDPR), GUUE L 119 del 4 maggio 2016.

Commissione europea (2021), Proposta di Regolamento che stabilisce regole armonizzate sull'intelligenza artificiale, COM(2021) 206 final, Bruxelles, 21 aprile 2021.

Parlamento europeo e Consiglio dell'Unione europea (2022), Regolamento (UE) 2022/2554 relativo alla resilienza operativa digitale per il settore finanziario (DORA), GUUE L 333 del 27 dicembre 2022.

Parlamento europeo e Consiglio dell'Unione europea (2024), Regolamento (UE) 2024/1689 del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale (regolamento sull'intelligenza artificiale – AI Act), GUUE L 2024/1689 del 12 luglio 2024.

Parlamento europeo e Consiglio dell'Unione europea (2026), Regolamento (UE) 2026/1744 dell'8 luglio 2026, che modifica i regolamenti (UE) 2024/1689, (UE) 2018/1139 e (UE) 2023/1230 per quanto riguarda la semplificazione dell'attuazione di regole armonizzate sull'intelligenza artificiale (Omnibus digitale sull'IA), GUUE L 2026/1744 del 24 luglio 2026.

Legge 30 luglio 1990, n. 218, e decreto legislativo 20 novembre 1990, n. 356, in materia di ristrutturazione e integrazione patrimoniale degli istituti di credito di diritto pubblico (cosiddetta legge Amato).

Banca d'Italia, Circolare n. 285 del 17 dicembre 2013, Disposizioni di vigilanza per le banche, Parte Prima, Titolo IV, Capitolo 3 – Il sistema dei controlli interni, e successivi aggiornamenti.

Legge 23 settembre 2025, n. 132, Disposizioni e deleghe al Governo in materia di intelligenza artificiale, GU Serie Generale n. 223 del 25 settembre 2025, in vigore dal 10 ottobre 2025.