



FONDAZIONE CENTESIMUS ANNUS PRO PONTIFICE

Etica e finanza sostenibile

*L'intelligenza artificiale al servizio
della persona e del bene comune*

RAPPORTO DI RICERCA
frutto della collaborazione tra
Fondazione Centesimus Annus Pro Pontifice e Deloitte

MMXXVI

Indice

	Introduzione <i>di Anna Maria Tarantola</i>	3
	Il rapporto in dieci punti	4
I	Dove siamo	5
	1. Oltre il rumore di fondo: i numeri dell'adozione	6
	2. La traiettoria: dall'AI che prevede all'AI che agisce	8
II	Dove l'AI crea valore	11
	3. Il valore lungo tutta la banca	12
	4. L'AI al servizio della finanza sostenibile	14
	5. Le poste in gioco per il sistema	16
III	La posta etica	19
	6. La bussola: tre riferimenti convergenti	20
	7. Il bivio: ottimizzare i risultati o servire lo sviluppo	22
	8. Cinque principi per un'AI al servizio della persona	24
IV	Governare l'AI etica	29
	9. La responsabilità di chi decide	30
	10. Il ciclo dell'AI-vigilanza	36
	11. La sicurezza come tema di Consiglio	38
	12. Equità, genere, inclusione	43
V	La responsabilità dei CEO	48
	13. Solo un'AI affidabile rende la finanza sostenibile	49
	Raccomandazioni al sistema finanziario	53
	Conclusioni — Una finanza che serve	54
	La mappa dell'enciclica per chi decide	55
	Note metodologiche e fonti	56
	Per approfondire — Bibliografia ragionata	58

INDICE DELLE FIGURE E DELLE TAVOLE

Tavola 1	<i>Dall'automazione all'autonomia: i casi d'uso dell'AI</i>	8
Figura 1	<i>Dove l'AI crea valore nella banca</i>	13
Figura 2	<i>La bussola: tre riferimenti convergenti</i>	20
Figura 3	<i>La difesa in profondità: i presidi a cerchi concentrici</i>	30
Tavola 2	<i>L'architettura dei presidi</i>	31
Figura 4	<i>Il ciclo dell'AI-vigilanza</i>	36
Figura 5	<i>L'orizzonte del bene comune</i>	54
Tavola 3	<i>La mappa dell'enciclica per chi decide</i>	55

INTRODUZIONE

La misura di ogni progresso è il volto della persona

di Anna Maria Tarantola · Fondazione Centesimus Annus Pro Pontifice

In oltre quarant'anni di vigilanza bancaria ho imparato una cosa semplice: ciò che regge una banca non sono i suoi attivi, ma la fiducia che le viene riconosciuta. È una convinzione che ho portato con me nel lavoro della Fondazione e nei *Dialoghi per una Finanza Integralmente Sostenibile*, da cui nasce questo rapporto. Al centro sta una domanda che la prima enciclica di Leone XIV, *Magnifica Humanitas*, ha reso non più rinviabile: l'intelligenza artificiale che sta trasformando la finanza saprà servire lo sviluppo integrale della persona, o si limiterà a ottimizzare i risultati?

La Fondazione lo ha sempre creduto: l'economia e la finanza migliorano quando finanza, imprese, mondo educativo e istituzioni accettano di dialogare invece di ignorarsi. Le pagine che seguono non offrono ricette. Offrono un quadro dei fatti, esperienze internazionali e una bussola: i criteri con cui la dottrina sociale della Chiesa, che l'enciclica rilegge davanti alle *res novae* dell'intelligenza artificiale, illumina le scelte di chi guida banche e istituzioni finanziarie. L'enciclica non è un testo tecnico e non pretende di esserlo: consegna a chi decide principi per pensare, criteri per giudicare, orientamenti per agire, con uno stile di ascolto, discernimento e responsabilità.

Su due punti desidero essere esplicita, perché vi ho dedicato molto del mio impegno. Il primo riguarda le donne. Le decisioni automatizzate apprendono dal passato e ne possono ereditare i pregiudizi: se la storia del credito e del lavoro ha penalizzato le donne, un algoritmo addestrato su quella storia rischia di riprodurre la discriminazione con l'apparenza dell'oggettività. Una finanza che si dice sostenibile non può permetterlo. Il secondo riguarda la finanza e la pace: le scelte di allocazione del capitale, che non sono mai moralmente neutre, possono basarsi su criteri tali da orientare gli investimenti in modo da garantire la competitività delle imprese e un'economia di pace, e l'intelligenza artificiale, che le rende più rapide e pervasive, rende la responsabilità di chi decide più grande, non minore.

La tecnologia, ricorda Leone XIV, non è mai neutrale: prende la forma delle intenzioni di chi la concepisce, la finanzia, la regola e la usa (cfr MH, §9). Per questo il rapporto si rivolge anzitutto a chi decide e dedica la sua ultima parte alla responsabilità personale dei vertici. A loro chiedo di custodire, dentro l'efficienza, ciò che nessun algoritmo potrà mai calcolare: la dignità di ogni volto.

Ciò che regge una banca non sono i suoi attivi, ma la fiducia che le viene riconosciuta.

DALLA INTRODUZIONE

Far dialogare finanza, imprese, educazione e istituzioni

Un'AI che eredita i pregiudizi del passato non è sostenibile

Allocare capitale non è mai moralmente neutro

SINTESI

Il rapporto in dieci punti

Dieci messaggi, una sola
direzione: la persona al centro

*Trasparenza, inclusione ed
equità non frenano
l'innovazione: la rendono
umana.*

CFR MAGNIFICA HUMANITAS,
§164

- 1 **Adozione diffusa, uso reale ancora raro.** quasi tutte le banche sperimentano l'AI, ma poche la usano nelle decisioni che contano (credito, pagamenti, controlli): è lì, non negli annunci, che si misura la trasformazione.
- 2 **Più l'AI agisce da sola, più chi la guida ne risponde.** si passa da sistemi che suggeriscono a sistemi che agiscono in autonomia: con l'autonomia la responsabilità di chi li governa cresce, non diminuisce.
- 3 **L'AI crea valore in ogni funzione della banca.** dal rapporto con il cliente ai controlli, dalle operazioni al credito. La finanza sostenibile è la frontiera più promettente, ma regge solo se i dati che la alimentano sono di qualità.
- 4 **Finanza sostenibile: AI affidabile e persona al centro.** nel tempo dell'AI una finanza è davvero sostenibile solo se i suoi sistemi sono affidabili e se la persona, con la sua dignità, resta al centro di ogni scelta, mai ridotta a un punteggio.
- 5 **Affidabile significa lecita, etica, robusta e controllabile insieme.** qualità che valgono solo se presenti tutte: rispetta i diritti e le leggi (lecita), incorpora criteri etici verificabili e non solo dichiarati (etica), è sicura e accurata nei modelli e nelle previsioni (robusta), è controllabile dalle funzioni e autorità di controllo. Nessuna, da sola, basta.
- 6 **Nessuna decisione che discrimina.** i sistemi imparano dal passato e rischiano di ripeterne le ingiustizie con apparente oggettività: serve attenzione costante a non discriminare, a tutela dei più vulnerabili e delle donne in particolare.
- 7 **La bussola esiste già, e i rischi riguardano tutti.** enciclica, principi internazionali e regole europee dicono le stesse cose: persona, equità, trasparenza, responsabilità; e rischi come l'instabilità, la dipendenza da pochi grandi fornitori e la perdita di fiducia sono di tutto il sistema.
- 8 **I principi contano se diventano controlli concreti.** una politica decisa dal Consiglio, una verifica indipendente prima di usare un sistema, una sorveglianza continua dei modelli nel tempo, la sicurezza discussa in Consiglio, verifiche regolari contro le discriminazioni.
- 9 **Il vertice guida l'intera banca con le decisioni e con l'esempio.** le scelte etiche dei vertici, e soprattutto l'esempio che danno, orientano chi ogni giorno progetta, sviluppa e usa l'AI: l'etica di un'organizzazione si decide in alto e si diffonde verso il basso.
- 10 **Spetta ai CEO, qui e ora.** mentre l'AI entra in tutti i processi, fare bene gli affari e fare il bene non sono in conflitto: la fiducia è il vero capitale di una banca, e custodirla è responsabilità personale di chi la guida.

I



Dove siamo

L'adozione dell'intelligenza artificiale nella finanza, oltre il rumore di fondo

1. Oltre il rumore di fondo: i numeri dell'adozione
2. La traiettoria: dall'AI che prevede all'AI che agisce

CAPITOLO 1

Oltre il rumore di fondo: i numeri dell'adozione

31%

degli operatori dei mercati finanziari
(intermediari e gestori del risparmio)
usa l'IA

OCSE – Banca d'Italia, 2026

39%

degli operatori finanziari italiani usa
l'AI nell'operatività quotidiana: 70%
nelle assicurazioni, 59% nelle banche

OCSE – Banca d'Italia, 2026

33%

degli intermediari ha istituito un
comitato o una funzione di
governance dell'IA

Banca d'Italia, Indagine FinTech 2025

**Il vantaggio non è la tecnologia:
sono dati, competenze e
governo**

1.1 Un quadro a due velocità

L'adozione dell'intelligenza artificiale nel sistema finanziario ha superato la soglia della sperimentazione. La rilevazione OCSE e Banca d'Italia dell'aprile 2026, condotta su 450 operatori, indica che il 39% degli operatori finanziari italiani utilizza già l'AI nell'operatività quotidiana, con punte del 70% nelle compagnie assicurative e del 59% negli istituti bancari, a fronte del 31% tra gli intermediari e le società di gestione del risparmio.¹ Il dato si inserisce in un quadro europeo più ampio: secondo la seconda *European Financial Services AI Survey* di EY, condotta nell'ottobre 2024, il 90% delle istituzioni finanziarie europee ha integrato l'AI in qualche misura.²

Dietro questi numeri, però, la maturità è ben più contenuta. In Italia solo il 33% degli intermediari ha istituito un comitato o una funzione dedicata alla governance dell'IA;³ in Europa, soltanto il 9% delle istituzioni si colloca «davanti alla curva». È in questo divario, tra l'adozione diffusa e l'uso effettivo nelle decisioni che contano, che si gioca la credibilità della trasformazione: i comitati e i laboratori non cambiano una banca; la cambiano i modelli che entrano nei processi e toccano i clienti.

1.2 Che cosa dicono davvero i numeri

Letti con attenzione, questi dati raccontano tre cose. La prima: il vantaggio non sta nell'accesso alla tecnologia, ormai diffusa, ma nella qualità dei dati, delle competenze e del governo: non a caso oltre la metà delle istituzioni indica trasparenza ed esplicabilità tra gli ostacoli principali.³ La seconda: i progetti che durano nascono da un problema del cliente, non da una tecnologia in cerca di applicazioni. La terza: dove l'adozione è guidata dalla sola logica di costo, i benefici si rivelano effimeri e i rischi reputazionali crescono.

1. OCSE e Banca d'Italia, L'intelligenza artificiale nei mercati finanziari italiani, aprile 2026 (indagine su 450 operatori); dati di governance dall'Indagine FinTech di Banca d'Italia.

2. EY, European Financial Services AI Survey (2ª ed.), rilevazione ottobre 2024, pubbl. dicembre 2024.

3. Deloitte, EMEA Model Risk Management Survey 2025 (uso dell'AI/ML tra banche e assicurazioni in EMEA).

CAPITOLO 1 · Oltre il rumore di fondo: i numeri dell'adozione

1.3 Un valore potenziale ingente, una distribuzione non scontata

Il potenziale economico è documentato in modo convergente. Per il McKinsey Global Institute, l'AI generativa potrebbe aggiungere ogni anno tra 200 e 340 miliardi di dollari al settore bancario globale, pari al 2,8–4,7% dei ricavi dell'industria, soprattutto attraverso guadagni di produttività.⁴ È una stima di potenziale, non un risultato acquisito: la sua realizzazione dipende dalle scelte di governo che le istituzioni compiono ora, quando le architetture sono ancora plastiche. E riguarda anzitutto i giacimenti di valore più maturi (prevenzione delle frodi e antiriciclaggio, gestione del rischio di credito, relazione con il cliente, operations), dove l'adozione, secondo le rilevazioni di settore, è già oggi più diffusa.⁵

200–340 mld \$

il valore potenziale annuo del gen AI per la banca globale: 2,8–4,7% dei ricavi

McKinsey Global Institute, 2023

1.4 Il banco di prova della finanza sostenibile

La finanza sostenibile è il banco di prova più interessante dell'intera trasformazione: nella rilevazione 2025 di Deloitte presso oltre 2.100 dirigenti, l'81% dichiara di usare già l'AI a questo fine.⁶ La tecnologia promette di colmare il divario informativo che da sempre limita la misura degli impatti. Ma proprio la fragilità del dato è la prima questione etica, prima ancora del modello: una misura inattendibile, più che un difetto tecnico, è una promessa non mantenuta verso il risparmiatore.

81%

dei dirigenti usa già l'AI a supporto della sostenibilità (survey multisettoriale)

Deloitte, 2025

1.5 Ciò che i numeri non dicono

Le statistiche di adozione non misurano la qualità delle scelte: non dicono se i modelli trattino con equità chi ha una storia atipica, né se il valore creato raggiunga il cliente. L'enciclica offre il criterio che il rapporto fa proprio: davanti a ogni innovazione, la domanda decisiva è se essa contribuisca alla crescita delle persone e dei popoli in umanità (cfr MH, §85); e questa i cruscotti non la catturano.

32%

delle imprese italiane usa l'AI a inizio 2026 (era il 27% nel 2025): un'adozione che diventa strutturale

Banca d'Italia, 2026

RIQUADRO 1

Che cos'è un «foundation model»

Le banche non costruiscono i sistemi di AI più avanzati: li adattano. Un «foundation model» è un modello molto grande, addestrato su enormi quantità di dati e adattabile a molti compiti, dal testo alle immagini al codice, senza essere riaddestrato da zero. Pochi fornitori globali li producono; tutti gli altri vi si appoggiano. Questa dipendenza, per sistemi il cui funzionamento interno neppure chi li costruisce sa spiegare del tutto, è essa stessa un fatto di governo.

PER CAPIRE

UNA DIPENDENZA DI GOVERNO

La qualità del dato è la prima questione etica

4. McKinsey Global Institute, The economic potential of generative AI: the next productivity frontier, 14 giugno 2023.

5. Deloitte, EMEA Model Risk Management Survey 2025: tra i casi d'uso prevalenti, antiriciclaggio/adequata verifica e prevenzione delle frodi.

6. Deloitte Global, 2025 C-suite Sustainability Survey (oltre 2.100 dirigenti in 27 Paesi; dato multisettoriale, non limitato alla finanza).

CAPITOLO 2

La traiettoria: dall'AI che prevede all'AI che agisce

2.1 La mappa del valore

I controlli guidano l'adozione; la sostenibilità è la frontiera

La mappa del valore è ormai consolidata: i controlli e la prevenzione delle frodi guidano l'adozione, seguiti dal rischio di credito e dalla relazione con il cliente; tra i casi d'uso più diffusi le rilevazioni di settore collocano proprio antiriciclaggio, adeguata verifica e *fraud detection*.⁷ La finanza sostenibile resta l'ambito più giovane, e insieme il più promettente, perché qui il divario informativo da colmare è massimo.

TAVOLA 1 | Dall'automazione all'autonomia: i casi d'uso dell'AI nella finanza (quote sul totale dei casi d'uso di AI dichiarati)

CARATTERISTICA DEL CASO D'USO	CHE COSA INDICA	QUOTA
Con automazione della decisione	una qualche decisione presa dal sistema	55%
di cui semi-autonomi	con supervisione umana	24%
Basati su <i>foundation model</i>	modelli generativi di base	17%
Ad alta materialità	impatto rilevante su persone o conti	16%
Pienamente autonomi	senza supervisione umana	2%

Fonte: Bank of England e Financial Conduct Authority, «Artificial intelligence in UK financial services 2024» (21 novembre 2024). «Con automazione» indica i casi con un qualche grado di decisione automatizzata.

2%

la quota di casi d'uso dell'AI pienamente autonomi: la delega del «comportamento» è appena cominciata

Bank of England – FCA, 2024

2.2 Dalla previsione all'azione

Non si delega più un calcolo: si delega un comportamento

La direzione di marcia è chiara: dai modelli che *prevedono* (una probabilità di insolvenza, un'anomalia, una preferenza) ai sistemi che *agiscono*, agenti capaci di eseguire processi completi con crescente autonomia. È il passaggio più importante del decennio, perché cambia la natura della delega: non si affida più alla macchina un calcolo, ma un comportamento. Oggi quel passaggio è ancora agli inizi, appena il 2% dei casi d'uso è pienamente autonomo⁷, ed è proprio ora, mentre l'autonomia è bassa, che vanno fissate le regole del suo aumento.

⁷ Bank of England e Financial Conduct Authority, Artificial intelligence in UK financial services 2024, 21 novembre 2024 (automazione delle decisioni e materialità dei casi d'uso); per i casi d'uso prevalenti, Deloitte, EMEA Model Risk Management Survey 2025.

CAPITOLO 2 · La traiettoria: dall'AI che prevede all'AI che agisce

2.3 L'autonomia che cresce, la responsabilità che non si delega

Con l'autonomia dei sistemi cresce la responsabilità di chi li governa. L'enciclica è netta su questo punto: affidare di fatto a un algoritmo il potere di selezionare chi merita e chi no, senza che nessuno porti più il peso della decisione, significa rivestire l'esclusione di una neutralità apparente contro cui non è possibile protestare (cfr MH, §103). Nelle decisioni che toccano la vita delle persone, dal credito al lavoro all'accesso ai servizi, l'ultima parola deve restare a chi può risponderne: è la condizione di umanità della delega, prima ancora che un requisito normativo.

Quando nessuno porta più il peso della decisione, l'ingiustizia si fa silenziosa.

CFR MAGNIFICA HUMANITAS, §103

2.4 La velocità della tecnica e il ritmo delle istituzioni

Esiste uno squilibrio strutturale tra la velocità con cui la tecnologia evolve e il ritmo con cui maturano consapevolezza, regole e controlli. L'enciclica osserva che chiedere prudenza, verifiche rigorose e, dove serve, un rallentamento non è ostilità al progresso, ma cura responsabile (cfr MH, §106). Per una banca questa osservazione ha un significato operativo preciso: la cadenza dell'adozione va commisurata alla capacità di governo effettiva (dati, competenze, presidi), non alla pressione competitiva del momento. Adottare più lentamente ciò che non si sa ancora governare non è un ritardo: è la forma che la prudenza del banchiere assume nel tempo dell'AI.

Prudenza e verifiche non sono ostilità al progresso, ma cura

Da qui in avanti il rapporto cambia passo: dalla descrizione alla valutazione. La seconda parte mostra dove il valore si crea; la terza discute a quali condizioni quel valore sia degno della persona.

La cadenza dell'adozione si commisura alla capacità di governo

RIQUADRO 2

Quarantacinque minuti, 440 milioni di dollari

Il 1° agosto 2012 un errore nel rilascio di un software fece sì che il sistema di trading automatico della Knight Capital, allora tra i maggiori operatori dei mercati azionari statunitensi, inviasse in circa quarantacinque minuti milioni di ordini errati su oltre centocinquanta titoli. Nessun controllo riuscì a fermarlo in tempo. La perdita fu di circa 440 milioni di dollari e l'impresa, fra le più solide del settore, sfiorò il fallimento nel giro di una mattinata. È la lezione più nuda dell'autonomia senza freni: un sistema che agisce da solo può distruggere in pochi minuti ciò che si è costruito in decenni.

Fonte: SEC, comunicazioni del 2 agosto 2012; SEC Order, 2013.

PRINCIPIO IN GIOCO

VIGILANZA

MAGNIFICA HUMANITAS · LEONE XIV · 2026

“

*La qualità di una civiltà si misura dalla capacità di
«riconoscere l'altro come volto e non come funzione».*

LETTERA ENCICLICA MAGNIFICA HUMANITAS, §114



*Vale per le civiltà, e vale per le banche: il cliente è un volto,
non una funzione di redditività.*

II



Dove l'AI crea valore

La banca, i clienti, la finanza sostenibile: la geografia del valore e le sue condizioni

3. Il valore lungo tutta la banca

4. L'AI al servizio della finanza sostenibile

5. Le poste in gioco per il sistema

CAPITOLO 3

Il valore lungo tutta la banca

Il valore nel front office è restituzione di attenzione

Chi progetta o finanzia modelli che prosperano sulla debolezza umana ne porta la responsabilità morale.

CFR MAGNIFICA HUMANITAS, §170

La macchina segnala, l'analista decide, il processo documenta

3.1 La relazione con il cliente

Nel front office l'AI promette ciò che la banca di prossimità sapeva fare e la banca industriale ha in parte smarrito: conoscere ciascun cliente. Consulenza più accessibile, servizio continuo, offerte coerenti con la vita reale delle persone: il valore qui non è l'automazione del contatto, ma la restituzione di attenzione. Quell'attenzione comprende anche il dovere di formare le persone e di offrire strumenti digitali semplici e sicuri per operare con la banca: oggi quegli strumenti sono spesso poco intuitivi e non sempre affidabili, e renderli facili e protetti è parte del servizio non meno della consulenza. La misura del successo, però, non può essere soltanto la conversione commerciale: un sistema che conosce a fondo il cliente può servirlo meglio o sfruttarne meglio le debolezze. L'enciclica chiama per nome questo crinale: quando un modello d'impresa prospera sulla fragilità delle persone, chi lo progetta e chi lo finanzia assume una responsabilità morale che non può essere elusa (cfr MH, §170).

3.2 Operations e controlli

Nelle operations l'AI riduce attese, errori e costi unitari; nei controlli moltiplica la capacità di vedere. La prevenzione delle frodi e dell'antiriciclaggio è l'ambito dove l'adozione è più matura e dove il beneficio è meno ambiguo: qui la macchina protegge il cliente e l'integrità del sistema insieme. È anche l'ambito che insegna il metodo giusto: la macchina segnala, l'analista decide, il processo documenta. Quella divisione del lavoro, capacità di calcolo alla macchina e responsabilità alla persona, è il modello da estendere a tutto il resto della banca.

CAPITOLO 3 · Il valore lungo tutta la banca

3.3 Il rischio di credito

Il credito è il luogo dove l'AI tocca più direttamente la vita delle persone e delle imprese. Modelli più capaci leggono segnali che i bilanci non catturano, allargano la base informativa, riducono i tempi. Ma è qui che l'enciclica colloca il suo avvertimento più preciso per la finanza: le decisioni che toccano il credito, il lavoro e l'accesso ai servizi rischiano di essere consegnate a sistemi che non conoscono la possibilità che una persona cambi, producendo nuove forme di scarto (cfr MH, §102). Per un algoritmo l'errore è un dato da correggere; per una persona può essere l'inizio di un cambiamento (cfr MH, §128). Una banca che incorpora questa consapevolezza nei propri modelli, prevedendo revisione umana, percorsi di riabilitazione creditizia e attenzione alle seconde possibilità, non fa beneficenza: fa credito nel senso pieno della parola, che è dare fiducia.

Per un algoritmo l'errore è da correggere; per una persona può essere l'inizio di un cambiamento.

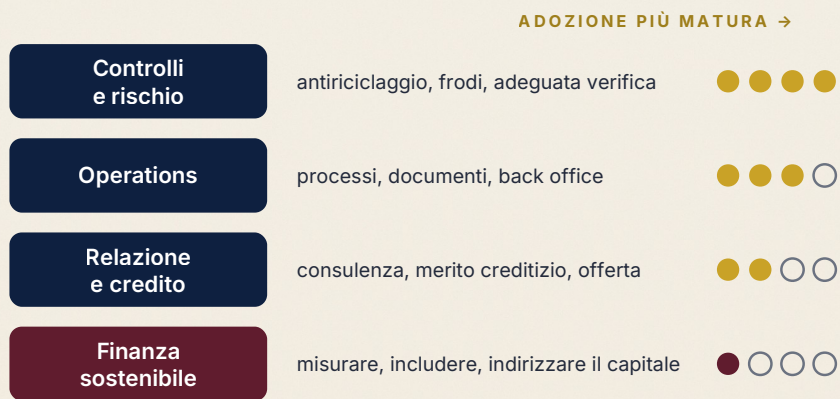
CFR MAGNIFICA HUMANITAS, §128

3.4 La condizione che attraversa tutto

In ciascun ambito la condizione del valore è la stessa: che la decisione resti comprensibile, contestabile e sottoposta a controllo, perché la persona non sia ridotta a profilo (cfr MH, §164). Non è un vincolo aggiunto al modello: è la specifica di progetto. La giustizia, ricorda l'enciclica, non è un correttivo da applicare dopo l'adozione delle tecnologie, ma una condizione da praticare nel loro stesso disegno (cfr MH, §109).

Fare credito è dare fiducia: i modelli devono saperlo

FIGURA 1 | Dove l'AI crea valore nella banca (per ambito, con la maturità relativa dell'adozione)



La giustizia non si applica dopo: si pratica nel disegno stesso dei sistemi.

CFR MAGNIFICA HUMANITAS, §109

CAPITOLO 4

L'AI al servizio della finanza sostenibile

Rendere osservabile ciò che era solo dichiarato

81%

dei dirigenti dichiara di usare già l'AI a supporto della sostenibilità

Deloitte Global, 2025

Leggere storie che i bilanci non raccontano

4.1 Misurare

La finanza sostenibile soffre da sempre di un deficit di misura: gli impatti ambientali e sociali, e anche di governance, sono difficili da osservare, costosi da verificare, esposti a dichiarazioni di comodo. L'AI aggredisce questo deficit alla radice: fonti alternative come le immagini satellitari, i sensori e i dati amministrativi rendono osservabile ciò che prima era solo dichiarato; l'analisi automatica del linguaggio rende confrontabili migliaia di rendicontazioni. Misurare meglio significa anche smascherare meglio: la stessa tecnologia che alimenta la rendicontazione può rivelarne le esagerazioni. Resta però una misura che la tecnologia da sola non offre: l'enciclica chiede di adottare indicatori di successo che includano la qualità e la dignità del lavoro (cfr MH, §156).

4.2 Includere

La seconda promessa è l'inclusione. Milioni di persone e di piccole imprese restano fuori dal credito non perché non meritino fiducia, ma perché la loro storia non è scritta nei formati che le banche sanno leggere. Modelli capaci di valorizzare dati alternativi, come i flussi di pagamento, la regolarità delle utenze o l'attività economica osservabile, possono leggere storie che i bilanci non raccontano. L'enciclica colloca questa promessa dentro un principio: i beni di cui dipende oggi la ricchezza delle nazioni (algoritmi, piattaforme, dati) sono soggetti anch'essi alla destinazione universale dei beni, e la loro concentrazione in poche mani alimenta il divario tra inclusi ed esclusi (cfr MH, §67). L'inclusione finanziaria non è un programma accessorio: è il modo in cui la finanza partecipa alla giustizia del proprio tempo.

CAPITOLO 4 · L'AI al servizio della finanza sostenibile

4.3 Indirizzare il capitale

La terza promessa riguarda l'allocazione: strumenti di analisi che rendono comparabili migliaia di emittenti, portafogli costruiti su evidenze d'impatto e non su etichette, monitoraggio continuo degli impegni presi. Qui l'AI può restituire alla finanza la sua funzione più alta, quella che l'enciclica richiama distinguendo la finanza per la finanza dalla finanza per lo sviluppo e ricordando che la funzione sociale del credito resta insostituibile (cfr MH, §160). Indirizzare il capitale verso ciò che fa vivere è il mestiere originario del banchiere; l'AI non lo cambia: lo riarma di strumenti all'altezza.

RIQUADRO 3

Credito agricolo e dati satellitari

In diverse economie emergenti, operatori di sviluppo e istituti specializzati estendono il credito a piccole aziende agricole prive di storia creditizia combinando immagini satellitari, dati meteorologici e segnali alternativi nella valutazione del merito, riducendo le garanzie reali richieste. È la prima condizione dell'alleanza tra AI e inclusione: il dato che dà voce a chi non ha bilanci. La stessa tecnica, applicata senza trasparenza, può però escludere su basi opache; e ciò che distingue le due strade è la possibilità di capire e contestare la decisione.

PRINCIPIO ILLUSTRATO

BENE COMUNE · INCLUSIONE

Anche algoritmi, piattaforme e dati sono soggetti alla destinazione universale dei beni.

CFR MAGNIFICA HUMANITAS,
§67

Il dato alternativo dà voce a chi non ha bilanci

La finanza per la finanza è cosa ben diversa dalla finanza per lo sviluppo.

CFR MAGNIFICA HUMANITAS,
§160

CAPITOLO 5

Le poste in gioco per il sistema

Più efficiente e insieme più fragile: il paradosso dell'omologazione

Un potere che si concentra in poche mani tende a farsi opaco e a sfuggire al controllo.

CFR MAGNIFICA HUMANITAS, §95

Presidiare ciò che è comune: dati, validazione, fornitori

5.1 Stabilità e nuove concentrazioni

Il valore ha un rovescio sistemico che la prospettiva di vigilanza vede con chiarezza. Se molte istituzioni adottano modelli simili, addestrati su dati simili, forniti da pochi grandi operatori tecnologici, il sistema può diventare più efficiente e insieme più fragile: errori correlati, comportamenti gregari nei momenti di stress, dipendenza operativa da terze parti che non rispondono alle regole bancarie. L'enciclica legge questa dinamica in termini di concentrazione del potere digitale, con infrastrutture, dati e capacità di calcolo nelle mani di pochi attori privati, spesso più forti degli Stati (cfr MH, §5, §95), e chiede che un potere di tale portata non resti opaco né fuori dal controllo pubblico.

5.2 Il rischio di modello come questione di sistema

Per il singolo istituto il modello che sbaglia è un costo; per il sistema, mille modelli che sbagliano insieme sono un rischio di stabilità. La risposta non è rallentare tutti, ma presidiare ciò che è comune: qualità e tracciabilità dei dati, validazione indipendente, diversificazione dei fornitori, piani di continuità che non presuppongano la disponibilità perpetua di un servizio esterno. Sono, tradotti nel linguaggio della vigilanza, gli stessi criteri di responsabilità e trasparenza che l'enciclica chiede per ogni uso dell'AI che tocchi beni pubblici e diritti fondamentali (cfr MH, §108).

RIQUADRO 4

Il crollo lampo del 6 maggio 2010

Il 6 maggio 2010, in pochi minuti, il trading automatico trascinò l'indice Dow Jones a perdere quasi mille punti, circa il 9%, prima di rimbalzare quasi altrettanto rapidamente. Per un breve intervallo evaporarono circa mille miliardi di dollari di valore. Nessun crollo «reale» lo giustificava: erano stati gli algoritmi a reagire ad altri algoritmi, amplificando una vendita localizzata fino a scuotere l'intero mercato. La lezione riguarda il sistema: quando molti operatori automatici condividono la stessa logica, la razionalità del singolo può diventare la fragilità di tutti.

Fonte: Rapporto congiunto SEC-CFTC sugli eventi del 6 maggio 2010 (30 settembre 2010).

PRINCIPIO IN GIOCO

STABILITÀ DEL SISTEMA

CAPITOLO 5 · Le poste in gioco per il sistema

5.3 La fiducia come capitale del sistema

La materia prima della banca non è il denaro: è la fiducia. Ogni episodio di opacità algoritmica, ogni discriminazione scoperta, ogni violazione di dati erode un capitale che appartiene al sistema intero, non al solo istituto che sbaglia. Per questo la qualità etica dell'AI bancaria è un bene comune in senso proprio: nessun operatore può custodirlo da solo, tutti possono comprometterlo. Ed è anche la ragione per cui la cooperazione di sistema, fatta di standard condivisi, segnalazione reciproca degli incidenti e ricerca comune sulla sicurezza dei modelli, non è ingenuità competitiva ma manutenzione dell'infrastruttura invisibile su cui tutti operano.

La materia prima della banca è la fiducia

5.4 Il quadro europeo come alleato

L'Europa ha scelto di disciplinare per prima l'intelligenza artificiale e la resilienza operativa del sistema finanziario. È una scelta che gli operatori possono vivere come un onere o riconoscere come un alleato: regole uguali per tutti trasformano la correttezza da svantaggio competitivo in condizione di mercato, e proteggono chi investe in qualità dalla concorrenza di chi taglia gli angoli. L'enciclica dà a questa intuizione il suo fondamento: non basta invocare l'etica, servono quadri giuridici adeguati, vigilanza indipendente e una politica che non abdichi al proprio compito (cfr MH, §106). Per il sistema italiano ed europeo, fare della conformità ben fatta una piattaforma di fiducia, e non un semplice adempimento, è un vantaggio strategico ancora largamente inespresso.

La qualità etica dell'AI bancaria è un bene comune

Non basta invocare l'etica: servono regole, vigilanza indipendente, politica che non abdichi.

CFR MAGNIFICA HUMANITAS, §106

5.5 Regole globali per un mondo digitale

Il limite del quadro europeo è la sua geografia. Come ha più volte osservato Anna Maria Tarantola, la realtà digitale è per sua natura mondiale, mentre le regole restano nazionali o continentali.¹ L'AI Act è un passo importante, fondato su un solido approccio per livelli di rischio, ma le sue norme non vincolano Stati Uniti e Cina, i due maggiori produttori di intelligenza artificiale. Da questa asimmetria nasce il rischio di una frammentazione normativa tra superpotenze tecnologiche: gli algoritmi non conoscono dogane, e nessuno Stato può proteggere da solo i propri cittadini. Di qui la sollecitazione a una gestione coordinata a livello locale e globale, guidata da principi antropocentrici ed etici condivisi e da regole davvero globali, applicabili a tutti i paesi, perché lo sviluppo tecnologico resti orientato al bene comune.

La realtà digitale è mondiale. Pertanto sarebbe opportuno che le norme siano globali, applicabili a tutti i paesi.

ANNA MARIA TARANTOLA

¹ Cfr. A. M. Tarantola, *L'impatto dell'IA sulle giovani generazioni* (Simposio «La Bibbia e l'AI»); *Le nuove generazioni nel mondo dell'intelligenza artificiale*, in *Dizionario della Dottrina Sociale della Chiesa*; e il volume da lei curato *Artificial Intelligence and Care of Our Common Home*.

DALLA RICERCA · PARTE SECONDA

“

*La finanza per la finanza è altra cosa dalla finanza
per lo sviluppo: la funzione sociale del credito resta
insostituibile.*

SINTESI DA MAGNIFICA HUMANITAS, §160



*Il risparmio che diventa credito all'economia reale
è il punto in cui la banca serve la vita.*



III



La posta etica

Servire lo sviluppo della persona: la bussola dell'enciclica per chi decide

6. La bussola: tre riferimenti convergenti
7. Il bivio: ottimizzare i risultati o servire lo sviluppo
8. Cinque principi per un'AI al servizio della persona

CAPITOLO 6

La bussola: tre riferimenti convergenti

Tre riferimenti, una sola
direzione: la persona come fine

Le norme dicono il «come»;
l'enciclica fonda il «perché».

SINTESI DEL RAPPORTO

L'etica comincia dove finisce
l'obbligo

6.1 Una convergenza che semplifica la vita a chi decide

Chi guida una banca non deve scegliere fra etiche concorrenti. I tre riferimenti disponibili, l'enciclica *Magnifica Humanitas*, i principi internazionali per un'AI affidabile e il quadro normativo europeo, convergono sugli stessi punti: la persona come fine, l'equità e la non discriminazione, la trasparenza delle decisioni, la responsabilità identificabile, la robustezza e la sicurezza dei sistemi. La convergenza non è casuale: nasce dal riconoscere lo stesso rischio, cioè che la potenza di calcolo, lasciata alla sola logica dell'ottimizzazione, tratti le persone come materiale del processo anziché come sua ragione.

Il contributo proprio dell'enciclica, dentro questa convergenza, è il fondamento: la dignità non dipende da ciò che la persona realizza o produce, e non deve essere guadagnata né dimostrata (cfr MH, §51, §53). Le norme proteggono la dignità; l'enciclica spiega perché la dignità precede le norme.

FIGURA 2 | La bussola: tre riferimenti convergenti (tutti orientano verso la stessa direzione: la persona come fine)



6.2 Oltre la conformità

Per chi decide, la differenza è pratica: la conformità dice fin dove si deve arrivare, il fondamento dice perché non fermarsi lì. C'è infatti uno spazio che nessuna regola può occupare: quello delle scelte lecite ma non per questo giuste, delle pratiche conformi che tradiscono lo spirito della norma, delle innovazioni che arrivano prima del legislatore. È lo spazio dell'etica propriamente detta, che comincia dove finisce l'obbligo; ed è lì che si forma, o si dissipa, la fiducia.

CAPITOLO 6 · La bussola: tre riferimenti convergenti

6.3 Che cosa chiede l'enciclica, in concreto

Letta con gli occhi di chi amministra, *Magnifica Humanitas* consegna indicazioni sorprendentemente operative. Sul metodo delle decisioni: criteri di discernimento chiari (dignità della persona, destinazione universale dei beni, attenzione ai più fragili, cura della casa comune) tradotti in prassi quali progettazione responsabile, valutazioni d'impatto umano e sociale, inclusione, alfabetizzazione digitale (cfr MH, §14). Sulla responsabilità: che sia sempre possibile identificare chi deve rendere conto delle decisioni, motivarle, controllarle e, quando necessario, contestarle e rimediare ai danni (cfr MH, §105). Sulla corresponsabilità: a ciascuno il suo tratto di muro, dagli scienziati agli imprenditori, dai lavoratori ai legislatori, nella logica della sussidiarietà (cfr MH, §13).

A ciascuno il suo tratto di muro: la corresponsabilità è la logica della ricostruzione.

CFR MAGNIFICA HUMANITAS, §13

6.4 La sussidiarietà nel digitale

Un passaggio dell'enciclica merita particolare attenzione da parte del sistema finanziario: nel contesto digitale, il livello «superiore» che assorbe competenze e capacità decisionale non è più soltanto lo Stato, ma ogni grande attore economico e tecnologico che fissa, di fatto, le condizioni della vita comune. La sussidiarietà chiede che questi processi non si impongano in modo opaco e unilaterale, ma siano resi partecipabili: verifiche indipendenti, trasparenza sugli algoritmi, accesso equo ai dati, strumenti di ricorso (cfr MH, §71). Una banca è, per i propri clienti, esattamente uno di quei livelli superiori. Il modo in cui rende contestabili le proprie decisioni automatizzate è la forma quotidiana della sua sussidiarietà.

Rendere conto, motivare, controllare, poter contestare

La banca è un «livello superiore»: la contestabilità è la sua sussidiarietà

CAPITOLO 7

Il bivio: ottimizzare i risultati o servire lo sviluppo

L'ideologia più insidiosa: attribuire più valore a chi è più efficiente e performante.

CFR MAGNIFICA HUMANITAS, §51

L'AI non ha inventato questa logica: può industrializzarla

Babele ottimizza le persone; Neemia le mette al lavoro insieme

7.1 L'ideologia più insidiosa

Al cuore dell'enciclica sta una diagnosi che riguarda da vicino la finanza. Tra le ideologie del presente, la più insidiosa è quella per cui ogni persona dovrebbe guadagnarsi il proprio valore, sicché varrebbe di più chi è più efficiente e performante: in questa prospettiva la persona diventa mezzo per ottenere risultati, risorsa da usare, e cessa di essere riconosciuta come fine (cfr MH, §51). Non è una descrizione astratta: è la logica implicita di ogni sistema che classifica le persone per redditività attesa e basta. L'AI non ha inventato questa logica; può però industrializzarla, applicandola a ogni cliente, in ogni istante, con una coerenza che nessuna organizzazione umana aveva mai raggiunto.

7.2 La sindrome di Babele, in versione finanziaria

L'enciclica apre con due immagini bibliche: la torre di Babele e la ricostruzione delle mura di Gerusalemme. Babele è il progetto grandioso che sacrifica la dignità delle persone all'efficienza e pretende di tradurre tutto, persino il mistero della persona, in dati e prestazioni (cfr MH, §7, §10). Esiste una versione finanziaria di questa sindrome: il sistema perfettamente ottimizzato in cui ogni cliente è un profilo, ogni decisione un punteggio, ogni relazione una transazione. Funziona, finché la fiducia regge. La via alternativa, quella di Neemia, è meno spettacolare: ricostruire pezzo per pezzo, con responsabilità condivisa, mettendo la persona al centro delle scelte (cfr MH, §8, §16).

CAPITOLO 7 · Il bivio: ottimizzare i risultati o servire lo sviluppo

7.3 Che cosa chiede l'enciclica a impresa e finanza

All'impresa, l'enciclica chiede di riconoscere nell'iniziativa economica una vocazione, capace di generare ricchezza e migliorare la vita di tutti, a condizione che il lavoro dignitoso sia parte essenziale del servizio reso alla società e non variabile dipendente dal profitto (cfr MH, §157); di includere la qualità e la dignità del lavoro tra gli indicatori di successo (cfr MH, §156); di accompagnare ogni introduzione di automazione con scelte verificabili di tutela, riqualificazione e partecipazione dei lavoratori (cfr MH, §156). Alla finanza chiede di ricordare la lezione delle crisi: un'intermediazione slegata da fondamenti antropologici e morali ha prodotto abusi e crisi sistemiche di portata mondiale (cfr MH, §160). E a entrambe chiede di superare la misura unica del prodotto interno lordo, adottando metriche che diano peso alla dignità del lavoro, alla prosperità condivisa, alla riduzione delle disuguaglianze (cfr MH, §159).

La dignità del lavoro tra gli indicatori di successo dell'impresa.

CFR MAGNIFICA HUMANITAS, §156

7.4 Il bivio non è tra profitto ed etica

Sarebbe un fraintendimento leggere l'enciclica come una requisitoria contro il profitto. Il bivio che essa indica è un altro: tra un profitto che nasce dal servire clienti, imprese e territori, e un profitto che nasce dall'estrarre. Il primo è compatibile con decenni; il secondo con qualche trimestre. La finanza sostenibile, prima ancora che una famiglia di prodotti, è questa scelta di origine del profitto. E l'AI, che amplifica qualunque logica le si consegna, rende la scelta più visibile e più urgente, non diversa.

Una finanza slegata da fondamenti morali ha già prodotto crisi sistemiche mondiali.

CFR MAGNIFICA HUMANITAS, §160

RIQUADRO 5

Babele e Neemia

L'enciclica si apre con due icone bibliche contrapposte. Babel è la grande opera che sceglie l'uniformità al posto della comunione e «sacrifica la dignità delle persone all'efficienza» (cfr MH, §7). Neemia è la ricostruzione delle mura di Gerusalemme: la città rinasce non per l'iniziativa di uno solo, ma per la responsabilità condivisa di tutto il popolo, che «ricostruisce i legami prima ancora delle pietre» (cfr MH, §8). Anche la finanza ha davanti questo bivio: la torre perfettamente ottimizzata in cui ogni cliente è un profilo, o la costruzione paziente che tiene la persona come fine.

LE IMMAGINI DELL'ENCICLICA

DUE CITTÀ

Profitto che serve o profitto che estrae: questo è il bivio

CAPITOLO 8

Cinque principi per un'AI al servizio della persona

La bussola del rapporto: i riferimenti convergenti dell'enciclica, dei principi internazionali e del quadro europeo, ricondotti a cinque parole che chiedono di diventare prassi.

Il progresso tecnico è autentico quando rende la vita più degna dell'uomo.

CFR MAGNIFICA HUMANITAS, §129

Cinque parole che chiedono di diventare prassi

Principi verificabili, non dichiarazioni d'intenti



PRINCIPIO I

Centralità della persona

Ogni sistema risponde a chi serve, non a ciò che ottimizza: nelle decisioni che toccano la vita delle persone, l'ultima parola resta a un essere umano responsabile (cfr MH, §103).



PRINCIPIO II

Equità e non discriminazione

I modelli sono verificati contro i pregiudizi di genere, di età e di territorio, prima del rilascio e durante l'esercizio, perché gli algoritmi opachi non riproducano discriminazioni (cfr MH, §80).



PRINCIPIO III

Trasparenza

Le decisioni algoritmiche sono comprensibili, contestabili, documentate: chi è toccato da una decisione ha il diritto di capirla e gli strumenti per ricorrere (cfr MH, §164, §71).



PRINCIPIO IV

Vigilanza continua

L'autorizzazione non esaurisce la responsabilità: la sorveglianza è permanente e si fonda su evidenze, non su dichiarazioni (cfr MH, §105-106).



PRINCIPIO V

Bene comune

Il valore creato dall'AI si misura anche su inclusione finanziaria, accesso al credito, coesione sociale: ciò che non entra nei cruscotti, entra nella storia (cfr MH, §85, §161).

CAPITOLO 8 · Cinque principi per un'AI al servizio della persona

8.1 L'uomo al centro non è uno slogan

«Centralità della persona» rischia di essere la formula più pronunciata e meno praticata della stagione. L'enciclica le restituisce un contenuto esigente articolandola in tre condizioni. La prima è la decisione: nelle scelte che incidono sulla vita delle persone, qualcuno deve poterne rispondere: la responsabilità non si dissolve nella macchina (cfr MH, §103, §105). La seconda è la comprensibilità: una decisione che non può essere spiegata a chi la subisce non è una decisione, è un verdetto (cfr MH, §164). La terza è l'apertura al cambiamento: i sistemi che inchiodano le persone al loro passato statistico negano ciò che le rende persone, la possibilità di diventare diverse (cfr MH, §102, §128).

Una decisione che non si può spiegare è un verdetto

8.2 Dall'etica dichiarata all'etica verificabile

Il passaggio decisivo, quello su cui si misura la serietà di un'istituzione, è dalla dichiarazione alla verifica. Un principio diventa reale quando è tradotto in requisiti che un sistema può soddisfare o violare, in soglie che un controllo può misurare, in evidenze che un revisore può ispezionare. L'equità diventa una metrica calcolata su ogni modello rivolto al cliente; la trasparenza, un obbligo di spiegazione con formati definiti; la vigilanza, un registro degli eventi avversi con responsabili e scadenze. Questa traduzione dell'etica in ingegneria non la impoverisce: la rende esigibile. Ed è esattamente ciò che distingue una carta dei valori da un sistema di governo.

Le persone non sono il loro passato statistico

Un principio è reale quando un controllo può misurarlo

RIQUADRO 6

Quando l'equità si misura: il caso COMPAS

Nel 2016 un'inchiesta della testata ProPublica analizzò COMPAS, un algoritmo usato nei tribunali statunitensi per stimare il rischio di recidiva: gli imputati neri venivano segnalati come ad alto rischio molto più spesso di imputati bianchi con profili simili. Il caso aprì il dibattito mondiale sull'equità algoritmica e mostrò una cosa che vale anche per il credito: l'equità non è una dichiarazione, ma una proprietà da misurare: su ogni modello, contro i dati, prima e dopo il rilascio.

Fonte: ProPublica, «Machine Bias» (J. Angwin et al.), 23 maggio 2016.

PRINCIPIO IN GIOCO

EQUITÀ

CAPITOLO 8 · Cinque principi per un'AI al servizio della persona

8.3 Disarmare l'AI

Disarmare non significa rinunciare alla tecnologia, ma impedirle di dominare l'umano.

CFR MAGNIFICA HUMANITAS,
§110

Competere sulla fiducia, non sulla potenza

Ogni scelta progettuale esprime una visione dell'umanità.

CFR MAGNIFICA HUMANITAS,
§111

L'enciclica consegna al lessico pubblico una parola destinata a durare: disarmare. Disarmare l'intelligenza artificiale significa sottrarla alla logica della competizione armata, non più solo militare ma economica e cognitiva: la corsa all'algoritmo più potente e alla banca dati più vasta per consolidare un vantaggio su tutti gli altri. Disarmare non è rinunciare alla tecnologia: è impedirle di dominare l'umano; è sottrarla ai monopoli e renderla discutibile, contestabile, e perciò abitabile (cfr MH, §110).

Per il sistema finanziario, la parola si traduce in scelte riconoscibili. Disarma l'AI la banca che non insegue la potenza del modello ma la sua affidabilità; che mantiene pluralità di fornitori per non consegnare a un solo attore le proprie decisioni; che rende contestabili gli esiti automatizzati invece di blindarli dietro la complessità; che condivide con il sistema le evidenze sugli incidenti invece di custodirle come segreti. Disarmare, in banca, significa scegliere la fiducia come terreno di competizione al posto della pura potenza computazionale.

8.4 Una richiesta anche a chi sviluppa

L'enciclica rivolge un appello diretto a chi sviluppa i sistemi: ogni scelta progettuale esprime una visione dell'umanità, e chi progetta porta un peso etico che non può scaricare sull'utilizzatore (cfr MH, §111). Per le banche, che sviluppano poco e acquistano molto, l'appello si trasforma in un criterio di acquisto: la visione dell'uomo incorporata in un sistema (come tratta gli errori, come gestisce i casi limite, che cosa ottimizza davvero) va valutata insieme al prezzo e alle prestazioni. Si chiama, da sempre, dovuta diligenza; l'enciclica chiede che diventi anche etica (cfr MH, §179).

CAPITOLO 8 · Cinque principi per un'AI al servizio della persona

8.5 Chi decide l'etica dei sistemi

Resta una domanda che l'enciclica pone con franchezza e che il sistema finanziario non può eludere: non basta «allineare» l'AI a valori umani se quei valori sono decisi da pochi; il codice etico incorporato nei sistemi deve poter essere discusso e sottoposto a criteri di giustizia condivisi (cfr MH, §107). Per una banca questo significa due cose. All'interno: le scelte etiche incorporate nei modelli (quali errori tollerare, quali categorie proteggere, dove fissare le soglie) sono decisioni di governo, da assumere negli organi competenti e da documentare, non parametri tecnici da lasciare ai team di sviluppo. All'esterno: la partecipazione ai luoghi dove gli standard si scrivono (comitati tecnici, consultazioni, associazioni di categoria) non è attività di rappresentanza, ma esercizio di responsabilità sul codice etico del sistema.

Non serve un'AI più morale, se quella morale è decisa da pochi.

CFR MAGNIFICA HUMANITAS, §107

8.6 Il limite come informazione

Un ultimo insegnamento della terza parte riguarda il limite. La cultura dell'ottimizzazione tratta ogni limite come un difetto da rimuovere; l'enciclica ricorda che l'umano matura spesso attraverso il limite, non malgrado esso (cfr MH, §118). Tradotto nel mestiere bancario: il cliente fragile, il caso che non rientra nel modello, l'eccezione che richiede giudizio non sono attriti da eliminare, ma informazioni sul mondo reale, e occasioni in cui la banca dimostra di essere qualcosa di più di un sistema di punteggi. Le organizzazioni che cancellano ogni attrito cancellano anche i momenti in cui la fiducia si costruisce.

Le soglie dei modelli sono decisioni di governo, non parametri tecnici

Dove c'è attrito, spesso si costruisce fiducia

DALLA RICERCA · PARTE TERZA

“

Disarmare l'intelligenza artificiale: sottrarla ai monopoli, renderla discutibile e contestabile, e perciò abitabile.

SINTESI DA MAGNIFICA HUMANITAS, §110



In banca, disarmare significa competere sulla fiducia anziché sulla pura potenza di calcolo.

IV



Governare l'AI etica

Controlli, sicurezza e inclusione: chi decide, come si controlla, quali rischi si presidiano

9. La responsabilità di chi decide
10. Il ciclo dell'AI-vigilanza
11. La sicurezza come tema di Consiglio
12. Equità, genere, inclusione

CAPITOLO 9

La responsabilità di chi decide

9.1 Nessuno è senza responsabilità

C'è chi governa, chi decide investimenti, chi guida istituzioni: nessuno è senza responsabilità.

CFR MAGNIFICA HUMANITAS, §212

L'enciclica osserva che non tutti hanno lo stesso potere di incidere sulla realtà, perché c'è chi governa, chi decide investimenti, chi guida istituzioni, ma nessuno è privo di un proprio ambito di azione e di responsabilità (cfr MH, §212). In banca questa osservazione disegna una mappa precisa. Il Consiglio di amministrazione risponde dell'indirizzo: quali usi dell'AI sono coerenti con la missione dell'istituto, quali rischi si accettano, quali si rifiutano. Il vertice esecutivo risponde dell'attuazione; le funzioni di controllo della verifica indipendente; e chi usa i sistemi ogni giorno risponde dell'uso: nessuna architettura di governo sostituisce il giudizio di chi preme il tasto.

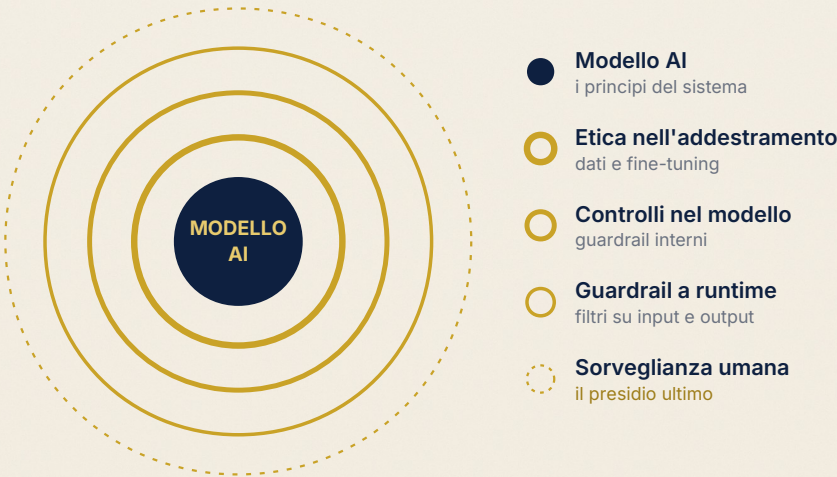
9.2 La difesa in profondità

Una difesa in profondità, non un muro unico

Sul piano tecnico, i controlli per un'AI affidabile non sono un muro unico, ma una difesa in profondità: si dispongono a cerchi concentrici attorno al modello, dall'etica incorporata nell'addestramento ai filtri che, durante l'esercizio, vigilano su ciò che entra e su ciò che esce. Quattro i presidi: proteggere l'input, controllare l'output, blindare modello e fornitori, limitare i poteri degli agenti autonomi. Nessuno strato, da solo, basta: l'anello più esterno resta la sorveglianza di una persona responsabile.

FIGURA 3 | La difesa in profondità: i presidi a cerchi concentrici (dal cuore del modello al perimetro della sorveglianza umana)

Nessuno strato basta da solo: l'anello esterno è la persona



CAPITOLO 9 · La responsabilità di chi decide

9.3 Il sapere necessario per deliberare

Una responsabilità esercitata senza comprensione è una firma, non una decisione. I Consigli non devono diventare collegi di tecnici, ma devono possedere, anche con il sostegno di esperti indipendenti, ciò che serve a porre le domande giuste: su quali dati il sistema è addestrato, che cosa accade quando sbaglia, chi se ne accorge e in quanto tempo.

Una responsabilità senza comprensione è una firma, non una decisione

9.4 L'architettura dei presìdi

Gli strati di difesa si traducono in un'architettura di governo, che il rapporto sviluppa presidio per presidio. Va proporzionata alla dimensione dell'istituto, e misura la propria efficacia sugli esiti raggiunti, non sulla propria esistenza.

TAVOLA 2 | L'architettura dei presìdi

PRESIDIO	RESPONSABILITÀ	STRUMENTO CHIAVE	FREQUENZA
Indirizzo strategico	Consiglio di Amministrazione	Politica AI e soglie di tolleranza al rischio	ANNUALE
Governo del portafoglio modelli	Comitato AI di Gruppo	Registro dei sistemi e classificazione del rischio	TRIMESTRALE
Validazione indipendente	Seconda linea di controllo	Collaudo etico-tecnico pre-rilascio	PER RILASCIO
Sorveglianza in esercizio	Funzione AI-vigilanza	Indicatori di deriva e registro incidenti	CONTINUA
Sicurezza e resilienza	CISO · Consiglio	Scenari avversi e piani di risposta	SEMESTRALE
Equità e inclusione	Comitato etico	Verifiche di equità sui modelli al cliente	SEMESTRALE

La forma organizzativa della responsabilità

I presìdi si misurano sugli esiti, non sull'esistenza

CAPITOLO 9 · La responsabilità di chi decide

9.5 Le tre linee di difesa applicate all'AI

Il sistema dei controlli interni che le banche conoscono da tempo vale anche per l'intelligenza artificiale, a patto di rileggerlo alla luce delle sue caratteristiche. La prima linea è chi costruisce e usa i modelli: a lei spetta documentare le scelte fatte, registrare i limiti del sistema, far emergere gli scostamenti appena si manifestano. La seconda linea, cioè la conformità, la gestione del rischio e il responsabile della protezione dei dati, verifica in modo indipendente che i modelli rispettino le regole, le soglie e i diritti delle persone. La terza linea, la revisione interna, accerta che questo impianto funzioni davvero. Presa da sola, nessuna delle tre basta: è il loro disegno d'insieme a fare la differenza, non la loro presenza sulla carta.

Tre linee di difesa, un solo disegno

9.6 Il secondo livello deve evolvere

Il cambiamento più impegnativo tocca la seconda linea. La conformità deve imparare a leggere un modello, non più soltanto un contratto: capire su quali dati è stato addestrato, quali decisioni produce, dove rischia di sbagliare. La gestione del rischio deve trattare il rischio di modello con lo stesso rigore che applica al rischio di credito, fissando metriche, soglie e limiti operativi. Il responsabile della protezione dei dati deve entrare nel ciclo di vita del sistema dalle prime scelte, non a cose ormai fatte. Tutto ciò richiede competenze nuove, accesso effettivo ai dati e ai modelli, un mandato esplicito del Consiglio. Quando queste condizioni mancano, la seconda linea finisce per apporre una firma su ciò che non comprende.

La conformità deve leggere un modello, non solo un contratto

RIQUADRO 7

Ventiseimila famiglie e un algoritmo

Nei Paesi Bassi un algoritmo antifrode del fisco classificò come frodatori circa 26.000 famiglie, penalizzando chi aveva nazionalità non olandese e basso reddito e sospendendo i benefici senza revisione umana. Migliaia di famiglie furono rovinate. Nel gennaio 2021 il governo si dimise. È la lezione che nessun Consiglio dovrebbe dimenticare: quando la responsabilità si dissolve nella macchina, il danno è reale ma la colpa non è di nessuno.

Fonte: Inchiesta parlamentare olandese e Autorità per la protezione dei dati (2020-2021).

PRINCIPIO IN GIOCO

RESPONSABILITÀ

CAPITOLO 9 · La responsabilità di chi decide

9.7 Una funzione che vede l'AI per intero

Le funzioni di controllo non possono più osservare l'AI a compartimenti separati. Un solo modello di credito tocca insieme la conformità, che presidia le regole; il rischio, che misura la bontà della previsione; la protezione dei dati, che custodisce le informazioni personali impiegate; l'equità, che pesa gli effetti sui diversi gruppi di clienti. Per questo molte banche stanno costituendo un presidio unico, un comitato o una funzione dedicata, che riunisce queste competenze e risponde direttamente al Consiglio. Non è l'aggiunta di un nuovo strato di burocrazia. È la creazione del luogo in cui responsabilità diverse si parlano prima che il modello entri in produzione, e non dopo, quando l'errore è già arrivato al cliente.

Le scelte che contano si prendono alla lavagna

9.8 Le garanzie si progettano, non si aggiungono

C'è un momento preciso in cui un sistema di AI diventa affidabile, oppure non lo diventa: il momento in cui lo si progetta. Le tutele che rendono un modello equo, comprensibile e sicuro non si possono applicare alla fine, come una vernice sopra un oggetto già costruito. Vanno pensate nel disegno, quando si scelgono i dati, l'architettura, i limiti d'uso. L'enciclica lo afferma senza ambiguità: la giustizia non è un correttivo da applicare in seguito, ma una condizione da praticare nel disegno stesso dei sistemi (cfr MH, §109). Per la banca questo sposta il baricentro delle decisioni. Le scelte che davvero contano non si prendono al collaudo: si prendono alla lavagna, mesi prima.

La giustizia si pratica nel disegno dei sistemi, non come correttivo successivo.

CFR MAGNIFICA HUMANITAS,
§109

CAPITOLO 9 · La responsabilità di chi decide

9.9 Quattro controlli da pensare nel disegno

Quattro famiglie di controlli vanno previste fin dall'inizio. I controlli sui dati di addestramento ne verificano la qualità, la rappresentatività e la liceità, perché un modello eredita i difetti di ciò che riceve. I controlli sul modello ne documentano il comportamento, ne fissano i limiti d'uso e stabiliscono in anticipo come fermarlo. I controlli sui dati in ingresso filtrano ciò che entra nel sistema, così che istruzioni o informazioni manipolate non lo possano deviare. I controlli sui dati in uscita esaminano gli esiti prima che raggiungano il cliente, con una persona che resta l'ultima a decidere. Sono i quattro punti esatti in cui una delibera del Consiglio si traduce, alla fine della catena, in righe di codice.

Quattro controlli: dati, modello, ingresso, uscita

9.10 Dalla delibera al codice

Fra la decisione del Consiglio e il comportamento del sistema corre una catena di traduzioni: la politica diventa requisito, il requisito diventa controllo, il controllo diventa misura osservabile. Una governance forte presidia l'intera catena, non soltanto il suo primo anello. È anche ciò che consente al Consiglio di rendere conto, davanti all'autorità di vigilanza e ai clienti, mostrando le evidenze invece delle buone intenzioni: quali controlli erano previsti, come hanno funzionato alla prova dei fatti, che cosa è stato corretto quando qualcosa è andato storto.

Dalla delibera del Consiglio alle righe di codice

CAPITOLO 9 · La responsabilità di chi decide

9.II La governance come vantaggio, non come peso

Si è abituati a vedere nella governance un costo, un freno alla velocità di adozione. Nel tempo dell'AI vale il contrario. La banca che sa spiegare e correggere i propri sistemi può adottarli più rapidamente, perché ne tiene sotto controllo i rischi. Quella che non lo sa fare procede al buio, e prima o poi ne paga il conto in perdite, sanzioni o fiducia perduta.

Una governance solida non rallenta l'innovazione: le offre le fondamenta su cui reggersi. È la distanza che separa il correre dal correre rischi. Ed è, in ultima analisi, ciò che il Consiglio può offrire di più prezioso a un'impresa che affida parti crescenti del proprio mestiere a sistemi che apprendono: non la promessa che nulla andrà storto, ma la capacità di accorgersene in tempo e di rispondere.

Spiegare e correggere permette di adottare più in fretta

RIQUADRO 11

La disciplina del rischio di modello

Molto prima dell'AI generativa, le autorità di vigilanza avevano compreso che un modello può sbagliare in modo sistematico e costoso. Negli Stati Uniti la guida SR 11-7 della Federal Reserve e dell'OCC, del 2011, ha fissato il quadro di riferimento per la gestione del rischio di modello: ogni modello va convalidato da chi non l'ha costruito, sorvegliato nel tempo, documentato in ogni scelta. L'intelligenza artificiale non rende obsoleto quel principio. Lo rende più urgente, perché i modelli che apprendono cambiano comportamento dopo il rilascio. La banca non parte da zero: parte da una disciplina che deve estendere.

Fonte: Federal Reserve e OCC, «Supervisory Guidance on Model Risk Management» (SR 11-7), 2011.

IL RIFERIMENTO

GOVERNANCE DEL RISCHIO

CAPITOLO 10

Il ciclo dell'AI-vigilanza

10.1 Dalla farmacovigilanza alla finanza

Un sistema autorizzato non è un sistema dimenticato. Il rapporto propone un ciclo permanente in quattro tempi, presidiato dal vertice aziendale, sul modello della sorveglianza che da decenni accompagna il farmaco lungo tutto il suo ciclo di vita.

FIGURA 4 | Il ciclo dell'AI-vigilanza



Il parallelo con il farmaco è esigente per scelta: come un principio attivo, un modello produce effetti desiderati ed effetti avversi, che si manifestano nell'uso e nel tempo. La vigilanza non è sfiducia nella tecnologia: è la forma matura della fiducia, quella che si fonda su evidenze e non su dichiarazioni.

La vigilanza è la forma matura della fiducia.

TESI DEL RAPPORTO

CAPITOLO 10 · Il ciclo dell'AI-vigilanza

10.2 Perché serve un nome nuovo

Il sistema finanziario conosce da decenni la validazione dei modelli. L'AI-vigilanza non la sostituisce: la prolunga nel tempo e la allarga negli oggetti. Nel tempo, perché i sistemi che apprendono cambiano comportamento con i dati che incontrano (la deriva non è un'eventualità, è la fisiologia), e perché molti effetti si manifestano solo nell'uso reale, su popolazioni che i collaudi non rappresentano. Negli oggetti, perché accanto alle metriche di prestazione vanno sorvegliate quelle di equità, di robustezza, di concentrazione: un modello può restare accurato in media e diventare ingiusto ai margini.

La deriva dei modelli non è un'eventualità: è la fisiologia

10.3 Il registro degli eventi avversi

Il cuore del ciclo è il registro degli eventi avversi: la raccolta sistematica, con responsabili e scadenze, di ciò che i sistemi sbagliano: decisioni errate, esiti anomali, reclami fondati, incidenti sfiorati. Tre regole lo rendono vivo. Chi segnala è protetto, perché un registro che punisce la sincerità raccoglie silenzio. Ogni evento produce una decisione tracciata: correzione, riaddestramento, ritiro. E gli esiti aggregati salgono fino al Consiglio con cadenza definita, perché la vigilanza che non informa il vertice non è governo, è archivio. L'enciclica fonda questo impianto nel principio di responsabilità: che sia sempre possibile identificare chi rende conto, motivare le decisioni, contestarle e rimediare ai danni (cfr MH, §105).

Accurato in media, ingiusto ai margini: ciò che la sola prestazione non vede

Un registro che punisce la sincerità raccoglie silenzio

RIQUADRO 8

Il talidomide e la nascita della farmacovigilanza

Commercializzato dalla fine degli anni Cinquanta come sedativo sicuro e prescritto anche in gravidanza, il talidomide si rivelò responsabile di gravi malformazioni nei neonati. Nel 1961 due medici, McBride in Australia e Lenz in Germania, collegarono indipendentemente il farmaco ai danni; fu ritirato nel dicembre di quell'anno. Da quella tragedia nacque la farmacovigilanza: la sorveglianza sistematica degli effetti avversi di un farmaco dopo la sua immissione in commercio. Il rapporto propone per l'AI la stessa logica, perché come un farmaco un modello rivela i suoi effetti nell'uso e nel tempo.

Fonte: McBride, «Lancet» (1961); Lenz (1961); Programma OMS di farmacovigilanza (1968).

L'ORIGINE DELLA VIGILANZA

FARMACOVIGILANZA

CAPITOLO 11

La sicurezza come tema di Consiglio

II.1 Perché la sicurezza è cambiata di natura

Scudo e arma: l'AI cambia la sicurezza due volte

L'intelligenza artificiale ha cambiato la sicurezza informatica due volte: come scudo e come arma. Come scudo, perché moltiplica la capacità di rilevare anomalie e di rispondere agli attacchi. Come arma, perché abbassa il costo dell'offesa: frodi su misura, impersonificazioni convincenti, attacchi che si adattano alle difese. A questo si aggiunge una superficie nuova: i sistemi di AI stessi diventano bersagli: dati avvelenati, modelli manipolati, istruzioni nascoste. Per una banca, la conseguenza è che la sicurezza ha smesso di essere una funzione tecnica ed è diventata una condizione dell'attività: si può perdere in un giorno la fiducia costruita in un secolo.

Si può perdere in un giorno la fiducia di un secolo

II.2 Dal piano scritto al piano provato

La risposta che il rapporto propone è di metodo: portare la sicurezza in Consiglio non come informativa periodica, ma come esercizio. Scenari avversi discussi almeno due volte l'anno con il vertice al tavolo; piani di risposta provati, non solo scritti; simulazioni che includano la dimensione pubblica: che cosa si dice ai clienti, ai mercati e alle autorità nelle prime ore. La resilienza operativa che il quadro europeo richiede trova qui la sua interpretazione più seria: non l'adempimento documentale, ma la capacità dimostrata di continuare a servire i clienti sotto attacco.

Piani provati, non solo scritti

CAPITOLO 11 · La sicurezza come tema di Consiglio

11.3 Le terze parti critiche

La sicurezza di una banca è ormai, in misura crescente, la sicurezza dei suoi fornitori. Modelli, infrastrutture di calcolo e dati risiedono spesso presso pochi grandi operatori esterni: una concentrazione che la prospettiva di vigilanza europea ha riconosciuto come rischio sistemico e che l'enciclica legge come questione di potere: quando infrastrutture decisive per la vita comune si concentrano in poche mani, la trasparenza e il controllo non sono opzioni (cfr MH, §95, §108). Per il singolo istituto la traduzione è prosaica e impegnativa: conoscere davvero le proprie dipendenze, negoziare diritti di verifica effettivi, mantenere alternative praticabili, provare periodicamente l'uscita. La sovranità operativa non è autarchia: è la capacità di cambiare strada senza fermarsi.

Quando infrastrutture decisive si concentrano in poche mani, trasparenza e controllo non sono opzioni.

CFR MAGNIFICA HUMANITAS, §95, §108

11.4 Il dovere di cooperare

Contro avversari che condividono strumenti e tecniche, la difesa frammentata è una difesa perdente. La cooperazione di sistema, fatta di condivisione tempestiva degli indicatori di minaccia, esercitazioni congiunte e canali stabili con le autorità, è la forma che la solidarietà assume in questo campo. L'enciclica ne dà la ragione di fondo: la sicurezza degli ambienti digitali è parte del bene comune, e il bene comune non si custodisce da soli (cfr MH, §187). Nessuna banca è un'isola; nel cyberspazio, meno che mai.

Sovranità operativa: cambiare strada senza fermarsi

RIQUADRO 9

La «Top 10» delle minacce all'AI

Il catalogo più autorevole dei rischi specifici dei sistemi di AI generativa, la OWASP Top 10 for LLM Applications, elenca minacce che il software tradizionale non conosceva: la *prompt injection* (manipolare il modello con istruzioni nascoste nei dati), l'avvelenamento dei dati di addestramento, l'eccesso di autonomia concesso agli agenti. Per una banca significa che la sicurezza non può più essere delegata alla sola funzione tecnica: la superficie d'attacco comprende ora i dati del modello, le sue istruzioni e la sua stessa autonomia.

Fonte: OWASP, «Top 10 for LLM Applications».

Nel cyberspazio nessuna banca è un'isola

IL RIFERIMENTO

SICUREZZA

CAPITOLO 11 · La sicurezza come tema di Consiglio

II.5 L'AI offensiva: un avversario nuovo

L'attacco non sfonda una porta:
persuade una persona

La stessa tecnologia che difende la banca arma anche chi la attacca. L'intelligenza artificiale abbassa il costo dell'offesa e ne alza la qualità. Scrive messaggi di phishing in un italiano impeccabile e cuciti sulla singola vittima. Imita la voce e il volto di un dirigente dentro una videochiamata. Costruisce programmi capaci di adattarsi alle difese che incontrano. Il salto non è soltanto di scala, è di natura: l'attacco non sfonda più una porta, persuade una persona. E contro la persuasione le difese tecniche tradizionali, fatte di firewall e filtri, possono fare ben poco.

Per una banca questo ridisegna la mappa del rischio. La frode più pericolosa non arriva più da un'intrusione nei sistemi, ma da una richiesta che sembra legittima e proviene, in apparenza, da una persona di fiducia. La formazione del personale, le procedure di verifica a più livelli, la sana abitudine di diffidare dell'urgenza diventano presidi di sicurezza tanto quanto i sistemi informatici.

RIQUADRO 12

Il direttore finanziario che non era il direttore

Nel gennaio 2024 un dipendente dell'ufficio di Hong Kong di Arup, società di ingegneria londinese, ricevette dal «direttore finanziario» la richiesta di alcuni trasferimenti riservati. Insospettito, chiese una videochiamata per verificare. Sullo schermo comparvero il direttore e alcuni colleghi: li riconobbe, parlavano e si muovevano in modo convincente. Erano tutti deepfake costruiti con l'AI a partire da video pubblici. Rassicurato, eseguì quindici bonifici per circa venticinque milioni di dollari. Nessun sistema informatico era stato violato: l'attacco non passò dal firewall, passò dalla fiducia.

Fonte: Polizia di Hong Kong (febbraio 2024); CNN (maggio 2024).

IL CASO

SICUREZZA

CAPITOLO 11 · La sicurezza come tema di Consiglio

11.6 AI contro AI: la difesa aumentata

Se l'attacco impiega l'intelligenza artificiale, anche la difesa deve impiegarla. Nessuna squadra umana può esaminare in tempo reale i milioni di eventi che attraversano i sistemi di una banca; un modello sì. L'AI difensiva riconosce gli schemi anomali in mezzo al rumore, segnala la transazione che non torna, isola il sistema colpito prima che il danno si propaghi, propone una risposta mentre l'attacco è ancora in corso. Il principio resta quello già incontrato per il credito e per la vigilanza: la macchina vede e segnala, la persona valuta e decide. Ma senza la macchina, oggi, la persona vedrebbe troppo tardi.

La macchina vede e segnala, la persona decide

11.7 Una corsa senza traguardo

La sicurezza, nel tempo dell'AI, ha smesso di essere un muro da innalzare una volta per tutte. È diventata una corsa che non finisce: ogni difesa nuova suscita un attacco nuovo, e ogni attacco nuovo richiede una difesa nuova. Per il Consiglio questo comporta un cambio di mentalità. Va abbandonata l'idea di una soluzione definitiva e adottata quella di una capacità da mantenere, finanziare e aggiornare di continuo, come si fa con la liquidità o con il capitale. La domanda giusta non è più «siamo sicuri?». È «in quanto tempo ci accorgiamo di un attacco, e in quanto tempo ci rialziamo?».

Non «siamo sicuri?» ma «quanto in fretta ci rialziamo?»

CAPITOLO 11 · La sicurezza come tema di Consiglio

II.8 Il mandato del Consiglio sulla sicurezza

La difesa cibernetica nel tempo dell'AI non si delega per intero alla funzione tecnica, perché chiama in causa scelte che spettano al vertice: quanto investire, quali rischi accettare consapevolmente, come comportarsi nelle prime ore di una crisi, quando ogni minuto conta. Il Consiglio fissa il mandato e ne verifica gli esiti, senza entrare nel dettaglio tecnico ma senza nemmeno disinteressarsene.

Tre domande aiutano a non illudersi di essere al sicuro. I nostri sistemi di AI sono stati messi alla prova contro attacchi che usano a loro volta l'AI? Sappiamo distinguere un dirigente vero da uno falso, prima di autorizzare un'operazione? Abbiamo deciso in anticipo che cosa diremo ai clienti, ai mercati e alle autorità se un attacco dovesse riuscire? Le risposte a queste domande misurano la sicurezza di una banca meglio di qualunque certificato appeso alla parete.

II.9 Il fattore umano resta decisivo

Per quanto sofisticata, nessuna difesa tecnologica sostituisce le persone. L'attacco subito da Arup fu fermato, alla fine, da una verifica umana: troppo tardi per il denaro, in tempo per capire l'accaduto. Investire nella consapevolezza di chi lavora, costruire procedure che rendano normale fermarsi e controllare, premiare chi solleva un dubbio invece di penalizzarlo sono scelte di governo, non di tecnica. E sono spesso le più efficaci, perché agiscono sull'anello che gli attaccanti hanno imparato a colpire per primo.

Distinguere un dirigente vero da uno falso

La sicurezza è una capacità da mantenere, come il capitale

CAPITOLO 12

Equità, genere, inclusione

12.1 Disuguaglianze ereditate, disuguaglianze automatizzate

I sistemi che apprendono dal passato imparano anche le sue ingiustizie. Se la storia del credito ha penalizzato le donne, i giovani, i territori marginali, un modello addestrato su quella storia tenderà a riprodurla, con l'aggravante della scala e dell'apparenza di oggettività. L'enciclica nomina il rischio senza giri di parole: gruppi sociali penalizzati da algoritmi opachi che riproducono pregiudizi e discriminazioni (cfr MH, §80); e ricorda che l'uguale dignità delle donne, finché non si traduce in accesso effettivo al credito, al lavoro e alle responsabilità, resta un'affermazione incompiuta (cfr MH, §57).

Algoritmi opachi possono riprodurre pregiudizi e discriminazioni.

CFR MAGNIFICA HUMANITAS, §80

12.2 Le verifiche di equità

La risposta tecnica esiste e va resa obbligatoria all'interno dell'istituto: ogni modello rivolto al cliente è sottoposto a verifiche di equità indipendenti, documentate e ripetute, prima del rilascio e durante l'esercizio. Le verifiche misurano gli esiti per gruppi rilevanti, indagano le disparità, distinguono le differenze giustificate dal merito di credito da quelle che il merito non spiega. Dove la disparità ingiustificata emerge, il modello si corregge o non si rilascia. Non è un appesantimento: è la stessa disciplina che la banca da sempre applica al rischio, estesa alla giustizia.

La scala e l'apparenza di oggettività aggravano l'ingiustizia

La disciplina del rischio, estesa alla giustizia

CAPITOLO 12 · Equità, genere, inclusione

12.3 L'inclusione come misura del valore

L'equità evita il danno; l'inclusione crea il valore. Una banca può avere modelli perfettamente equi e continuare a servire soltanto chi era già servito. L'enciclica sposta l'asticella: la giustizia, nel tempo del digitale, passa anche attraverso l'accesso ai benefici dell'innovazione (cfr MH, §161); e il valore creato dall'AI va misurato anche su inclusione finanziaria, accesso al credito, coesione sociale. Il rapporto propone perciò che gli istituti integrino nel cruscotto direzionale indicatori di inclusione (nuovi clienti prima esclusi, credito a soggetti senza storia tradizionale, divari di accesso per genere e territorio) con la stessa dignità degli indicatori di redditività.

La giustizia passa anche attraverso l'accesso ai benefici dell'innovazione.

CFR MAGNIFICA HUMANITAS,
§161

12.4 Le persone dentro la banca

L'equità riguarda infine chi in banca lavora. L'enciclica avverte che, mentre l'AI promette di sollevare dalle mansioni ripetitive, i lavoratori rischiano di doversi adattare al ritmo delle macchine, di essere dequalificati e sottoposti a sorveglianza automatizzata (cfr MH, §150); e chiede che ogni introduzione di automazione sia accompagnata da scelte verificabili di tutela, riqualificazione e partecipazione (cfr MH, §156). Per il sistema bancario, che attraversa una trasformazione profonda delle professioni, questo è il banco di prova interno della coerenza: non si può promettere ai clienti una tecnologia che serve la persona e praticare con i propri collaboratori una tecnologia che la sorveglia.

Indicatori di inclusione con la stessa dignità di quelli di redditività

La coerenza si misura prima di tutto dentro la banca

RIQUADRO 10

«È solo l'algoritmo»: il caso Apple Card

Nel novembre 2019 uno sviluppatore raccontò pubblicamente di aver ottenuto dalla Apple Card un limite di credito venti volte superiore a quello della moglie, che pure aveva un punteggio di credito migliore. L'assistenza sapeva ripetere solo: «è l'algoritmo». L'autorità finanziaria di New York aprì un'indagine e non accertò violazioni di legge, ma il caso restò emblematico: anche quando il genere non è tra i dati in ingresso, un esito che nessuno sa spiegare è, per chi lo subisce, indistinguibile dall'arbitrio.

Fonte: New York State Dept. of Financial Services (2021); stampa, novembre 2019.

PRINCIPIO IN GIOCO

EQUITÀ E TRASPARENZA

CAPITOLO 12 · Equità, genere, inclusione

12.5 Le donne e l'accesso al credito

Tra i gruppi che un modello mal disegnato rischia di penalizzare, le donne occupano un posto particolare, perché la disparità che le riguarda è antica e si è depositata nei dati. Per generazioni il credito è stato intestato, garantito e misurato attorno alla figura maschile. Le carriere femminili, più spesso interrotte dal lavoro di cura, appaiono agli occhi di un modello come storie irregolari, e dunque più rischiose, anche quando non lo sono affatto. Un sistema che apprende da quel passato tende a chiedere alle donne garanzie più alte, o a riservare loro condizioni peggiori, e lo fa senza che nessuno l'abbia deciso, dietro l'apparenza rassicurante di un calcolo oggettivo.

L'enciclica è netta: l'uguale dignità delle donne resta un'affermazione incompiuta finché non si traduce in accesso effettivo al credito, al lavoro e alle responsabilità (cfr MH, §57). Per una banca la traduzione è concreta e verificabile. Si misurano separatamente gli esiti dei modelli per le donne e per gli uomini. Si indaga ogni differenza che emerge. Si distingue ciò che il merito di credito spiega da ciò che non spiega in alcun modo. E dove la disparità non ha giustificazione, si corregge il modello prima di affidargli anche una sola decisione.

*L'uguale dignità delle donne
attende ancora di tradursi in
accesso effettivo.*

CFR MAGNIFICA HUMANITAS,
§57

**Misurare separatamente gli esiti
per donne e uomini**

CAPITOLO 12 · Equità, genere, inclusione

12.6 I più esposti

Lo stesso vale per chiunque si presenti al credito con una storia fuori dagli schemi consueti: i giovani senza un passato finanziario, i lavoratori autonomi dal reddito discontinuo, i migranti privi delle garanzie tradizionali, gli abitanti di territori che i dati descrivono come marginali. Sono le persone per cui l'AI promette di più, perché può leggere segnali che i bilanci ignorano. E sono le stesse per cui rischia di sbagliare di più, perché i dati che le riguardano sono pochi e spesso distorti. Davanti all'incertezza, su chi è già fragile, un modello tende alla prudenza eccessiva: nel dubbio, esclude.

Qui si misura tutta la distanza tra un'AI che allarga l'accesso e una che lo restringe. La prima viene disegnata sapendo dove può sbagliare: prevede una revisione umana per i casi limite, lascia alla persona la possibilità di spiegare la propria storia, non chiude la porta soltanto perché i dati scarseggiano. La seconda si limita a ottimizzare un numero, e scarica sui più deboli il costo della propria incertezza. La scelta tra le due strade, ancora una volta, non la compie l'algoritmo: la compiono le persone che lo progettano e il Consiglio che ne fissa le regole.

Nel dubbio, un modello incerto
esclude i più fragili

Non chiudere la porta solo
perché i dati scarseggiano

DALLA RICERCA · PARTE QUARTA

“

*Le decisioni che toccano le persone devono restare
comprensibili, contestabili e controllate: nessuno sia
ridotto a profilo.*

SINTESI DA MAGNIFICA HUMANITAS, §164



*È il requisito di progetto di ogni sistema degno
di entrare in una banca.*



V



La responsabilità dei CEO

*Il bilanciamento tra obiettivi d'impresa ed etica della finanza, nel tempo in cui
l'AI pervade ogni processo*

CAPITOLO 13

Solo un'AI affidabile rende la finanza sostenibile

13.1 Perché ora, e perché loro

Questo capitolo si rivolge a un destinatario preciso: chi guida le banche. Non per cortesia istituzionale, ma per una ragione di fatto: l'intelligenza artificiale sta entrando in tutti i processi aziendali nello stesso momento (credito, investimenti, pagamenti, relazione, controlli, persone), e le scelte architetturali che si compiono ora resteranno nei sistemi per anni. Quando la tecnologia pervade tutto, la sua direzione non è più delegabile a una funzione: coincide con la direzione dell'impresa. L'enciclica lo dice del nostro tempo in generale: la tecnologia assume il volto di chi la pensa, la finanzia, la regola e la usa (cfr MH, §9); per una banca, quel volto è in ultima istanza quello del suo vertice.

C'è poi una ragione che riguarda la natura del prodotto. La banca non vende oggetti: vende promesse: custodire, restituire, sostenere. Una promessa vale quanto la fiducia in chi la fa. Se i sistemi che decidono il credito, muovono i risparmi e parlano con i clienti non sono affidabili, cioè esatti, equi, comprensibili e sicuri, la promessa si svuota dall'interno, qualunque cosa dicano i bilanci. Da qui la tesi che dà il titolo al capitolo: nel tempo dell'AI, solo un'intelligenza artificiale affidabile, che ponga al centro l'individuo, può rendere la finanza effettivamente sostenibile. Non è un'aggiunta etica al modello di business: è la sua condizione di durata.

Quando l'AI pervade tutto, la sua direzione coincide con quella dell'impresa

La banca vende promesse: l'affidabilità dei sistemi è la sostanza del prodotto

La tecnologia assume il volto di chi la pensa, la finanzia, la regola e la usa.

CFR MAGNIFICA HUMANITAS, §9

CAPITOLO 13 · Solo un'Al affidabile rende la finanza sostenibile

13.2 Il bilanciamento, visto dal vertice

Il rapporto non ignora la condizione reale di chi guida: obiettivi trimestrali, concorrenti che annunciano, investitori che confrontano, tecnologie che promettono riduzioni di costo immediate contro benefici reputazionali differiti. Sarebbe ingeneroso, oltre che inutile, chiedere ai vertici di ignorare questa pressione. La tesi del rapporto è diversa: il bilanciamento tra obiettivi d'impresa ed etica della finanza non si risolve scegliendo un piatto della bilancia, ma cambiando l'orizzonte temporale su cui la bilancia è letta. Sul trimestre, tagliare gli angoli paga; sul decennio, la fiducia è l'unico attivo che compone. Le crisi finanziarie lo hanno dimostrato al sistema; l'enciclica lo eleva a giudizio storico: l'intermediazione slegata da fondamenti morali ha già prodotto crisi sistemiche mondiali (cfr MH, §160).

Non scegliere un piatto della bilancia: cambiare l'orizzonte su cui la si legge

Sul trimestre tagliare gli angoli paga; sul decennio compone solo la fiducia

13.3 Tre obiezioni, tre risposte

«I controlli rallentano l'innovazione.» Rallentano le innovazioni mal concepite, che sono la maggioranza dei costi: il collaudo etico-tecnico costa meno del ritiro di un sistema dal mercato. «Se non lo facciamo noi, lo farà un concorrente.» È l'argomento per cui esistono le regole comuni: il quadro europeo, vissuto come alleato, trasforma la correttezza in condizione di mercato uguale per tutti. «I clienti guardano il prezzo, non l'etica.» Guardano il prezzo finché la fiducia è data per scontata; la guardano eccome il giorno in cui non lo è più; e quel giorno, per qualcuno, arriva sempre.

Il collaudo costa meno del ritiro

CAPITOLO 13

Solo un'AI affidabile rende la finanza sostenibile

13.4 L'agenda personale del CEO

Sette impegni che nessuno può assumere al posto di chi guida.

- I **Firmare la politica AI, non solo approvarla**
portare in Consiglio finalità, limiti e tolleranze al rischio; nominare un responsabile con accesso diretto al vertice.
- II **Chiedere le evidenze, non le assicurazioni**
una pagina al mese: che cosa hanno sbagliato i sistemi, chi se n'è accorto, che cosa è cambiato.
- III **Mettere l'inclusione nel cruscotto**
accanto a costi e ricavi: chi è entrato, chi resta fuori, dove i modelli escludono senza giustificazione.
- IV **Provare la crisi prima che arrivi**
due esercitazioni l'anno con il vertice al tavolo: attacco, errore di modello, fornitore che si ferma.
- V **Difendere la decisione umana**
nelle scelte che toccano la vita delle persone, l'ultima parola a chi può risponderne, per architettura e non per eccezione.
- VI **Trattare i collaboratori come i clienti**
nessuna automazione senza riqualificazione e partecipazione: la coerenza interna precede quella esterna.
- VII **Rendere conto pubblicamente**
un capitolo della relazione annuale sull'AI: sistemi in uso, incidenti, correzioni, verifiche di equità.

Sette impegni che non si delegano

Imprese che riconoscano nel lavoro e nella dignità un criterio di successo.

CFR MAGNIFICA HUMANITAS,
§181

CAPITOLO 13 · Solo un'AI affidabile rende la finanza sostenibile

13.5 Che cosa risponderemo

L'enciclica chiude la sua riflessione sull'intelligenza artificiale riprendendo la domanda che Giovanni Paolo II poneva a ogni tecnica: rende la vita umana più umana, più degna dell'uomo? (cfr MH, §129). Per chi guida una banca, la domanda ha una forma quotidiana e verificabile. Il cliente respinto dal modello ha potuto capire perché, e qualcuno ha potuto riconsiderare? La famiglia senza storia creditizia ha trovato una porta o un punteggio? Il collaboratore la cui mansione è cambiata è stato accompagnato o sorvegliato? Il risparmiatore che ha creduto a un prodotto sostenibile ha comprato una misura o un'etichetta?

Sono domande a cui i sistemi non rispondono: rispondono le persone che li hanno voluti, comprati, configurati e mantenuti. È questa, in ultima istanza, la responsabilità dei CEO che il rapporto ha inteso mettere a fuoco: non una responsabilità aggiuntiva, da esercitare dopo quelle gestionali, ma la stessa responsabilità gestionale guardata alla sua profondità reale, là dove le scelte di architettura diventano destini di persone.

13.6 Una parola sul tempo

Tutto, in questa materia, congiura a favore del rinvio: la tecnologia non è matura, le regole non sono stabili, i concorrenti non si muovono. Ma le architetture si stanno scrivendo ora, e ciò che entra nei sistemi in questi anni vi resterà a lungo. L'enciclica ricorda che ogni generazione riceve in eredità il compito di dare forma al proprio tempo (cfr MH, §1). Per la generazione che guida oggi le banche, la forma del tempo passa dai sistemi che sta scegliendo. «Fare bene, e fare il bene»: le due metà di un'unica responsabilità, e di un'unica firma.

Ai sistemi non si chiede conto:
si chiede a chi li ha voluti

*Rende la vita più umana,
più degna dell'uomo? La
domanda vale per ogni
tecnica.*

CFR MAGNIFICA HUMANITAS,
§129

Le architetture si scrivono ora, e
resteranno

RACCOMANDAZIONI

Dieci raccomandazioni al sistema finanziario

I Una politica AI deliberata dal Consiglio

che dichiari finalità, limiti e tolleranze al rischio, e nomini un responsabile con accesso diretto al vertice.

II La funzione di AI-vigilanza

con registro dei sistemi, indicatori di deriva e obbligo di segnalazione degli eventi avversi, riferita periodicamente al Consiglio.

Dieci impegni verificabili

III Validazione indipendente prima del rilascio

collaudo etico-tecnico di ogni sistema che tocca i clienti: nessun modello in produzione senza verifica di seconda linea.

IV Verifiche di equità documentate e ripetute

su ogni modello rivolto al cliente, con correzione o ritiro dove la disparità non è giustificata dal merito.

La giustizia si pratica nel disegno stesso dei sistemi, non dopo la loro adozione.

V La decisione umana per architettura

nelle scelte che toccano la vita delle persone, revisione e ultima parola a chi può risponderne.

VI Trasparenza esigibile verso il cliente

spiegazioni comprensibili delle decisioni automatizzate e strumenti effettivi di ricorso.

CFR MAGNIFICA HUMANITAS,
§109

VII La sicurezza come esercizio di vertice

scenari avversi due volte l'anno in Consiglio; piani provati; presidio delle terze parti critiche e delle vie d'uscita.

VIII L'inclusione nel cruscotto direzionale

indicatori di accesso al credito e ai servizi per chi era escluso, con la stessa dignità degli indicatori economici.

IX La dignità del lavoro come criterio

nessuna automazione senza tutela, riqualificazione e partecipazione dei collaboratori coinvolti.

X Rendicontazione pubblica annuale

un capitolo della relazione su sistemi in uso, incidenti, correzioni e verifiche: la fiducia si nutre di conto reso.

**La fiducia si alimenta di
trasparenza**

«Fare bene, e fare il bene»: le due metà di un'unica responsabilità.

CONCLUSIONI

Una finanza che serve

Una convinzione e una consegna

Rimettere l'essere umano al centro delle nostre scelte.

CFR MAGNIFICA HUMANITAS,
§16

La ricerca consegna al sistema finanziario una convinzione e una consegna. La convinzione: l'intelligenza artificiale può rendere la finanza più capace di servire: più precisa nel misurare, più larga nell'includere, più lungimirante nell'indirizzare il capitale. La consegna: nulla di tutto questo accadrà da sé. Accadrà se chi decide governerà la tecnologia con la stessa serietà con cui governa il capitale e il rischio, e se accetterà di misurarla sulla domanda che l'enciclica ripropone al nostro tempo: rende la vita più degna della persona? (cfr MH, §129).

È la prospettiva che *Magnifica Humanitas* affida a chi opera nell'economia: scegliere, dentro il cantiere del proprio tempo, tra la torre che ottimizza le persone e la città che le fa fiorire (cfr MH, §9-10, §16). A questa misura, l'unica che non si lascia automatizzare, il rapporto consegna le sue pagine.

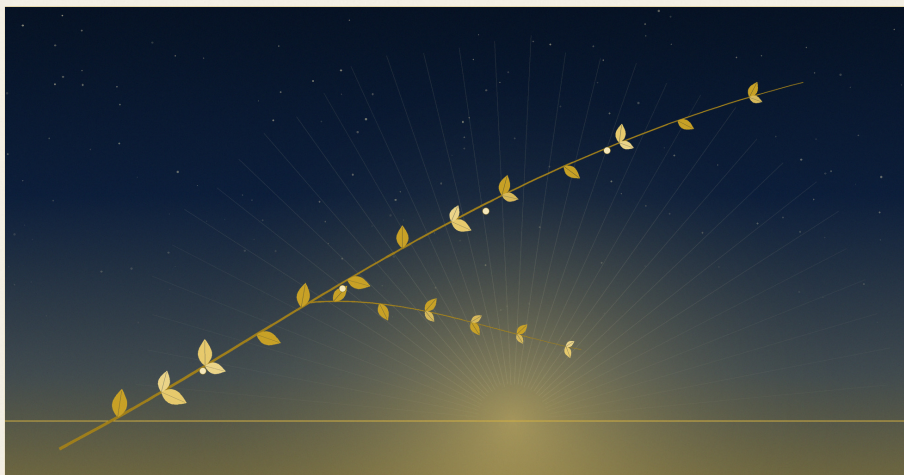


FIGURA 5 | L'orizzonte del bene comune

RINGRAZIAMENTI

Si ringrazia la società di consulenza Deloitte per l'ausilio fornito nella preparazione del rapporto e, in particolare: Alessandra Ceriani (Engineering, AI & Data), Elio Santoro (Climate & Sustainability), Tommaso Stranieri (Data & AI Ethics), Andrea Leuzzi (Engineering, AI & Data), Andrea Clerici (Climate & Sustainability).

APPENDICE

La mappa dell'enciclica per chi decide

La tavola raccoglie, per tema, i passaggi di *Magnifica Humanitas* richiamati nel rapporto: una guida alla lettura diretta del testo, che nessuna sintesi sostituisce.

TAVOLA 3 | La mappa dell'enciclica per chi decide

TEMA	CONTENUTO RICHIAMATO	PARAGRAFI
Decidere eticamente	criteri di discernimento e prassi; verifica e contestabilità delle decisioni; prudenza e ritmo dell'adozione	§14 · §104-106
Responsabilità di chi fa impresa	vocazione dell'iniziativa economica; dignità del lavoro come indicatore; automazione con tutele e partecipazione	§156-157 · §150 · §181
Responsabilità della finanza	funzione sociale del credito; lezione delle crisi; criteri di trasparenza, inclusione, equità; oltre il PIL	§160-164 · §159
L'uomo al centro	dignità non condizionata alla prestazione; decisione umana; la persona non è il suo profilo; possibilità di cambiare	§51-53 · §102-103 · §128 · §164
Disarmare l'AI	sottrarre l'AI alla corsa di potenza; renderla discutibile, contestabile, abitabile; chi decide l'etica dei sistemi	§110 · §107
Un'AI affidabile	responsabilità identificabile e rimedi; vigilanza, regole e politica che non abdica; appello agli sviluppatori	§105-106 · §111 · §209
Equità e inclusione	algoritmi e discriminazioni; diritti delle donne; accesso ai benefici dell'innovazione; destinazione universale dei beni digitali	§80 · §57 · §161 · §67
Potere digitale e sistema	concentrazione di dati e infrastrutture; sussidiarietà digitale; dati come bene comune	§5 · §95 · §71 · §108
Lavoro e trasformazione	ritmo delle macchine e dequalificazione; criteri sociali per l'innovazione; famiglia e giovani	§150-156 · §165-169
Limite e fiducia	il limite come luogo di maturazione; cura e volto dell'altro; corresponsabilità di tutti	§118 · §114 · §13 · §212

Una guida alla lettura diretta del testo

Nessuna sintesi sostituisce l'originale

APPENDICE

Note metodologiche e fonti

Il rapporto sintetizza gli esiti di analisi delle fonti internazionali sull'adozione dell'AI nella finanza, analisi dei documenti ufficiali dei vertici di istituzioni finanziarie, analisi delle pubblicazioni degli studiosi di etica della tecnologia e con la prospettiva delle autorità di vigilanza.

MAGISTERO PONTIFICIO E SANTA SEDE

- 1 Leone XIV, *Lettera enciclica Magnifica Humanitas. Sulla custodia della persona umana nel tempo dell'intelligenza artificiale*, Città del Vaticano, 15 maggio 2026.
vatican.va/content/leo-xiv/it/encyclicals
- 2 Francesco, *Discorso ai partecipanti ai «Dialoghi per una Finanza Integralmente Sostenibile»*, Città del Vaticano, 3 giugno 2024.
vatican.va/content/francesco/it/speeches/2024/june

FONDAZIONE CENTESIMUS ANNUS PRO PONTIFICE

- 3 Fondazione Centesimus Annus Pro Pontifice – SACRU, *Artificial Intelligence and Care of Our Common Home: A Focus on Industries, Finance, Education and Communication*, Vita e Pensiero, 2025 (introd. card. J. Tolentino de Mendonça).
centesimusannus.org
- 4 Fondazione Centesimus Annus Pro Pontifice, *Dialoghi per una Finanza Integralmente Sostenibile*, atti e materiali del percorso di ricerca.
centesimusannus.org
- 5 P. Garonna, M. Freda (a cura di), *La questione del debito al tempo del Giubileo*, Fondazione Centesimus Annus Pro Pontifice, 2025.
- 6 A. M. Tarantola, *L'economia della pace: il ruolo delle banche*, in «Dizionario di dottrina sociale della Chiesa», fasc. 2/2025.
dizionariodottrinasociale.it
- 7 A. M. Tarantola, *Le nuove generazioni nel mondo dell'intelligenza artificiale*, ivi, fasc. 1/2025.

AUTORITÀ E ISTITUTI DI VIGILANZA

- 8 Banca d'Italia, *L'impatto economico dell'adozione dell'intelligenza artificiale: evidenza dalle imprese italiane*, Questioni di Economia e Finanza n. 1005, marzo 2026.
bancaditalia.it
- 9 OCSE, in collaborazione con la Banca d'Italia, *L'intelligenza artificiale nei mercati finanziari italiani*, 2026.
bancaditalia.it
- 10 Banca d'Italia, *Adoption and Expected Impact of Generative AI: Evidence from Italian Households*, Questioni di Economia e Finanza n. 929, 2025.
bancaditalia.it
- 11 Banca d'Italia, *Intelligenza artificiale e rapporti bancari basati sulle relazioni banca-impresa*, Temi di Discussione n. 1476, 2025.
bancaditalia.it
- 12 Banca d'Italia, *Embracing the Digital Transition: the Adoption of Cloud Computing and AI by Italian Firms*, Questioni di Economia e Finanza n. 946, 2025.
bancaditalia.it
- 13 *Indagine FinTech Biennale di Banca d'Italia* (Edizione 2025).
bancaditalia.it

NOTE METODOLOGICHE E FONTI · fonti citate

ISTITUZIONI EUROPEE E GRUPPI DI ESPERTI

14 Commissione europea — Gruppo di esperti ad alto livello sull'IA (AI HLEG), *Ethics Guidelines for Trustworthy AI*, 8 aprile 2019.
digital-strategy.ec.europa.eu

15 Parlamento europeo e Consiglio, *Regolamento (UE) 2024/1689 — Artificial Intelligence Act*, 13 giugno 2024.
eur-lex.europa.eu/eli/reg/2024/1689

16 Parlamento europeo e Consiglio, *Regolamento (UE) 2016/679 — GDPR*, 27 aprile 2016.
eur-lex.europa.eu/eli/reg/2016/679

ORGANIZZAZIONI INTERNAZIONALI

17 OCSE, *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449), 22 maggio 2019, agg. 3 maggio 2024.
legalinstruments.oecd.org

RICERCA ACCADEMICA

18 L. Floridi, J. Cowsls, *A Unified Framework of Five Principles for AI in Society*, in «Harvard Data Science Review», 1(1), 2019.
doi.org/10.1162/99608f92.8cd550d1

19 L. Floridi et al., *AI4People — An Ethical Framework for a Good AI Society*, in «Minds and Machines», 28(4), 2018.
doi.org/10.1007/s11023-018-9482-5

ANALISI E STUDI DI SETTORE

20 McKinsey Global Institute, *The Economic Potential of Generative AI: The Next Productivity Frontier*, 14 giugno 2023.
mckinsey.com

21 McKinsey & Company, *Capturing the Full Value of Generative AI in Banking*, dicembre 2023.
mckinsey.com

22 EY, *European Financial Services AI Survey* (2ª ed.), ottobre–dicembre 2024.
ey.com

23 Deloitte, *EMEA Model Risk Management Survey 2025*, 2025.
deloitte.com

24 Deloitte Global, *2025 C-suite Sustainability Survey*, 2025.
deloitte.com

25 World Economic Forum, in collaborazione con Deloitte, *The New Physics of Financial Services: How Artificial Intelligence is Transforming the Financial Ecosystem*, 2018.
weforum.org/reports

CRITERIO SUI DATI

Ogni dato quantitativo riportato è attribuito alla fonte primaria nel punto in cui compare; non sono utilizzati dati privi di fonte verificabile. I valori percentuali si riferiscono agli universi delle rispettive rilevazioni, ossia società finanziarie del Regno Unito (Bank of England–FCA), istituzioni finanziarie europee (EY) e campione multisettoriale (Deloitte), e non sono perfettamente omogenei tra loro: vanno letti come ordini di grandezza convergenti.

AVVERTENZA SULLE CITAZIONI DELL'ENCICLICA

I contenuti della Lettera enciclica *Magnifica Humanitas* sono resi in forma indiretta, con rimando puntuale ai paragrafi del testo ufficiale (vatican.va). L'unica citazione testuale è riportata a pagina 10.

PER APPROFONDIRE

Bibliografia ragionata

Una selezione di Documenti ufficiali, Libri, Articoli e Rapporti di rilievo e di agevole reperibilità, scelti come ausilio per chi desidera approfondire i temi del rapporto.

DOTTRINA SOCIALE E CURA DELLA CASA COMUNE

- Pontificio Consiglio della Giustizia e della Pace, *Compendio della Dottrina Sociale della Chiesa*, Libreria Editrice Vaticana, 2004.

Il quadro organico dei principi su cui poggia l'intero rapporto.
vatican.va

- Francesco, *Laudato si'* (2015) e *Fratelli tutti* (2020); Benedetto XVI, *Caritas in veritate* (2009).

Le encicliche sociali richiamate da Magnifica Humanitas: casa comune, fraternità, economia.
vatican.va

ETICA E GOVERNO DELL'INTELLIGENZA ARTIFICIALE

- Luciano Floridi, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, Raffaello Cortina Editore, 2022.

I pilastri per un'IA antropocentrica: testo di riferimento, anche universitario.

- Luciano Floridi, *La quarta rivoluzione. Come l'infosfera sta trasformando il mondo*, Raffaello Cortina Editore, 2017.

«Infosfera» e vita «onlife»: l'identità umana nell'era dei dati.

- Luciano Floridi, *La differenza fondamentale. Artificial Agency: una nuova filosofia dell'intelligenza artificiale*, Mondadori, 2025.

Perché l'AI è una forma nuova di agire, distinta dall'agente umano.

- Luciano Floridi, *Il nodo etico. Informazione e valori nella società digitale*, Raffaello Cortina Editore, 2026.

Responsabilità e valori nell'infosfera: chi risponde delle decisioni automatizzate.

- Paolo Benanti, *Human in the loop. Decisioni umane e intelligenze artificiali*, Mondadori Università, 2022.

L'algoritica e il valore irrinunciabile della decisione umana.

- Paolo Benanti, *L'uomo è un algoritmo? Il senso dell'umano e l'intelligenza artificiale*, Castelveccchi, 2025.

Il confine tra computazione e coscienza, contro ogni riduzionismo.

- Paolo Benanti, *La nuova logica del dominio. Potere computazionale, democrazia e condizione umana*, Laterza, 2026.

Il potere computazionale che plasma opinioni e identità: la sfida alla democrazia.

- Paolo Benanti, *Le macchine sapienti. Intelligenze artificiali e decisioni umane*, Marietti 1820, 2018.

Uno dei primi testi italiani sull'algoritica e sulle decisioni delegate alle macchine.

- Massimo Chiriatti, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, LUISS University Press, 2021.

La differenza tra calcolo statistico e comprensione consapevole.

- Massimo Chiriatti, *#Humanless. L'algoritmo egoista*, Hoepli, 2019.

Come l'economia dei dati cambia lavoro, management e competenze.

- Rita Cucchiara, *L'intelligenza non è artificiale. La rivoluzione tecnologica che sta già cambiando il mondo*, Mondadori, 2021.

Dalla voce di una grande scienziata italiana: rigore tecnico e responsabilità etica.

- Stuart Russell, *Human Compatible: AI and the Problem of Control*, Viking, 2019.

Perché il controllo dei sistemi più capaci è il problema centrale.

- Mustafa Suleyman, con Michael Bhaskar, *The Coming Wave*, The Bodley Head, 2023.

Governare una tecnologia che corre più veloce delle sue regole.

PER APPROFONDIRE · bibliografia ragionata

DATI, POTERE E DIRITTI NELL'ERA DIGITALE

- › Shoshana Zuboff, *The Age of Surveillance Capitalism*, PublicAffairs, 2019.
L'economia fondata sulla raccolta dei dati personali.
- › Cathy O'Neil, *Weapons of Math Destruction*, Crown, 2016.
Quando gli algoritmi, con apparente neutralità, amplificano la discriminazione.
- › Kate Crawford, *Atlas of AI*, Yale University Press, 2021.
I costi umani e ambientali dell'AI: una lettura per la «casa comune».
- › Madhumita Murgia, *Code Dependent: Living in the Shadow of AI*, Picador, 2024.
Il volto umano dell'AI: chi ne subisce le decisioni, spesso i più fragili.

ECONOMIA, LAVORO E VALORE

- › Michael J. Sandel, *The Tyranny of Merit: What's Become of the Common Good?*, Farrar, Straus and Giroux, 2020.
I limiti morali di una società che misura le persone sulla sola prestazione.
- › Mariana Mazzucato, *The Value of Everything: Making and Taking in the Global Economy*, Allen Lane, 2018.
La differenza tra creare ed estrarre valore, nell'economia e nella finanza.

SOSTENIBILITÀ E FINANZA RESPONSABILE

- › Rebecca Henderson, *Reimagining Capitalism in a World on Fire*, PublicAffairs, 2020.
Ripensare l'impresa perché serva la società senza rinunciare al profitto.
- › Alex Edmans, *Grow the Pie: How Great Companies Deliver Both Purpose and Profit*, Cambridge University Press, 2020.
L'evidenza che scopo e profitto non sono in conflitto: servire conviene.
- › Francesco Morace, Marzia Tomasin, *L'alfabeto della sostenibilità. 26 modi di essere sostenibili*, Egea, 2023.
La transizione ecologica e sociale in 26 storie d'impresa.
- › Riccardo Milano, *La grande transizione. Finanza etica e futuro sostenibile*.
La nascita della finanza etica in Italia, dai modelli speculativi ai criteri ESG.
- › AIAF, *La finanza responsabile ed etica in Italia*.
Una ricognizione della finanza responsabile e dell'ESG nel contesto italiano.



Fondazione
Centesimus Annus
Pro Pontifice

Cortile San Damaso, 00120 Città del Vaticano
www.centesimusannus.va · centannus@foundation.va