



IVASS
ISTITUTO PER LA VIGILANZA
SULLE ASSICURAZIONI



Quaderno n. 16

La governance dell'Artificial Intelligence nel
settore assicurativo tra principi etici,
responsabilità del board e cultura aziendale

Diana Capone



Febbraio 2021



(decreto legge 6 luglio 2012 n. 95 convertito con legge 7 agosto 2012 n. 135)

La serie Quaderni intende promuovere la diffusione di studi e contributi originali sui temi assicurativi al fine di suscitare commenti critici e suggerimenti.

Le opinioni espresse nei lavori sono attribuibili ai soli autori e non impegnano in alcun modo la responsabilità delle Istituzioni di appartenenza.

via del Quirinale 21 - 00187
ROMA telefono +39 06 42133.1

Tutti i diritti riservati.

È consentita la riproduzione a fini didattici e non commerciali,
a condizione che venga citata la fonte

La serie è disponibile online nel sito www.ivass.it

ISSN 2421-4671 (online)

La governance dell'Artificial Intelligence nel settore assicurativo tra principi etici, responsabilità del board e cultura aziendale

Diana Capone

Abstract

L'impiego dell'Artificial Intelligence (AI) nel settore finanziario e, in particolare, in quello assicurativo rappresenta un'importante opportunità per accrescere la capacità degli operatori di: soddisfare bisogni innovativi del mercato, produrre reddito e, contestualmente, favorire il benessere della società in generale. L'utilizzo dei nuovi strumenti di AI, se inconsapevole dei rischi legali, tecnici ed etici che ad essi si accompagnano, può, d'altro canto, compromettere la fiducia del mercato, accrescere discriminazioni e diseguaglianze, ostacolare l'inclusività nell'offerta. Il rispetto di principi etici, ben definiti ma nello stesso tempo flessibili e adattabili ai diversi contesti, è la chiave di volta per garantire uno sviluppo responsabile dell'AI. La valenza teoretica di tali principi va però tradotta in vita pratica, facendo leva sulla capacità delle imprese di auto-organizzarsi, adottando efficaci e consapevoli modelli di governo dell'AI. Con il presente lavoro, si propone l'utilizzo di un paradigma di governo, gestione e controllo dei rischi - ispirato a quello elaborato per il Misconduct Risk dal Financial Stability Board - che renda concreta l'applicazione di tali principi etici attraverso i tre fondamentali pilastri organizzativi dell'impresa: la governance societaria (comprensiva anche del sistema dei controlli), i processi interni e le politiche del personale, la cultura aziendale. L'adozione di tale modello di governo ha le potenzialità per assicurare un uso responsabile dell'AI, soprattutto se accompagnato dalla disclosure al mercato dei presidi adottati dall'impresa. Si suggerisce, infine, come ulteriore ambito di riflessione, la possibilità di includere la sostenibilità digitale – anche in termini di governo responsabile dell'AI - nell'ambito dei principi Environmental, Social, Governance (ESG), come strumento per rendere efficacemente azionabile il controllo del mercato.

*Our technology, our machines, is part of our humanity.
We created them to extend ourselves, and that is what is unique about human beings.*

Ray Kurzweil

Premessa¹

Il dibattito sull'utilizzo e sui rischi posti dall'intelligenza artificiale nel sistema finanziario è vivo e partecipato. I contributi sono molteplici e alcuni principi generali e regole di comportamento sono già stati elaborati da standard setter e regolatori internazionali². Anche organismi del settore finanziario, sia internazionali³ sia nazionali, si stanno muovendo per dettare proprie linee guida in materia; alcune autorità hanno già dato indicazioni specifiche a cui dovranno attenersi le imprese vigilate che utilizzano sistemi di intelligenza artificiale⁴.

La ragione di questo diffuso interesse da parte dei supervisori per l'intelligenza artificiale discende non solo dalla necessità di vigilare sugli equilibri prudenziali e sulla compliance dei modelli di business ad elevato contenuto tecnologico adottati dagli intermediari regolamentati, ma anche dall'impatto che l'uso di questi strumenti, per lo più ancora poco conosciuti, può avere sui livelli di fiducia del pubblico nel mercato finanziario, più in generale.

La fiducia del mercato verso le istituzioni finanziarie condiziona la capacità di svolgere la funzione economico-sociale che è loro affidata, rappresenta l'elemento fondante della loro resilienza a fronte di eventi negativi e incide sulla loro potenzialità di generare reddito per gli stakeholders. La fiducia, soprattutto se riferita a prodotti e servizi innovativi, è di lenta costruzione ed è complessa da mantenere. In termini di perdita di fiducia i comportamenti non coerenti con le regole di settore, le violazioni di regole di condotta non scritte, l'operatività non in linea con i principi etici o con standard professionali possono essere molto più rilevanti delle stesse sanzioni irrogate dalle autorità di vigilanza di settore.

In un contesto in continua evoluzione quale quello attuale, in cui le forme di interazione con il potenziale cliente si evolvono a ritmi diversi e, a volte, incompatibili con i tempi del legislatore, la cura degli interessi del potenziale utente dei servizi finanziari va perseguita anche al di là di ciò che è formalmente prescritto dalle norme, per loro natura più lente nell'adeguarsi alle nuove realtà.

Con l'avvento dell'intelligenza artificiale le istituzioni finanziarie possono offrire servizi ulteriori e possono soddisfare al meglio, a costi ridotti o invariati, le richieste del pubblico. Un intervento normativo affrettato potrebbe, da un lato, limitare lo sviluppo del settore e, dall'altro, non essere efficace nel cogliere effettive esigenze di tutela, data la ridotta esperienza applicativa sinora maturata.

Dal lato del consumatore, infatti, gli algoritmi intelligenti e l'utilizzo sofisticato dei big data consentono di intercettare esigenze finanziarie o assicurative nuove e di modellare i prodotti e servizi sulle necessità del cliente. Questi strumenti possono essere un formidabile volano di benessere e sviluppo, ma nello stesso tempo potrebbero confermare o rafforzare forme di discriminazione nei confronti delle fasce più deboli o prestarsi a utilizzi che massimizzino i benefici esclusivamente di chi li impiega

¹ Le opinioni espresse non impegnano in alcun modo l'istituzione di appartenenza. Ringrazio Stefano De Polis, Piero Franchini, Valentina Langasio e Alfonso Rosolia per i commenti. Ogni errore od omissione è imputabile esclusivamente a me.

² Commissione Europea, World Economic Forum, Organizzazione per la Cooperazione e lo Sviluppo Economico.

³ Bank for International Settlements; European Banking Authority; European Insurance and Occupational Pensions Authority.

⁴ Cfr per una infografica – aggiornata al 2019 - sui diversi interventi il Report di Hermes dell'aprile 2019, *Investors' expectations on responsible artificial intelligence and data governance*.

o ancora, come tutte le tecnologie di cui ancora non si conoscono appieno le potenzialità, possono produrre conseguenze non volute, perché non pienamente prevedibili.

Dal lato delle imprese, i principali vantaggi si ritrovano in una diversa gestione dei rischi, nella riduzione di costi (di gestione e di personale) e di rischi operativi (errori e frodi) nonché nella possibilità di attrarre e fidelizzare nuovi clienti, offrendo una customer experience ottimizzata e su misura. Gli svantaggi sono quelli connessi alla complessità tecnica dei nuovi prodotti e servizi, che possono porre rischi di natura diversa rispetto a quelli tradizionali (si pensi al cyber risk) oppure possono dare origine a forme differenti di emersione di rischi già conosciuti, tipici delle attività finanziarie (reputazionali, legali, di condotta, operativi, di compliance).

Queste potenziali criticità richiedono forme di mitigazione e gestione, a tutela sia dell'integrità e stabilità del sistema finanziario sia degli interessi del pubblico, che siano strutturate in modo da accompagnare lo sviluppo delle nuove tecnologie in un contesto di tutela degli interessi della clientela e, più in generale, della fiducia del pubblico nel settore finanziario.

Scopo del presente lavoro è evidenziare il ruolo che possono avere i modelli di governo e organizzativi delle imprese nel contribuire efficacemente alla tutela del consumatore dei servizi assicurativi e nel mitigare i rischi per le imprese medesime legati all'utilizzo dell'intelligenza artificiale.

La struttura del lavoro è articolata come segue. Nel primo paragrafo, si analizzano brevemente i principali limiti legati all'utilizzo dell'intelligenza artificiale, così come evidenziati dalla letteratura; nel secondo paragrafo, si riassumono in principali impieghi dell'intelligenza artificiale nel settore assicurativo; nel terzo paragrafo si tenta una ricostruzione ragionata delle linee guida adottate dalle autorità intervenute sulla tematica dell'intelligenza artificiale; nel quarto paragrafo, si propone l'utilizzo di un paradigma di gestione dei rischi posti dall'intelligenza artificiale fondato su presidi di governo societario, procedurali ed etici apprestati dalle imprese; nel quinto paragrafo si traggono le conclusioni e si indicano ulteriori ambiti di riflessione.

1. Intelligenza artificiale: definizioni, potenzialità e limiti applicativi

Esistono molteplici definizioni del termine intelligenza artificiale. La prima e più ampia è quella elaborata dall'High-Level Expert Group on Artificial Intelligence della Commissione Europea⁵ secondo la quale: *“Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal.”*

Questa definizione presenta un livello di astrattezza e flessibilità tale da includere le attuali diverse tipologie di AI e i possibili futuri sviluppi e, nello stesso tempo, è sufficientemente precisa per offrire la necessaria certezza nell'identificazione del fenomeno, in prospettiva anche a fini regolamentari, individuandone gli elementi di base: algoritmi e dati⁶.

⁵ Cfr. documento: *A definition of AI: Main capabilities and scientific disciplines* - High-Level Expert Group on Artificial Intelligence, Aprile 2019.

⁶ Cfr. Commissione Europea, White Paper 2020, in cui viene anche precisato che: *“In any new legal instrument, the definition of AI will need to be sufficiently flexible to accommodate technical progress while being precise enough to provide the necessary legal certainty. For ... any possible future discussions on policy initiatives, it seems important to clarify the main elements that compose AI, which are ‘data’ and ‘algorithms’. AI can be integrated in hardware. In case of machine learning techniques, which constitute a subset of AI, algorithms are trained to infer certain patterns based on a set of data in order to determine the actions needed to achieve a given goal. Algorithms may continue to learn when in use.”*

Nella sostanza, dunque, i sistemi di intelligenza artificiale consistono in programmi e applicazioni informatiche che utilizzano le informazioni raccolte in modo da: i) comunicare, verbalmente o per iscritto; ii) interpretare immagini o leggere documenti; iii) assumere decisioni con modalità sovrapponibili a quelle di un operatore umano. Queste capacità si basano, prevalentemente, su algoritmi che consentono di elaborare, in tempi molto ridotti, grandi quantità di dati e che possono evolversi sulla base dei feed back – risposte della clientela, dati aggiornati, etc. - ricevuti nel corso della propria operatività (machine learning)⁷. A tal fine i sistemi di intelligenza artificiale possono utilizzare *“symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.”*⁸

Sistemi informatici dotati di simili potenzialità possono essere applicati in tutti i settori, compreso quello finanziario, e consentono di rimodulare i modelli di business tradizionali rendendoli più efficienti - con risparmi di costo per gli utenti e aumenti dei margini di redditività per le imprese - e più rispondenti alle richieste di servizi evoluti da parte della clientela.

A fronte dei vantaggi descritti, l'utilizzo di questi strumenti di intelligenza artificiale pone rischi peculiari legati ai limiti insiti nelle tecnologie che sono alla base del loro funzionamento.

Un importante limite può discendere dai dati utilizzati nelle elaborazioni: se non sono rappresentativi dei diversi segmenti della popolazione o la loro qualità non è adeguata ne discende un livello scarso degli output prodotti dagli algoritmi. Considerati il numero elevatissimo di dati e le varie modalità con cui gli stessi vengono acquisiti, la loro rappresentatività dipende da fattori non sempre controllabili a priori. In particolare, un elemento specifico che incide sulla rappresentatività dei dati è costituito dalla composizione del digital footprint⁹, che spesso costituisce la principale fonte di informazioni. Contribuiscono, infatti, a costruire tale impronta digitale solo coloro che sono presenti sul web e lasciano, volontariamente o involontariamente (ad esempio attraverso l'utilizzo di social media o di siti che raccolgono informazioni sugli utenti), le proprie tracce. Ne consegue che i risultati delle elaborazioni degli algoritmi tendono a riflettere, e talvolta anche amplificare, il digital divide¹⁰ tra chi ha la capacità e le possibilità di utilizzare internet e i canali virtuali in generale e chi invece non opera in questi ambiti (per età, collocazione geografica, preferenze individuali, disponibilità economiche).

Da un lato, quindi, i sistemi di intelligenza artificiale possono estendere le possibilità di accesso ai servizi anche a soggetti a basso reddito, consentendo sostanziali risparmi nei costi; dall'altro, l'uso dei big data potrebbe amplificare il gap economico, escludendo di fatto dall'accesso ai servizi più innovativi i soggetti scarsamente digitalizzati¹¹, che hanno avuto poche possibilità di incidere sulla definizione dei servizi.

Sempre con riferimento alle modalità di alimentazione degli algoritmi, rileva il fatto che i dati disponibili sono fortemente storicizzati: il digital footprint si riferisce, infatti, sempre al passato e valutazioni meccaniche di questi dati potrebbero non produrre risultati adeguati, quando questi si proiettino in un futuro in cui gli eventi non si ripetono in modo analogo al passato. Da questo punto di vista le valutazioni/decisioni degli strumenti di intelligenza artificiale presentano le medesime

⁷ Cfr. Commissione Europea, High Level Expert Group on AI, cit. *“As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)”*. Ancora secondo Jordan e Mitchell: *“Machine learning is a type of AI that provides computers with the ability to learn without being explicitly programmed”* in Jordan et al. 2015.

⁸ Commissione Europea, High Level Expert Group on AI, cit.

⁹ Il digital footprint viene qui inteso come tutte le informazioni che sono generate dall'attività svolta da ciascun individuo on line. Può includere sia ciò che viene direttamente creato dagli utenti di internet (email, social media, post), sia le informazioni che sono lasciate dalla sola interazione passiva con il web (es. dati di accesso IP).

¹⁰ Blumenstock, JE, et al 2012

¹¹ Soto, et al. 2011.

limitazioni di quelle decisioni umane fallaci perché basate esclusivamente sull'esperienza e su quanto avvenuto e che prescindono da considerazioni logiche approfondite e da valutazioni più ampie dell'evoluzione del contesto di riferimento.

Le valutazioni prodotte dei sistemi di intelligenza artificiale restano comunque molto più precise rispetto a quelle umane, perché si basano sulla capacità di analizzare una quantità infinitamente superiore di elementi e di coglierne i legami. Proprio questa capacità di individuare i legami tra diversi dati è, però, un ulteriore elemento di potenziale vulnerabilità dei sistemi di intelligenza artificiale. Infatti, il meccanismo di funzionamento degli algoritmi, che basa il proprio output su legami ricorrenti tra dati, potrebbe riflettere ed amplificare discriminazioni implicite nella nostra società (ad esempio nomi femminili che risultano legati a un tipo di professione diversa rispetto a quella che ricorre con nomi maschili: insegnante nel primo caso, amministratore esecutivo nel secondo)¹² o distorsioni, derivanti dai pregiudizi, consapevoli o inconsapevoli, di chi è incluso nella rilevazione delle informazioni e alimenta con le sue informazioni la base dati.

Per ovviare a questo limite, generalmente, si agisce su due fronti: sul disegno stesso dell'algoritmo o su come le decisioni vengono prese. L'algoritmo infatti si limita ad analizzare le informazioni fornite/acquisite e sulla base di quelle formula previsioni e prende decisioni; se viene impostata una regola nell'algoritmo che impedisce decisioni discriminatorie, si potrebbe eliminare l'effetto non voluto¹³. Questo risultato potrebbe anche essere raggiunto utilizzando strumenti incorporati nell'algoritmo medesimo che consentano di verificare (ed eventualmente correggere) se nella decisione finale hanno avuto peso determinante caratteristiche il cui rilievo è in contrasto con i valori etici della società (es. razza, religione, sesso).

Resta da considerare se l'accuratezza dell'output dell'algoritmo venga alterata da questi interventi, che avrebbero comunque l'indubitabile vantaggio di rendere l'algoritmo meno discriminatorio di una decisione umana. È molto difficile, infatti, che un'intelligenza umana sia effettivamente in grado di liberarsi di pregiudizi e preferenze individuali, mentre un'intelligenza artificiale applicherà indefettibilmente le regole che gli sono state assegnate.

In letteratura è stato mostrato che, anche con procedure formalizzate, gli incaricati umani tendono comunque ad esercitare un certo grado di discrezionalità, in cui possono sempre manifestarsi bias individuali¹⁴. L'AI, invece, prende le decisioni esclusivamente sulla base dei dati acquisiti e delle regole codificate, senza esercitare alcuna discrezionalità; se quindi i dati non contengono bias e le regole sono chiare, il risultato sarà più imparziale rispetto a quello delle decisioni umane¹⁵.

Anche la stessa elevata possibilità di personalizzazione consentita dagli algoritmi, può in alcuni casi tradursi in un potenziale limite. La differenziazione del prezzo, che è una delle manifestazioni più immediate della personalizzazione, può essere percepita soggettivamente come ingiusta oppure essere effettivamente ingiustificata, perché basata non sull'individuo ma sulle caratteristiche di altri individui assimilabili. Ad esempio l'algoritmo intelligente di alcuni siti di viaggio è in grado di proporre

¹² Caliskan, A. et al. mostrano come "applying machine learning to ordinary human language results in human-like semantic biases. We replicated a spectrum of known biases, as measured by the Implicit Association Test, using a widely used, purely statistical machine-learning model trained on a standard corpus of text from the World Wide Web. Our results indicate that text corpora contain recoverable and accurate imprints of our historic biases, whether morally neutral as toward insects or flowers, problematic as toward race or gender, or even simply veridical, reflecting the status quo distribution of gender with respect to careers or first names."

¹³ Rambachan et al, 2020

¹⁴ Cfr. Gates et al: "Formal procedures can limit opportunities to exercise prejudicial discretion or fall victim to implicit bias. Automated underwriting increased approval rates for minority and low-income applicants by 30% while improving the overall accuracy of default predictions"

¹⁵ Cfr. Barocas et al. : "Advocates of algorithmic techniques like data mining argue that these techniques eliminate human biases from the decision-making process. But an algorithm is only as good as the data it works with. Data is frequently imperfect in ways that allow these algorithms to inherit the prejudices of prior decision makers. In other cases, data may simply reflect the widespread biases that persist in society at large"

soluzioni di pernottamento commisurate al budget di chi fa la ricerca, sulla base di informazioni non soltanto acquisite dall'algoritmo nel corso di precedenti esperienze con il medesimo cliente, ma anche inferite dai dati che ha a disposizione (es. disponibilità di carte di credito, massimali di spesa delle carte medesime)¹⁶. Questa differenziazione, benché da un punto di vista economico sia pienamente logica e sensata¹⁷, può essere lesiva dell'equità nei rapporti contrattuali, soprattutto se si manifesta in modo da ledere gli interessi del potenziale cliente, come nell'esempio fatto in precedenza, offrendo a prezzi più alti la medesima soluzione alloggiativa a clienti ritenuti, sulla base di elementi inferiti dall'AI, propensi – per scelte soggettive o capacità economiche - a spendere di più.

Nell'applicazione dell'AI un'ulteriore potenziale criticità è la difficoltà di ricostruire i meccanismi di elaborazione dei dati alla base delle decisioni prese (spiegabilità) e, in particolare, dalla impossibilità di distinguere tra correlazione, che si manifesta quando due elementi si presentano costantemente insieme e si muovono in modo simmetrico, e legame causale. Di norma l'algoritmo non è in grado di distinguere tra i due fenomeni e soltanto di recente sono stati elaborati metodi di costruzione e analisi dei dataset che sembrano promettenti nel risolvere questa difficoltà¹⁸.

Questa caratteristica applicativa degli algoritmi, che utilizzano dati del passato e prevedono il futuro sulla base di ciò che è accaduto senza utilizzare alcuna narrativa di causalità tra gli eventi, ha portato a sostenere l'inutilità di teorie che spieghino causalmente fenomeni che siano comunque correttamente previsti dagli algoritmi ("the end of theory"¹⁹). In realtà, sistemi di AI non supportati da una teoria funzionano, come detto, sempre correttamente soltanto se il futuro si manifesta in modo simile al passato e i dati sono analizzati in modo da evitare la confusione tra correlazione e causalità.

Le limitazioni descritte, insite nei meccanismi di funzionamento dell'AI, possono anche combinarsi tra loro e diventare ancora più complesse da gestire e risolvere. È ormai di scuola il caso dell'algoritmo intelligente che, nel selezionare i candidati da invitare alle interviste per l'accesso ad una nota università americana, individuò tra i criteri di esclusione la provenienza della domanda da circoscrizioni in cui erano più alti i tassi di criminalità e più elevata la presenza di comunità latine ed afroamericane, perché dai dati del passato emergeva che candidati di successo non provenivano da quelle aree. In questo caso la spiegazione poteva risiedere nel fatto che i dati utilizzati dall'algoritmo incorporavano un bias (base dati discriminatoria) oppure era stato l'algoritmo stesso ad aver estrapolato dagli eventi del passato un criterio predittivo non significativo, o ancora, per l'effetto combinato delle due circostanze, si era prodotta un'evidente ed involontaria discriminazione²⁰. In questi casi non è agevole escludere l'elemento di discriminazione: anche se lo sviluppatore pone la regola secondo la quale determinate caratteristiche devono essere ignorate (provenienza da determinate regioni), l'algoritmo può individuare una combinazione di altri fattori correlati, che si accompagnano in modo ricorrente alla caratteristica proibita, e prenderli a base delle proprie decisioni che continuano così ad essere discriminatorie.

¹⁶ Cfr. Hannak et al, 2014.

¹⁷ In questo senso Barocas et al, 2016: *"data mining is always a form of statistical (and therefore seemingly rational) discrimination. Indeed, the very point of data mining is to provide a rational basis upon which to distinguish between individuals and to reliably confer to the individual the qualities possessed by those who seem statistically similar. Nevertheless, data mining holds the potential to unduly discount members of legally protected classes and to place them at systematic relative disadvantage"*.

¹⁸ Cfr. Dhir, et al. 2019

¹⁹ *"The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all."* Anderson, C (2008). The end of theory: the data deluge makes the scientific method obsolete. Wired Magazine 16(7).

²⁰ Cfr. Barocas et al *"data mining can discover surprisingly useful regularities that are really just preexisting patterns of exclusion and inequality. Unthinking reliance on data mining can deny historically disadvantaged and vulnerable groups full participation in society. Worse still, because the resulting discrimination is almost always an unintentional emergent property of the algorithm's use rather than a conscious choice by its programmers, it can be unusually hard to identify the source of the problem or to explain it to a court"*. Nello stesso senso Pasquale, F. 2015.

L'uso di strumenti di AI può rendere anche più difficoltosa la tutela giurisdizionale: è infatti molto complessa la prova dell'avvenuta discriminazione nelle sedi giudiziarie. A fini predittivi l'AI rileva la compresenza di migliaia di caratteristiche e, come detto, non utilizza legami univoci di causa/effetto: il tentativo di ricostruire la causalità nelle decisioni così prese rischia di tradursi molto spesso in una razionalizzazione ex post²¹.

2. Intelligenza artificiale e attività assicurativa: campi di applicazione, benefici e rischi

Con il termine intelligenza artificiale nel presente lavoro ci si intende riferire a sistemi di calcolo che sono in grado di portare a termine compiti che implicano un livello di comprensione, ragionamento e decisione simile a quello umano.

Secondo Agrawal et al, tutte le attività umane possono suddividersi in 5 elementi di base: acquisizione di dati, previsioni, giudizio, azione e risultati²². L'intelligenza artificiale, in particolare nelle forme del machine learning e dell'analisi predittiva, è in grado di gestire efficacemente i primi due elementi (l'acquisizione di dati e le previsioni), a condizione che si adottino cautele per gestire i limiti insiti nei suoi meccanismi di funzionamento.

L'utilizzo dell'AI nel settore assicurativo, e in quello finanziario più in generale, si concentra su questi due momenti, mentre più complesso è l'utilizzo dell'intelligenza artificiale nel terzo elemento, quello del giudizio. Il giudizio implica, infatti, la valutazione della situazione nella sua unicità, il tener conto delle reazioni della controparte umana e, infine, il ricorso al pensiero laterale se la soluzione da adottare è diversa da quella ordinaria (si pensi a forme di consulenza evolute su prodotti di investimento complessi). Nella sostanza l'utilità e la sicurezza dell'intelligenza informatica finisce dove inizia l'insostituibilità delle capacità della mente umana nel prendere decisioni²³.

Nella fase del giudizio, quindi, soprattutto se si deve uscire da schemi ripetitivi e automatici, è al momento difficile ipotizzare una sostituzione dell'AI al contributo umano. Nelle altre fasi dell'attività, invece, l'impiego dell'AI consente di accrescere la flessibilità nell'offerta, di ridurre i costi, in particolare dell'intermediazione, e di definire prodotti e servizi personalizzati, costruiti sulle esigenze della clientela.

Nell'industria assicurativa sono da tempo utilizzati algoritmi complessi che permettono di definire in modo accurato il premio che deve essere corrisposto dal cliente in relazione al rischio di cui è portatore. Questi strumenti, di fatto statistici, sono considerati da molti una prima forma di machine learning applicata alle assicurazioni.

Non sorprende, quindi, che proprio nel settore assicurativo si stia assistendo a una sempre maggiore diffusione dell'utilizzo dell'AI. Secondo l'EIOPA ²⁴ *"in 2018 already 31% of European insurance undertakings where using ML and another 24% are at a proof of concept stage, and in some jurisdictions (UK n.d.r.) the level of adoption is already 100%"*.

Strumenti e tecniche di intelligenza artificiale vengono applicate nelle fasi di: a) acquisizione del cliente attraverso chatbox; 2) consulenza; 3) personalizzazione dei prodotti offerti e della quantificazione del premio; 4) gestione dei sinistri, compreso il contrasto delle frodi.

²¹ Swedloff, 2019

²² A. Agrawal, J et al., 2016

²³ Un esempio di questa insostituibilità lo si trova per i veicoli a guida automatica. Se un pedone decide di attraversare all'improvviso, il veicolo potrebbe non essere in grado di decidere se frenare (dando priorità all'incolumità del pedone ma creando potenzialmente un danno per il guidatore), sterzare (esponendo a un rischio sia pedone che guidatore) o proseguire (con pericolo per il pedone).

²⁴ Big Data and Analytics in motor and health insurance: a thematic review. EIOPA, 2019

1. *Chatbox*

Sono assistenti virtuali che supportano i clienti rispondendo alle domande sul web oppure in chat, anche via applicazioni su smartphone. Questa forma di interazione con il cliente è, ovviamente, meno costosa di quella con una controparte umana ed è più veloce. Soprattutto per le conversazioni caratterizzate da un elevato livello di ripetitività può rivelarsi anche più efficace, liberando risorse per le attività a maggior contenuto decisionale e discrezionale. Ad esempio gli assistenti virtuali possono seguire il cliente nella lettura di tutti gli aspetti dei prospetti e guidarli negli adempimenti richiesti per assicurare la compliance nel collocamento del prodotto, senza correre il rischio di trascurare passaggi apparentemente meno rilevanti, come potrebbe accadere per un consulente umano alle prese con attività ripetitive e prive di evidente valore aggiunto. Inoltre, ricerche recenti mostrano che i consumatori potrebbero avere maggiore fiducia nell'interfaccia di AI rispetto a quella umana, poiché ritengono che gli assistenti virtuali siano più imparziali e non portatori di interessi egoistici²⁵. È noto che ciascun agente del mercato (umano) persegue il proprio esclusivo interesse, la mancanza di umanità delle chatbox risulta quindi rassicurante per il cliente che ne presume la neutralità. Tutto ciò a condizione che, ovviamente, il programmatore della chatbox non inserisca una distorsione nell'algoritmo di funzionamento, prevedendo ad esempio che le indicazioni dell'assistente virtuale debbano avere come fine ultimo l'indurre il cliente ad acquistare il prodotto. In questo caso la neutralità della chatbox viene meno e si pone anche un problema di trasparenza nei confronti del cliente sulle modalità di programmazione e di interazione con l'assistente virtuale.

2. *Consulenza*

Strumenti di AI stanno prendendo sempre più piede anche nell'ambito della consulenza nella scelta dei prodotti, in favore della potenziale clientela. Spesso l'interazione con il consulente umano può essere resa più complessa dalle dinamiche tipiche delle relazioni interpersonali: il discutere di denaro o di abitudini di spesa può essere difficoltoso dalla paura del giudizio altrui o, più semplicemente, da un desiderio di riservatezza. L'interazione con un consulente virtuale può, invece, essere percepita come rassicurante, perché priva delle implicazioni psicologiche che possono impedire al cliente di discutere in modo trasparente dei propri orizzonti di investimento o delle esigenze assicurative. Recenti ricerche²⁶ mostrano che sono soprattutto i millennials a volersi avvalere del supporto di consulenti virtuali. Questo strumento ha, pertanto, le potenzialità per coinvolgere, in futuro, le fasce di popolazione relativamente più giovane²⁷. La facilità di accesso al servizio, in termini di tempo e di luogo, e la possibilità di colloquiare con un consulente virtuale, non giudicante e imparziale, potrebbe facilitare la pianificazione del risparmio a fini previdenziali anche a chi non entrerebbe volentieri in contatto con un intermediario umano oppure, per limiti di tempo o di collocazione geografica, non è in grado di farlo. Il rischio dell'attività di consulenza svolta con strumenti di AI è che l'offerta dei prodotti possa variare in modo discriminatorio in relazione alla tipologia di soggetti che interagiscono con l'assistente virtuale o che le modalità di utilizzo dei dati sia lesivo della privacy. Altro potenziale rischio, soprattutto per i prodotti assicurativi, è che all'offerta di una polizza si accompagnino anche tentativi di indirizzare o modificare il comportamento del consumatore, in una forma di condizionamento occulto se questi non ne

²⁵ Scopino G., 2015

²⁶ Confronta Chatbot Report 2019: Global Trends and Analysis, <https://chatbotsmagazine.com/chatbot-report-2019-global-trends-and-analysis-a487afec05b>

²⁷ L'impiego dell'AI in quest'ambito potrebbe essere d'aiuto nell'affrontare il problema del gap previdenziale, che, come noto, a causa dell'aumentata longevità e dell'insufficienza del risparmio investito con finalità previdenziali, sta progressivamente aumentando. Secondo il World Economic Forum (2019): "For the eight countrieswhich ...are some of the most populated nations, the gap at 2015 was already at \$70 trillion. If measures are not taken to increase overall levels of savings, we project this gap to grow to \$400 trillion by 2050". http://www3.weforum.org/docs/WEF_Investing_in_our_Future_report_2019.pdf

è consapevole. Anche se non si tratta di esperimenti diffusi al momento, sono possibili schemi di distribuzione di polizze attraverso l'utilizzo di AI che configurano forme di condizionamento comportamentale, che, da un lato, limitano, in modo indiretto, la libertà di scelta dell'assicurando, e, dall'altro, possono svolgere una meritoria attività di prevenzione, migliorando la qualità della vita e la sicurezza degli assicurati. Ad esempio, vengono offerte polizze scontate se accompagnate dalla sottoscrizione di un abbonamento ad una palestra, oppure suggerite strade più sicure rispetto ad altre collegandovi sconti per le polizze auto, o indicati sistemi di allarme particolari per beneficiare di riduzioni di prezzo sulle assicurazioni sulla casa. Anche il meccanismo della scatola nera²⁸, molto diffuso in Italia, ha la potenzialità per essere utilizzato per modificare il comportamento della clientela a fronte di un sostanzioso sconto sulla polizza auto. In tutti questi casi il cliente potrebbe essere condizionato nella libera espressione delle proprie preferenze dall'offerta di ritorni economici per una serie di comportamenti potenzialmente molto ampia. Le modalità di intervento delle assicurazioni ne potrebbero essere radicalmente mutate: da reattive (al verificarsi dell'evento di rischio provvedo al ristoro dell'eventuale danno ove dovuto) divenire proattive (suggerisco i comportamenti più sicuri per evitare di dover pagare il sinistro quando si verifica).

3. *Personalizzazione del prodotto e/o del premio* – L'AI ha potenzialità enormi soprattutto nell'identificare nuovi collegamenti tra caratteristiche degli individui e pattern di rischio (anche apparentemente difficili da spiegare: ad esempio tra merito di credito e stile di guida)²⁹. In prospettiva sarà anche possibile ricavare previsioni sui rischi sanitari individuali dalle informazioni rilevate e condivise tramite personal device (es. che rilevano peso, pressione cardiaca, livello di movimento, apporti calorici etc.). Questa possibilità si potrà poi tradurre in una personalizzazione a livello di singolo individuo del premio per le polizze salute, sulla base della condivisione volontaria dei dati raccolti. Livelli di personalizzazione così elevati potrebbero mettere a rischio l'inclusività dei prodotti assicurativi, rendendo eccessivamente costoso per alcuni assicurarsi. Inoltre, con il tempo, la base di riferimento degli assicurati potrebbe divenire talmente granulare da rendere non praticabile il risk pooling, sul quale si basa, almeno in parte, la capacità di produrre reddito di un'impresa di assicurazione. Ancora l'individualizzazione del premio potrebbe rendere non assicurabili i gruppi più vulnerabili, o perché non è possibile monitorare nei loro confronti i comportamenti che possono portare ad un maggiore risparmio (ad esempio non hanno un digital footprint specifico), o perché sono più rischiosi o considerati tali perché rientranti in una categoria così classificata dagli algoritmi. Il punto nodale è trovare il giusto bilanciamento tra un'auspicabile precisa corrispondenza tra rischio e premio, con il conseguente beneficio per utenti meno rischiosi, e la necessità di non escludere dalla platea dei potenziali assicurabili altre categorie più deboli. Nella personalizzazione del premio si annida, poi, il maggior rischio di discriminazione, che, pur trasversale a tutte le ipotesi di utilizzo dell'AI, potrebbe avere in questo caso un impatto economico e sociale diretto poiché si tradurrebbe in una diversa quantificazione del prezzo da pagare per poter beneficiare dell'assicurazione.
4. *Gestione dei sinistri e prevenzione delle frodi* – L'AI può migliorare la gestione dei sinistri, offrendo strumenti che ne velocizzano la trattazione e che consentono di individuare tentativi di frode. Sotto il primo profilo, attraverso strumenti di comprensione ed elaborazione del linguaggio naturale (c.d. Natural Language Processing) è possibile estrarre, dalla mole dei documenti inviati in caso di sinistri (ricette mediche, referti di pronto soccorso, verbali dei vigili

²⁸ La scatola nera auto è un dispositivo elettronico mobile dotato di rilevatore GPS, e collegato alla batteria della vettura, capace di registrare una grande quantità di dati inerenti la condotta e lo stile di guida dell'automobilista, sinistri inclusi

²⁹ Cfr. Marwick, 2014.

urbani), le informazioni necessarie per determinare il se e il quanto del risarcimento dovuto³⁰. Per identificare e scongiurare i tentativi di frode, sono stati sviluppati strumenti di AI in grado di utilizzare i dati ricavabili dai principali social media per verificare la coerenza tra questi e le circostanze rappresentate dall'assicurato nella richiesta di risarcimento. In positivo questi strumenti offrono la possibilità di ridurre i tempi e i costi procedurali dei sinistri nonché di prevenire comportamenti fraudolenti, consentendo l'abbassamento del premio per gli assicurati. In negativo, la capacità dell'AI di elaborare quantità notevolissime di dati può portare a una sistematica sottovalutazione dei danni dei sinistri in favore degli assicurati. Il meccanismo potrebbe funzionare in modo iterativo: inizialmente l'AI potrebbe stabilire che alcune stime di danno sono sopravvalutate, applicando le istruzioni dell'impresa di assicurazione che richiede sia corrisposto un ristoro adeguato e ragionevole ma il più basso possibile, e fare un'offerta più bassa. Se il feedback ricevuto all'offerta di ristoro è negativo (l'assicurato reclama, fa causa all'impresa o cambia assicurazione), l'AI offrirà una cifra più alta al successivo assicurato che presenterà un sinistro con caratteristiche simili. Se, al contrario, il feedback è positivo - perché l'assicurato non è in grado di reclamare, non è consapevole dei suoi diritti o perché la differenza tra la cifra offerta e quella attesa è minima - la sottostima del danno può divenire strutturale perché l'algoritmo tenderà a replicare sempre le stesse offerte. In più, attraverso questo processo iterativo, l'AI potrebbe individuare il punto esatto di sottostima del sinistro, che non porta l'assicurato a reclamare ma, nello stesso tempo, non ristora pienamente il danno, con uno squilibrio contrattuale, ripetuto e sistemico, in danno degli assicurati e in favore dell'assicurazione che estrae il massimo beneficio possibile dall'utilizzo dell'AI³¹.

Questa veloce ricognizione dei principali ambiti in cui il settore assicurativo può sperimentare l'utilizzo dell'AI ne evidenzia i molti potenziali vantaggi, ma allo stesso tempo permette di metterne in luce i rischi, anch'essi potenzialmente rilevanti.

I vantaggi possono essere sintetizzati come: 1) riduzione del premio per gli assicurati in conseguenza della maggiore efficienza in tutta la filiera assicurativa di distribuzione e produzione; 2) più accurata quantificazione dei premi per classi di rischio degli assicurati, anche per effetto della riduzione delle frodi; 3) estensione dell'assicurabilità a nuove categorie di persone, i cui dati divengono disponibili grazie alle nuove fonti; 4) proattività delle coperture assicurative che divengono anche un ausilio nella prevenzione dei rischi o nella gestione delle malattie; 5) semplificazione dell'accesso alle polizze, con potenziale riduzione dei gap assicurativi soprattutto nei settori di recente emersione.

Tra le complessità da gestire emerge in primo luogo il tema di coniugare la capacità degli algoritmi di personalizzare i premi con il principio di mutualità che fonda la gestione assicurativa dei rischi. Assume poi rilevanza la condivisione, più o meno volontaria dei dati personali degli assicurati, attuali o potenziali. Da sempre le imprese di assicurazioni basano le proprie interazioni con la clientela sulle informazioni di cui sono in possesso. Prima della diffusione dei big data, la gran parte delle informazioni venivano volontariamente e consapevolmente fornite dal cliente, oggi, con l'ampliamento generalizzato dell'impronta digitale di ciascun individuo, a dati forniti su richiesta specifica (ad esempio quante volte si va in palestra) si affiancano dati che vengono direttamente osservati (es. numero di accessi in palestra ricavabile dalle app di geolocalizzazione) o che vengono inferiti da altri dati disponibili (es. acquisto di materiale sportivo con la carta di credito) o ancora che vengono acquistati dalle piattaforme di uso quotidiano (social media, strumenti di geolocalizzazione etc.). L'individuo rischia, quindi, di perdere le sue prerogative di dominus dei dati, anche soltanto per superficialità nella loro gestione, scarsa consapevolezza dell'uso che altri soggetti ne possono fare, non adeguata

³⁰ In UK è stata sviluppata una soluzione di AI che consente di produrre una stima immediata dei costi di riparazione sulla base delle fotografie che vengono scattate dalle auto coinvolte in un incidente.

³¹ Swedloff, 2020.

conoscenza della quantità di propri dati in circolazione. Va precisato, in proposito, che, da un punto di vista legale³², non è necessaria l'autorizzazione espressa dell'assicurato per l'utilizzo di determinate tipologie di dati da parte delle imprese assicurative, nella misura in cui queste abbiano un legittimo e fondato interesse a raccogliervi ed elaborarli, ad esempio in vista della prevenzione delle frodi. A questo fine, ad esempio, potrebbero legittimamente essere raccolte le informazioni che gli assicurati pubblicano sui social media, a condizione che, ovviamente, la quantità di dati processati sia coerente con la finalità per cui sono raccolti: una sovrabbondanza di dati, acquisiti e/o processati, potrebbe essere, infatti, non giustificata alla luce della finalità medesima.

A prescindere dalle considerazioni sulla loro astratta legalità, le modalità di utilizzo dell'AI nel settore assicurativo fanno sorgere comunque interrogativi di coerenza con più generali principi etici. In sintesi, dalle iniziative sopradescritte, emergono i seguenti profili di criticità: 1) accresciute asimmetrie informative tra imprese e consumatori e conseguente squilibrio di potere per effetto della capacità di raccolta ed elaborazione di grandissime quantità di dati da parte dell'AI; 2) iperpersonalizzazione dei premi e rischi di discriminazione/esclusione di categorie di persone nonché riduzione dei meccanismi di solidarietà sociale; 3) impiego dell'AI e limitazione della libertà di scelta degli individui.

3. Soft law, principi etici e disciplina dell'uso dell'AI

Va premesso che l'utilizzo dell'AI nel settore finanziario è soggetto alle medesime disposizioni di legge che regolano le modalità operative più tradizionali³³. Pertanto, la consulenza prestata tramite chatbox, ad esempio, deve assicurare il pieno rispetto delle previsioni di legge o regolamentari che si applicano a tale tipologia di operatività se prestata da una persona fisica. Poiché le previsioni attualmente vigenti sono state pensate per modalità operative più tradizionali, il quadro normativo potrebbe essere inadeguato a disciplinare l'utilizzo di strumenti tecnologicamente avanzati. Può infatti, da un lato, verificarsi il caso che norme esistenti siano sovradimensionate rispetto alle modalità operative dell'AI e richiedano, in prospettiva, modifiche per tener conto di questi modelli di business avanzati, una volta che se ne definiranno e comprenderanno appieno le caratteristiche; d'altro lato, nessuna norma di legge prende direttamente in considerazione i rischi³⁴ posti dall'AI, proprio a causa dell'innovatività del fenomeno e la limitata conoscenza delle sue potenzialità applicative³⁵.

³² Il trattamento dei dati personali resi pubblici dall'interessato, anche tramite social network, deve intendersi legittimo se fondato su una delle condizioni di liceità di cui all'art. 6 del Regolamento generale per la protezione dei dati personali n. 2016/679 - GDPR. Pertanto, il trattamento deve intendersi consentito – anche in assenza di formale consenso dell'interessato ma in presenza dell'azione positiva inequivocabile della pubblicazione sui social del dato - se “è necessario per il perseguimento del legittimo interesse del titolare del trattamento o di terzi, a condizione che non prevalgano gli interessi o i diritti e le libertà fondamentali dell'interessato che richiedono la protezione dei dati personali, in particolare se l'interessato è un minore”. Cfr. tra gli altri, Caggiano, 2018.

³³ Cfr. Commissione Europea, White Paper, 2020: “The EU has a strict legal framework in place to ensure inter alia consumer protection, to address unfair commercial practices and to protect personal data and privacy. In addition, the acquis contains specific rules for certain sectors (e.g. healthcare, transport). These existing provisions of EU law will continue to apply in relation to AI, although certain updates to that framework may be necessary to reflect the digital transformation and the use of AI”. Un'opportunità di regolamentazione specifica viene riconosciuta dalla Commissione, secondo un risk based approach, per quelle applicazioni di AI che vengono qualificate ad alto rischio sulla base di criteri la cui definizione dovrà essere “clear and easily understandable and applicable for all parties concerned”.

³⁴ La direttiva europea in materia di *Insurance Distribution* (2016/97) e le disposizioni di attuazione nazionali (articolo 119-bis del Codice delle assicurazioni private) richiedono ai distributori di prodotti assicurativi di agire con equità, professionalità, onestà, correttezza e trasparenza nel miglior interesse del contraente. Queste clausole di carattere generale devono trovare applicazione anche alle modalità di distribuzione che prevedano l'utilizzo dell'AI e rappresentano una possibile prima, anche se indiretta, risposta alle esigenze di tutela richiamate nel testo.

³⁵ *Compliance with the law is merely necessary (it is the least that is required), but significantly insufficient (it is not the most than can and should be done)* Floridi, 2018.

Al momento, regolatori, standard setter e la letteratura³⁶ in materia sono concordi nel ritenere che interventi prescrittivi del legislatore devono essere attentamente valutati, sia perché potrebbero ostacolare ingiustificatamente l'AI senza ancora conoscerne realmente le potenzialità, sia per evitare di compromettere l'assetto tecnologicamente neutrale della regolamentazione, funzionale allo sviluppo equilibrato del settore.

Resta però ferma la necessità di indirizzare l'uso dell'AI nel settore assicurativo (e finanziario più in generale) in modo da massimizzare i benefici sia per i consumatori sia per le imprese, costruire solide relazioni tra assicuratori e assicurati, nonché evitare che il gap di conoscenza e capacità tecnologiche tra consumatore e compagnie venga da queste utilizzato a proprio esclusivo vantaggio.

Si tratta di obiettivi imprescindibili per tutelare la fiducia del pubblico nell'AI e nel settore assicurativo nel suo complesso, per raggiungerli è necessario un quadro di riferimento che individui principi etici condivisi per l'uso dell'AI e detti raccomandazioni di base (soft law).

Negli ultimi anni lo sforzo di definizione di questi principi e linee guida è stato intenso e pervasivo: Jobin et al in uno studio³⁷ del 2019 hanno identificato ben 84 documenti contenenti "soft-law or non-legal norms" per l'AI, di cui l'88% pubblicati dopo il 2016³⁸, con un tasso di produzione in rapida crescita. I contributi offerti sono trasversali, in termini sia di natura dei loro autori (settore privato, settore pubblico) sia di provenienza geografica; l'analisi di questi documenti consente di individuare e mettere a fattor comune alcuni principi etici di base, da seguire nello sviluppo ed adozione dell'AI e da incorporare nei processi relativi alla sua applicazione.

Floridi et al. (2018) sistematizzano i criteri etici fondamentali, a cui disegno, sviluppo e applicazione dell'AI dovrebbero ispirarsi, prendendo a riferimento i principi basilari della bioetica, adattandoli e integrandoli per tenere in considerazione le peculiarità dell'AI. Sono, quindi, individuati i seguenti sei principi etici di base:

- 1) *principio di Beneficenza*: l'AI deve essere impiegata in modo da promuovere il benessere, preservando la dignità umana e la sostenibilità ambientale
- 2) *principio di Non Maleficenza* o del *primum non nocere*: l'AI deve garantire il rispetto della privacy, della sicurezza dei dati e deve prevenire gli effetti indesiderati derivanti dal suo utilizzo. La privacy viene tutelata riconoscendo al singolo il potere di avere accesso e di controllare i propri dati individuali e il loro uso; la sicurezza attraverso presidi che impediscano la sottrazione dei dati o il loro utilizzo illecito; la prevenzione degli effetti dannosi indesiderati è un obiettivo in divenire - che riguarda sia la componente umana dell'AI sia l'AI in senso proprio - ed è legato alle potenzialità, non ancora pienamente prevedibili, degli strumenti di intelligenza artificiale (compreso il condizionamento delle scelte individuali conseguente all'azione degli strumenti di AI, si pensi ad esempio agli algoritmi di ricerca del web che selezionano le informazioni da riportare).
- 3) *principio di Autonomia*: nell'utilizzo dell'AI è implicita la rinuncia a parte del potere decisionale della componente umana in favore di strumenti non umani, ciò al fine di rendere più efficienti i meccanismi decisionali, a livello individuale e collettivo. Per rispettare il principio di autonomia

³⁶ Al riguardo Floridi et al. evidenziano che "fear, ignorance, misplaced concerns or excessive reaction may lead a society to underuse AI technologies below their full potential, for what might be broadly described as the wrong reasons. This may cause significant opportunity costs. It might include, for example, heavy-handed or misconceived regulation, under-investment, or a public backlash".

³⁷ Jobin, et al 2019.

³⁸ Secondo Jobin: "most documents were produced by private companies (n = 19; 22.6%) and governmental agencies respectively (n= 18; 21.4%), followed by academic and research institutions (n =9; 10.7%), intergovernmental or supranational organizations (n =8; 9.5%), non-profit organizations and professional associations/scientific societies (n = 7 each; 8.3% each), private sector alliances (n = 4; 4.8%), research alliances (n = 1; 1.2%), science foundations (n = 1; 1.2%), federations of worker unions (n = 1; 1.2%)".

occorre individuare il punto di equilibrio tra il potere di scelta riservato alla componente umana e quello delegato a strumenti di AI; alla componente umana del processo dovrebbe essere sempre riservato il potere di individuare quali decisioni delegare per ragioni di efficacia ed efficienza e su quali mantenere il pieno controllo. Resta fermo che la decisione umana di delegare all'AI deve sempre poter essere revocata e la decisione dell'AI poter sempre essere modificata dall'intervento umano.

- 4) *principio di Giustizia*: l'uso dell'AI deve promuovere il benessere preservando la solidarietà tra diverse categorie di persone. In questo senso, l'AI deve essere utilizzata in modo da: correggere errori del passato, anche superando decisioni discriminatorie; assicurare che i benefici che apporta siano condivisibili dalla società in generale e non riservati ad alcuni; evitare che ne possano essere minate le basi delle strutture sociali esistenti (es. solidarietà nell'assicurazione).
- 5) *principio dell'Esplicabilità dell'AI* attraverso trasparenza e accountability. La spiegabilità dell'AI va intesa come risposta a due domande complementari: come funziona (trasparenza) e chi è responsabile per il modo in cui funziona (accountability). Si tratta di un principio imprescindibile per dare attuazione a tutti i precedenti: per comprendere se l'AI è usata per il bene o per il male deve essere chiaro come è costruita e quali effetti produce sul singolo e sulla società in generale; per valutare l'estensione della delega deve essere chiaro quale centro di imputazione umano risponde delle decisioni dell'AI. Per costruire e preservare la fiducia del pubblico nell'AI, è fondamentale che i risultati a cui perviene siano spiegabili.

Anche le più puntuali raccomandazioni adottate da policymakers, standard setters o associazioni di categoria dei diversi settori di applicazione dell'AI, sia a livello globale (OECD)³⁹ sia europeo (Commissione Europea)⁴⁰ possono essere ricondotte a questi principi di carattere generale.

Minimo comune denominatore di queste raccomandazioni è richiedere che le modalità di configurazione e di governo dell'AI siano tali da assicurare: sostenibilità sociale ed ambientale (deve apportare benefici nei confronti della società in generale e delle generazioni a venire); accountability (possibilità di individuare un processo per l'adozione delle tecniche di AI e soggetti che ne rispondano e lo guidino); equità (nel doppio significato di giustizia e non discriminazione); trasparenza (che include un concetto di comprensibilità proporzionata al soggetto nei cui confronti va garantita)⁴¹; sicurezza tecnologica e protezione dei dati personali (inclusa la possibilità di garantire al proprietario di avere accesso ai propri dati).

Coerentemente alle indicazioni della Commissione Europea e con specifico riguardo al settore assicurativo, l'EIOPA - insieme alle altre Autorità di Vigilanza europee del settore bancario e finanziario⁴² - riferendosi all'uso dei Big Data ma con un ragionamento pienamente applicabile a tutte le forme dell'AI - ha invitato le imprese a sviluppare ed applicare *good practices* volte a promuovere un trattamento equo, trasparente e non discriminatorio nei confronti dei consumatori. A questo fine, l'EIOPA individua una lista, indicativa e non esaustiva, di cautele e presidi che possono agevolare il formarsi di queste buone prassi, raccomandando nel contempo di non limitarsi ad essere compliant soltanto con tali indicazioni, che hanno carattere esemplificativo. Ad avviso dell'EIOPA, l'adozione di queste buone prassi - unitamente al pieno rispetto delle previsioni applicabili all'uso dell'AI e contenute nella legislazione sulla disciplina per la protezione dei dati e in quella a tutela del consumatore - contribuirà ad assicurare che le strategie di utilizzo dei Big Data (e dell'AI più in

³⁹ OECD Council Recommendation on Artificial Intelligence, <https://www.oecd.org/going-digital/ai/principles/>

⁴⁰ Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, "Ethics Guidelines for Trustworthy AI" (Apr. 2019): <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

⁴¹ Cath, 2018.

⁴² Joint Committee of the European Supervisory Authorities EBA, ESMA, EIOPA, *Joint Committee Final Report on Big Data, JC/2018/0415* (Mar. 2018)

generale) siano sviluppate responsabilmente e in modo da tenere in considerazione gli interessi della clientela.

Per alcuni aspetti è la stessa tecnologia di AI, opportunamente sviluppata ed impiegata, a poter offrire strumenti che agevolano il rispetto di queste raccomandazioni. A titolo di esempio si richiamano: la definizione di algoritmi che, fin dalla loro progettazione, incorporano i richiesti requisiti di tutela della privacy, evitando che siano limitate ex post le opportunità di analisi ed utilizzo dei big data⁴³; gli audit algoritmici⁴⁴, che permettono di utilizzare algoritmi di controllo per verificare se le tecniche di AI utilizzate producono risultati discriminatori, così da migliorarne le capacità predittive e rendere più semplice individuare e dimostrare i casi di discriminazione⁴⁵; infine, le tecniche di analisi controfattuale, che consentono di verificare se, modificando la caratteristica relativamente alla quale si presume una discriminazione (es. genere), la decisione adottata sarebbe stata la stessa⁴⁶.

4. Governance, politiche del personale e cultura aziendale: tre pilastri per un utilizzo responsabile dell'AI

E' indubitabile che, per effetto dell'utilizzo delle tecniche di AI, i modelli di business delle assicurazioni cambieranno fortemente, con importanti vantaggi per il mercato ma anche con rischi nuovi per i consumatori e, in futuro, ciò potrà portare anche ad un cambio di prospettiva dei regolatori del settore⁴⁷.

Le legislazioni relative alla protezione dei dati e alla tutela dei consumatori, in particolare del settore finanziario, sono sostanzialmente adeguate a prevenire i rischi più immediati e – allo stato attuale – manifesti per gli interessi del pubblico⁴⁸. Non è ancora sufficientemente chiaro, invece, come prevenire rischi e potenziali inefficienze derivanti dalle rilevanti asimmetrie informative che l'impiego dell'AI - da parte delle imprese finanziarie in generale e delle assicurazioni più in particolare - genera nei confronti del pubblico. Queste asimmetrie sono tali da poter causare fallimenti di mercato, causando uno squilibrio strutturale di conoscenza e limitando l'efficienza del libero scambio di prodotti e servizi, e potrebbero giustificare, in astratto, l'intervento del legislatore⁴⁹.

⁴³ Per maggiori riferimenti cfr. Monreale, et al. 2014.

⁴⁴ Sandvig, et al, 2014.

⁴⁵ Cfr. Barocas, s. et al.: "... data mining can discover surprisingly useful regularities that are really just preexisting patterns of exclusion and inequality. Unthinking reliance on data mining can deny historically disadvantaged and vulnerable groups full participation in society. Worse still, because the resulting discrimination is almost always an unintentional emergent property of the algorithm's use rather than a conscious choice by its programmers, it can be unusually hard to identify the source of the problem or to explain it to a court."

⁴⁶ Russell, et al.

⁴⁷ "The introduction of these new technologies is disrupting the premises of insurance regulation. AI is fundamentally different from the statistical techniques that once drove these same functions. Humans and human intuition do not underlie AI processes. AI works as an unstructured, iterative process to find correlations or solutions to problems posed by data scientists. These solutions are often unexpected and may not have easy or obvious causal stories. That is itself a big change. Regulating in this new era requires some thought. These new algorithmic methods could... undermine protections built into insurance regulation, but they could also minimize some of the concerns that drove regulation in the first place". Swedloff, 2020.

⁴⁸ Come precisato dalla Commissione Europea: "Developers and deployers of AI are already subject to European legislation on fundamental rights (e.g. data protection, privacy, non-discrimination), consumer protection, and product safety and liability rules. Consumers expect the same level of safety and respect of their rights whether or not a product or a system relies on AI. However, some specific features of AI (e.g. opacity) can make the application and enforcement of this legislation more difficult. For this reason, there is a need to examine whether current legislation is able to address the risks of AI and can be effectively enforced, whether adaptations of the legislation are needed, or whether new legislation is needed. Given how fast AI is evolving, the regulatory framework must leave room to cater for further developments. Any changes should be limited to clearly identified problems for which feasible solutions exist." Commissione Europea, White Paper 2020.

⁴⁹ Cfr. Swedloff cit. "Insurance regulation exists both to correct market failures and to protect insurance markets in light of the important role that insurance plays in society. To do this work, regulators—including state insurance commissioners, legislatures, and courts—have erected a series of structures designed to prevent opportunism, correct the behavioral quirks of insurance purchasers, and socialize risks more broadly".

Ad avviso di chi scrive, la velocità nello sviluppo e la rapida estensione dei potenziali ambiti di utilizzo degli strumenti di intelligenza artificiale non consentono, al momento, la definizione di un framework normativo statico che imponga obblighi, codifichi divieti e irroghi sanzioni.

Le sperimentazioni sul campo non sono, infatti, sempre compatibili con un sistema fatto di regole rigide ed eteroimposte⁵⁰. Nella fase attuale di sviluppo è importante assicurare flessibilità al quadro regolamentare di riferimento, ferma restando la necessaria, piena tutela dei diritti fondamentali individuali, ad oggi garantita dalla normativa sui dati personali e da quella a tutela del cliente, da tenere costantemente aggiornate e adeguate alle nuove esigenze.

Non va, però, trascurato il potenziale disequilibrio delle parti in causa (imprese da un lato, clientela attuale e potenziale dall'altro) che può essere colmato solo con la piena aderenza a principi etici, di cui al precedente paragrafo, e con un agire delle imprese che non sia guidato dal solo obiettivo della massimizzazione dei vantaggi ma ad esso affianchi quello del miglioramento del servizio prestato, recuperando margini di efficienza e rafforzando i modelli di business. In questo modo l'uso sistematico dell'AI nel settore finanziario non rappresenterà un'ulteriore minaccia alla fiducia del pubblico nei confronti di imprese di assicurazione, banche, e altri attori del mercato.

La sola enunciazione di principi etici, che rischiano di rimanere astratti, teorici e suscettibili di variabile applicazione, non è però sufficiente per ottenere il risultato descritto; occorre anche far leva sulla capacità delle imprese di autoregolarsi e auto organizzarsi. Ragioni economiche, e non soltanto etiche, sono alla base di questa necessaria riflessione organizzativa; come efficacemente rilevato dai consulenti indipendenti del Governo dell'UK⁵¹, la capacità di costruire una relazione solida e di fiducia con il mercato potrà rappresentare un vantaggio competitivo importante in un quadro fortemente deregolamentato e, in prospettiva, reso ancora più complesso dall'ingresso di nuovi player (quali le high tech companies) dotati di innovatività, grandi capacità di elaborazione dei dati e consistenti mezzi patrimoniali.

Per facilitare le imprese nel compito di definire presidi per l'utilizzo etico dell'AI, possono applicarsi paradigmi rivelatisi efficaci per la gestione e mitigazione di rischi che presentano caratteristiche simili a quelli posti dall'AI e per i quali la regolamentazione di per sé non è sufficiente a prevenire il verificarsi di danni rilevanti per la clientela e per il sistema finanziario più in generale.

Si fa riferimento in particolare al Misconduct Risk, venuto alla ribalta soprattutto negli ultimi anni come potenziale causa di instabilità sistemica, di vulnus alla fiducia del pubblico e di uscita dal mercato di alcune imprese soprattutto del settore bancario⁵². Detta tipologia di rischio presenta profili di somiglianza con quelli posti dalle nuove tecnologie, soprattutto per le molteplici forme in cui può manifestarsi e per le difficoltà che la regolamentazione incontra nel gestirlo e prevenirlo⁵³. In particolare, il Misconduct Risk viene definito come il rischio che nasce da condotte che non siano coerenti con gli standard del settore, siano essi legali, professionali o etici⁵⁴. Questa ampia definizione

⁵⁰ Cfr. Anche Commissione Europea supra.

⁵¹ *"With most insurers still in the midst of scoping out potential applications of AI in their product and service ranges, there is a window of opportunity to develop a blueprint for a credible governance regime – one that sets out a vision for the responsible use of AI but also the practical steps needed to achieve it. Before long, the industry will face fresh challenges, among them the potential entry of large tech companies into the market and the growth of new verticals such as cyber security insurance. It is in the industry's interests to put in place the necessary safeguards for AI's deployment before these new players and products materialise."* Centre for Data Ethics and Innovation, 2019

⁵² Si richiama a titolo di esempio il caso di misconduct rappresentato dalle *"Danske Bank's violations of the AML legislation involved substantial amounts, and the shortcomings in the bank's overall governance went on for several years. The case has had serious consequences for the bank, and it has damaged public confidence in the financial system as well as Denmark's reputation"* Fonte: Finalistynet, Report on the Danish FSA's supervision of Danske Bank as regards the Estonia case, 2019

⁵³ Financial Stability Board (FSB), 2018.

⁵⁴ Come precisato dal Financial Stability Board il Misconduct Risk si differenzia dal compliance risk perché è maggiormente connesso a *"day-to-day risk decisions of the different businesses as a first line of defence, or breaches related to codes of conduct, whereas compliance risk is seen as abiding by laws, regulations and rules."* FSB, 2017.

potrebbe consentire di ricondurre i rischi posti da un utilizzo non etico dell'AI ad una delle possibili manifestazioni del Misconduct Risk, ma, più che un esercizio astratto di tassonomia dei rischi, quello che interessa in questa sede è descrivere un paradigma di gestione dei rischi dell'AI che sia efficace, anche perché già conosciuto e sperimentato.

Il framework di presidio del Misconduct Risk, incentrato su misure di carattere preventivo e non soltanto reattivo, si presta a questo scopo, proprio perché trova il suo fondamento in un agire dell'impresa che sia coerente, non soltanto con le norme ad essa direttamente applicabili, ma anche con principi etici e regole soft di comportamento non scritte.

Il paradigma di gestione e controllo dei rischi posti dall'AI che si intende proporre, costruito sulla falsariga di quello adottato per il Misconduct Risk, passa quindi attraverso tre fondamentali pilastri organizzativi dell'impresa: la governance societaria (comprensiva anche del sistema dei controlli), i processi interni e le politiche del personale, la cultura aziendale.

a) *Governance*

La valorizzazione del board delle imprese finanziarie, in termini di accresciute responsabilità e di corrispondenti competenze tecniche, è stata la risposta fondamentale di standard setter e regolatori alle criticità riscontrate nella corporate governance delle imprese medesime, durante e per effetto dell'ultima crisi finanziaria.

Oggi questi nuovi board, più consapevoli e più coinvolti nella gestione aziendale, possono rappresentare un pilastro importante anche per la gestione dei rischi posti dall'impiego dell'AI, a condizione che siano fortemente coinvolti nella definizione delle strategie per uno sviluppo ed utilizzo etico delle nuove tecnologie.

In questo senso, è innanzitutto fondamentale che le strategie di utilizzo delle tecniche di AI siano decise in modo collegiale e pienamente informato, nella profonda consapevolezza dei risvolti etici che la loro adozione comporta. A questo fine, la capacità specifica richiesta al board, di assumere collettivamente la responsabilità per le implicazioni dell'uso dell'AI e di valutarne i relativi rischi, dovrebbe essere supportata e rafforzata attraverso un'attività di formazione adeguata e da flussi informativi, il più possibile veloci ed informali, tra il board nel suo complesso e le funzioni tecniche⁵⁵.

Nelle realtà in cui l'AI rappresenta un elemento importante del modello di business potrebbe essere funzionale ad una migliore valutazione collettiva dei profili di delicatezza delle strategie, la costituzione all'interno del board di un comitato - composto da amministratori che abbiano specifiche competenze in materia e che siano indipendenti dal business - a cui siano attribuite funzioni consultive nei confronti del collegio su tematiche di impiego responsabile dell'AI. La possibilità di riporto diretto, sia da parte delle funzioni di controllo (internal audit) sia da quelle di sviluppo dell'AI, a questo comitato rappresenterebbe una naturale conseguenza del ruolo che gli viene attribuito.

Come per gli altri comitati endoconsiliari, frequenti nei consigli di amministrazione delle imprese finanziarie, il comitato non sostituirebbe il board nelle valutazioni strategiche relative all'AI - che restano di esclusiva competenza di quest'ultimo - ma sarebbe chiamato a contribuire, con specifico know how tecnico e, soprattutto, informativo, all'efficacia delle decisioni del collegio. Comitati con analoghe caratteristiche e funzioni sono stati istituiti in alcune imprese non finanziarie, particolarmente attive nel campo dell'AI, e hanno avuto alterne vicende; quelli che si sono rivelati di maggiore impatto avevano struttura simile ai comitati tradizionalmente conosciuti

⁵⁵ Floridi et al. citato

nel settore finanziario (composti da membri del board con adeguati requisiti di fitness and propriety)⁵⁶.

Sempre con riguardo alla governance, un ulteriore importante intervento organizzativo potrebbe essere quello di individuare una figura all'interno del board responsabile per lo sviluppo, l'applicazione e l'utilizzo dell'AI all'interno dell'azienda.

La definizione di responsabilità individuali all'interno dell'impresa, facilmente azionabili anche in termini reputazionali, che partano dal board per proseguire a cascata nell'intera organizzazione aziendale, rappresenta un presidio molto rilevante della coerenza dell'operato dell'impresa con le regole e i principi etici che dovrebbero guidarne l'attività⁵⁷. In diversi ordinamenti, quale ad esempio quello inglese con il Senior Management Regime, la definizione di una dimensione individuale della responsabilità ha rappresentato, insieme alla valorizzazione del ruolo del board, la risposta regolamentare agli scandali finanziari - susseguiti a partire dalla crisi finanziaria globale (Libor, Forex per fare qualche esempio) - finalizzata a indurre un cambiamento culturale e a favorire un contesto eticamente sano nel settore. Oggi questi stessi principi di responsabilità individuale possono essere applicati per favorire un maggior coinvolgimento della componente umana nell'uso dell'AI (c.d. *human-in-the-loop systems*).

In particolare, l'individuazione di un componente del board, come responsabile per l'uso dell'AI nell'operatività aziendale, comporta come possibili ulteriori corollari che: a) il medesimo soggetto sia coinvolto in una due-diligence - preliminare e periodica - sull'utilizzo dell'AI nel business aziendale, che comprenda una ricognizione completa di tutte le caratteristiche dell'AI, inclusa la qualità dei dati⁵⁸; b) e nei suoi confronti sia attuata e documentata una piena spiegabilità e trasparenza degli strumenti utilizzati. Sia la due diligence sia la spiegabilità degli strumenti consentono, da una lato, di rendere concreta l'accountability del componente del board incaricato di sovrintendere all'uso dell'AI e, dall'altro, di tenerlo indenne da responsabilità ove dimostri di aver pienamente adempiuto ai propri obblighi. A loro volta, i responsabili delle funzioni tecniche interne all'organizzazione competenti per l'AI (ad esempio gli sviluppatori) potrebbero costituire il supporto operativo del componente del board incaricato, senza che ne venga meno l'accountability ultima⁵⁹.

⁵⁶ Secondo Zetzsche et al. *"In order to address the information asymmetry as to AI's functions and limits, regulators should take advantage of an important practice emerging in some non-financial companies. These companies have created independent AI review committees to provide cross-disciplinary and impartial expertise to such companies developing and utilising AI. Some of these committees or boards have been quite impactful, such as Axon's management and board accepting the recommendation of its AI and Policing Ethics Board to impose a moratorium on the use of facial recognition in Axon's body cameras. In any case, these boards are designed to augment decision-making and do not detract from the ultimate responsibility vested in management and the board regarding AI governance"*

⁵⁷ Sulla base degli esiti di un'analisi condotta nel 2018, che mostrava che in diverse imprese di assicurazione non era possibile individuare chi fosse l'owner dei risk assessment condotti dall'AI per quantificare i premi, la Financial Conduct Authority (2018) ha invitato le imprese a considerare *"whether their organisational structures are fit for purpose, and whether they need to allocate individual Board members responsibility for overseeing uses of AI and other forms of data-driven technology. In doing so, insurers should refer to the FCA's Senior Managers and Certification Regime, which requests a Statement of Responsibilities for senior managers"*.

⁵⁸ Ovviamente la responsabilità del componente del board non si sostituisce a quella delle singole funzioni che sono responsabili della qualità dei dati, ma le completa riconducendole ad unità. Proprio con riferimento alla qualità dei dati un'importante azione di adeguamento degli assetti di governance emerge dal report EIOPA del 2019 in materia di Big Data and Analytics in motor and health insurance

⁵⁹ Secondo Zetzche et al. *"These key function holder rules, particularly if enhanced by specific due diligence and explainability requirements, will assist core staff of financial services firm to ensure that the AI under their control is performing in ways consistent with their personal responsibilities. If it is not, they will nonetheless be responsible. That is the nature of personal responsibility systems: the manager etc. in charge is responsible for themselves, their area, their staff, their third party contractors, and their IT, including AI. This encourages – as a result of direct personal responsibility – due diligence in investigating new technologies, its uses and its impact and on requiring explainability systems as part of any AI system – or IT system for that matter. This is necessary from the standpoint of an individual who has potential direct responsibility in the event of a regulatory action for any failure: due diligence and explainability will be the key to a personal defence."*

In questo complesso compito di governo dell'AI, anche il sistema di controllo delle imprese ha un ruolo fondamentale.

In virtù degli sviluppi normativi, ma anche della ormai acquisita cultura del controllo, tutte le imprese del settore finanziario sono ormai pienamente consapevoli⁶⁰ dell'importanza del modello delle tre linee di difesa per la gestione di qualsiasi tipologia di rischio, anche estremamente innovativo come è il caso dell'utilizzo degli strumenti di AI. Le strutture di risk governance esistenti potrebbero quindi vedere ampliato il proprio mandato per ricomprendervi le nuove responsabilità, in particolare chiarendo la suddivisione di ruoli e prerogative delle key functions (attuariale, risk management, compliance e audit) rispetto all'AI.

Potrebbe anche essere opportuno creare nuovi ruoli, in relazione alla complessità organizzativa o dimensionale del modello di business della singola impresa, per ricondurre ad unità le responsabilità di controllo in materia, ferma restando la necessità che, ove costituite, queste nuove figure siano integrate e coordinate con le esistenti, per favorire sinergie, contaminazioni culturali ed evitare di creare silos di competenze, estremamente pericolosi soprattutto per i settori innovativi.

Sicuramente in questo articolato sistema di presidi integrati, il coinvolgimento dei revisori esterni, dotati di competenze specialistiche ulteriori, potrebbe aggiungere un ulteriore livello di confronto, che accrescerebbe le potenzialità preventive del sistema di controllo interno⁶¹.

Possibile soluzione, che agevolerebbe questo obiettivo di integrazione, è la costituzione di team multidisciplinari, interni all'organizzazione, in cui collaborino e si confrontino persone con background tecnico e provenienza diversa (esperti di compliance, data scientist, informatici, attuari, responsabili del personale etc.). In momenti di emersione di esigenze e modelli nuovi di controllo, infatti, assume rilievo cruciale la contaminazione di esperienze tra chi sviluppa le tecnologie e chi sviluppa le regole interne all'azienda: non è, infatti, sufficiente rispettare le norme per come sono oggi, occorre chiedersi se regole non scritte, che si fondano su principi etici riconosciuti e condivisi, impongano limiti o richiedano cautele in ambiti in cui l'AI venga impiegata o sia in via di sviluppo.

b) *Processi interni e politiche del personale*

Il coinvolgimento della componente umana nello sviluppo dell'AI è centrale per assicurare la piena comprensibilità del fenomeno e per individuare centri di imputazione delle responsabilità, in modo da cogliere con tempestività eventuali criticità e porvi altrettanto tempestivo rimedio. Per ottenere questo risultato, oltre che una chiara suddivisione delle responsabilità di governo come descritte nel precedente punto, i processi e le politiche del personale sono determinanti, poiché possono supportare l'effettivo empowerment della componente umana nella gestione e sviluppo dell'AI.

⁶⁰ L'EIOPA - nel report del 2019 citato - prende atto della acquisita consapevolezza da parte di molte imprese di assicurazione dell'importanza di questi aspetti: *"Solvency II's governance requirements and the role played by the different key functions (i.e. audit, actuarial, compliance and risk management functions) were mentioned by many insurance undertakings as providing several 'lines of defence' to address potential issues arising from BDA."*

⁶¹ Cfr. Arndorfer et al. sia per una ricostruzione del three lines of defence model sia per l'importanza dell'external auditor: *"regulated financial institutions require a sound four-lines-of-defence model with an emphasis on the relationship between internal audit (third line of defence) and external audit and supervisors (both comprising the fourth line of defence). Despite their external status, the entities forming the fourth line of defence should be active in supervising and monitoring control issues in the organisation. This means that close interaction between the internal audit function, external audit and supervisors is crucially important."*

Un fondamentale punto di partenza è la definizione di una strategia e di conseguenti policy aziendali in materia di AI, da condividere, con adeguati flussi informativi e strumenti di comunicazione, a tutti i livelli dell'organizzazione aziendale.

Specifico focus sull'AI potrà essere inserito nel già previsto piano strategico sulla tecnologia, ma è fondamentale che il processo di adozione di linee guida e strategie in materia di AI sia autonomamente strutturato e preveda il coinvolgimento di più attori. All'interno dell'organizational ladder aziendale, infatti, la responsabilità di prevenire danni alla clientela e l'uso distorto degli algoritmi risiede senz'altro nel vertice (governance alta), ma chi può influire in modo rilevante su come l'AI venga concretamente usata sono anche coloro che l'implementano, la testano e la utilizzano nel quotidiano. Solo un processo biunivoco strutturato di comunicazione, tra chi opera con la clientela e chi al vertice decide le strategie di distribuzione e l'impiego dell'AI nell'operatività aziendale, può consentire nel continuo il presidio dei connessi rischi.

Le parti coinvolte in questo processo strutturato potranno essere in grado di svolgere efficacemente la propria funzione soltanto se in possesso della necessaria competenza professionale e di soft skill ulteriori, anche sotto il profilo della consapevolezza delle sfide etiche che l'AI può porre.

Una risposta a questa domanda di professionalità e caratteristiche personali potrà essere ricercata nelle politiche del personale, sia con riferimento all'hiring che alla formazione. Sotto il primo profilo, le politiche di assunzione potrebbero mirare a personale che sia in grado di comprendere rischi e potenzialità dell'AI da un punto di vista tecnico, ma anche che abbia consapevolezza dei rischi legali ed etici associati all'AI. L'interazione e il dialogo costante tra sviluppatori di software ed esperti di profili legali ed etici consente, infatti, di prevenire i rischi prima che si materializzino.

Anche l'attività di training potrebbe essere mirata sia sugli aspetti tecnici dell'AI sia sulle sue implicazioni di carattere etico.

In relazione alla dimensione e pervasività dell'uso dell'AI potrebbe anche essere opportuno avvalersi della collaborazione di figure specializzate⁶² nei profili etici dell'uso delle nuove tecnologie, che siano di ausilio ai board nella definizione delle relative strategie aziendali e siano di supporto al disegno e alle esecuzioni delle attività formative.

La necessità di valutare l'adeguatezza delle competenze tecniche riguarda non soltanto le figure inserite nell'organizzazione aziendale ma anche il board. Un'attività di formazione in materia di AI dei componenti è una misura importante per porre rimedio, da un lato, ai limiti strutturali della composizione della classe dirigente italiana, più anziana della media europea e non sempre al passo con l'evolversi del contesto operativo⁶³, e, dall'altro, per consentirgli di svolgere il loro ruolo di guida strategica dell'impiego etico dell'AI⁶⁴.

⁶² Cfr. West et al. *"It is important for companies to have respected ethicists on their staffs to help them think through the ethics of AI development and deployment. Giving these individuals a seat at the table will help to ensure that ethics are taken seriously and appropriate deliberations take place when ethical dilemmas arise, which is likely to happen on a regular basis. In addition, they can assist corporate leadership in creating an AI ethics culture and supporting corporate social responsibility within their organizations. These ethicists should make annual reports to their corporate boards outlining the issues they have addressed during the preceding year and how they resolved ethical aspects of those decisions."*

⁶³ Gruppo di Esperti sull'intelligenza artificiale – Ministero dello Sviluppo Economico

⁶⁴ Cfr. Alekseeva et al. *"The significantly higher premium for AI skills in management compared to other occupations suggests that the real shortage is of managers who are able to create and capture value with AI. If AI indeed is a general-purpose technology, then the real productivity gains will only come slowly over time and only as a consequence of deeper changes in how businesses operate. Hence, the evidence suggests that AI is at least as much a managerial challenge as it is a technological challenge."*

c) *Cultura aziendale*

Un board, pienamente coinvolto nella definizione delle strategie in materia di AI e consapevole dei connessi rischi, è poi chiamato a comunicare alle persone, inserite nell'organizzazione aziendale a tutti i livelli, i valori condivisi e i comportamenti attesi nell'utilizzo concreto dell'AI. Nella sostanza il board si fa parte attiva nella costruzione di una sana cultura aziendale di pieno rispetto dei principi che devono guidare l'impiego delle nuove tecnologie (*tone from the top*)⁶⁵.

A questo fine i principi etici astratti, così come descritti al par. 3, dovrebbero essere concretamente calati nella realtà di ciascuna impresa attraverso l'elaborazione di codici di condotta specifici. Anche per l'AI, come per molti aspetti che guidano l'operatività delle imprese regolamentate, non esiste una soluzione unica che vada bene per tutti. Ciascuna impresa è chiamata a definire concretamente, in coerenza con il proprio ambiente aziendale e con la propria operatività, principi quali l'equità, la trasparenza e la spiegabilità degli algoritmi, in modo da tradurre concetti generali in comportamenti concretamente azionabili dal personale.

Solo concentrandosi sul proprio modello di business e sugli specifici rischi che comporta, è infatti possibile graduare l'attenzione del codice di condotta, e conseguentemente del personale, ai concreti punti di attenzione che l'utilizzo dell'AI richiede, senza limitarsi a fare riferimento a cautele di carattere generale che riguardano l'intera industria assicurativa. Potrebbero essere utili anche codici di condotta elaborati a livello più generale per l'intero settore - con il coinvolgimento ampio di tutti gli stakeholder comprese le associazioni di categoria degli utenti dei prodotti e servizi assicurativi⁶⁶ - che potrebbero rappresentare un importante riferimento e una base di lavoro preziosa per le singole imprese di assicurazione.

Non è sufficiente però la mera astratta adozione del codice di condotta, è importante che il suo contenuto sia reso azionabile attraverso la capillare comunicazione all'interno dell'organizzazione aziendale, attraverso l'enforcement delle singole previsioni, in termini di contrasto alle condotte non coerenti, attraverso l'adeguamento nel continuo alla luce degli sviluppi nell'uso dell'AI.

In un contesto soggetto a rapidissima evoluzione, accanto a misure di carattere formale che perseguano le chiare violazioni del codice di condotta e che incoraggino l'emersione dei comportamenti irregolari (es. whistleblowing), assumono rilievo anche le azioni informali (quali lo stimolo al dialogo e al confronto, l'apertura alle critiche e l'accettazione dei possibili errori) che i manager possono avviare per favorire l'emersione delle criticità e lo sviluppo dei codici medesimi in parallelo alle modalità di applicazione delle nuove tecnologie.

La pubblicazione dei codici di condotta da parte delle imprese nel proprio sito web può, inoltre, rappresentare un passo importante per costruire la fiducia del pubblico del settore finanziario nell'utilizzo etico dell'AI e nell'effettiva condivisione dei benefici che da essa derivano. In questi documenti, infatti, ciascuna impresa renderà manifesti i principi che segue, i processi che utilizza per darvi attuazione e, più in generale, le modalità con le quali gestisce le principali criticità che l'uso dell'AI può comportare⁶⁷.

5. Conclusioni

L'utilizzo dell'AI nel settore finanziario e, in particolare, in quello assicurativo sta aprendo nuove opportunità di crescita, sia in termini di prodotti e servizi offerti alla clientela sia come modelli di business. Queste opportunità sono accompagnate da nuove potenziali criticità, tra queste, a livello

⁶⁵ Financial Stability Board, 2014

⁶⁶ Floridi et al. *"a multistakeholder approach... is the most effective way to ensure that AI will serve the needs of society, by enabling developers, users and rule-makers to be on board and collaborating from the outset"*

⁶⁷ Cfr. Darrell M. West et al, 2018

sistemico, il rischio che l'uso dell'AI sia percepito come produttivo di vantaggi e benefici economici esclusivamente per le imprese e non per la clientela, con conseguente perdita di fiducia nel settore finanziario.

Inoltre i benefici derivanti dall'impiego dell'AI nell'attività assicurativa possono realizzarsi appieno solo se ne è garantita l'affidabilità per i profili tecnici, legali, etici⁶⁸. In un contesto in cui l'intervento prescrittivo del regolatore – al di fuori dei campi della data privacy e cyber security - può risultare prematuro e alterare o limitare le potenzialità di sviluppo, è fondamentale la valorizzazione delle capacità delle imprese di auto organizzarsi.

Governance e organizzazione aziendale - intese in senso ampio come insieme di ruoli, persone, processi, strutture - sono le leve da attivare per evitare che l'utilizzo dei nuovi strumenti di intelligenza artificiale comprometta la relazione di fiducia tra istituzioni finanziarie e clientela.

La chiara imputazione delle responsabilità, manageriali e individuali, per lo sviluppo e applicazione dell'AI insieme alla cultura aziendale, improntata a principi etici condivisi, consentono di dare concreta attuazione all'esigenza di un coinvolgimento costante della componente umana per la gestione e mitigazione dei rischi dell'AI, siano essi attuali e potenziali, prevedibili o non ancora ipotizzabili.

La comunicazione all'esterno dei presidi adottati – in termini di governance, organizzazione interna, codici di condotta – ha poi le potenzialità per sostituire un enforcement affidato a regole prescrittive, che, per sua natura, ha carattere reattivo e non preventivo e, soprattutto, postula una piena conoscenza e maturità del fenomeno da regolare.

Un'adeguata disclosure dei presidi adottati dall'impresa per un'applicazione etica dell'AI e l'attenzione anche alla sostenibilità digitale nell'ambito dei principi ESG (Environmental, Social, Governance)⁶⁹ potrebbero rappresentare una prima risposta alle asimmetrie informative implicite nell'uso dell'AI e affiancare la disciplina del mercato alle raccomandazioni dei regolatori per raggiungere l'obiettivo della massimizzazione e condivisione dei benefici dell'AI per tutti gli stakeholders.

In prospettiva, saranno utili studi e analisi che rechino proposte per integrare la tassonomia dei criteri ESG⁷⁰, con criteri di valutazione della sostenibilità sociale, ulteriori rispetto a quelli che oggi genericamente richiamano soglie di salvaguardia minime in ambito sociale⁷¹, e, in particolare, con criteri riferiti alla sostenibilità digitale che potrebbero rendere effettivamente azionabile il controllo del mercato sull'uso responsabile dell'AI.

⁶⁸ Cfr. Commissione Europea, 2019: *“Un'AI affidabile si basa su tre componenti: 1) legalità, l'AI deve ottemperare a tutte le leggi e ai regolamenti applicabili; 2) eticità, l'AI deve assicurare l'adesione a principi e valori etici e 3) robustezza, dal punto di vista tecnico e sociale poiché, anche con le migliori intenzioni, i sistemi di IA possono causare danni non intenzionali. Ciascuna componente è necessaria ma non sufficiente per realizzare un'IA affidabile. Idealmente le tre componenti operano armonicamente e si sovrappongono.”*

⁶⁹ See Tang, 2020

⁷⁰ In particolare, il rapporto finale, curato dal Technical Expert Group (TEG) on Sustainable Finance e pubblicato nel marzo 2020, contiene *“recommendations relating to the overarching design of the Taxonomy, as well as guidance on how companies and financial institutions can make disclosures using the taxonomy.”*

⁷¹ In particolare, si chiede l'allineamento alle linee guida dell'Organizzazione per la Cooperazione e lo Sviluppo Economico per le multinazionali e ai Guiding Principles on Business and Human Rights delle Nazioni Unite.

BIBLIOGRAFIA

Agrawal A., Gans J. & Goldfarb A., *"The Simple Economics of Machine Intelligence"*, 2016

Alekseeva, L, Azar, J, Gine, M, Samila, S and Taska, B. 2020. *'The Demand for AI Skills in the Labor Market'*. London, Centre for Economic Policy Research, 2020.

Arndorfer, I, Minto A., *The Four Lines of Defence Model for Financial Institutions*, Financial Stability Institute, Occasional Paper 22, 2015

Aniko Hannak, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson, *Measuring Price Discrimination and Steering on E-commerce Web Sites*. In Proceedings of the 2014 Conference on Internet Measurement Conference (IMC '14). Association for Computing Machinery, New York, NY, USA, 305–318.

Bank of England/Financial Conduct Authority - *Machine learning in UK financial services*, October 2019

Barocas, Solon and Selbst, Andrew D., *Big Data's Disparate Impact* (2016). 104 California Law Review 671 (2016).

Blumenstock, JE and Eagle. *Divided We Call: Disparities in Access and Use of Mobile Phones in Rwanda*, Information Technology and International Development, 8(2), 1-16, (2012).

Caggiano, I. *Il consenso al trattamento dei dati personali nel nuovo Regolamento europeo. Analisi giuridica e studi comportamentali*, in Osservatorio del diritto civile e commerciale, Rivista semestrale, 1/2018, pp. 67-106.

Caliskan, A, Joanna J. B and Arvind N *"Semantics derived automatically from language corpora contain human-like biases."* Science 356.6334 (2017): 183-186.

Cath C. *Governing artificial intelligence: ethical, legal and technical opportunities and challenges*. Philos Trans A Math Phys Eng Sci. 2018 Oct 15; 376.

Centre for Data Ethics and Innovation (CDEI), *AI and Personal Insurance*, 2019

Coglianesi, Cary and Lehr, David, *Transparency and Algorithmic Governance* (November 9, 2018). Administrative Law Review, Vol. 71, P. 1, 2019, U of Penn Law School, Public Law Research Paper No. 18-38

Dhir, A. & Lee, C. M. *Integrating overlapping datasets using bivariate causal discovery*, 2019

EIOPA, *Big Data Analytics in motor and health insurance*, 2019

European Commission -Directorate-General for Financial Stability, Financial Services and Capital Markets Union - Expert Group on Regulatory Obstacles to Financial Innovation (ROFIEG): *30 Recommendations on Regulation, Innovation and Finance -Final Report to the European Commission*, December 2019

European Commission, *Ethics guidelines for trustworthy AI*, 2019

European Commission, *WHITE PAPER on Artificial Intelligence - A European approach to excellence and trust*, 2020

Financial Stability Board, *Stocktake of efforts to strengthen governance frameworks to mitigate Misconduct Risks*, 2017

Financial Stability Board, *Strengthening Governance Frameworks to Mitigate Misconduct Risk: a Toolkit for Firms and Supervisors*, 2018

Financial Stability Board, *Guidance on Supervisory Interaction with Financial Institutions on Risk Culture*, 2014

Floridi L. *Soft ethics and the governance of the digital*. Philosophy & Technology. 2018; 31(1):1–8.

Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E. *AI4People-An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations*. Minds Mach (Dordr). 2018; 28(4):689-707.

Gates, S & Perry, V & Zorn, P. *Automated underwriting in mortgage lending: Good news for the underserved? Housing Policy Debate - HOUSE POLICY DEBATE*. 13. 369-391. 2002

Gruppo di Esperti sull'intelligenza artificiale – Ministero dello Sviluppo Economico, *Strategia italiana per l'Intelligenza Artificiale*, 2020

Hannak, A., Soeller G., Lazer D., Mislove, A., and Wilson, C., *Measuring Price Discrimination and Steering on E-commerce Web Sites*. In Proceedings of the 2014 Conference on Internet Measurement Conference (IMC '14). Association for Computing Machinery, New York, NY, USA, 305–318.

Hermes Investment Management, *Investors' expectations on responsible artificial intelligence and data governance* Aprile 2019.

High-Level Expert Group on Artificial Intelligence- European Commission, *Policy and Investment Recommendations for Trustworthy AI*, 2019

High-Level Expert Group on Artificial Intelligence – European Commission, *A definition of AI: Main capabilities and scientific disciplines*, Aprile 2019

Jobin, A., Ienca, M. & Vayena, E. *The global landscape of AI ethics guidelines*. Nat Mach Intell 1, 389–399 (2019).

Joint Committee of the European Supervisory Authorities EBA, ESMA, EIOPA, *Joint Committee Final Report on Big Data*, JC/2018/0415 (Mar. 2018)

Jordan M. - Mitchell T., *Machine learning: Trends, perspectives, and prospects*, Science 349(6245): 255–260 - 2015

Marwick A, *How Your Data Are Being Deeply Mined*, N.Y. REV. BOOKS, Jan. 9, 2014,

McKay, McDaniel (2006) *Machine learning as pinnacle of formal decision-making?*

Monreale, A., Rinzivillo, S., Pratesi, F., *Privacy-by-design in big data analytics and social mining*. EPJ Data Sci. 3, 10 (2014).

Pasquale, F. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge, MA, USA, 2015.

Prince, Anya and Schwarcz, Daniel B., *Proxy Discrimination in the Age of Artificial Intelligence and Big Data* (August 5, 2019). 105 Iowa Law Review 1257 (2020).

Rambachan A, Kleinberg J, Mullainathan S, and Ludwig J - *An Economic Approach to Regulating Algorithms* NBER Working Paper No. 27111 - May 2020

Russell Chris, J. Kusner Matt, Loftus Joshua, Silva Ricardo, *When Worlds Collide: Integrating Different Counterfactual Assumptions in Fairness*, 31st Conference on Neural Information Processing Systems (NIPS 2017)

Sandvig, Christian, *Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms.* (2014).

Scopino g., *Do Automated Trading Systems Dream of Manipulating the Price of Futures Contracts? Policing Markets for Improper Trading Practices by Algorithmic Robots* [2015]

Soto V., Frias-Martinez, V., Virseda, J., & Frias-Martinez, E. (2011). *Prediction of Socioeconomic Levels Using Cell Phone Records. User Modeling, Adaption and Personalization*, 377-388

Tang, B “*Independent AI Ethics Committees and ESG Corporate Reporting on AI as emerging corporate and AI governance trends*” in Ivana Bartoletti, Susanne Chishti, Anne Leslie and Shan M. Millie, *The AI Book: The Financial Technology Handbook for Investors, Entrepreneurs and Visionaries in Regulation* (Wiley, 2019)

Technical Expert Group on Sustainable Finance, *Taxonomy: Final report*, Marzo 2020

Swedloff, Rick, *The New Regulatory Imperative for Insurance* (March 4, 2019). Boston College Law Review, 61 B.C.L. Rev 2031 (2020), Rutgers Law School Research Paper.

West Darrell M. and Allen John R., *How Artificial Intelligence is Transforming the World*, Brookings Institution report, April 24, 2018

Zetzsche, Dirk Andreas and Arner, Douglas W. and Buckley, Ross P. and Tang, Brian, *Artificial Intelligence in Finance: Putting the Human in the Loop* (February 1, 2020). CFTE Academic Paper Series: Centre for Finance, Technology and Entrepreneurship, no. 1, University of Hong Kong Faculty of Law Research Paper No. 2020/006

QUADERNI PUBBLICATI

- N. 1. [Il Ramo r.c. auto: raffronto tra l'Italia e alcuni paesi della UE su premi, sinistri e sistemi risarcitori del danno alla persona](#), di Lino Matarazzo (ottobre 2014).
- N. 2. [La riforma della CARD: costi dei sinistri e incentivi all'efficienza nel risarcimento diretto r.c. auto](#), di Riccardo Cesari, Marina Mieli e Arturo Valerio (maggio 2015).
- N. 3. [Dal chain ladder al modello di Merz e Wüthrich: derivazione completa del modello di volatilità della riserva sinistri in orizzonte annuale](#), di Stefano Cavastracci (giugno 2015).
- N. 4. [Redditività e ciclo del settore assicurativo italiano prima e durante la crisi](#), di Fabio Farabullini (novembre 2015).
- N. 5. [Seminari per i dieci anni del codice delle assicurazioni private](#), di Riccardo Cesari (Prefatore), Enrico Galanti (Coordinatore), Sandro Amorosino, Roberto Caponigro, Stefania Ceci, Luigi Farenga, Antonio Longo, Francesco Mauro, Gustavo Olivieri, Andrea Pezzoli, Salvatore Providenti, Umberto Santosuosso, Dario Zamboni (maggio 2016).
- N. 6. [Il diritto nella società contemporanea](#), di Ottavio De Bertolis (novembre 2016).
- N. 7. [Duration, convexity and the optimal management of bond portfolios for insurance companies](#), di Riccardo Cesari e Vieri Mosco (febbraio 2017).
- N. 8. [Il nuovo Regolamento IVASS sull'accesso agli atti - La distribuzione Assicurativa - Il gruppo dopo Solvency II](#), di E. Galanti, M. Binda, M. L. Cavina, M. Fodale, N. Gentile, R. Giay, P. Marano, P. Mariano, S. Marzucchi, A. Police, A. Serino, V. Troiano (Aprile 2017).
- N. 9. [Modello overdispersed Poisson: formula chiusa per la stima GLM della volatilità one year della riserva sinistri](#), di Stefano Cavastracci e Agostino Tripodi (giugno 2017).
- N. 10. [No news is good news: moral hazard in oligopolistic insurance markets](#), di Marco Cosconati (aprile 2018).
- N. 11. [Riforma del sistema europeo delle Autorità di controllo. Governance imprese assicurative. Gestione crisi nel settore bancario e assicurativo](#), di S. Butera, F. Buzzichelli, R. Cercone, A. Corinti, S. De Polis, C. Di Noia, S. Fortunato, E. Galanti, F. Montemaggiori, M. Morvillo, M. O. Perassi, P. Rosatone, V. Santoro, S. Scarcello, E. Serata, M. Siri (maggio 2018).
- N. 12. [Valore aggiunto e profittabilità delle compagnie italiane: cosa conta realmente?](#), di Leandro D'Aurizio (marzo 2019)
- N. 13. [Calamità naturali e coperture assicurative: valutazione dei rischi e policy options per il caso italiano](#), di Riccardo Cesari e Leandro D'Aurizio (luglio 2019)
- N. 14. [Defiscalizzare la r.c. auto: come e perché](#), di Riccardo Cesari e Antonio R. De Pascalis (ottobre 2019)

N. 15. [Two simple models of insurance fraud](#), di Riccardo Cesari (gennaio 2021)

