Quaderni FinTech

L'intelligenza artificiale nell'*asset* e nel *wealth management*

N. Linciano, V. Caivano, D. Costa, P. Soccorso, T.N. Poli, G. Trovatore

9

giugno 2022

CONSOB
COMMISSIONE NAZIONALE
PER LE SOCIETÀ E LA BORSA

Nella collana dei Quaderni FinTech
sono raccolti lavori di ricerca relativi
al fenomeno «FinTech» nei suoi molteplici aspetti
al fine di promuovere la riflessione e
stimolare il dibattito su temi attinenti
all'economia e alla regolamentazione
del sistema finanziario.

CONSOB

00198 Roma - Via G.B. Martini, 3
t +39.06.84771 centralino
f +39.06.8477612

20121 Milano - Via Broletto, 7
t +39.02.724201 centralino
f +39.02.89010696

h www.consob.it
e studi_analisi@consob.it

ISBN 9788894369762

L'intelligenza artificiale nell'*asset* e nel *wealth management*

N. Linciano, V. Caivano, D. Costa, P. Soccorso, T.N. Poli, G. Trovatore (*)

Sintesi del lavoro

L'impiego di sistemi di intelligenza artificiale (IA) nell'intermediazione mobiliare è oggetto di un intenso dibattito istituzionale volto ad analizzarne opportunità e rischi per investitori e operatori di mercato. Nel presente lavoro, dopo aver delineato le principali caratteristiche e le possibili applicazioni dell'IA nell'ambito dell'*asset* e del *wealth management* anche alla luce del dibattito in corso, si espongono i risultati di una ricognizione sull'utilizzo delle relative tecniche da parte delle principali società di gestione del risparmio operanti nel mercato italiano. Dalle evidenze raccolte emerge che i gestori riconoscono nello sviluppo di sistemi di IA una priorità strategica e ne prevedono una applicazione crescente nell'ambito delle diverse fasi del processo di investimento, sebbene sotto la continua supervisione umana di tutti i processi decisionali. Il lavoro si conclude con una riflessione sulle implicazioni dell'utilizzo delle tecniche di automazione nell'intermediazione finanziaria, con particolare riferimento ai temi di *governance* degli algoritmi, responsabilità dei *providers* e tutela della *privacy*.

(*) CONSOB, Divisione Studi.

Si ringraziano: Tommaso Vessio e Sergio Fernando Larregola Peron per il contributo fornito allo sviluppo del progetto e a una prima versione del lavoro; il prof. Andrea Perrone, il dott. Isacco Girardi e il dott. Giovanni Re Garbagnati per il sostegno alle attività di ricerca; la prof.ssa Fabiana di Porto per gli spunti offerti in diverse occasioni di confronto sui temi del lavoro; Giovanna Frati e Matteo Modena per gli utili commenti; Assogestioni, in particolare Roberta D'Apice, per la collaborazione prestata; le società di gestione del risparmio che hanno partecipato all'indagine, tra cui Axa Investment Managers, Arca Fondi Sgr, Anima Sgr, Eurizon Sgr, Fideuram Asset Management Sgr Spa, Generali Insurance Asset Management Sgr spa. Eventuali errori e imprecisioni sono imputabili esclusivamente agli autori. Le opinioni espresse nel lavoro sono attribuibili esclusivamente agli autori e non impegnano in alcun modo la responsabilità dell'Istituto. Nel citare il presente lavoro non è, pertanto, corretto attribuire le argomentazioni ivi espresse alla CONSOB o ai suoi Vertici.

Artificial intelligence in the asset and wealth management

N. Linciano, V. Caivano, D. Costa, P. Soccorso, T.N. Poli, G. Trovatore (*)

Abstract

The use of artificial intelligence techniques in asset and wealth management is under increasing scrutiny by regulators and supervisors interested in examining its benefits and costs.

The International Organization of Securities Commission (IOSCO) Report 'The use of artificial intelligence and machine learning by market intermediaries and asset managers', published in 2021, provides interesting insights into the scale and characteristics of the phenomenon. According to the IOSCO document, the use of artificial intelligence (AI) and machine learning (ML) by financial intermediaries and asset managers is increasingly widespread globally. The availability of new technologies together with increasing data and computing power could significantly influence the business models of operators with specific regard to activities such as financial advice, risk management, client identification and monitoring, trading, and portfolio management. The use of AI technologies by financial intermediaries and asset managers can lead to both significant benefits for firms and investors, who may benefit from better conditions for investment services, and new risks or the amplification of existing risks with negative impacts on the efficiency of financial markets and the investor protection.

This paper takes up and develops the elements of the current institutional debate on the risks and opportunities arising from the use of new technologies in financial intermediation, also in the wake of evidence on the degree of use by the main asset management companies operating in the Italian market.

(*) CONSOB, Research Department.

Thanks to: Tommaso Vessio and Sergio Fernando Larregola Peron for their contribution to the development of the project and to an initial version of the work; Professor Andrea Perrone, Dr Isacco Girardi and Dr Giovanni Re Garbagnati for their support for the research activities; Professor Fabiana di Porto for the insights offered on various occasions to discuss the topics of the work; Giovanna Frati and Matteo Modena for their useful comments; Assogestioni, in particular Roberta D'Apice, for their collaboration; the asset management companies that participated in the survey, including Axa Investment Managers, Arca Fondi Sgr, Anima Sgr, Eurizon Sgr, Fideuram Asset Management Sgr Spa, Generali Insurance Asset Management Sgr Spa. The authors are the only responsible for errors and imprecisions. The opinions expressed in the Report are the authors' personal views and are in no way binding on CONSOB.

The first part discusses some defining issues concerning both the data universe and the available AI techniques and illustrates the potential of the applications of new technologies in the various phases of the portfolio management value chain, as well as the possible limits and risks associated with them. The main features of AI systems, classes of algorithms and the most popular learning methods are briefly illustrated in an Appendix.

The second part of the paper presents the results of a survey on the use of AI systems in asset management. The survey was carried out in July 2021 in collaboration with Assogestioni, and involved eight large asset management companies, belonging to groups representing more than 60% of the assets under management in Italy as of the first quarter of 2022, and collected information and data relating to all stages of the value chain, as well as managers' opinions on the expected benefits and potential risks associated with new technologies. From the evidence gathered, it emerges that managers indicate the development of AI systems as a strategic priority due to the expected benefits in terms of innovation in management strategies, maintaining a competitive position and increasing operational efficiency. All the companies participating in the survey believe that over the next five years, AI systems will find increasing application in the various phases of the investment process. Looking ahead, managers indicate the need for stringent human oversight of decision-making processes based on the use of AI systems, although control functions and governance procedures for the underlying algorithms and data used are still evolving.

The third part of the paper investigates the implications concerning the use of automation techniques in financial intermediation, from which both benefits and risks may arise for the investor. With particular reference to risks, issues concerning the governance of algorithms, the responsibility of providers and the protection of privacy are examined. The Appendix provides, more in general, a preliminary review of the current legal and institutional debate on technological innovation, also with regard to the proposed EU regulation on artificial intelligence.

Sommario

III Profili di *investor protection*

(G. Trovatore, T.N. Poli; Appendice di T.N. Poli)

INTRODUZIONE

N. Linciano e R. D'Apice (*)

L'utilizzo di tecniche di intelligenza artificiale nell'*asset* e nel *wealth management* è all'attenzione crescente di regolatori e autorità di vigilanza interessati a vagliarne benefici e costi.

Il Rapporto della International Organization of Securities Commission (IOSCO) 'The use of artificial intelligence and machine learning by market intermediaries and asset managers', pubblicato nel 2021[1], fornisce interessanti indicazioni sulle proporzioni e sulle caratteristiche del fenomeno. Secondo il documento IOSCO, l'uso dell'intelligenza artificiale (IA) e del *machine learning*

(ML) da parte degli intermediari finanziari e dei gestori patrimoniali è sempre più diffuso a livello globale. La disponibilità delle nuove tecnologie unitamente a quella di dati e potenza di calcolo potrebbero influenzare notevolmente i modelli di business degli operatori con specifico riguardo al servizio di consulenza, alla gestione del rischio, all'identificazione e monitoraggio dei clienti, all'attività di trading e alla gestione del portafoglio. Dall'uso delle tecnologie IA da parte di intermediari finanziari e gestori patrimoniali possono discendere sia vantaggi significativi per imprese e investitori, che potrebbero beneficiare di migliori condizioni dei servizi di investimento, sia nuovi rischi ovvero l'amplificazione di rischi esistenti con ricadute negative per l'efficienza dei mercati finanziari e la protezione degli investitori.

Il presente Quaderno riprende e sviluppa gli elementi del dibattito istituzionale in corso, relativo ai rischi e alle opportunità derivanti dall'utilizzo di nuove tecnologie nell'intermediazione finanziaria, anche alla luce dell'evidenza sul grado di utilizzo da parte delle principali società di gestione del risparmio (Sgr) operanti nel mercato italiano.

La prima parte affronta alcune questioni definitorie concernenti sia l'universo dati sia le tecniche di IA disponibili e illustra le potenzialità delle applicazioni delle nuove tecnologie nelle varie fasi della catena del valore della gestione di portafoglio nonché i possibili limiti e rischi a esse legati. Le principali caratteristiche dei sistemi di IA, le classi di algoritmi e i metodi di apprendimento più diffusi sono illustrati sinteticamente in una Appendice.

(*) Nadia Linciano - CONSOB, Responsabile Divisione Studi (n.linciano@consob.it);
Roberta D'Apice - Assogestioni, Direttore Affari Legali (Roberta.Dapice@assogestioni.it).

1 IOSCO (2021), The use of artificial intelligence and machine learning by market intermediaries and asset managers, https://www.iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf.

La seconda parte del Quaderno espone i risultati di una ricognizione sull'utilizzo di sistemi IA nell'*asset management*. La ricognizione è stata realizzata a luglio 2021 in collaborazione con Assogestioni, ha coinvolto otto grandi Sgr, appartenenti a gruppi che rappresentano oltre il 60% delle masse gestite in Italia al primo trimestre 2022, e ha raccolto informazioni e dati riferibili a tutte le fasi della catena del valore, nonché le opinioni dei gestori su benefici attesi e rischi potenziali legati alle nuove tecnologie. Dalle evidenze raccolte emerge che i gestori riconoscono nello sviluppo di sistemi di IA una priorità strategica per i benefici attesi in termini di innovazione delle strategie di gestione, mantenimento della posizione competitiva e incremento dell'efficienza operativa. Tutte le società partecipanti all'indagine ritengono che nei prossimi cinque anni i sistemi di IA troveranno crescente applicazione nell'ambito delle diverse fasi del processo di investimento. In prospettiva, i gestori riconoscono la necessità di una stringente supervisione umana sui processi decisionali basati sull'utilizzo di sistemi di IA, sebbene le funzioni di controllo e le procedure di *governance* degli algoritmi sottostanti e dei dati utilizzati siano ancora in fase evolutiva.

La terza parte del Quaderno indaga le implicazioni riguardanti l'utilizzo delle tecniche di automazione nell'intermediazione finanziaria, dalle quali possono derivare per l'investitore sia benefici sia rischi. Con particolare riferimento ai rischi, si esaminano profili concernenti la *governance* degli algoritmi, la responsabilità dei *providers* e la tutela della *privacy*. Segue l'Appendice che, in linea più generale, riporta una rassegna preliminare della dottrina giuridica e del dibattito istituzionale maturato sinora in materia di innovazione tecnologica, anche con riguardo alla proposta di regolamentazione UE sull'intelligenza artificiale.

L'automazione nella gestione di portafoglio: opportunità e rischi nel dibattito in corso

1 L'universo dei dati disponibili e i sistemi di intelligenza artificiale

1.1 *Big data* e *alternative data*: definizioni e caratteristiche

Negli ultimi anni, grazie allo sviluppo delle nuove tecnologie, i dati provenienti da fonti non tradizionali hanno assunto un 'valore' autonomo. Tali dati possono essere distinti in due tipologie, secondo una terminologia spesso utilizzata (erroneamente) in modo intercambiabile: *big data* e *alternative data*.

L'espressione *big data* fa riferimento a un grande volume di dati strutturati, semi-strutturati e non strutturati che richiedono l'utilizzo di capacità computazionali e algoritmi innovativi[2] al fine di identificarli, archiviarli e analizzarli[3]. I *big data* sono comunemente caratterizzati secondo le cosiddette quattro 'V': *volume*, *variety*, *velocity, value*[4]. Nel corso del tempo, sono state individuate ulteriori caratteristiche tra cui la cosiddetta *veracity*, relativa ad accuratezza e affidabilità delle fonti e integrità dei dati[5].

L'espressione *alternative data* fa riferimento a dati innovativi, non strutturati ed eterogenei. A differenza dei *big data*, essi non hanno necessariamente dimensioni tali da richiedere nuove tecnologie di analisi[6].

2 Con l'espressione algoritmo, in termini generali, ci si riferisce a «*encoded procedures for transforming input data into a desired output, based on specified calculations*», Gillespie, T. (2013), The relevance of algorithms, in Gillespie, T., Boczkowski, P. e Foot, K., Media technologies: essays on communication, materiality, and society, Cambridge, 167–194.

3 Onukwugha, E. (2016), *Big Data and Its Role in Health Economics and Outcomes Research: A Collection of Perspectives on Data Sources, Measurement, and Analysis*, PharmacoEconomics 34, 91–93, https://doi.org/10.1007/s40273-015-0378-4.

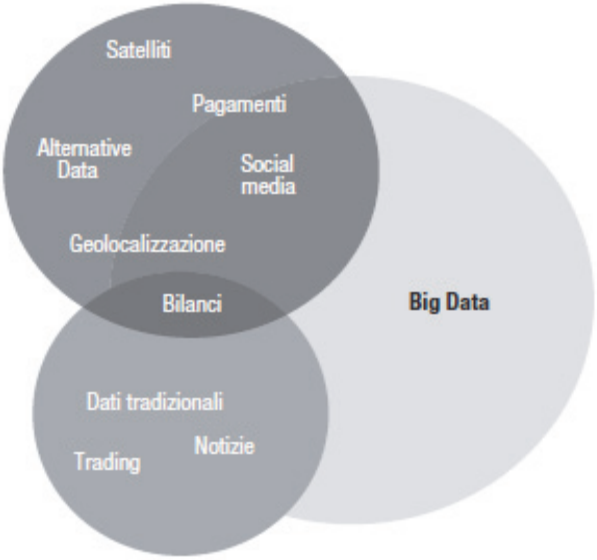
4 Gantz, J. e Reinsel D. (2011), *Extracting value from chaos*, IDC iView, 1, http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf. *Volume* fa riferimento all'ingente massa di informazioni che le tecnologie tradizionali non sono in grado di raccogliere; *variety* indica le differenti tipologie di dati disponibili e l'eterogeneità delle fonti da cui possono essere acquisiti; *velocity* si riferisce alla rapidità di generazione dei dati; *value* rinvia alla possibilità, attraverso metodologie di Big Data Analytics, di estrarre dai dati un contenuto informativo che ha un valore perché consente di prendere decisioni più informate, tempestive e consapevoli.

5 Piva, A. (2019), *Le 5V dei Big data: dal volume al valore*, www.osservatori.net.

6 In altri termini, l'espressione *big data* viene utilizzata con riguardo non tanto al carattere innovativo dei dati quanto soprattutto alla capacità della scienza dei dati di estrarne contenuto informativo (che può risultare innovativo). È il caso, ad esempio, dei dati ad altissima frequenza relativi alle negoziazioni di titoli che, anche se non sono particolarmente innovativi o di recente disponibilità, sono *big data* in senso dimensionale e offrono un nuovo contenuto informativo grazie ai progressi nella capacità computazionale disponibile. Vi sono invece database recenti e innovativi, che non sono grandi in relazione alle nuove tecnologie ma hanno un elevato valore. Si pensi, ad esempio, ai piani di viaggio aereo e di vacanza dei *top manager* delle aziende quotate che sono considerati innovativi e di valore, per il potere

La relazione tra le due categorie di dati è rappresentata nella Fig. 1.

Fig. 1 – La relazione delle fonti alternative e tradizionali di dati con i *big data*



Fonte: Guidolin, M., Magnani, M. e Mazza, P. (2021), Big data e sentiment analysis. Il futuro dell'asset management, Egea.

1.2 I metodi di machine learning

Le tecniche di *machine learning* (ML) rappresentano lo strumento attraverso il quale è possibile estrarre contenuto informativo di valore dall'insieme sempre più ampio dei dati disponibili. Il ML è considerato un sottoinsieme dell'intelligenza artificiale. In realtà, molte delle tecniche di ML si basano su modelli statistico-matematici lontani dai vecchi approcci di intelligenza artificiale, principalmente per la capacità di trattare grandi moli di dati e tipi di dati diversi (non strutturati, testuali, immagini), grazie all'impiego di *hardware* e tecniche matematiche di ottimizzazione che consentono di raggiungere un'elevata efficienza computazionale.

I metodi di ML sono classificabili a grandi linee in tre principali categorie.

La prima è nota come *apprendimento supervisionato* (*supervised learning*). L'algoritmo viene addestrato per apprendere la relazione che lega un insieme di osservazioni input a una variabile target output, in modo da poter riprodurre la variabile output per nuovi insiemi di input[7]. Questo metodo 'esperienziale' si distingue in approcci di regressione, che 'prevedono' l'output sulla base degli input, e approcci di classificazione, che classificano l'output in categorie.

predittivo di rendimenti e volatilità delle azioni, e che non sono di grandi dimensioni pur essendo complesso raccoglierli e assemblarli (Guidolin et al., *op. cit.* (2021)).

7 A titolo di esempio, si immagini un bambino che deve imparare a distinguere cani e gatti, sulla base di una preliminare educazione all'immagine che poggia su numerosi esempi di quadrupedi etichettati come cani e gatti.

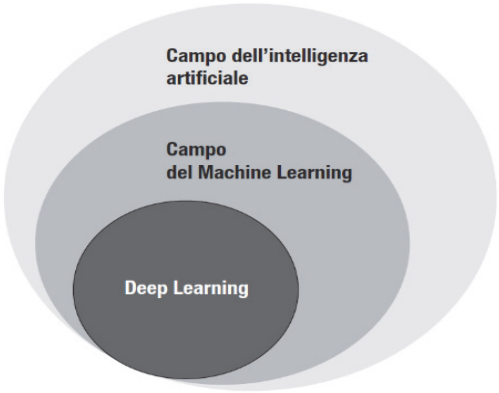
La seconda è nota come *apprendimento non supervisionato* (*unsupervised learning*) e permette di individuare direttamente la struttura dei dati. In questo caso non si dispone di un pregresso insieme di valori target, ma solo degli input. Lo scopo è raggruppare gli esempi in classi internamente coese, i cui elementi sono simili tra di loro sulla base dell'ottimizzazione di un criterio preassegnato, ma senza disporre delle etichette tipiche di un apprendimento supervisionato[8]. Il *clustering* individua sottogruppi di dati simili, mentre l'analisi fattoriale permette di identificare i principali *drivers* sottostanti ai dati.

Infine, nell'*apprendimento con rinforzo* (*reinforced learning*), l'ambiente fornisce segnali di rinforzo in termini di 'ricompensa', ma non indica esplicitamente la scelta corretta. L'algoritmo impara a scegliere in autonomia per raggiugere determinati obiettivi, interagendo con l'ambiente circostante. Si tratta di un metodo adatto a problemi dinamici di decisione su istanti di tempo successivi in cui si deve massimizzare la ricompensa complessiva della strategia decisionale.

Nell'ambito del ML, il *deep learning* è un sistema basato su architetture di reti neurali, secondo un approccio estremamente sofisticato che mira a replicare il funzionamento del sistema nervoso umano. L'apprendimento delle caratteristiche dei dati avviene in autonomia, anche senza istruzioni esplicite, e richiede un volume enorme di dati per l'addestramento. Le reti neurali profonde, complesse e multistrato generano una trasformazione non lineare e sempre più astratta dei dati. Il *deep learning* trova applicazioni nell'ambito del riconoscimento di immagini, della comprensione e della traduzione delle lingue e dell'automazione di compiti complessi (ad esempio, guidare un'auto).

La figura che segue illustra la relazione tra IA, ML e *deep learning* (Fig. 2; per una descrizione più dettagliata si rimanda all'Appendice '*I sistemi di intelligenza artificiale*').

Fig. 2 – La relazione tra intelligenza artificiale, *machine learning* e *deep learning* e il ciclo di vita di un sistema di IA



Fonte: Guidolin et al., *op. cit.* (2021), OECD (2021) Business and Finance Outlook: AI in Business and Finance, https://www.oecd.org/finance/oecd-business-and-finance-outlook-26172577.htm.

8 Tornando all'esempio di cui alla nota precedente, il bambino deve apprendere a classificare cani e gatti sulla base delle loro similitudini e differenze, ma senza il supporto di un maestro che fornisce le etichette corrette nella preliminare fase di educazione all'immagine.

2 Le applicazioni dei sistemi di IA nella gestione di portafoglio tra opportunità e rischi

2.1 Le applicazioni delle nuove tecnologie nella gestione di portafoglio

Nell'ambito dell'*asset management*, i *big data* e gli *alternative data* si stanno progressivamente affiancando (e in alcuni casi sostituendo) ai dati tradizionali e strutturati alla base delle strategie di investimento 'tradizionali' (dall'analisi fondamentale alle strategie quantitative)[9]. I dati tradizionali sono ritenuti oramai una *commodity* accessibile a tutti gli investitori; per contro, *big* e *alternative data* sono fonte di nuovo vantaggio informativo in grado di generare alfa[10].

Le applicazioni dell'IA e del ML per estrarre contenuto informativo e migliorare il processo decisionale nell'ambito della gestione di portafoglio possono avere ricadute positive su efficienza dei flussi di lavoro operativi, prestazioni, gestione del rischio e rapporto con i clienti[11]. Le prime evidenze raccolte con riferimento al comparto degli *hedge funds* mostrerebbero un impatto significativo sulla *performance* dei fondi che si dichiarano '*AI powered*'[12], confermando le aspettative dei gestori nei confronti delle nuove tecnologie e rivelate da numerose indagini campionarie pubblicate negli ultimi anni. I guadagni di efficienza potrebbero anche trasferirsi agli investitori *retail*, come emerge da alcune rilevazioni relative alle commissioni di gestione applicate dagli ETF gestiti in modo automatizzato.

Secondo CFA, le tecnologie più frequentemente utilizzate sono riferibili al ML e includono reti neurali, *cluster analysis* e *natural language processing* (NLP; Tav. 1).

9 Si vedano, tra gli altri: OECD, *op. cit.* (2019); Blackrock e Deloitte (2018), *Alternative data adoption in investing and finance. InFocus: Collective intelligence investing creates new rewards and risks*, https://www2.deloitte.com/us/en/pages/financial-services/articles/infocus-adopting-alternative-data-investing.html. Il Rapporto passa in rassegna anche i rischi legati all'uso crescente di tali dati, legati anche all'affidabilità e all'accuratezza delle fonti.

10 Tale circostanza potrebbe invertire l'aumento dell'incidenza delle strategie di investimento passivo; si veda Blackrock e Deloitte, *op. cit.* (2018).

11 Si vedano: Blackrock (2019), *Artificial intelligence and machine learning in asset management background*, https://www.blackrock.com/corporate/literature/whitepaper/viewpoint-artificial-intelligencemachine-learning-asset-management-october-2019.pdf; Deloitte (2019), *Artificial intelligence The next frontier for investment management firms*, https://www2.deloitte.com/global/en/pages/financial-services/articles/ai-next-frontier-in-investment-management.html. La maggior parte delle applicazioni delle nuove tecnologie nell'ambito della gestione di portafoglio riguardano il cosiddetto ML classico, ossia i metodi supervisionati e non. Il *deep learning*, anche per la necessità di utilizzare campioni di grandi dimensioni per l'addestramento, è verosimile che troverà applicazione (almeno inizialmente) nell'ambito del *trading* ad alta frequenza piuttosto che nell'*asset management*. Ovviamente ciò non esclude il fatto che i gestori già oggi possano farne indirettamente uso, ad esempio acquistando da *providers* esterni i dati di *sentiment* basati sull'applicazione di tecniche di *deep learning* ai testi pubblicati nei *social media* (Guidolin et al., *op. cit.* (2021)).

12 Si veda CFA, *op. cit.* (2020). Gli indici dei fondi *hedge* basati sull'IA forniti dal settore privato mostrano una sovraperformance rispetto agli indici dei fondi *hedge* convenzionali. Sebbene interessanti, tali evidenze sono da leggere con cautela, visto che si basano su indici elaborati dall'industria soggetti a una serie di distorsioni, come il *survivorship bias* o il *self-selection bias*. Inoltre, i gestori degli *hedge funds* (indubbiamente tra i primi a utilizzare sistemi di IA) non sono direttamente comparabili rispetto all'intensità e alla maturità di impiego delle nuove tecnologie, che avviene in gradi diversi, e alle metodologie sviluppate, che sono riservate. Di conseguenza, è difficile confrontare la *performance* dei prodotti che vengono definiti 'AI powered' con quelli tradizionali (OECD, *op. cit.* (2021)).

Tav. 1 – Le tecnologie di IA più frequentemente utilizzate nella gestione di portafoglio

Reti neurali artificiali	Modelli di regressione non lineare
	Rete di nodi connessi che modella vagamente i neuroni di un cervello
	Riceve un training set di coppie di input e output desiderati ed è in grado di apprendere la relazione tra di essi
	Possono essere utilizzate per prevedere l'output di input non noti in precedenza
	Applicazione tipica: previsioni
Alberi decisionali e foreste casuali	L'albero decisionale classifica le unità in base alle loro caratteristiche
	La classificazione viene effettuata percorrendo un albero logico dalla radice alle foglie, a ogni ramo si sposta a sinistra o a destra in base alle caratteristiche dell'unità; tali alberi possono essere interpretati dall'uomo
	Costruito automaticamente sulla base di un insieme di coppie di input e output desiderati
	Le foreste casuali fanno semplicemente la media dei risultati di diversi modelli di albero decisionale per produrre previsioni più affidabili
	Applicazione tipica: classificazione e previsione
Vector machine di supporto	Può essere utilizzata per la classificazione o la regressione
	Può gestire relazioni non lineari mappando gli ingressi in uno spazio a più dimensioni
	Più veloce da addestrare rispetto alle reti neurali artificiali
	Applicazione tipica: previsioni
LASSO	Modello di regressione ordinario con un termine di penalizzazione aggiuntivo che garantisce la scelta del sottoinsieme di variabili esplicative più piccolo possibile
	Riduce a zero le stime dei coefficienti spuri, migliorando in modo significativo le prestazioni *out-of-sample* del modello
	Applicazione tipica: Previsioni
Analisi *cluster*	Raggruppa i dati in gruppi in modo che le unità di ciascun gruppo abbiano caratteristiche simili
	Il numero di cluster può essere definito dall'utente o determinato automaticamente dall'algoritmo
	Applicazione tipica: classificazione dei beni
Algoritmi (genetici) evolutivi	Tecnica di ottimizzazione in grado di individuare le soluzioni preferite nell'ambito di grandi e complessi insiemi di soluzioni non lineari
	Processo ispirato all'evoluzione naturale
	Applicazione tipica: varianti dell'ottimizzazione di portafoglio che non possono essere risolvibili con gli algoritmi di ottimizzazione classici
Natural language processing	Gamma di tecniche utilizzate per elaborare dati in linguaggio naturale (ad esempio, testuali, audio)
	Particolarmente utile per estrarre informazioni dai media testuali (ad es., *social media*, siti web, articoli
	Applicazione tipica: analisi automatica di relazioni annuali aziendali, articoli

Fonte: CFA (2020), Artificial Intelligence in Asset Management, Bartram, S.M., Branke, J., e Motahari, M., (a cura di), https://www.cfainstitute.org/en/research/foundation/2020/rflr-artificial-intelligence-in-asset-management.

Le applicazioni delle nuove tecnologie possono riguardare diverse fasi della catena del valore, che la figura seguente schematizza in quattro momenti, indicando anche le sottofasi del processo di investimento e due aree che possono essere trasversali (Tav. 2).

Nell'ambito della raccolta, gestione e analisi dei dati, le tecniche di NLP possono estrarre contenuti informativi rilevanti da fonti eterogenee, che includono bilanci societari, articoli della stampa specializzata, informazioni reperibili nei *social media*[13].

13 Si veda CFA, *op. cit.* (2020); per dettagli si vedano tra gli altri: Azimi, M. e Agrawal, A. (2019), *Is Positive Sentiment in Corporate Annual Reports Informative? Evidence from Deep Learning*, https://ssrn.com/abstract=3258821; Schumaker,

Tav. 2 – Fasi della catena del valore della gestione di portafoglio in cui possono trovare applicazione sistemi di IA

identificazione del *target market* (definizione prodotto e clientela tipo)	raccolta, gestione e analisi dei dati
processo di investimento	classificazione dell'universo investibile (definizione *asset classes*)
	asset allocation strategica e ottimizzazione
	asset allocation tattica e ribilanciamenti
	selezione singoli strumenti *(stock/bond picking)*
	gestione del rischio
marketing, vendita e interazione con la clientela	
attività trasversali	gestione delle infrastrutture (diagnostica e monitoraggio del *cyber risk*)
	compliance

Fonte: elaborazioni CONSOB.

A differenza delle tecniche di analisi testuale più tradizionali, le tecniche di IA consentono di identificare automaticamente i fattori con il più alto potere predittivo dei rendimenti azionari[14], utilizzando variabili relative al contesto macroeconomico e aziendale. I modelli di IA possono inoltre essere proficuamente addestrati non solo sulla base di dati storici ma anche di dati prospettici, come le raccomandazioni di acquisto o vendita degli analisti[15].

Nell'ambito del processo di investimento, i risultati delle analisi dei dati possono essere incorporati nel processo di ottimizzazione del portafoglio per individuare i pesi delle attività che lo compongono anche in funzione del valore alfa e comunque degli obiettivi di *performance* dei gestori (ad esempio, replicare un *benchmark* o massimizzare il rapporto di Sharpe). Come emerge dalle prime evidenze empiriche disponibili, i sistemi di IA consentono di superare alcuni problemi legati all'applicazione del modello media-varianza di Markowitz[16] attraverso stime di rendimento e di rischio più

R.P. e Hsinchun, C. (2006), *Textual Analysis of Stock Market Prediction Using Financial News Articles*, AMCIS 2006 Proceedings, 185, 1431–40. https://pdfs.semanticscholar.org/db74/80f28a68b95ed35701b84a282d6ebd8eb366.pdf; Ke, Z.T., Kelly, B.T. e Xiu, D. (2019), *Predicting Returns with Text Data*, https://ssrn.com/abstract=3389884; Sprenger, T.O., Sandner, P.G., Tumasjan, A. e Welpe, I.M. (2014), *News or Noise? Using Twitter to Identify and Understand Company-Specific News Flow*, Journal of Business Finance & Accounting 41 (7–8): 791–830, https://doi.org/10.1111/jbfa.12086.

14 Si veda CFA, *op. cit.* (2020); per dettagli si vedano tra gli altri: Feng, G., Giglio, S. e Xiu, D. (2017), *Taming the Factor Zoo: A Test of New Factors*, Fama-Miller Working Paper; Chicago Booth Research Paper n. 17-04, https://ssrn.com/abstract=2934020; Freyberger, J., Neuhierl, A. e Weberm M. (2018), *Dissecting Characteristics Nonparametrically*, University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2018-50, https://ssrn.com/abstract=3223630.

15 Si veda CFA, *op. cit.* (2020); per dettagli si vedano tra gli altri Bew, D., Campbell, R.H., Ledford, A., Radnor, S. e Sinclair, A. (2019), *Modeling Analysts' Recommendations via Bayesian Machine Learning*, Journal of Financial Data Science 1 (1): 75–98.

16 Il modello media-varianza di Markowitz incontra nella prassi due sfide principali. In primo luogo, i pesi ottimali delle attività sono molto sensibili alla stima dei rendimenti attesi, che può presentare un *noise* tale da erodere i benefici della diversificazione (si vedano CFA, *op. cit.* (2020) e Kolm, P.N., Tütüncü, R. e Fabozzi, F.J. (2014), *60 Years of Portfolio Optimization: Practical Challenges and Current Trends*, European Journal of Operational Research 234 (2): 356–71, https://doi.org/10.1016/j.ejor.2013.10.060). In secondo luogo, la stima della matrice di varianza-covarianza richiede un'elevata profondità storica dei dati e l'assunzione di correlazioni stabili tra i rendimenti delle attività; la matrice diventa instabile quando la correlazione tra gli *assets* aumenta, ossia quando la diversificazione è più importante e al contempo più difficile da raggiungere (de Prado, L.M. (2016), *Building Diversified Portfolios that Outperform Out of Sample*, Journal of Portfolio Management 42 (4): 59–69, https://doi.org/10.3905/jpm.2016.42.4.059).

accurate di quelle prodotte dai metodi tradizionali e offrono approcci di ottimizzazione del portafoglio alternativi a quelli generati dalle tecniche lineari tradizionali[17].

L'uso dei *big data* e dei sistemi di IA si rivela utile anche nelle fasi precedenti all'investimento, nell'ambito delle operazioni di *back-office* e amministrazione degli investimenti e nell'ambito del marketing e della distribuzione dei prodotti finanziari. Queste ultime fasi beneficiano anzitutto di una migliore segmentazione della clientela, realizzabile sulla base di un ampio insieme di informazioni su preferenze e scelte individuali (raccolte ad esempio dai *social media*[18]), e dalla progettazione di modelli distributivi in linea con le caratteristiche di ciascun segmento e in grado di individuare opportunità di vendita combinata di più prodotti al medesimo cliente (*cross-selling*).

La gestione del rischio è un'altra attività che può trarre notevoli benefici dall'applicazione di sistemi di IA, con particolare riferimento al rischio di mercato e al rischio di credito. L'IA può, ad esempio, migliorare la modellizzazione del rischio di mercato attraverso l'uso di informazioni qualitative, estratte da fonti testuali o da immagini (come quelle satellitari), che permettono di stimare variabili finanziarie o economiche a livello aggregato e aziendale in modo più accurato rispetto ai dati tradizionali[19]. Tra le tecniche di IA che hanno conosciuto un impiego crescente negli ultimi anni, affiancandosi o sostituendosi agli approcci tradizionali, si ricordano l'analisi discriminante multivariata e i modelli logit e probit[20].

17 Si vedano, tra gli altri: CFA, *op. cit.* (2020); de Prado, *op. cit.* (2016); Snow, D. (2019), *Machine Learning in Asset Management*, JFDS: https://jfds.pm-research.com/content/2/1/10 , https://ssrn.com/abstract=3420952; Xiaoqiang, Z. e Ying, C. (2017), *An artificial intelligence application in portfolio management, Advances in Economics, Business and Management Research* (AEBMR), vol. 37, International Conference on Transformations and Innovations in Management (ICTIM-17), https://www.atlantis-press.com/proceedings/ictim-17/25885114; Zhang, Z., Zohren, S., e Stephen, R. (2020), *Deep Learning for Portfolio Optimization*, The Journal of Financial Data Science.

18 Nello specifico, è possibile individuare alcune modalità di applicazioni denominate, rispettivamente, *customer as a whole*, *sentiment analytics*, *customer segmentation*, *next best offer* e *channel journey*. La metodologia *customer as a whole* '*è finalizzata a individuare il comportamento passato e presente di un determinato soggetto, nella condizione che tramite queste informazioni sarà possibile determinare le sue future scelte e preferenze*'. Diversamente, la *sentiment analytics* '*consente di ascoltare e individuare, grazie alle conversazioni online, che cosa dicono, quali sono le opinioni dei customer a proposito di un determinato prodotto o servizio, grazie ad una raccolta costante di dati (tramite blog-forum, piattaforme, social network ecc.) prendendo in considerazione una o più porzioni di testo, quali i nomi di un brand o il nome di una persona*'. Si veda Mattassoglio F. (2016), *La profilazione dell'investitore nell'era dei big data. I rischi dell'esternalizzazione della regola del 'Know your customer rule'*; in Rivista trimestrale del diritto dell'economia, Vol. 4. Alle evidenti opportunità associate allo sviluppo di modelli con potenzialità conoscitive e predittive si associano i rischi illustrati nella Sezione '*Profili di investor protection*'.

19 Si veda Financial Stability Board (2017), *Artificial Intelligence and Machine Learning in Financial Services*, http://www.fsb.org/2017/11/artificial-intelligence-andmachine-learning-in-financial-service. Con riferimento al contenuto informativo delle fonti di dati testuali (come gli articoli di notizie, i post *online*, i contratti finanziari, i verbali e le dichiarazioni delle banche centrali e i *social media*) si veda Groth, S. e Muntermann, J. (2011), *An Intraday Market Risk Management Approach Based on Textual Analysis*, Decision Support Systems 50 (4); https://doi.org/10.1016/j.dss.2010.08.019. 2011; con riguardo al contenuto informativo delle immagini satellitari (analizzate ad esempio per prevedere le vendite nei supermercati o i futuri raccolti) si veda Zsolt, K., Painter, M., Panos, N. Patatoukas e Zeng, J. (2018), *On the capital market consequences of alternative data: evidence from outer space*, 9th Miami Behavioral Finance Conference 2018, https://ssrn.com/abstract=3222741. Per ulteriori riferimenti bibliografici e per una rassegna delle tecnologie utilizzabili si rimanda a CFA, *op. cit.* (2020).

20 Le due tecniche più utilizzate sono ANNs, impiegate sin dai primi anni '90 del secolo scorso, e SVMs, che sembrerebbero fornire previsioni leggermente più accurate rispetto alle prime risultando meno esposte a problemi di overfitting. Esiste comunque un'ampia gamma di approcci di IA utilizzabile per la modellizzazione del rischio di credito, le cui previsioni possono anche essere combinate per ottenere migliori prestazioni. Si veda CFA, *op. cit.* (2020).

Nel complesso, il valore generato dalle strategie basate su *big data* e dati alternativi sembra destinato a crescere e a compensare i costi di sviluppo e utilizzo di sistemi di IA. In tale ambito, una decisione chiave riguarda le modalità di acquisizione delle nuove tecnologie, ossia il grado di esternalizzazione delle attività. Ad esempio, nell'ambito della raccolta, archiviazione e analisi dei dati, per ogni singola fase il gestore può scegliere di operare *in house* ovvero rivolgersi a soggetti esterni. Nella fase della raccolta, in particolare, i dati potranno essere forniti da *data providers* specializzati (ad esempio, BigTech). Nell'ambito della successiva archiviazione dei dati (*storage*), il gestore potrà dotarsi di una propria piattaforma informatica[21] oppure utilizzare la piattaforma di un *provider* esterno (*cloud computing*). Anche la fase più delicata finalizzata a garantire l'affidabilità dei dati (cosiddetto *data-cleaning*) può usare come modello un database esterno.

L'esternalizzazione totale presenta il vantaggio di non richiedere investimenti in infrastrutture e competenze e lo svantaggio della non esclusività delle analisi e dei progetti acquisiti. Viceversa, la totale internalizzazione permette di raccogliere, conservare e analizzare i dati grezzi in modo rispondente alle proprie esigenze, a fronte di elevati investimenti nella ricerca e nell'acquisizione di tecnologie e risorse umane adeguate. Soluzioni soluzioni intermedie di esternalizzazione parziale consentono di raggiungere un compromesso, potendo coniugare ad esempio l'acquisto dei dati (grezzi, processati o semiprocessati) con l'analisi interna dei dati stessi[22].

Lo sviluppo di modelli proprietari fornisce un vantaggio competitivo che al momento possono cogliere solo gli attori di maggiori dimensioni. Proprio per questo motivo alcuni osservatori paventano la possibilità che l'automazione della gestione di portafoglio diventi appannaggio dei grandi *players*, rafforzando la tendenza alla concentrazione già in corso, almeno finché l'*outsourcing* non diventerà una modalità di accesso ai modelli di IA sufficientemente diffusa[23]. Tuttavia, anche quest'ultimo sviluppo presenta potenziali criticità legate all'eventualità che il numero di *providers* di modelli di IA sia ristretto e che ciò alimenti comportamenti imitativi da parte dei gestori, con un conseguente innalzamento dei rischi sistemici e della volatilità, soprattutto nelle fasi di turbolenza dei mercati.

2.2 Limiti e rischi delle nuove tecnologie

L'utilizzo delle nuove tecnologie nella gestione di portafoglio è foriero non solo di benefici ma anche di rischi. Rimandando ai successivi contributi di questo Quaderno l'approfondimento di alcuni temi con particolare riguardo alla gestione di portafoglio, in questa sede si passano in rassegna i profili di criticità legati rispettivamente a: qualità dei dati, distorsioni e discriminazioni; intepretabilità e robustezza degli algoritmi; sicurezza cibernetica. In ultimo si illustrano talune criticità

21 BaFin (2018), *Big data meets artificial intelligence: challenges and implication for the supervision and regulation of financial services*, www.bafin.de, 55.

22 Guidolin et al., *op. cit.* (2021).

23 OECD (2015), *Data-Driven Innovation: Big data for growth and well-being*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264229358-en.

riferibili ai profili di *privacy* che, pur non incidendo direttamente sui risultati della gestione di portafoglio, rilevano sul piano della tutela dell'investitore.

Data quality, bias *e discriminazione*

I dati sono il cuore di qualsiasi sistema di IA e per questo la *data quality* è centrale. Essa può essere valutata tenendo conto di cinque dimensioni: la *availability*, ossia la semplicità di accesso ai dati disponibili; la *usability*, ossia l'utilizzabilità del dato in riferimento al singolo utente; la *reliability*, intesa come affidabilità; la *relevance*, ossia l'importanza dei dati per gli obiettivi del business aziendale o comunque rispetto al fenomeno di interesse; la *presentation quality*, rispetto ai metodi con cui si possono illustrare i dati agli utenti[24].

L'affidabilità della fonte ha ovviamente un ruolo determinante nella *data quality*. Essa può risultare compromessa da dati che non sono rappresentativi della popolazione da analizzare e/o sono affetti da distorsioni e discriminazioni sistematiche che vanno a detrimento di specifici gruppi e/o di singoli individui fino a determinarne l'esclusione finanziaria[25]. Questa tipologia di *bias* è alla base della cosiddetta discriminazione involontaria. Essa si distingue dalla discriminazione volontaria, che fa riferimento all'ipotesi in cui l'algoritmo viene elaborato *ex ante* per selezionare secondo specifiche categorie di dati sensibili (ad esempio, rispetto a razza, etnia e orientamento sessuale), generando così un database basato su una segmentazione distorta. La discriminazione involontaria è un effetto *ex post* dei *bias* soggettivi riconducibili agli sviluppatori o dei *bias* oggettivi derivanti dalle distorsioni ovvero dalla scarsa rappresentatività dei dati con cui viene alimentato un algoritmo addestrato con dati di buona qualità[26].

Nei modelli di ML ad apprendimento supervisionato sono fondamentali le fasi di pulizia (*data cleaning*) ed etichettatura (*labelling*) dei dati utilizzati per l'addestramento[27]. L'industria comincia a sperimentare l'utilizzo di *data set* sintetici opportunamente costruiti attraverso sistemi di IA, in modo da mitigare talune

24 Cai, L. e Zhu, Y. (2015), *The challenges of data quality and data quality assessment in the Big data era*, Data Science Journal, 14.

25 Si pensi, ad esempio, al caso in cui, con riguardo a sistemi di accesso al credito, i modelli di *credit scoring* penalizzano sistematicamente alcune caratteristiche individuali; lo stesso può accadere con riguardo all'acquisto di una polizza sanitaria. I modelli di IA basati su *big data*, inoltre, possono non riflettere la realtà in modo accurato. Si pensi ad esempio al caso in cui un individuo, che ha stipulato un'assicurazione sanitaria, ha acquistato un'attrezzatura per praticare uno sport estremo e tale acquisto sia censito dal modello di IA della compagnia assicurativa per determinare il premio della polizza. L'algoritmo però non sarebbe in grado di distinguere la circostanza in cui l'acquisto è effettuato per sé dal caso in cui si tratti di un regalo e in quest'ultimo caso darebbe un suggerimento sbagliato, associando all'acquisto un innalzamento del rischio di danni all'integrità fisica dell'acquirente.

26 La proposta di Regolamento della Commissione europea del 21 aprile 2021, che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione (maggiori dettagli nel seguito), vieta esplicitamente l'utilizzo di sistemi di IA per alcune finalità tra cui quella del *social scoring*, in quanto in grado di portare a risultati discriminatori e all'esclusione di determinati gruppi (per ulteriori dettagli si rimanda alla Sezione '*Profili di* investor protection' e all'Appendice della Terza parte).

27 In particolare, nel *data cleaning* uno dei profili più complessi è il trattamento dei cosiddetti *outliers*, la cui individuazione e rimozione migliorano la capacità di apprendimento dell'algoritmo, ma possono anche modificare la rappresentatività dei dati e limitare le potenzialità di addestramento dell'algoritmo con riferimento a situazioni anomale. Si veda Bogroff, A. e Guégan, D. (2019), *Artificial intelligence, data, ethics. A holistic approach for risks and regulation*, University Ca' Foscari of Venice, Dept. of Economics Research Paper Series, 19/WP/2019, 11.

distorsioni dovute all'intervento umano o ai *bias* presenti nelle serie storiche dei dati reali utilizzabili per l'addestramento.

Interpretabilità dei modelli di IA

L'impiego di sistemi di IA si associa a una criticità nota come assenza di interpretabilità (*explainability*) di alcuni modelli, ossia alla difficoltà di spiegarne il risultato in funzione degli input sottostanti. Le reti neurali 'profonde' sono intrinsecamente complesse e quindi opache, tanto da produrre risultati difficili da decifrare; viceversa, gli algoritmi di ML classici sono quasi sempre interpretabili (si pensi, ad esempio, alle regressioni lineari e logistiche e agli alberi decisionali; per ulteriori dettagli si veda l'Appendice '*I sistemi di intelligenza artificiale*').

La crescente richiesta di risultati spiegabili in termini umani, da parte sia dei regolatori sia del mercato[28], sta spingendo molti operatori ad adoperarsi per innalzare l'intepretabilità dei modelli utilizzati, anche attraverso la cosiddetta *explainability by design*, ossia incorporata nel sistema di IA. Tale soluzione non è un'alternativa possibile quando la trasparenza è incompatibile con la tutela del diritto di proprietà intellettuale ovvero con la stessa tecnologia, come nel caso prima menzionato delle reti neurali.

La rilevanza concreta dell'assenza di interpretabilità varia caso per caso. Essa può tradursi in una perdita monetaria a carico di alcune categorie più che compensata dai vantaggi derivanti dall'impiego del modello di IA ovvero può generare effetti avversi non mitigabili, come nel caso di errori nella determinazione dell'affidabilità creditizia di un individuo. La non interpretabilità unitamente al ricorso diffuso all'*outsourcing* per l'acquisizione di modelli 'chiavi-in-mano' sviluppati da pochi fornitori esterni possono inoltre accentuare comportamenti gregari e prociclici e generare quindi un impatto sistemico, a fronte dell'incapacità degli operatori che utilizzano i modelli di comprenderne la logica sottostante. L'opacità degli algoritmi di IA potrebbe favorire anche l'adozione non intenzionale di pratiche scorrette (come nel caso dei comportamenti collusivi su cui si tornerà a breve).

I rischi appena menzionati giustificherebbero l'introduzione di un obbligo di interpretabilità degli algoritmi, che in settori fortemente regolati come quelli bancario e assicurativo risulterebbe in linea con gli obblighi di trasparenza dei processi decisionali imposti dal legislatore[29]. In alcuni ordinamenti, l'attenzione si è concentrata soprattutto sui rischi a cui un determinato modello espone il suo utilizzatore e sulla possibile mitigazione dei rischi stessi attraverso una declinazione dell'interpretabilità più o meno intensa a seconda di una serie di circostanze individuate[30]. Alcuni

28 Alcuni studi mostrano che l'interpretabilità influenza la percezione di accuratezza del sistema di IA da parte dei suoi utenti, a prescindere dall'accuratezza effettiva; si veda Nourani, M., Kabir, S., Mohseni, S. e Ragan, E.D. (2020), *The Effects of Meaningful and Meaningless Explanations on Trust and Perceived System Accuracy in Intelligent Systems*, http://www.aaai.org.

29 Si pensi, ad esempio, nell'ambito della vigilanza micro-prudenziale all'adempimento da parte delle banche degli obblighi di patrimonializzazione e alla tutela della *privacy* nell'ambito della disciplina dettata dal regolamento europeo in materia di decisioni di erogazione del credito o di pricing delle assicurazioni (GDPR).

30 In questa direzione si è mosso l'Information Commissioner's Office britannico, che ha indicato cinque criteri utili per comprendere il tipo di interpretabilità necessaria (dominio, impatto, dati usati, *urgency* e *audience*; si veda Explaining decisions made with AI, https://ico.org.uk/for-organisations/guide-to-data-protection/key-dp-themes/explaining-decisions-made-with-ai/).

commentatori invitano a riflettere sulle ragioni per le quali l'approccio all'intepretabilità dei modelli di IA debba essere diverso da quello adottato rispetto a modelli statistici usati in finanza da tempo, paventando il rischio di disparità di trattamento e di ripercussioni negative per l'innovazione (per approfondimenti sul tema anche con riguardo alla gestione di portafoglio si rimanda alla Sezione '*Profili di* investor protection').

Robustezza dei modelli di IA

Le prestazioni dei modelli di IA incontrano importanti limiti rispetto a eventi estremi o shock idiosincratici. Questi creano infatti discontinuità nei dati e nelle relazioni tra variabili economiche e finanziarie e rendono necessaria una ricalibrazione di algoritmi addestrati con serie storiche riferite a condizioni dei mercati 'normali'[31]. È sempre più avvertita l'esigenza di aumentare la robustezza dei modelli, ad esempio utilizzando per l'addestramento i già ricordati set sintetici, opportunamente costruiti, oppure sviluppando modelli basati sull'apprendimento rinforzato (che impara generando e contrassegnando i dati durante l'addestramento con una strategia *trial and error*, invece di riceverli prima).

I modelli di IA, inoltre, richiedono nel continuo un'attività di verifica e validazione per intercettare cambiamenti nelle proprietà statistiche dei dati (*data drift*) dovute al sopravvenire di nuove condizioni che possono riguardare, ad esempio, modifiche nelle preferenze e nei comportamenti degli individui o novità legislative.

I modelli di ML, poi, possono presentare problemi di *over-fitting*, ossia essere eccessivamente adattati ai dati di *training* tanto da non essere generalizzabili ad altri *data set*. Diversi accorgimenti nella fase di test e validazione possono mitigare questa eventualità (si veda l'Appendice '*I sistemi di intelligenza artificiale*'); in alcuni ordinamenti è inoltre richiesta la predisposizione di accorgimenti automatici che segnalino situazioni di criticità (sino ai cosiddetti *kill switches*, che interrompono le funzioni di un modello se necessario).

Un altro limite è legato alla presenza di regressioni spurie, in cui i risultati sono influenzati dalla compresenza di trend nelle serie utilizzate nelle analisi di regressione. In questa circostanza il modello può segnalare relazioni di causalità di fattori inesistenti che solo il controllo umano può individuare ed escludere.

In generale, gli algoritmi basati sulle tecnologie di ML non possono sostituire totalmente l'intuito e l'esperienza umani, come mostrano anche alcuni esempi riferiti a diversi episodi (a partire dalla crisi finanziaria del 2008), quando seguire le indicazioni dei modelli avrebbe significato esacerbare gli andamenti avversi dei mercati[32].

31 È quanto avvenuto, ad esempio, in occasione dell'emergenza sanitaria da Covid-19. Una survey condotta nel Regno Unito mostra che circa il 35% delle banche ha sperimentato un peggioramento delle *performance* dei modelli di ML durante la pandemia (Bholat, D., Gharbawi, M. e Thew, O. (2020), *The impact of Covid on machine learning and data science in UK banking*, Bank of England, Bank of England Quarterly Bulletin Q4 2020, https://www.bankofengland.co.uk/quarterly-bulletin/2020/2020-q4/the-impact-of-covid-onmachine-learning-and-data-science-in-uk-banking).

32 Guidolin et al., *op. cit.* (2021) riportano le risultanze di alcuni studi che confrontano le strategie seguite dai gestori nel 2008-2010 negli Stati Uniti (crisi *subprime*) e in Europa (crisi del debito sovrano) con le decisioni simulate che avrebbero preso sistemi di *trading* basati sull'IA e sul *sentiment* dai *social media*.

Quest'ultima considerazione rimanda anche alla necessità che lo sviluppo di sistemi di IA sia affidato a un *team* con conoscenze diversificate, in grado di affiancare gli esperti informatici e della scienza dei dati che, pur essendo essenziali per la progettazione e lo sviluppo di tali modelli, possono non avere una conoscenza specifica del settore finanziario.

Impatti su concorrenza e cybersecurity

Un altro tema all'attenzione del dibattito concerne, come già ricordato, la concentrazione di dati e di modelli di IA nelle mani di pochi *providers* specializzati (tipicamente Big Tech). Tale concentrazione si riflette nella creazione di posizioni dominanti che generano effetti distorsivi sulla concorrenza[33], limitazioni all'accessibilità ai dati e alle nuove tecnologie da parte degli operatori più piccoli, impatti avversi sulla stabilità dei mercati nella misura in cui l'affermarsi di pochi operatori di dimensioni sistemiche può alimentare fenomeni di *herding*[34].

La convergenza verso l'utilizzo di dati e modelli simili potrebbe inoltre aumentare il rischio di violazioni della sicurezza cibernetica, poiché diventa più facile sferrare attacchi informatici su soggetti che agiscono nello stesso modo[35]. Inoltre, in caso di utilizzo di algoritmi che imitano la strategia di altri operatori, un attacco informatico a uno di questi potrebbe compromettere a catena gli eventuali sistemi connessi, generando impatti sistemici[36].

I *cyber attacks* potenzialmente in grado di compromettere la gestione di portafoglio automatizzata possono essere molteplici[37]. L'attività illecita potrebbe essere

33 Il potere dei cosiddetti TechFins può rappresentare una forma di concorrenza sleale. Poiché queste società operano in un mercato non regolamentato, sono soggette a minori requisiti rispetto agli intermediari finanziari tradizionali e non sostengono i costi (ad esempio di licenza, registrazione e *compliance*) sostenuti dagli *incumbents*. Questi ultimi vedono colpita anche la loro capacità di reagire alle evoluzioni di mercato e, quindi, la loro adattabilità ai cambiamenti (sul tema si rinvia a Buckley, R.P, Arner, D.W., Veidt, R. e Zetzsche, D.A. (2019), *Building FinTech Ecosystems: Regulatory Sandboxes, Innovation Hubs and Beyond*, University of Hong Kong Faculty of Law Research Paper n. 2019/100). L'emergere di nuovi operatori di mercato, che operano al di fuori del perimetro regolamentare, è un fenomeno complesso per il legislatore e per le autorità di vigilanza, che non possono esercitare i controlli necessari per salvaguardare la stabilità del mercato finanziario e proteggere i partecipanti più deboli. Allo stesso tempo, le caratteristiche di flessibilità e di evoluzione costante dei nuovi operatori rendono estremamente difficile l'applicazione di norme di *hard law*, che si fondano su definizioni e categorie rigide. Si veda Zetzsche, D.A. (2012), *Investment Law as Financial Law: From Fund Governance over Market Governance to Stakeholder Governance?* in Birkmose, H.S., Neville, M. ed Engsig Sørensen, K. (a cura di), *The European Financial Market in Transition* (Kluwer Law International), 339-343.

34 Un altro elemento di criticità riguarda la caratteristica degli algoritmi di ML riferibile a una certa 'autonomia' di comportamento, in virtù del processo di auto-apprendimento e di conseguente perfezionamento, che potrebbe portare anche a fenomeni di collusione tacita (ad esempio, nell'ambito di modelli di *pricing* dinamico). È stato argomentato, in particolare, che i modelli riconoscono le interdipendenze reciproche e possono adattarsi alle azioni di altri partecipanti al mercato o altri modelli di IA, portando a una collusione anche 'inconsapevole' tra gli operatori che gli utilizzano.

35 In merito si rinvia a Linciano, N. e Soccorso, P. (2017), *FinTech e RegTech: approcci di regolamentazione e di supervisione*, in Paracampo, M.T. (a cura di), *Fintech. Introduzione ai profili giuridici di un mercato unico tecnologico dei servizi finanziari*, Torino, 32; Financial Stability Board, *op. cit.* (2017), 29 e ss.; BaFin, *op. cit.* (2018), 157.

36 Tale circostanza potrebbe realizzarsi anche qualora l'attacco informatico fosse lanciato a sistemi di *cloud computing*, tipicamente in mano a pochi *outsourcers*, per interromperne l'operatività. ACPR (2020), Governance of Artificial Intelligence in Finance, https://acpr.banque-france.fr/en/governance-artificial-intelligence-finance.

37 Su possibili esempi di *cyber attacks*, si rinvia a Tina, A. (2020), *Cybersicurezza: integrità dei processi e dei dati*, in Cian, M. e Sandei, C. (a cura di), Diritto del FinTech, Milano, 16.

volta ad 'avvelenare' la qualità dei dati utilizzati dagli algoritmi manipolandoli o distruggendoli: è questo ad esempio il caso di un attacco nei confronti di una piattaforma di *cloud-computing*[38] che potrebbe pregiudicare di riflesso l'attività dei gestori che esternalizzano l'acquisizione e l'archiviazione dei dati usati dall'algoritmo di gestione. Una diversa ipotesi, ugualmente idonea a compromettere la corretta prestazione del servizio, potrebbe verificarsi nel caso di intrusioni finalizzate a colpire le infrastrutture e i *software* utilizzati nell'elaborazione e nell'analisi dei dati. Del resto, la vulnerabilità dei *software* è ontologica[39] e peraltro cresce con l'aumentare della complessità dei sistemi informatici utilizzati dai gestori.

Impatto aggregato dei sistemi di IA sulla gestione di portafoglio

Potrebbe essere opportuna una breve riflessione sull'impatto aggregato del ricorso ai sistemi di IA nel *portfolio management*. Nella misura in cui i sistemi di IA sfruttano database dai contenuti eterogenei, operano una selezione dello specifico subset informativo da utilizzare (*data cleaning*) nei processi di investimento e applicano metodologie e algoritmi auto-educanti e auto-costituenti, condizioni di partenza relativamente simili per un certo numero di intermediari potrebbero risultare in decisioni di output e scelte di investimento significativamente differenti.

A livello aggregato, segnali di investimento/disinvestimento dissimili influenzerebbero i meccanismi di *price formation* e potrebbero esercitare impatti aleatori sui mercati, eventualmente mitigandone la volatilità e rendendo più complessa l'interpretazione dei movimenti di mercato. L'eterogeneità di comportamenti individuali sarebbe, tuttavia, limitata in uno scenario caratterizzato dall'ampio ricorso all'*outsourcing*, laddove una limitata varietà di tecniche IA – rese disponibili da relativamente pochi provider – potrebbe uniformare e standardizzare l'azione collettiva dei diversi *players* di mercato (come già accennato anche nel precedente sottoparagrafo '*Interpretabilità dei modelli di IA*'). Potrebbe, pertanto, sussistere un duplice *trade-off* tra originalità dell'azione e probabilità di propagazione degli shock, scenari, rispettivamente, associati a sviluppo delle competenze interne e ricorso all'esternalizzazione.

In un simile contesto, non si pone soltanto il tema di vantaggio competitivo inteso quale obiettivo di processare l'informazione in modo tempestivo (rapidità di esecuzione che consente di anticipare l'azione del *competitor*), bensì assume maggior rilievo il vantaggio informativo inteso quale *modus operandi* che poggia sulla solidità e sull'affidabilità dei sistemi adottati (processare l'informazione nel modo corretto).

Impatti sulla privacy

La raccolta e l'uso di dati relativi a caratteristiche, preferenze e comportamenti degli individui nell'ambito di tecnologie che si avvalgono di *big data* pongono problemi di tutela della *privacy* ulteriori rispetto a quelli legati a raccolta e uso di dati personali secondo modalità 'tradizionali'.

38 Sulle opportunità e i rischi cibernetici per i sistemi di *cloudcomputing*, si rinvia a Lillà, M. e Cavallo, M.A. (2018), *Cybersecurity and Liability in a Big Data World*, in Market and Competition Law Review, 77 ss..

39 Sul punto si veda Walch, A. (2015), *The Bitcoin Blockchain as Financial Market Infrastructure: a consideration of operational risk*, 18 NYU *Journal of Legislation and Public Policy*, 856 e ss., https://ssrn.com/abstract=2579482.

Un primo profilo di criticità è legato alla molteplicità di soggetti che possono acquisire i dati generati da un individuo senza che questi ne sia consapevole, come accade ad esempio nel caso di un utente di un sito internet i cui dati diventano disponibili sia allo sviluppatore del sito sia all'operatore di cui quest'ultimo si avvale per il tracciamento dell'attività dell'utente. Un ulteriore profilo di attenzione riguarda l'operatività dei cosiddetti *data broker*, che aggregano e organizzano dati rivenienti da diverse fonti per trasferirli a soggetti terzi, alimentando un mercato poco trasparente soprattutto per gli utenti finali, ignari dell'utilizzo dei propri dati acquisiti dai siti internet e dalle piattaforme *online* frequentati[40].

La combinazione di più fonti di dati tra loro eterogenee, oltre a potersi riflettere negativamente sulla qualità delle informazioni risultanti dall'aggregazione[41], aumenta i rischi di sicurezza informatica, a detrimento della tutela della riservatezza.

L'industria sta proponendo nuove soluzioni tecniche in grado di mitigare il rischio di violazione della *privacy* e sta sperimentando la generazione e l'uso dei già ricordati *data set* sintetici, assemblati ai fini dell'addestramento di un modello di ML senza che ciò comporti una significativa riduzione delle prestazioni.

La mitigazione del rischio di violazioni della *privacy* può trarre beneficio anche da adeguate regole di *data governance*, sebbene in tale ambito un limite significativo derivi dalla frammentazione delle norme in materia di gestione e controllo dei dati, in ambito sia domestico sia internazionale (per approfondimeni si rimanda alla Sezione '*Profili di* investor protection').

3 Il dibattito istituzionale in corso

Le applicazioni dell'IA sono all'attenzione crescente delle istituzioni internazionali. L'OECD ha pubblicato cinque principi, che identificano gli interessi la cui tutela è imprescindibile[42], e altrettante raccomandazioni destinate ai governi, per favorire lo sviluppo di sistemi di IA innovativi e affidabili nel rispetto dei diritti umani e dei principi democratici[43]. L'applicazione delle raccomandazioni viene promossa da un Osservatorio, istituito per supportare le singole giurisdizioni.

40 AGCM, AGCOM e Garante per la protezione dei dati personali (2019), Indagine conoscitiva sui big data, https://www.agcm.it/dotcmsdoc/allegati-news/IC_Big%20data_imp.pdf.

41 Bareinboim, E. e Pearla, J. (2016), *Causal inference and the data-fusion problem*, Proceedings of the National Academy of Sciences of the United States of America, Vol. 113/27, pp. 7345-7352, http://dx.doi.org/10.1073/pnas.1510507113.

42 In particolare, i principi statuiscono che: l'IA deve favorire crescita inclusiva, sviluppo sostenibile e benessere; i sistemi di IA devono rispettare i diritti umani, i valori democratici e l'equità; devono essere soddisfatti trasparenza e interpretabilità, affinché le persone capiscano i risultati basati sull'intelligenza artificiale; i sistemi di IA devono essere robusti e sicuri lungo il loro ciclo di vita e i potenziali rischi dovranno essere continuamente valutati e gestiti; organizzazioni e individui che sviluppano, distribuiscono o gestiscono sistemi di IA sono responsabili per il loro corretto funzionamento (OECD, 2019; https://oecd.ai/en/ai-principles).

43 In particolare, le raccomandazioni esortano i governi a: facilitare gli investimenti pubblici e privati nella ricerca e nello sviluppo per stimolare un'innovazione affidabile; promuovere ecosistemi di IA accessibili con infrastrutture digitali e meccanismi per condividere dati e conoscenze; creare un quadro politico in grado di favorire la diffusione di tali tecnologie; fornire alle persone le competenze per beneficiare dei vantaggi che derivano dall'utilizzo dell'IA e sostenere

L'OECD propone inoltre un approccio di *policy* per l'identificazione e la mitigazione delle criticità riferibili alle applicazioni dell'IA, declinabile in funzione del ciclo di vita del modello e della classificazione del modelli di IA, che può trovare applicazione anche nell'ambito dei modelli di IA usati in finanza[44].

Gli approcci nazionali alle applicazioni di IA al settore finanziario sono ad oggi eterogene, anche quanto a fase di sviluppo (Tav. 3).

Tav. 3 – Approcci nazionali alla regolamentazione delle applicazioni di IA nel settore finanziario

Principali applicazioni dell'IA nell'ambito dell'attività delle autorità di regolamentazione del settore finanziario
Mappare e raccogliere informazioni sull'utilizzo dell'IA nei servizi finanziari
Offrire orientamenti per l'utilizzo dell'IA nei servizi finanziari
Istituire *sandboxes* regolamentari
Regolamentare specifiche applicazioni ad alto rischio dell'IA nel settore finanziario
Utilizzare l'IA per l'attività di vigilanza e regolamentazione del settore finanziario

Fonte: OECD, *op. cit.* (2021).

In particolare, molte giurisdizioni si sono dotate di strategie dedicate all'IA, mentre rimane ancora contenuto il numero di Paesi che ha definito una disciplina dettagliata in materia[45]. Nella maggior parte dei casi, infatti, regolatori e autorità di controllo hanno dettato principi generali, ispirati all'approccio di neutralità tecnologica, riferiti soprattutto alla validazione degli algoritmi prima della loro distribuzione/utilizzo e al controllo continuativo delle relative prestazioni durante tutto il loro ciclo di vita.

Il principio di neutralità tecnologica può essere messo alla prova dalla crescente complessità degli algoritmi (e in particolare dai modelli di *deep learning*), dalle modalità e dagli ambiti di applicazione alla finanza e dalla crescente pervasività di utilizzo[46].

Inoltre, come si è già ricordato, alcune tecniche avanzate di IA potrebbero essere incompatibili con previsioni normative esistenti. Si pensi ad esempio alla non

i lavoratori al fine di garantire una transizione equa; cooperare tra giurisdizioni e settori per la condivisione di informazioni, lo sviluppo di standard e di una gestione responsabile (https://oecd.ai/en/ai-principles).

44 In particolare, il ciclo di vita comprende sei fasi: progettazione; raccolta e analisi dei dati; definizione del modello e interpretazione; verifica e validazione; sviluppo; applicazione e controllo. La classificazione tiene conto di una serie di fattori, ossia il contesto in cui il sistema di IA è applicato; i dati e gli input usati dal sistema; il modello di IA; l'obiettivo assegnato al sistema di IA, che a sua volta genera un impatto sul contesto. L'intersezione tra questi due profili (fase del ciclo di vita e fattore di classificazione) si associa a criticità specifiche. Nell'ambito della finanza, un esempio di applicazione di questo approccio può essere rappresentato dalle applicazioni di *credit scoring* nel settore finanziario (contesto) che raccolgono la storia dei pagamenti e altri dati personali (dati/input) al fine di elaborare una raccomandazione (compito/output) utilizzando una rete neurale per determinare se una persona ha la probabilità di non onorare un prestito. Un altro esempio è riferibile a un sistema di trading che sulla base delle preferenze dell'utente e i dati di mercato (dati/input) raccomanda ed eventualmente esegue ordini di acquisto e vendite (task/output) usando l'apprendimento automatico. Questi esempi presentano criticità specifiche, da trattare come tali. Si veda OECD, *op. cit.* (2021).

45 OECD (2019b), *Business and Finance Outlook: Strengthening Trust in Business*, OECD Publishing, Paris, https://doi.org/10.1787/af784794-en.

46 Gensler, G. e Bailey, L. (2020), *Deep Learning and Financial Stability*, http://dx.doi.org/10.2139/ssrn.3723132.

interpretabilità di alcuni modelli e ai requisiti di trasparenza, oppure alle disposizioni in materia di *privacy* che incidono su raccolta e gestione dei dati confliggendo con esigenze di addestramento e revisione degli algoritmi[47]; infine, l'opacità di alcuni modelli potrebbe anche rendere complesso identificare e provare violazioni delle norme esistenti. Queste circostanze rendono dunque indispensabile riflettere sulla necessità di adeguare norme e prassi di vigilanza ai nuovi paradigmi introdotti dall'IA[48].

Chiarezza e omogeneità del quadro normativo di riferimento tra settori e tra giurisdizioni sono esigenze fortemente avvertire dagli attori di mercato, che indicano nell'arbitraggio regolamentare e nella concorrenza da parte di soggetti non regolati come le Big Tech importanti ostacoli all'innovazione. Né possono considerarsi risolutivi i numerosi principi, linee guida e *best practices* in materia di IA pubblicati sinora, date le difficoltà percepite di dare concretezza operativa a tali iniziative[49].

Nell'Unione europea, il legislatore comunitario ha proposto di recente nuove regole e azioni tese a definire un quadro giuridico sull'IA organico e un nuovo piano coordinato con gli Stati membri. Si fa riferimento, in particolare, alla proposta di Regolamento della Commissione europea del 21 aprile 2021, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica alcuni atti legislativi dell'Unione.

Tale iniziativa si affianca a due ulteriori proposte. La prima è rappresentata dal Piano coordinato di revisione dell'intelligenza artificiale 2021, che pone le basi affinché la Commissione e gli Stati membri collaborino nell'attuazione di azioni congiunte ed eliminino la frammentazione dei programmi di finanziamento, delle iniziative e delle azioni intraprese a livello dell'UE e dei singoli Stati membri. La seconda concerne il Regolamento del Parlamento europeo e del Consiglio sui prodotti macchine, che dovrebbe sostituire la direttiva 2006/42/CE del 17 maggio 2006 relativa alle macchine, che garantisce la libera circolazione delle macchine all'interno del mercato UE e assicura un alto livello di protezione per gli utenti e altre persone esposte.

Queste proposte sono il frutto di un lungo dibattito sulla necessità di costruire un mercato UE dell'IA affidabile, sicuro e rispettoso dei diritti fondamentali, per la cui disamina si rimanda alla Sezione '*Profili di* investor protection'.

* * * * *

Con particolare riguardo alle applicazioni all'*asset management*, la IOSCO ha pubblicato a settembre 2021 un rapporto finale sull'uso dell'IA e del ML da parte di intermediari e gestori di portafoglio[50].

47 Il GDPR dell'UE per la protezione dei dati impone vincoli temporali nella conservazione dei dati individuali, ma le regole relative all'AI potrebbero richiedere alle aziende di tenere traccia dei set di dati utilizzati per addestrare gli algoritmi a fini di revisione.

48 ACPR, *op. cit.* (2020).

49 Si veda Bank of England e FCA (2020), *Minutes: Artificial Intelligence Public-Private Forum-First meeting*, https://www.bankofengland.co.uk/minutes/2020/artificial-intelligence-public-private-forum-minutes. Un ultimo ma non meno fondamentale punto, che riguarda trasversalmente operatori di mercato, legislatori e autorità di vigilanza, concerne lo sviluppo di conoscenze e competenze diversificate e multidisciplinari, secondo un approccio innovativo e in linea con la complessità dei sistemi di IA e delle relative applicazioni alla finanza.

50 Già nel 2019 il Board IOSCO aveva individuato IA e ML come priorità.

Il rapporto IOSCO intende fornire alle autorità di regolamentazione indicazioni per fronteggiare i potenziali rischi che l'utilizzo delle nuove tecnologie potrebbe creare per l'efficienza del mercato finanziario e la tutela dei risparmiatori. In particolare, la IOSCO suggerisce ai regolatori sei aree cui dedicare particolare attenzione, con riferimento a standard di condotta che gli operatori dovrebbero rispettare nell'uso dell'IA e del ML.

I regolatori dovrebbero richiedere a intermediari e gestori di dotarsi di procedure di *governance*, controlli e supervisione chiari e adeguati in ragione dello sviluppo, delle fasi di test, dell'utilizzo e del monitoraggio dei risultati rilasciati dai sistemi di IA e ML prevedendo la responsabilità, la supervisione e il coinvolgimento diretto dell'alta dirigenza con competenze e conoscenze adeguate in materia. Ciò richiede un quadro di *governance* interna documentato, con chiare linee di responsabilità e con la designazione da parte dell'alta dirigenza di un individuo (o gruppi di individui) con competenze, conoscenze, abilità ed esperienza adeguate all'approvazione dell'applicazione iniziale e degli aggiornamenti successivi e sostanziali della tecnologia.

Un'altra importante area di attenzione riguarda lo sviluppo, il test e il monitoraggio continuo degli algoritmi. Il test dovrebbe essere condotto in un ambiente separato dall'ambiente reale e prima dell'applicazione effettiva per garantire che le tecnologie di IA e ML operino in modo conforme agli obblighi normativi previsti dalle società e reagiscano come previsto anche in condizioni di stress dei mercati.

Sono anche necessarie conoscenze, competenze ed esperienze adeguate in materia di sviluppo, test e supervisione delle tecnologie di IA e ML. A tale riguardo le funzioni di conformità e *risk management* dovrebbero essere in grado di comprendere pienamente gli algoritmi e i risultati prodotti dagli stessi e di condurre la *due diligence* su qualsiasi fornitore esterno.

In materia di esternalizzazione, si raccomanda un'adeguata e approfondita conoscenza delle relazioni instaurate con i fornitori esterni, incluso il grado di dipendenza dalle stesse, l'affidabilità e il livello di supervisione necessario. Per garantire una chiara attribuzione di responsabilità, le società dovrebbero disporre di accordi/contratti chiari su qualità del servizio prestato (misurata anche tramite indicatori), ambito delle funzioni esternalizzate e responsabilità in capo al fornitore dei servizi.

La IOSCO affronta anche il tema della trasparenza sugli algoritmi a beneficio delle Autorità di regolamentazione, degli investitori e degli altri soggetti/parti interessate. Le Autorità di regolamentazione dovrebbero definire nel dettaglio la natura e la tipologia di informazioni che le società dovrebbero fornire, anche al fine di garantire un'adeguata supervisione sulle stesse, prestando altresì particolare attenzione a quelle che influiscono in misura maggiore sugli interessi dei clienti.

Per presidiare la qualità dei dati, si raccomandano adeguati controlli diretti sui *data set*, al fine di prevenire distorsioni e consentire applicazioni robuste e affidabili.

Con riguardo agli aspetti etici, il rapporto IOSCO pone l'accento sulla definizione di criteri volti a ridurre al minimo il rischio di risultati discriminatori. In particolare, secondo la IOSCO i partecipanti al mercato dovrebbero controllare attentamente

lo sviluppo di tecnologie IA e ML che utilizzano ampi *data set* alternativi come dati satellitari o *feeds* di Twitter, al fine di garantire che i modelli sviluppati non discriminino un determinato segmento della popolazione e che le decisioni guidate dall'IA siano eque e imparziali[51].

La cybersecurity

Negli ultimi anni, il tema della *cybersecurity* ha ricevuto crescente attenzione nel dibattito istituzionale, poiché rappresenta un elemento essenziale dell'affidabilità di qualsiasi sistema software e pertanto anche dei sistemi di IA.

Data la natura potenzialmente transnazionale del fenomeno, il tema della *cybersecurity* è stato oggetto di un report IOSCO[52] che ha sottolineato la necessità di un coordinamento internazionale tra regolatori di settore per mitigare il rischio cibernetico.

In ambito europeo, il Piano d'azione sul Fintech[53] ha sin da subito individuato nella sicurezza tecnologica e informatica un presidio da preservare e potenziare per la fiducia e la stabilità del sistema finanziario. La Commissione ha in tale circostanza evidenziato l'importanza per i servizi digitali di «*adottare un approccio basato sulla 'sicurezza sin dalla progettazione' (Punto 3)*».

Il sempre maggior numero di potenziali «*minacce informatiche*», definite come «*qualsiasi circostanza, evento o azione*»[54] in grado di pregiudicare il corretto funzionamento delle tecniche di automazione determinando danni agli investitori, impone quindi la necessità di predisporre misure funzionali a prevenire e mitigare i rischi che ne derivano.

Dal punto di vista regolamentare, la direttiva 2016/1148/UE (cosiddetta NIS)[55], in considerazione del ruolo vitale svolto dalle reti, dai sistemi e servizi informativi nella società e della necessità che essi siano «*affidabili e sicuri per le attività economiche e sociali e in particolare ai fini del funzionamento del mercato int*erno», nel tentativo di offrire una maggiore protezione dai rischi e dagli «*incidenti a carico della*

51 A tal fine il Fintech Network di IOSCO ha identificato cinque temi principali che potrebbero essere alla base dell'uso etico delle tecniche di IA e ML: garantire che il modello utilizzato agisca in buona fede ossia nel migliore interesse degli investitori e preservando l'integrità del mercato; comprendere e interpretare le decisioni basate su IA e ML per identificare condotte errate; prevedere l'intervento umano su ciò che il modello può e non può decidere; responsabilizzare la dirigenza con riguardo alle azioni dei modelli applicati; garantire la spiegabilità/comprensibilità dei risultati originati dai modelli utilizzati.

52 IOSCO (2016), *Cyber Security in Securities Markets – An International Perspective*, www.iosco.org.

53 Comunicazione della Commissione Europea COM(2018) 109/final dell'8 marzo 2018, Piano d'azione per le tecnologie finanziarie: per un settore finanziario europeo più competitivo e innovativo. Tra le iniziative legislative che costituiscono il Piano si ricorda la proposta di Regolamento in materia di resilienza operativa digitale per il settore finanziario (cosiddetto DORA – Digital Resilience Operation Act).

54 Si veda l'art. 2, n. 8, Regolamento 2019/881/UE che definisce le minacce informatiche come «*qualsiasi circostanza, evento o azione che potrebbe danneggiare, perturbare o avere un impatto negativo di altro tipo sulla rete e sui sistemi informativi, sugli utenti di tali sistemi e altre persone*».

55 L'art. 4 della Direttiva NIS definisce la «*sicurezza della rete e dei sistemi informativi*» come la «*capacità di una rete e dei sistemi informativi di resistere, a un determinato livello di riservatezza, a ogni azione che comprometta la disponibilità, l'autenticità, l'integrità o la riservatezza dei dati conservati o trasmessi o trattati e dei relativi servizi offerti o accessibili tramite tale rete o sistemi informativi*».

sicurezza»[56], ha previsto una serie di obblighi in tema di *cybersecurity* per tutti gli operatori, anche nei servizi finanziari[57]. A tal proposito si rammenta che nella medesima prospettiva muove la proposta di Regolamento DORA (Digital operational resilience for the financial sector), del 24 settembre 2020 e attualmente in fase di trilogo, che ha l'obiettivo di innalzare gli standard di sicurezza delle infrastrutture e il monitoraggio dei fornitori ICT operanti nel settore finanziario.

Il Gruppo indipendente di esperti ad alto livello sull'IA istituto dalla Commissione Europea, ha evidenziato che la «*robustezza tecnica*» e la sicurezza costituiscono una «*componente cruciale per ottenere un'IA affidabile*» essendo strettamente connesse alla prevenzione dei danni[58]. In tal senso, è fondamentale la resilienza ai possibili attacchi informatici e la sicurezza, poiché «*i sistemi di IA, come tutti i sistemi software, dovrebbero essere protetti contro le vulnerabilità che li espongono allo sfruttamento da parte degli avversari, ad esempio l'hacking*»[59]. Per tali ragioni, il Gruppo suggerisce l'adozione di un «*piano di emergenza e sicurezza generale*» per chi utilizza i sistemi di IA, che preveda «*misure di salvaguardia che attivino un piano di emergenza in caso di problemi*». In particolare, i sistemi di IA «*potrebbero quindi passare da una procedura statistica a una procedura basata su regole, oppure richiedere un operatore umano prima di continuare la loro azione*»[60]. La possibilità di intervento umano in tal caso può fungere da effettivo presidio operativo idoneo a prevenire potenziali danni in caso di malfunzionamento del sistema (per ulteriori dettagli sui contenuti degli Orientamenti etici per un'IA affidabile si veda la successiva Appendice '*Intelligenza artificiale e tutela della persona*').

Alla luce di quanto sopra, ben si comprende l'importanza strategica che riveste la sicurezza informatica nell'ottica della vigilanza[61] in chiave di tutela degli investitori e delle stesse istituzioni finanziarie.

L'attività di supervisione oltre che sugli aspetti tradizionali dovrà considerare anche la capacità di resilienza dei sistemi informatici alle minacce informatiche, valutando se i presidi organizzativi previsti in caso di attacco cibernetico siano adeguati a proteggere i sistemi di intelligenza artificiale e i dati analizzati.

In merito, la Direttiva NIS, ispirandosi al principio di proporzionalità, impone agli operatori di servizi essenziali e ai fornitori di servizi digitali l'adozione di «*misure*

56 Si vedano i primi due considerando della direttiva (UE) 2016/1148.

57 La direttiva è stata recepita nell'ordinamento italiano con il d.lgs. n. 65 del 18 maggio 2018. Per una completa ricostruzione in ambito domestico degli interventi normativi in materia si rinvia a Bassini, M. (2017), *Cybersecurity*, in Paracampo, M.T. (a cura di), *op. cit.*, 247 e ss..

58 Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale (2018), *Orientamenti etici per un'IA affidabile,* punto 66; www.ec.europa.eu.

59 Così il punto 67, degli *Orientamenti etici* (2018), che prosegue sottolineando come «*gli attacchi possono colpire i dati (avvelenamento dei dati), il modello (model leakage) o l'infrastruttura sottostante, sia software che hardware. Se il sistema di IA è attaccato, ad esempio con input ingannevoli (adversarial attack), possono essere modificati sia i dati che il comportamento del sistema, il quale può essere indotto ad adottare decisioni diverse o ad arrestarsi completamente*».

60 Così il punto 67, degli Orientamenti etici (2018).

61 Sul punto, si rinvia a CONSOB (2019), La digitalizzazione della consulenza in materia di investimenti finanziari, in Quaderni FinTech, 3, www.consob.it, 99 e BaFin, *op. cit.* (2018), 60.

tecniche e organizzative adeguate e proporzionate alla gestione dei rischi posti alla sicurezza della rete e dei sistemi informativi» che, tenendo conto «*delle più aggiornate conoscenze in materia*», assicurano «*un livello di sicurezza della rete e dei sistemi informativi adeguato al rischio esistente*» e «*la continuità del servizio*»[62].

Tali obblighi invero non costituiscono una novità per il settore finanziario, dove la sicurezza dei dati rientra fra gli specifici requisiti organizzativi disciplinati dal Regolamento delegato MiFID II, secondo il quale, nello specifico, le imprese di investimento devono istituire, applicare e mantenere «*procedure e sistemi idonei a tutelare la sicurezza, l'integrità e la riservatezza delle informazioni (art. 21, par. 2)*».

Con riferimento alla gestione dei problemi connessi all'interruzione dell'operatività dovuta a minacce informatiche, è altresì previsto per le imprese la necessità di dotarsi «*un'idonea politica di continuità dell'attività che consenta loro*» di garantire la continuità del servizio[63].

Per gli intermediari automatizzati, infatti, *a fortiori* il *cyber risk* costituisce uno degli elementi identificativi del rischio operativo[64] che deve essere mitigato e gestito primariamente attraverso procedure adottate a livello organizzativo.

62 Si vedano l'art. 12, cc. 1 e 2, e l'art. 14, cc 1 e 3, del d. lgs. 65/2018 di recepimento della Direttiva NIS.

63 Si veda l'art. 21, par. 3, Reg. 565/2017/UE (cosiddetto Regolamento delegato MiFID II).

64 Sul punto si veda il tredicesimo considerando della Direttiva NIS secondo cui «*il rischio operativo rappresenta un elemento cruciale della regolamentazione e vigilanza prudenziali nel settore bancario e in quello delle infrastrutture dei mercati finanziari. Copre tutte le operazioni comprese la sicurezza, l'integrità e la resilienza delle reti e dei sistemi informativi*».

Appendice

P. Brandimarte (*)

I sistemi di intelligenza artificiale

Classi di algoritmi e metodi di apprendimento

L'apprendimento supervisionato

A questa classe appartengono i classici metodi di regressione lineare, che risalgono al 1800, nei quali si cerca di legare una variabile Y a una variabile X, secondo un modello lineare:

$$Y=\alpha+\beta\_1\ X\_1+\beta\_2\ X\_2+\cdots+\beta\_m\ X\_m$$

Nei modelli di questo tipo, ben familiari a chi si occupa di finanza, α e β hanno un significato preciso; essi sono alla base anche di modelli econometrici piuttosto sofisticati. In econometria Y è la variabile regredita e X sono i regressori (o variabili esplicative); nel *machine learning* si usano, rispettivamente, i termini equivalenti *target* e *feature*.

Questi modelli sono ampiamente utilizzati nel contesto di *predictive analytics*, e sono estesi in diverse direzioni, che comprendono l'uso di *features* non lineari che coinvolgono diverse *features* elementari. Inoltre, la variabile *target*, come le *features*, non è necessariamente numerica, ma può essere categorica e qualitativa. Si pensi, ad esempio, ai sistemi di *fraud detection* per transazioni su carta di credito. In questo caso si parla di classificatori.

Nel seguito, si farà riferimento ad esempi di classificazione binaria, per semplicità di visualizzazione. Tuttavia, molti dei metodi accennati si applicano a più classi e a casi in cui il *target* è quantitativo.

Classificatori lineari

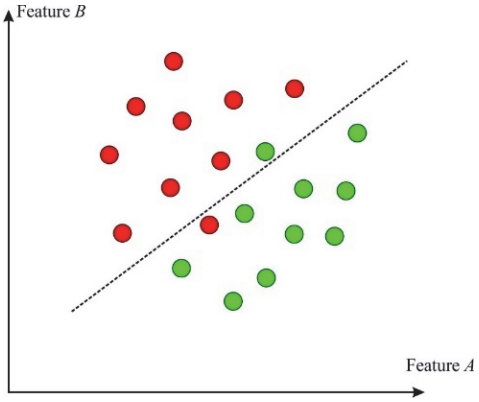
Un semplice classificatore lineare si basa su una funzione del tipo $Y=\alpha+\beta\_1\ X\_1+\beta\_2\ X\_2$. Quando essa assume un valore superiore a una certa soglia, l'oggetto è classificato in una classe (ad esempio come rosso), altrimenti nell'altra classe (ad esempio, come verde). Si tratta quindi di trovare opportuni coefficienti della funzione lineare, oltre ai valori di soglia, in modo da minimizzare gli errori di classificazione.

Un esempio di classificatore lineare è illustrato nella figura seguente, dove la linea tratteggiata rappresenta una linea di demarcazione tra gli elementi di due classi (Fig. 3).

(*) Paolo Brandimarte, professore ordinario Politecnico di Torino.

Ovviamente, non è sempre possibile separare linearmente due o più classi, e possono emergere errori di classificazione, come i due esempi riportati in figura. Essi sono simili ai noti falsi positivi e falsi positivi in medicina.

Fig. 3 – Esempio di classificatore lineare

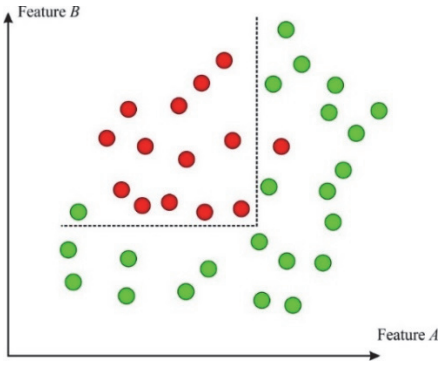


A volte non è sufficiente una classificazione, ma occorre la probabilità che un elemento appartenga a una classe (ad esempio, si vuole stimare la forza di un segnale di trading). A questo scopo, sono stati proposti modelli di regressione logistica, in cui la variabile *target* in uscita da un modello di regressione lineare è forzata ad assumere valori compresi tra 0 e 1.

Alberi di classificazione e random forest

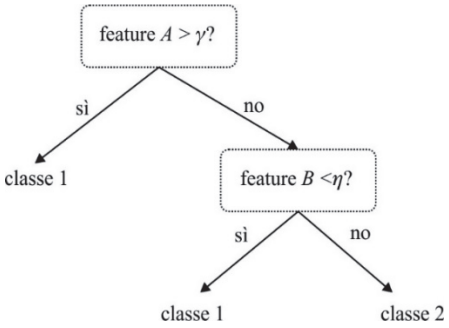
Vi sono situazioni in cui non è possibile separare classi in modo semplicemente lineare. Parimenti, nel caso di *target* quantitativi, vi possono essere interazioni non-lineari tra variabili che è difficile catturare. Un esempio è illustrato nella figura seguente (Fig. 4).

Fig. 4 – Esempio di classificatore non lineare



In questo caso, una possibilità è quella di ricorrere ad alberi di classificazione o regressione, come ad esempio nella figura seguente (Fig. 5).

Fig. 5 – Alberi di classificazione o regressione

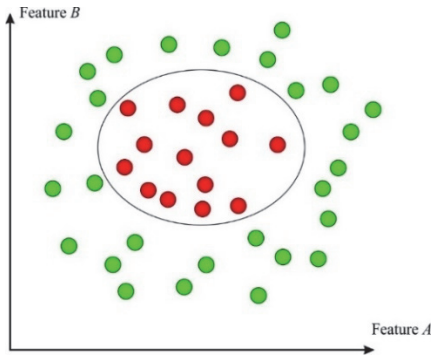


Tale metodo si basa su '*splitting features*', ovvero sulla classificazione sequenziale sulla base di valori di soglia per diverse *features*. Un semplice albero di classificazione fornisce spesso prestazioni non completamente soddisfacenti, ma si possono costruire diversi alberi e utilizzare le diverse previsioni per raggiungere un consenso. Questo approccio è noto come *random forest*.

Support vector machines

Le *support vector machines* (SVM) sono classificatori mirati a eliminare alcuni potenziali problemi della regressione logistica e dei classificatori lineari, ovvero la non separabilità lineare e la possibile fragilità della regola di classificazione a fronte di nuovi esempi (Fig. 6).

Fig. 6 – Support vector machine



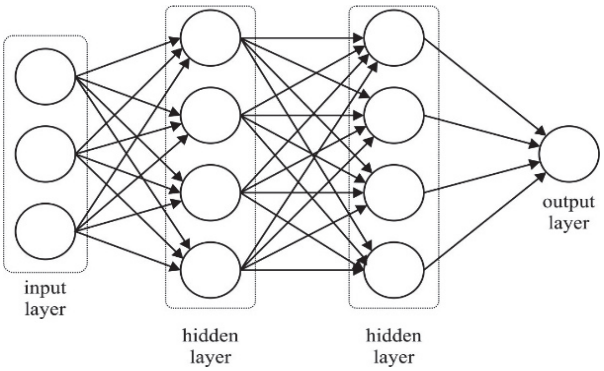
La figura qui sopra illustra come, in casi difficili, si possono ricavare *features* diverse (in questo caso la distanza da un centro) che permettono di separare le classi. Le SVM, a differenza della regressione logistica, non si basano sull'estensione di classici modelli statistici (come la regressione lineare), bensì sull'applicazioni di modelli di ottimizzazione.

Reti neurali

Un problema comune ai modelli di regressione con *feature* non lineare e alle SVM è la difficoltà nel trovare le *features* che permettono di ottenere prestazioni soddisfacenti. In molti casi, la conoscenza dello specifico dominio applicativo fornisce idee utili. Tuttavia, può valere la pena di immaginare approcci che siano in grado di ricavare automaticamente le *features* utili.

Le reti neurali si basano su un'architettura descritta nella figura riportata di seguito (Fig. 7). I dati in ingresso alimentano 'neuroni' che applicano una trasformazione non lineare. L'uscita di ogni neurone, moltiplicata per opportuni pesi, diventa l'ingresso di altri neuroni, disposti in strati (*layer* successivi), fino ad alimentare il *layer* di uscita.

Fig. 7 – Reti neurali



Gli strati interni sono detti *hidden* (nascosti) e hanno lo scopo di costruire, mediante processi di apprendimento supervisionato, le *features* necessarie. Si parla di *deep learning*, sia perché la rete ha diversi strati sia perché si creano *features* non esplicite nei dati iniziali.

Va notato come il termine *deep learning* sia fuorviante, in quanto la rete non crea davvero concetti nuovi, ma semplicemente impara a ricreare un output. Tuttavia, l'approccio è potente e ha dimostrato il suo valore specialmente in problemi di riconoscimento di immagini.

Over-fitting *e validazione out of sample*

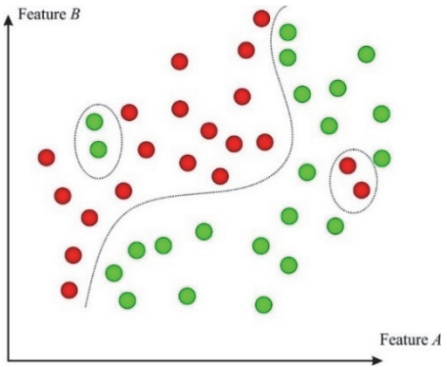
Qualsiasi metodo di apprendimento supervisionato si decida di utilizzare, si pone una questione rilevante: cosa garantisce che il metodo sia in grado di generalizzare, ovvero di 'prevedere' correttamente al di fuori dell'insieme di esempi su cui ha effettuato il *learning*? La questione è ben nota in finanza. Un consulente finanziario che tesse le lodi di una strategia di gestione, che avrebbe ottenuto ottimi risultati sullo scenario passato, dovrebbe evitare di ottimizzare tale strategia sulla base di esso.

Ci si riferisce al pericolo dell'*over-fitting*, illustrato nella figura seguente, in cui si costruisce una superficie di classificazione molto mirata all'insieme di esempi

usati per apprendere, ma che difficilmente si adatterà a dati esterni al campione (Fig. 8).

Tale rischio è tanto più forte quanto più il metodo è sofisticato. Per ovviare al problema, si usano approcci di validazione: non tutti i dati vengono usati nel *training sample*, ovvero il set di casi usati per apprendere; parte di essi (*validation* o *test sample*) viene riservata per una fase di validazione in cui si verifica la previsione per casi noti, ma non presentati all'algoritmo di apprendimento.

Fig. 8 – Over-fitting

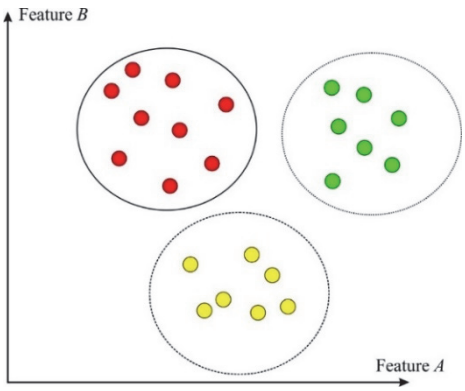


A parte il rischio di over-fitting, approcci di ML presentano il potenziale problema della non interpretabilità. Mentre i coefficienti di un semplice modello di regressione hanno un significato preciso, i pesi di una rete neurale non sono chiaramente interpretabili.

Apprendimento non supervisionato

In questo caso si cerca di raggruppare un insieme di oggetti in gruppi coesi. Tipico esempio sono i metodi di *clustering*, in cui non abbiamo etichette (valori *target*) ma solo oggetti caratterizzati da *features*, e che vogliamo raggruppare in *clusters* tali che la distanza tra oggetti del medesimo *cluster* sia minimizzata. L'idea è illustrata nella figura seguente (Fig. 9).

Fig. 9 – Metodi di clustering



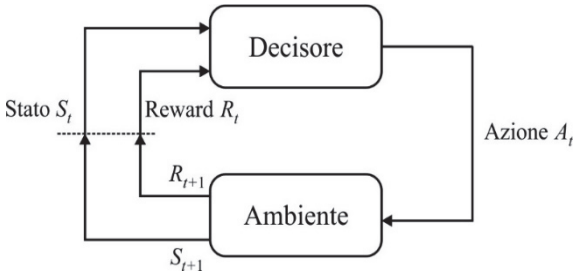
Le *features* utilizzate nel *clustering* possono essere complesse e riguardare anche segnali nel tempo, che si vogliano raggruppare per similitudine.

Reinforcement learning

Le strategie di *reinforcement learning* sono adatte a problemi di decisione dinamica (su un orizzonte di tempo) in condizioni di incertezza.

Il sistema (o ambiente) è caratterizzato da uno stato, che possiamo parzialmente influenzare con le nostre azioni. Lo stato può essere la ricchezza di un portafoglio e le azioni sono le decisioni di investimento. Alcuni stati sono sotto il nostro controllo, almeno parziale, altri no (come ad esempio i prezzi degli *assets* in un mercato liquido). Quando al tempo t, effettuiamo l'azione A_t nello statoS_t, otterremo un *reward* (profitto) R_(t+1) nell'istante di tempo successivo. Inoltre, lo stato del sistema verrà aggiornato a S_(t+1) e il ciclo si ripeterà (Fig. 9).

Fig. 10 – *Reinforcement learning*



Lo scopo è apprendere una politica che associa a ogni stato una decisione, in modo da massimizzare (in valore atteso) la somma attualizzata dei *rewards* futuri (introducendo un opportuno fattore di sconto).

Le strategie di *reinforcement learning* si applicano quando non si ha un modello affidabile dell'ambiente, per cui si può solo tentare di apprendere la politica applicando azioni e misurando le informazioni di *reward* e stato futuro. A differenza dei metodi di apprendimento supervisionato, non abbiamo un'informazione esplicita riguardo all'azione che dovremmo applicare in ciascuno stato ma abbiamo comunque un'informazione circa l'impatto delle nostre scelte. Pertanto, si può affermare che tali metodi si pongono a un livello intermedio rispetto agli altri due.

L'automazione nella gestione di portafoglio in Italia

1 Introduzione

Questa Sezione espone i risultati di una ricognizione (nel seguito anche *survey*), realizzata in collaborazione con Assogestioni, che ha coinvolto le principali società di gestione del risparmio (Sgr) operanti nel mercato italiano. L'indagine ha raccolto evidenze sul grado di utilizzo di sistemi di intelligenza artificiale (IA) nell'*asset management* in tutte le fasi della catena del valore, anche alla luce delle iniziative pianificate e in corso di sviluppo da parte dei gestori, nonché le opinioni dei gestori su benefici attesi e rischi potenziali legati all'utilizzo di sistemi di IA.

Le aree indagate nella *survey* riguardano i seguenti profili: gli obiettivi strategici legati allo sviluppo e all'utilizzo di sistemi di IA; l'uso di sistemi di IA e le tecnologie prevalenti; l'organizzazione e la *governance*; i benefici attesi e i rischi percepiti; le prospettive evolutive.

Ai fini della *survey*, per sistemi di intelligenza artificiale si intendono sistemi *software* (e se pertinente anche *hardware*) progettati da esseri umani che, dato un obiettivo complesso, agiscono attraverso l'acquisizione e l'interpretazione di dati strutturati o non strutturati, elaborando le informazioni derivate da tali dati e decidendo l'azione o le azioni migliori da intraprendere per raggiungere l'obiettivo dato. I sistemi di IA possono usare regole simboliche[65] o imparare un modello numerico nonché adattare il loro comportamento analizzando come il contesto è influenzato dalle loro azioni precedenti. Come disciplina scientifica, l'IA include diversi approcci e tecniche, come ad esempio l'apprendimento automatico (*machine learning* o ML)[66].

2 Caratteristiche delle società partecipanti alla *survey*

Le società partecipanti alla *survey* sono otto, rappresentative a livello di gruppo del 60% del patrimonio gestito a fine marzo 2022, riferito sia alle gestioni collettive sia alle gestioni di portafoglio su base individuale in Italia. Sette società sono autorizzate alla commercializzazione di Oicr, sei alla gestione su base individuale e alla

65 L'intelligenza artificiale simbolica è il termine che indica l'insieme dei sistemi di IA basati su rappresentazioni simboliche dei problemi, della logica e della ricerca. L'intelligenza artificiale simbolica utilizza strumenti come la programmazione logica, le regole di produzione, le reti semantiche e i *frames*.

66 La definizione di IA è ispirata a quella fornita dal HLEG sull'intelligenza artificiale della Commissione europea (https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines).

consulenza in materia di investimenti e solo due al servizio di ricezione e trasmissione di ordini.

La composizione del business risulta piuttosto variegata. La quota di masse gestite sotto forma di fondi UCITs varia, infatti, dal 2% al 95% del totale delle masse gestite a fine 2020 (28% in media). Conseguentemente anche la quota di gestioni di portafoglio su base individuale ha un intervallo di variazione molto ampio tra lo 0% e il 90% (64% in media). Le masse gestite in fondi pensione vanno dallo 0% al 13% del totale (4% in media), mentre quelle gestite in fondi di investimento alternativi (cosiddetti FIA) vanno dallo 0% al 10% del totale (4% in media)[67].

Tav. 4 – Distribuzione delle masse gestite a fine 2020 dalle Sgr partecipanti all'indagine

	media semplice	media ponderata	massimo	minimo
fondi di investimento UCITs	44%	28%	95%	2%
gestioni di portafoglio su base individuale	48%	64%	90%	0%
fondi di investimento alternativi (FIA)	3%	4%	10%	0%
fondi e gestioni pensionistiche	5%	4%	13%	0%

Fonte: elaborazioni CONSOB su dati di *survey*.

3 Gli obiettivi strategici legati allo sviluppo e all'uso di sistemi di IA

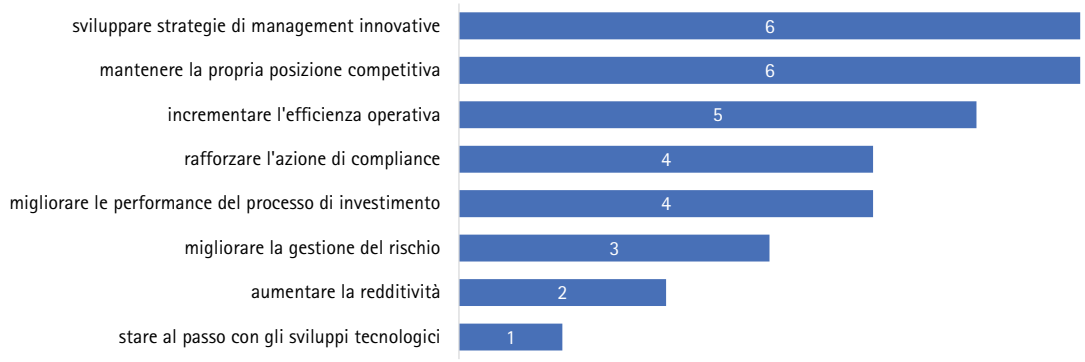
Per tutte le società partecipanti alla *survey* lo sviluppo di sistemi di IA rappresenta una priorità strategica o diventerà tale in un prossimo futuro. Nella quasi totalità dei casi, ricerca, sviluppo e diffusione di sistemi di IA sono parte di una generale strategia di innovazione tecnologica; solo in un caso è stata definita una strategia dedicata.

L'importanza dello sviluppo di sistemi di IA quale obiettivo strategico si poggia su diverse motivazioni, tra cui le più frequentemente indicate riguardano: lo sviluppo di strategie di gestione innovative; il mantenimento della propria posizione competitiva; l'incremento dell'efficienza operativa. Anche il miglioramento delle *performance* dei processi di investimento e il rafforzamento dell'azione di *compliance* risultano motivazioni rilevanti, mentre sono ritenuti meno importanti il miglioramento della gestione dei rischi, l'aumento della redditività e la necessità di stare al passo con gli sviluppi tecnologici[68] (Fig. 11).

67 Le categorie di UCITs sono rappresentate principalmente da fondi obbligazionari (dall'8% al 41%) e flessibili (dal 7% al 53%) seguiti dai bilanciati (da 0% a 45%) e azionari (da 2% a 22%), mentre i fondi monetari sono gestiti solo da 4/7 Sgr con una quota del 83%, 6%, 3% e 2% rispettivamente (una Sgr non ha fornito i dati). Le categorie di FIA sono residuali in termini di masse gestite (presenti in 6 su 8 Sgr) e in due casi rappresentate da fondi *real estate*, in un caso da fondi di *private equity*, in tre casi da altre tipologie di fondi e in un caso da *real estate* (52%), *private equity* (20%) e altre tipologie (28%).

68 Delle due società per le quali l'IA diventerà una priorità strategica in futuro, una ha indicato anche le relative motivazioni in limiti di natura operativa e organizzativa, costi allo stato attuale superiori ai benefici attesi e difficoltà a trovare *use case* applicabili.

Fig. 11 – Lo sviluppo di sistemi di IA quale obiettivo strategico: principali motivazioni
(numero di società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

4 L'uso di sistemi di IA e le tecnologie prevalenti

4.1 Grado e ambiti di applicazione dei sistemi di IA

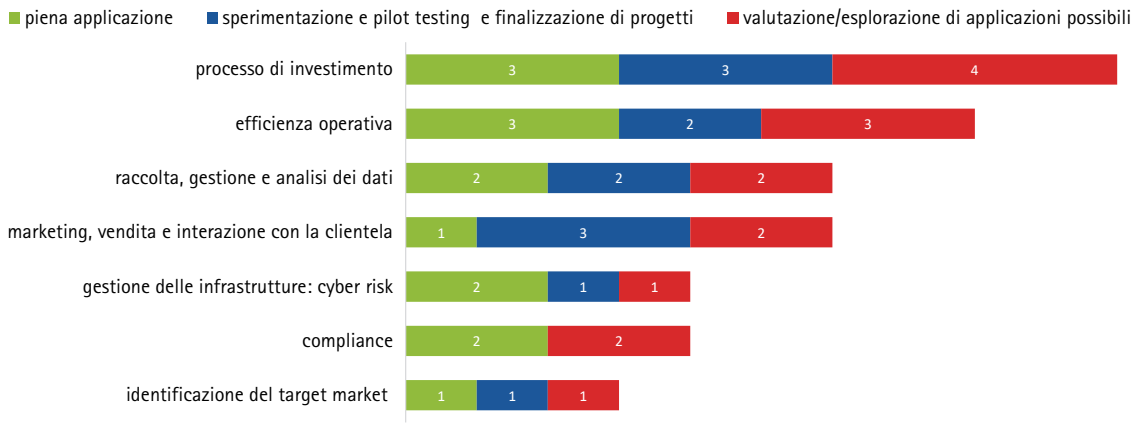
La quasi totalità (sette su otto) delle società partecipanti all'indagine dichiara di avere attualmente in uso sistemi di IA da almeno un anno: nello specifico, da uno a tre anni in cinque casi, da tre a cinque anni in un caso e da oltre cinque anni in un altro caso.

La piena applicazione di sistemi di IA in almeno un ambito viene dichiarata da cinque delle otto Sgr coinvolte nella survey. Tra queste, due società dichiarano di aver raggiunto la piena applicazione di tecniche di IA in cinque distinti ambiti. Cinque società riferiscono inoltre di essere in fase di sperimentazione, test o finalizzazione dei progetti in almeno un ambito, mentre le restanti Sgr sono ancora in fase di studio e di valutazione delle possibili applicazioni. Il grado di sviluppo di tecnologie di IA risulta ovviamente più maturo per le rispondenti che se ne avvalgono da più tempo.

Gli ambiti di utilizzo dei sistemi di IA presentano una grande eterogeneità, risultando prevalenti il processo di investimento e le attività volte all'incremento dell'efficienza operativa, che registrano la piena applicazione in tre società. Altre aree rilevanti sono rappresentate dalle attività di gestione e analisi dei dati e di gestione delle infrastrutture e *compliance*, per le quali la piena applicazione è stata raggiunta da due società.

Il processo di investimento è anche l'area interessata dal maggior numero di sperimentazioni in atto e di ricerca di possibili applicazioni future, seguito dalle attività di marketing, vendita e interazione con la clientela e dalle attività mirate all'incremento dell'efficienza operativa (Fig. 12).

Fig. 12 – Stato di applicazione dei sistemi di IA per ambito/area
(numero di società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

4.2 I prodotti e servizi coinvolti

L'utilizzo di sistemi di IA si riferisce nella maggior parte dei casi alla gestione di fondi UCITs (cinque Sgr su sette) e, a seguire, al servizio di consulenza in materia di investimenti (tre casi); risulta invece meno diffusa l'applicazione alla gestione di portafoglio su base individuale e di fondi di investimento alternativi (due casi) e alla gestione di fondi pensione aperti (un caso)[69]. La maggior parte delle società ritiene che l'utilizzo di sistemi di IA non sia più utile nell'ambito della gestione di fondi alternativi (FIA) rispetto ai fondi UCITs. Due Sgr ritengono al contrario che tali tecnologie siano meno utili nella gestione di fondi alternativi e motivano tale opinione, rispettivamente, facendo riferimento alla minore disponibilità di dati relativi a mercati privati rispetto ai mercati quotati ovvero al fatto che la gestione di FIA richiede al momento analisi basate su un intenso coinvolgimento umano.

Tutte le Sgr dichiarano di gestire fondi cosiddetti ESG, ossia fondi la cui politica di investimento tiene conto di criteri di sostenibilità ambientale, sociale e di *governance*. Tra queste, solo due affermano di applicare le nuove tecnologie nell'ambito degli investimenti sostenibili, in particolare per la verifica e l'analisi dei dati e delle informazioni contenuti nella rendicontazione di sostenibilità pubblicata dalle società, e per la verifica della coerenza degli investimenti rispetto alla strategia volta all'integrazione dei fattori ESG nella propria politica di investimento. Una società dichiara inoltre di utilizzare l'IA anche per il monitoraggio e la gestione dei rischi per la sostenibilità.

La metà delle società rispondenti alla *survey* ritiene che i sistemi di IA possano assicurare la coerenza delle scelte di investimento con la politica del fondo sebbene non in piena autonomia, poiché il controllo umano è ritenuto necessario.

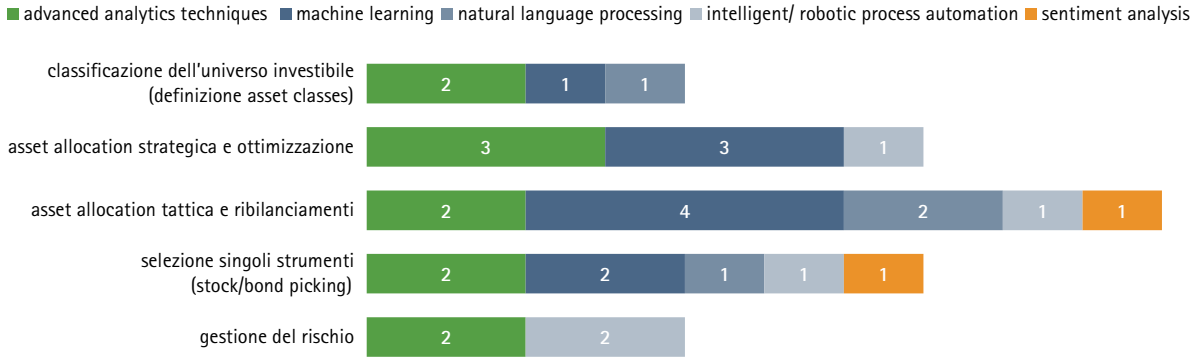
69 Una società ha inoltre dichiarato di non aver circoscritto l'utilizzo di sistemi di IA a un unico prodotto o servizio.

4.3 Le tecnologie utilizzate

Come anticipato, l'ambito in cui l'IA trova un maggiore utilizzo nelle società intervistate è quello del processo di investimento. Laddove l'applicazione è piena, le nuove tecnologie vengono usate sia nella fase di allocazione strategica, ossia nella scelta degli investimenti e nell'ottimizzazione dei portafogli con una visione di lungo periodo, sia nella fase di allocazione tattica, ossia per i ribilanciamenti di portafoglio connessi a dinamiche di breve periodo[70].

Le tecnologie maggiormente citate dalle società intervistate sono il Machine Learning e le Advanced Analytics Tecniques, che trovano (quasi sempre) applicazione in tutte le fasi del processo di investimento; seguono le tecnologie di Robotic Process Automation e Natural Language Process, mentre risultano meno diffuse le applicazioni di *sentiment analysis*. In quasi tutte le società intervistate l'utilizzo di sistemi di IA si avvale di diverse tecnologie; solo in un caso, infatti, viene usato esclusivamente il *machine learning* (Fig. 13).

Fig. 13 – Tecnologia utilizzata per fasi del processo di investimento
(numero di società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

Le tecnologie utilizzate sono di origine mista, ossia sviluppate in parte internamente e in parte in *outsourcing* o in *partnership* con soggetti esterni all'azienda. Tra le principali ragioni connesse all'origine mista, si annoverano in primo luogo la presenza di tempi di ingegnerizzazione e implementazione troppo lunghi (sei casi) e (a distanza) la mancanza di adeguate professionalità interne in alcuni ambiti (due casi), la possibilità di concentrarsi su un solo un aspetto della tecnologia (due casi) e la presenza di precedenti accordi o collaborazioni con soggetti esterni (un caso)[71]. Nessuna società

70 In due casi i sistemi di IA vengono utilizzati anche per la gestione del rischio, in un caso anche per la classificazione dell'universo di asset su cui investire e in un altro caso per l'attività di *stock* o *bond picking*. Tra i rischi più citati da coloro che utilizzano sistemi di IA anche con finalità di *risk management* emergono il rischio di liquidità, di mercato, di assorbimento patrimoniale per i fondi pensione e il *drow-down risk* nell'attività di trading (il *draw-down* indica la quantità di denaro, in percentuale rispetto al capitale complessivo, persa nel trading e la conseguente riduzione del proprio capitale iniziale; esso misura quindi il costo in termini di perdite di un sistema di trading e permette quindi di valutarne il livello di rischio e l'efficienza).

71 Una società ha inoltre indicato la possibilità di massimizzare l'efficienza del processo di innovazione secondo le logiche dell'*open science/open source*.

ha invece indicato i costi eccessivi tra le ragioni che possono scoraggiare lo sviluppo interamente *in-house* delle applicazioni tecnologiche.

Con riguardo ai dati utilizzati congiuntamente con i sistemi di IA, prevalgono quelli strutturati (indicati da tutti le rispondenti), mentre il ricorso anche a dati semi-strutturati viene segnalato da due Sgr e quello a dati non strutturati da due diverse società. In nessun caso vengono utilizzare tutte e tre le tipologie di dati menzionate. Come per le tecnologie, anche la fonte dei dati utilizzati è in parte interna e in parte esterna (cinque società). Solo in due casi la fonte esclusiva è quella esterna.

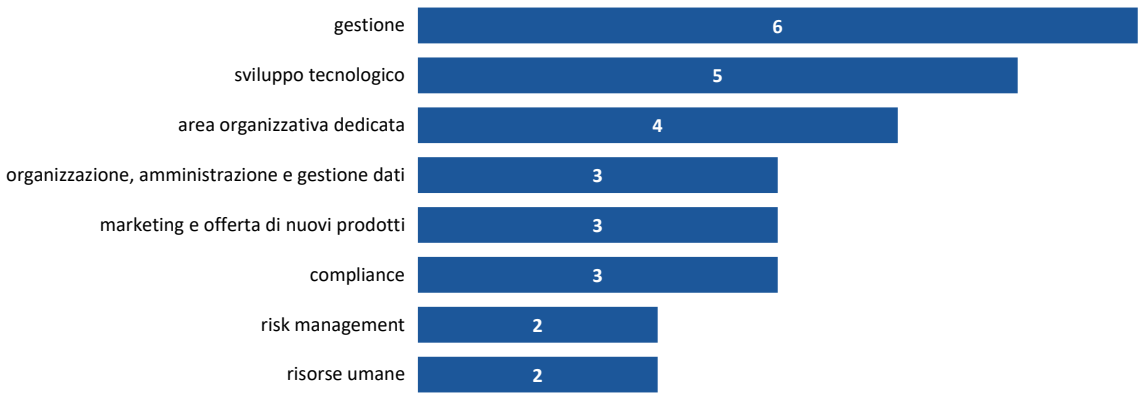
Infine, le applicazioni di IA sono eseguite esclusivamente su *server* locali in quattro casi su sette, unicamente in *cloud* in due casi e mediante entrambe le modalità in un caso.

5 Organizzazione e *governance*

5.1 Profili organizzativi

La *survey* ha indagato se le società intervistate avessero adottato un assetto organizzativo specifico al fine di integrare e gestire al meglio le nuove tecnologie nella struttura aziendale. Le aree funzionali coinvolte nell'utilizzo di sistemi di IA sono numerose, anche in quelle società che hanno dichiarato di aver costituito una o più specifiche unità organizzative dedicate (quattro società). Tra le aree più frequentemente coinvolte nell'utilizzo di sistemi di IA rientrano la gestione di portafoglio e lo sviluppo tecnologico, seguite da organizzazione, amministrazione e gestione dei dati e marketing; le aree dedicate alla *compliance*, alla gestione delle risorse umane e al *risk management* risultano invece meno citate (Fig. 14).

Fig. 14 – Aree funzionali coinvolte nell'utilizzo di sistemi di IA
(numero di società; risposta multipla)



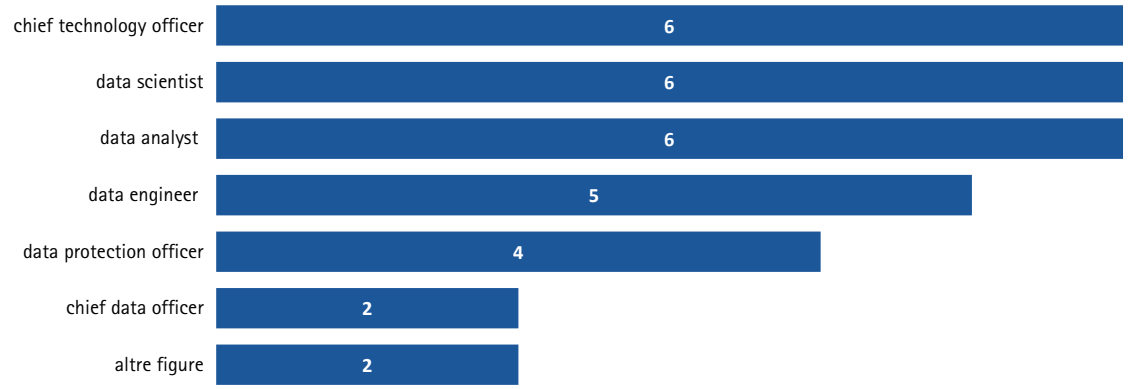
Fonte: elaborazioni CONSOB su dati di *survey*.

Le risposte delle Sgr alla *survey* mettono in evidenza che le società, considerano importante investire in tecnologie di intelligenza artificiale che possono rendere

più efficiente il lavoro svolto dalla funzione di *compliance* e *risk management* (si veda il par. 3).

Le figure professionali presenti nella maggior parte delle società sono numerose e variegate. In particolare, la maggior parte delle società indica la presenza nel proprio organico di esperti sviluppatori di sistemi di IA (*data scientist*) e di esperti di analisi dei dati (*data analyst*) e ha individuato un responsabile per lo sviluppo tecnologico (*Chief Technology Officer*). Cinque Sgr indicano di avvalersi anche di esperti nel trattamento dei dati (*data engineer*), quattro società hanno individuato un responsabile per la protezione dei dati (*Data Protection Officer*) e due indicano di aver individuato un responsabile della qualità dei dati (*Chief Data Officer*). La presenza di altre figure professionali coinvolte nello sviluppo e gestione di sistemi di IA è indicata da due società; tra le figure citate vi sono specialisti in specifiche tecnologie di IA (*intelligent/robotic automation*) e responsabili dei sistemi informatici e del rischio tecnologico (Fig. 15).

Fig. 15 – Figure professionali coinvolte in sviluppo e gestione di sistemi di IA presso le Sgr partecipanti all'indagine
(numero società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

5.2 *Governance* dei dati e degli algoritmi

Un altro ambito indagato dalla *survey* ha riguardato il sistema di *governance* dei dati e degli algoritmi utilizzati. L'impressione generale che si trae dalle risposte raccolte è che le società riconoscano la necessità di una stringente supervisione umana sui processi decisionali basati sull'utilizzo di sistemi di IA, sebbene il ruolo rivestito dalle funzioni di controllo e le procedure di *governance* degli algoritmi sottostanti e dei dati utilizzati siano ambiti ancora in fase evolutiva.

In particolare, l'applicazione di tecnologie di IA risulta essere sempre sottoposta a supervisione umana, poiché nessuna delle rispondenti ha indicato di utilizzare sistemi che agiscono in piena autonomia. In quattro società su sette il controllo umano risulta molto forte poiché il sistema di IA non ha alcuna autonomia o non è mai utiliz-

zato direttamente per prendere decisioni di business rilevanti. Due Sgr riferiscono invece di una autonomia parziale, poiché sono gli input umani a influenzare le decisioni in misura preponderante. Solo in un caso, il sistema di IA in uso viene qualificato come totalmente autonomo: tuttavia le decisioni finali che questo suggerisce sono poi sottoposte a controllo umano.

Tre società delle sette che fanno ricorso a sistemi di IA ritengono che i processi decisionali basati su tali sistemi non dovrebbero essere trattati diversamente rispetto ai normali processi decisionali che coinvolgono esclusivamente il fattore umano. Le restanti quattro società sono invece di opinione contraria, sulla base di motivazioni piuttosto eterogenee ma tutte riconducibili alla necessità di sottoporre sempre le scelte indicate da un sistema di IA a un controllo umano finalizzato a valutarne il corretto funzionamento, l'interpretabilità dei risultati, gli effetti delle scelte indicate e i relativi rischi potenziali. Una società indica tra le motivazioni la considerazione che la natura dei mercati ha una componente non razionale che cambia in modo repentino e che non sempre può essere correttamente identificata da una regola codificata. Un'altra società sottolinea anche la necessità di monitorare il funzionamento dei sistemi di IA al fine di intercettare eventuali funzionamenti anomali o discrepanze tra la qualità dell'output ottenuto e la qualità attesa.

Come evidenziato in precedenza, le procedure di *governance* dei sistemi di IA sono ancora in fase evolutiva: le società che hanno introdotto specifiche procedure di *governance* degli algoritmi sottostanti a tali tecnologie sono solo tre, mentre risultano più numerose le Sgr che hanno predisposto specifiche procedure di *governance* dei dati (cinque casi su sette).

6 Benefici attesi e rischi percepiti

Un ulteriore profilo indagato dalla *survey* ha riguardato i benefici e i rischi derivanti dall'utilizzo di tecnologie di IA percepiti dalle società.

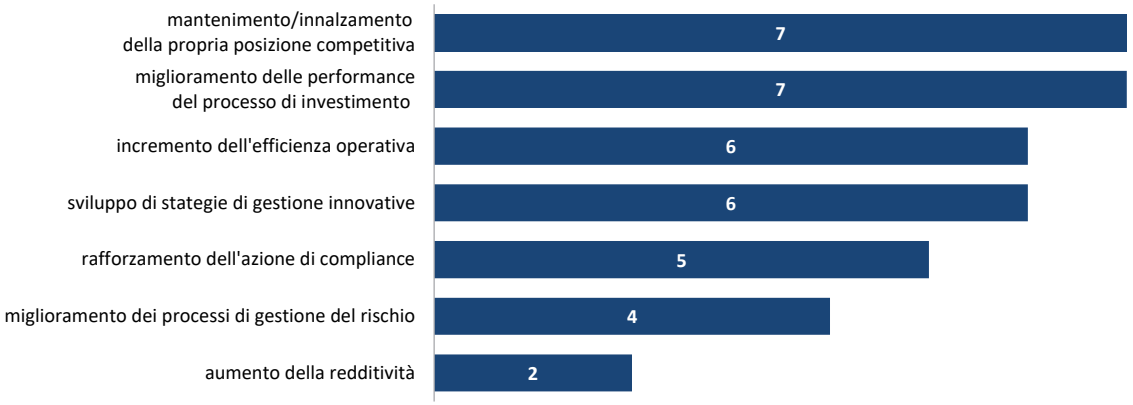
La maggior parte delle rispondenti (cinque su otto) ritiene che i benefici possano essere pienamente apprezzati sin dalle prime fasi di applicazione di tali tecnologie o al massimo entro tre anni, mentre tre società ritengono che sia necessario un arco temporale più lungo. La quasi totalità delle Sgr (sette su otto) identifica i benefici nel mantenimento o innalzamento della propria posizione competitiva e nel miglioramento delle *performance* del processo di investimento; la maggioranza (sei società) menziona l'incremento dell'efficienza operativa e il contributo allo sviluppo di strategie di gestione innovative; più della metà (cinque società) indica il rafforzamento dell'azione di *compliance* e la metà menziona il miglioramento dei processi di gestione del rischio. La possibilità che l'utilizzo di sistemi di IA possa determinare un aumento della redditività è indicata solo da due società (Fig. 16).

L'ambito a cui si associano i maggiori benefici attesi dall'utilizzo dei sistemi di IA è rappresentato dal processo di investimento (sette società), seguito dalla raccolta, gestione e analisi dei dati (sei società), l'efficienza operativa (cinque società) e l'attività di marketing, vendita e interazione con la clientela (quattro società). Minori benefici

sono invece associati alla gestione delle infrastrutture (diagnostica e monitoraggio del *cyber risk*) e alla definizione del *target market*, menzionate solo da due Sgr, e al rafforzamento dell'azione di *compliance*, segnalato da una Sgr.

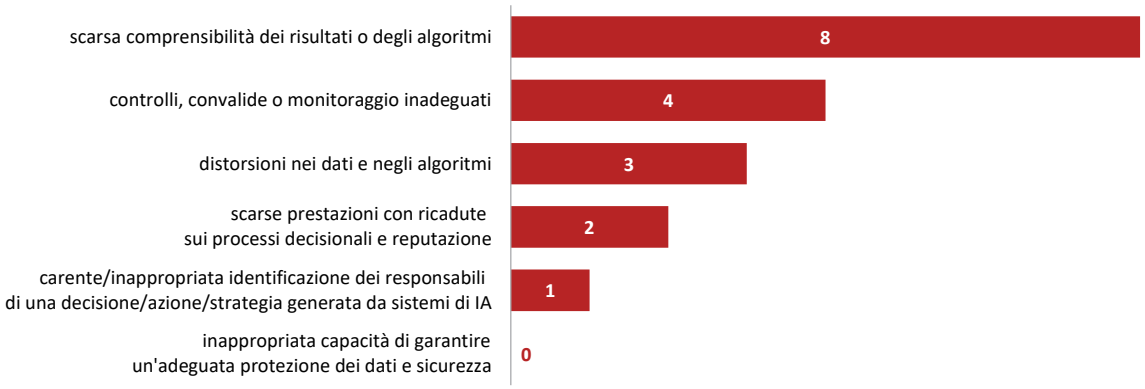
Tra i rischi più rilevanti derivanti dall'utilizzo dei sistemi di IA nell'*asset management* tutti i partecipanti all'indagine indicano la scarsa comprensibilità degli output e/o degli algoritmi, mentre la metà del campione segnala l'inadeguatezza di controlli, convalide o monitoraggio. Tra i rischi potenziali, inoltre, tre società evidenziano la presenza di distorsioni nei dati e negli algoritmi. Sono invece meno avvertiti i rischi collegati a scarse prestazioni del sistema di IA con ricadute su processi decisionali e reputazione (due società), e il rischio di carente o inappropriata identificazione dei responsabili di una decisione, azione o strategia generata da sistemi di IA (una società). Nessuna rispondente ha indicato quale fonte di rischio la difficoltà di garantire una adeguata protezione dei dati e la sicurezza (Fig. 17).

Fig. 16 – Percezione dei benefici associati all'utilizzo di sistemi di IA
(numero società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

Fig. 17 – Percezione dei rischi associati all'utilizzo di sistemi di IA
(numero società; risposta multipla)



Fonte: elaborazioni CONSOB su dati di *survey*.

7 Prospettive evolutive

Investire nello sviluppo di sistemi di IA, sin da subito e a maggior ragione nell'ambito di un orizzonte temporale di cinque anni, è ritenuto molto importante dalla quasi totalità delle società intervistate. Sette Sgr (su otto) hanno infatti intrapreso la realizzazione di progetti tesi all'applicazione delle nuove tecnologie nell'ambito del processo di investimento, sei società stanno sviluppando progetti finalizzati all'incremento dell'efficienza operativa e cinque società hanno indicato quale area interessata da progetti di sviluppo di sistemi di IA la raccolta, gestione e analisi dei dati.

Con specifico riferimento al processo di investimento, tutti i progetti in essere coinvolgono la fase di *asset allocation* tattica e di ribilanciamento del portafoglio. Sei società indicano l'attività di selezione degli strumenti in cui investire tra le fasi maggiormente interessate da nuovi progetti, mentre l'*asset allocation* strategica è indicata da quattro Sgr.

Lo stadio di avanzamento dei progetti in corso è eterogeneo sia tra società (che in tre casi prevedono di ultimare nell'arco di un anno e in cinque casi entro tre anni) sia tra i diversi progetti intrapresi da una medesima società.

In prospettiva, tutti le rispondenti ritengono che nei prossimi cinque anni i sistemi di IA troveranno maggiore applicazione sicuramente nell'ambito del processo di investimento e con riferimento a diverse fasi, tra cui la selezione degli strumenti in cui investire, l'analisi finanziaria, la gestione e analisi dei dati e l'individuazione di nuove strategie di investimento grazie alla capacità di elaborare grandi quantità di dati strutturati e non strutturati. Molte società inoltre menzionano l'attività di gestione dei rischi e delle infrastrutture (diagnostica e monitoraggio del *cyber risk*), la *compliance*, il marketing e lo sviluppo di nuovi prodotti. Un ulteriore ambito indicato come rilevante è l'incremento dell'efficienza operativa che le nuove tecnologie renderanno possibile migliorando i processi interni legati alle attività di *backoffice* e i processi esterni legati ai rapporti con clienti e *outsourcers*. Una società indica tra le applicazioni future l'analisi dei fattori ESG.

Con riferimento agli elementi di contesto che nei prossimi cinque anni potranno incidere maggiormente sull'intensità di applicazione dei sistemi di IA da parte dell'industria dell'*asset management*, le società partecipanti alla *survey* individuano fattori prevalentemente di tipo tecnologico e regolamentare.

Da un lato, infatti, si ritiene che lo sviluppo tecnologico possa rendere l'utilizzo dell'IA progressivamente più attrattivo per l'industria, garantendo, tra le altre cose, una crescente capacità computazionale, una maggiore interpretabilità dei dati e degli algoritmi, un calo dei costi delle componenti *hardware* e dello *storage* dei dati, un ampliamento dell'insieme e della qualità dei dati accessibili, una maggiore disponibilità di componenti *software* in logica *open-source*.

Dall'altro lato, le società ritengono essenziale che l'evoluzione della regolamentazione sia tale da supportare lo sviluppo e la diffusione di innovazioni tecnologiche. In questa prospettiva, l'istituzione della *regulatory sandbox* di cui all'art. 36 del

d.l. 34/2019 (d.l. Crescita) viene considerata un elemento di facilitazione delle sperimentazioni basate anche sull'utilizzo di tecnologie basate sull'IA, garantendo un'adeguata tutela dei consumatori e dei mercati finanziari e creando un contesto competitivo omogeneo; un possibile fattore di sviluppo di prodotti e servizi finanziari innovativi e di revisione dei processi e dei servizi esistenti da parte dell'industria; uno strumento in grado di favorire l'apertura dell'industria italiana a collaborazioni con soggetti esteri, consentendo al Paese di stare al passo con gli sviluppi tecnologici in atto a livello internazionale. Alcune società inoltre giudicano positivamente l'istituzione di una *regulatory sandbox* poiché essa consente agli operatori finanziari di sperimentare tecnologie e processi innovativi in un ambiente protetto e di conformarsi in modo più efficiente e a costi inferiori alle normative vigenti.

TERZA
PARTE

Profili di *investor protection*

1 Tecniche di automazione e contenuto dell'obbligo di diligenza dovuta dall'intermediario

1.1 Tecniche di automazione e regole di condotta: la valutazione di adeguatezza...

L'utilizzo di tecniche di automazione per l'analisi dei *big data* consente al gestore di effettuare la valutazione di adeguatezza prevista dall'art. 25, par. 2, della Direttiva 2014/65/UE (MiFID II) secondo nuove modalità[72].

Nei paragrafi seguenti saranno illustrati i diversi momenti, nella declinazione degli obblighi di condotta vigenti, nei quali possono essere utilizzate dai gestori le nuove tecniche di automazione (d'ora in poi TA), descrivendone le potenzialità e i rischi connessi e gli eventuali riflessi sul contenuto dell'obbligo di diligenza in capo al gestore.

La profilazione[73] del cliente costituisce il primo *step* attraverso il quale gli intermediari, tradizionalmente mediante il questionario, acquisiscono le informazioni «*di cui necessitano per comprendere le caratteristiche essenziali dei clienti e disporre di una base ragionevole per determinare*»[74] l'adeguatezza dell'investimento offerto.

Sul processo di raccolta di tali informazioni la letteratura in ambito di *behavioural economics* ha da tempo evidenziato il problema relativo ai limiti cognitivi e comportamentali degli investitori[75]. Come noto, infatti, gli investitori non professionali

72 L'art. 25, par. 2, MiFID II prevede espressamente che «*(q)uando effettua consulenza in materia di investimenti o gestione di portafoglio, l'impresa di investimento ottiene le informazioni necessarie in merito alle conoscenze ed esperienze del cliente o potenziale cliente in materia di investimenti riguardo al tipo specifico di prodotto o servizio, alla sua situazione finanziaria, tra cui la capacità di tale persona di sostenere perdite e ai suoi obiettivi di investimento, inclusa la sua tolleranza al rischio, per essere in grado di raccomandare i servizi di investimento e gli strumenti finanziari che siano adeguati al cliente o al potenziale cliente e siano in particolare adeguati in funzione della sua tolleranza al rischio e della sua capacità di sostenere perdite*». *Nell'ordinamento italiano la regola di adeguatezza è riconducibile all'art. 21, lett. b), del d. lgs. 58/1998 (Testo Unico della Finanza, d'ora in poi 'TUF') secondo cui gli intermediari devono «b) acquisire, le informazioni necessarie dai clienti e operare in modo che essi siano sempre adeguatamente informati*». Sul tema, in maniera più diffusa, si rinvia alle trattazioni di Perrone, A. (2018), *Il diritto del mercato dei capitali*, Milano, 216 ss. e di Annunziata, F. (2015), *La disciplina del mercato mobiliare*, Torino, 146 e ss..

73 La profilazione è attualmente definita dell'art. 4 del Regolamento (UE) 2016/679, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE «*qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati personali per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica*».

74 Si veda l'art. 54, par. 2, Regolamento delegato MiFID II.

75 Sul concetto di *behavioural finance* si rinvia a Ricciardi, V. e Simon, H.K. (2008), *What is Behavioural Finance*, in Business, Education & Technology Journal, 2, 1. Con riferimento alla percezione degli italiani della conoscenza finanziaria si rinvia a CONSOB (2021), *Report on financial investments of Italian households*, www.consob.it, e CONSOB (2016),

hanno solitamente una conoscenza limitata in merito al funzionamento del mercato dei capitali e alle caratteristiche degli strumenti finanziari offerti. A tale riguardo, il Regolamento delegato 565/2017/UE (Regolamento delegato MiFID II) ha previsto che gli intermediari devono adottare «*le misure ragionevoli per assicurare che le informazioni raccolte sui clienti o potenziali clienti siano attendibili*»[76], specificando una serie di indicazioni tra cui la necessità di verificare che «*le domande utilizzate nel processo siano atte ad essere comprese dai clienti*»[77] e di «*intraprendere azioni per assicurare la coerenza delle informazioni sul cliente*»[78].

Oltre ai limiti cognitivi, l'azione dell'investitore può essere caratterizzata da *bias* comportamentali. Si pensi all'ipotesi in cui un investitore, per una eccessiva sicurezza sulle proprie conoscenze in merito alle dinamiche del mercato finanziario (cosiddetta *overconfidence*), ritenga di comprendere il funzionamento degli strumenti finanziari derivati senza avere una sua specifica preparazione in materia, ma solo per aver realizzato dei profitti ottenuti da operazioni aventi ad oggetto tali strumenti.

In relazione a tali descritti problemi, una possibile opportunità di perfezionamento dell'attività di profilatura sembra arrivare dalle nuove tecniche di *data mining*, associate alla presenza dei *big data*, che forniscono al gestore un potente strumento di supporto nell'adempimento della cosiddetta '*know your customer rule*'. In particolare, l'intermediario avrebbe a disposizione un maggior numero di informazioni, avendo accesso a tutta una serie di dati ulteriori di estrema utilità, senza la necessità di rivolgersi al cliente, così ottenendo un quadro più dettagliato sulla situazione dell'investitore. Inoltre, attraverso tali tecniche di supporto anche il gestore potrebbe valutare l'attendibilità dei dati ottenuti durante la fase di profilatura. In tal senso si potrebbe immaginare una gerarchia tra dati provenienti da fonti eterogenee[79] e il gestore potrebbe fondare la profilatura sulle informazioni ottenute tramite il questionario, controllandone la veridicità attraverso l'analisi di ulteriori dati acquisiti da terzi soggetti[80].

Financial advice seeking, financial knowledge and overconfidence, Gentile, M., Linciano, N. e Soccorso (a cura di), in Quaderni di finanza, 83, www.consob.it, 28.

76 Così l'art. 54, par. 7, Regolamento delegato MiFID II.

77 Si veda l'art. 54, par. 7, lett. c), Regolamento delegato MiFID II che segnala la necessità di «*assicurarsi che le domande utilizzate nel processo siano atte a essere comprese dai clienti, procurino un'immagine accurata degli obiettivi e delle esigenze del cliente e veicolino le informazioni necessarie a condurre la valutazione dell'idoneità*».

78 Si veda l'art. 54, par. 7, lett. d), Regolamento delegato MiFID II che impone l'adozione di «*azioni, laddove opportuno, per assicurare la coerenza delle informazioni sul cliente, per esempio analizzando se nelle informazioni da questi fornite vi siano delle evidenti imprecisioni*».

79 Le vere potenzialità del cosiddetto *data mining* consistono infatti nel poter elaborare gli *unstructured data*, cioè quei dati che non sono già stati elaborati in tabelle e che non hanno tipicamente un valore oggettivo (ad esempio, le preferenze dei clienti che possono essere dedotte dai comportamenti sui *social networks*). Con riguardo invece ai dati strutturati (come i bilanci societari), normalmente utilizzati dai consulenti, la big data analytics consente una maggiore potenza di analisi, si veda Sabbatini, R. (2019), *La rivoluzione dei big data al servizio dell'investimento*, www.wewealth.com. In tale ottica si colloca l'indicazione dell'ESMA di sfruttare dei *risk-profiling software* in grado di svolgere controlli di coerenza dei dati acquisiti dal cliente in sede di profilatura, ESMA (2012), *Guidelines on certain aspects of the MiFID suitability requirements*, www.esma.europa.eu, 10.

80 Si veda Giorgi, M. (2018), *Automazione, Big data e integrazione funzionale: la necessità di una nuova ermeneutica giuridica dei servizi di consulenza finanziaria*, in Lener, R. (a cura di), *Fintech: Diritto, tecnologia e finanza*, Roma, 68.

Attraverso l'utilizzo di tali tecniche di automazione e la conseguente acquisizione di indicazioni più specifiche sull'investitore, il gestore potrebbe realizzare un servizio sempre più personalizzato e adeguato alle caratteristiche ed esigenze del cliente, così superando la prassi delle linee di gestione che uniformano il servizio di investimento[81]. In tale ambito, una delle tecniche utilizzabili consiste nella cosiddetta '*customer as a whole*' in base alla quale, tenendo conto dei comportamenti del cliente, come gli acquisti passati e presenti[82], si riescono a prevedere il comportamento futuro che, per esempio, potrebbero incidere sull'attitudine al rischio e/o sugli obiettivi di investimento[83].

Un'ulteriore ipotesi di utilizzo simile dei *big data* è ipotizzabile per la valutazione di adeguatezza in chiave dinamica[84]. In particolare, la gestione di portafoglio, infatti, richiede una valutazione costante della coerenza dell'investimento al profilo del cliente per cui le tecniche di automazione possono consentire il relativo svolgimento in tempo reale. A tali fini, si possono utilizzare le tecniche nell'ambito della cosiddetta '*customer relationship management*'[85] (d'ora in poi, 'CRM') per valutare scostamenti tra il comportamento dell'investitore e il suo profilo finanziario. In tal modo, ad esempio, si potranno utilizzare i dati per verificare la necessità di ribilanciare il portafoglio e adeguare l'investimento a un mutato livello di rischio, risultante dall'analisi delle informazioni riguardanti il comportamento del cliente. L'utilizzo di questa tecnica di automazione permetterebbe così di verificare l'affidabilità o la validità temporale dei dati acquisiti attraverso modalità tradizionali.

Allo stesso tempo, oltre a offrire una soluzione ai problemi cognitivi e comportamentali che possono influenzare i clienti, i *big data* possono anche fornire una risorsa importante per far fronte all'adempimento degli obblighi stabiliti dal Regolamento delegato MiFID II. Infatti, l'uso di questa tecnologia permetterebbe una verifica efficiente dei dati dichiarati dai clienti, in conformità con l'obbligo di «*intraprendere azioni per assicurare la coerenza delle informazioni sul cliente*». Il controllo e il monitoraggio con tecniche di *big data analytics* possono permettere una verifica efficiente e

81 Una tale prassi viene segnalata, *ex multis*, da Annunziata, *op. cit.* (2015), 100; Costi, R. ed Enriques, L. (2004), *Il mercato mobiliare*, in *Trattato di diritto commerciale* (diretto da Cottino, G.), Padova, 250 ss., spec. 373. Sulla presenza di line di gestione, si veda Mastrangelo, G. (1998), *L'evoluzione della disciplina delle attività di gestione dei patrimoni mobiliari*, in Ferrarini, G. e Marchetti, P. (a cura di), *La riforma dei mercati finanziari*, Milano, 147 e ss.

82 In tal senso, «*the EBA acknowledged that, among other types of consumer data, the use of payment data in the market segment of retail payments is of particular interest to the EBA because, unlike other, one-off, types of data provided by consumers, payments data provide financial institutions with a continuous and extensive insight into consumers' purchasing habits, preferences and, therefore, lifestyle more generally*» (EBA (2017), *Report on innovative uses of customer data by financial institutions*, www.eba.eu, 4).

83 Si veda Mattasoglio, F. (2016), *La profilazione dell'investitore nell'era dei big data. Rischi dell'estremizzazione della regola del 'know your customer'*, in Riv. trim. dir. ec., suppl. 4, 247.

84 L'art. 25 par. 6, MiFID II prevede che: «*Se un'impresa di investimento offre la gestione di portafoglio o ha informato il cliente che effettuerà periodicamente una valutazione di adeguatezza, la relazione periodica conterrà una dichiarazione aggiornata che spieghi perché l'investimento corrisponde alle preferenze, agli obiettivi e alle altre caratteristiche del cliente*».

85 La CRM viene definita come «*the opportunity to contact the right costumer at the right time through the right marketing medium*», Neslin, A., Taylor, G.A., Grantham, D. e McNeil, R. (2012), *Overcoming the 'Recency Trap' in Customer Relationship Management*, in *J. Of the Acad. Mark. Sci.*, 320. Attraverso la CRM, il venditore di un prodotto «*tracks what, when and how fast the client buys a product, and which products he or she looks or does not look at*».

un'azione rapida in caso di discrepanze tra le caratteristiche del portafoglio e i requisiti degli investitori[86].

1.2 ...(segue) Profili di *product governance*

In materia di *product governance*[87] uno dei doveri di condotta in capo all'intermediario è la cosiddetta '*target market assessment*' (d'ora in poi, 'TMA'), ovvero l'individuazione del mercato di riferimento prima di commercializzare gli strumenti finanziari. Nello specifico, l'intermediario è tenuto a elaborare un processo di approvazione per ogni strumento finanziario precisando per ciascuno di questi «*il determinato mercato di riferimento di clienti finali all'interno della pertinente categoria di rischio*»[88].

Anche tale obbligo in capo all'intermediario potrà essere adempiuto attraverso una TA per l'analisi dei dati riguardanti la clientela ottenuti a vario titolo. In particolare, il *distributor*[89] può sfruttare le tecnologie dei *big data* al fine di svolgere una precisa determinazione del cosiddetto *target market*, usufruendo del contatto diretto con la clientela attraverso cui acquisire le informazioni personali per specificare il «*mercato di riferimento reale*»[90]. A tale riguardo, l'ESMA sottolinea che il *distributor*, diversamente dal *manufacturer*[91], avendo informazioni più dettagliate sulla clientela, dovrà svolgere in maniera più specifica la TMA considerando quella svolta dal produttore solamente un punto di partenza al fine di ridurre il rischio di offrire prodotti inadeguati ai clienti.

In questo ambito, il gestore potrà effettuare la cosiddetta '*customer segmentation*'[92], attività con cui, attraverso appositi algoritmi, viene suddivisa la clientela a cui si rivolge sulla base di caratteristiche comuni.

86 Si vedano: Perrone, A. (2021), *Intelligenza artificiale e servizi di investimento*, in Costa, C., Mirone, A., Pennisi, R., Sanfilippo, P.M. e Vigo, R. (a cura di), *Studi di diritto commerciale per Vincenzo Di Cataldo*, Vol. II, Torino, 712; Tina, *op. cit.* (2020), *Cybersicurezza: integrità dei processi e dei dati*, in Cian, M. e Sandei, C. (a cura di), *Diritto del Fintech*, Milano, 92, nota 17, sulla scorta di Panetta, F., *Il punto di vista del regolatore. Intervento al consegno, La diffusione della cultura assicurativa in Italia e l'impatto dell'innovazione tecnologica*, http://ivass.it, 9.

87 Per una diffusa trattazione della *product governance* si veda Troiano, V. (2016), *La product governance*, in Troiano, V. e Motroni, R. (a cura di), La MiFID II, Padova, 213 e ss.; Perrone, *op. cit.* (2018), 221 e ss.; Annunziata, *op. cit.* (2015), 158 e ss.

88 L'articolo 16, par. 3, Dir. MiFID II recita così: «*Il processo di approvazione del prodotto precisa per ciascuno strumento finanziario il determinato mercato di riferimento di clienti finali all'interno della pertinente categoria di clienti e garantisce che tutti i rischi specificamente attinenti a tale target siano stati analizzati e che la prevista strategia di distribuzione sia coerente con il target stesso*».

89 Con il termine *distributor* si identificano le imprese di investimento che hanno il compito di collocare i prodotti finanziari alla clientela. Si veda Annunziata, *op. cit.* (2015), 159.

90 Sul punto si veda ESMA (2018), *Orientamenti ESMA sugli obblighi di governance dei prodotti ai sensi della MiFID II*, www.esma.europa.eu, 6, parr. 17 e 11, parr. 35 e 36, secondo cui il distributor deve «*specificare il mercato di riferimento reale*», poiché il *manufacturer* «*determinerà il mercato di riferimento senza alcuna conoscenza specifica dei clienti*».

91 Con *manufacturer* si intendono le «*imprese di investimento che realizzano strumenti finanziari da offrire in vendita alla clientela*», Annunziata, *op. cit.* (2015), 159.

92 La segmentazione della clientela è uno degli elementi fondamentali del *Business Model Canvas* in quanto si focalizza sulle specifiche esigenze di ogni cluster di clienti. Il *Business Model Canvas* di Osterwalder è stato utilizzato per analizzare il fenomeno della consulenza automatizzata, si veda CONSOB, *op. cit.* (2019), 12.

Un ulteriore utilizzo delle TA riguarda la revisione periodica dei prodotti finanziari offerti secondo quanto previsto dalla MiFID II[93]. Più precisamente, da un lato, il *manufacturer* potrà, attraverso l'analisi delle informazioni sui clienti, valutare periodicamente la compatibilità del prodotto finanziario con il mercato di riferimento individuato anche a prescindere dalle informazioni fornite dal distributore[94]; dall'altro il *distributor* potrà svolgere più precisamente la trasmissione dei dati aggregati[95] al *manufacturer*, atto funzionale alla revisione del prodotto finanziario distribuito, avendo valutato puntualmente la propria esperienza riguardo quest'ultimo[96].

In tal caso, potrebbe essere utilizzata la *sentiment analysis*, ossia la metodologia che dalle informazioni rinvenibili nei *social network* consente di ottenere oppure ricostruire le opinioni e gli orientamenti dei clienti riguardo un prodotto o servizio[97]. In questo modo, si può chiarire la comprensione dei dati aggregati che l'intermediario *distributor* deve trasmettere all'intermediario *manufacturer* ai fini della revisione dello strumento finanziario distribuito.

1.3 Condizioni e limiti di utilizzo dei *big data*

Dalla breve disamina offerta dei possibili utilizzi di TA emergono le enormi potenzialità e i relativi benefici sia per il gestore nell'adempimento degli obblighi di condotta previsti sia per il cliente nel poter ricevere un servizio più efficiente e personalizzato. Al contempo però, esistono taluni rischi associati alla qualità dei dati processati. Infatti, costituendo l'elemento indispensabile per la gestione di portafoglio, diviene fondamentale che questi presentino standard qualitativi elevati[98]. Diversamente,

93 Le norme riguardanti la revisione periodica degli strumenti finanziari sono l'art. 16, co. 3, MiFID II e gli artt. 9 e 10 della Direttiva 593/2017/UE (Direttiva Delegata MiFID II).

94 Nello specifico, l'art. 9, co. 14, Direttiva Delegata MiFID II recita: «*Gli Stati membri prescrivono che le imprese di investimento riesaminino regolarmente gli strumenti finanziari da esse prodotti, tenendo conto di qualsiasi evento che possa influire materialmente sui rischi potenziali per il mercato di riferimento determinato. Le imprese di investimento valutano se lo strumento finanziario permanga coerente con le esigenze, le caratteristiche e gli obiettivi del mercato di riferimento e se sia distribuito al mercato di riferimento, o se raggiunga clienti per le cui esigenze, caratteristiche e obiettivi lo strumento finanziario non è compatibile*».

95 Sul punto, l'art. 10, co. 9, Direttiva Delegata MiFID II stabilisce: «*Gli Stati membri assicurano che i distributori forniscano ai produttori le informazioni sulle vendite e, se del caso, le informazioni sui riesami di cui sopra per corroborare i riesami dei prodotti svolti dai produttori*».

96 Per quanto riguarda i *distributors*, l'art. 10, co. 5, Direttiva delegata MiFID II recita: «*Gli Stati membri prescrivono che le imprese di investimento riesaminino regolarmente i prodotti di investimento da esse offerti o raccomandati e i servizi prestati, tenendo conto di qualsiasi evento che possa incidere materialmente sui rischi potenziali per il mercato di riferimento determinato. Le imprese valutano almeno se il prodotto o il servizio resti coerente con le esigenze, le caratteristiche e gli obiettivi del mercato di riferimento e se la prevista strategia di distribuzione continui ad essere appropriata. Le imprese riconsiderano il mercato di riferimento e/o aggiornano i dispositivi di governance dei prodotti qualora rilevino di avere erroneamente identificato il mercato di riferimento per un prodotto o servizio specifico o qualora il prodotto o il servizio non soddisfi più le condizioni del mercato di riferimento determinato, ad esempio quando il prodotto non è più liquido o diviene molto volatile a causa delle oscillazioni del mercato*».

97 Sul tema si rinvia a Mattasoglio, *op. cit.* (2016), 247.

98 Analoghe considerazioni sulla rilevanza della qualità del dato si segnalano da parte di BaFin, *op. cit.* (2018), 53. Nello stesso senso, European Parliament, *Risoluzione del Parlamento europeo del 17 maggio 2017 sulla tecnologia finanziaria: l'influenza della tecnologia sul futuro del settore finanziario*, in www.europarl.europa.eu, 2016, 29; CONSOB (2018), *Il Fintech e l'economia dei dati. Considerazioni su alcuni profili civilistici e penalistici. Le soluzioni del diritto vigente ai rischi per la clientela e gli operatori*, in Quaderni FinTech, 2, www.consob.it, 13.

ne sarebbero compromesse le prestazioni stesse della gestione automatizzata eliminando ogni beneficio illustrato[99].

La qualità del dato dipende da una pluralità di fattori (l'adeguatezza del campione, le modalità di raccolta dei dati e il processo di *data cleaning*[100]) e comporta conseguenze di evidente delicatezza. La scarsa qualità del dato impatta particolarmente sui dati non strutturati, i quali, non solo rappresentano la maggioranza dei dati rientranti nella più ampia categoria di *big data*, ma sono anche ciò che consente di effettuare o verificare una corretta profilazione granulare del cliente.

Al riguardo bisogna rilevare *in primis* che tali dati possono essere affetti dal cosiddetto *selection bias*[101]. A quest'ultimo, poi, si possono affiancare ulteriori *bias* nelle diverse fasi della *big data analytics*[102], con la conseguenza che la gestione di portafoglio sia fondata su dati sempre più distorti.

Il processo descritto derivante da dati compromessi produrrebbe un potenziale danno agli investitori generando scelte fondate su cosiddetti falsi positivi o negativi[103]. Nel primo caso l'investitore non sarebbe inserito all'interno di un profilo che, invece, sarebbe adatto alle sue caratteristiche. All'investitore potrebbe essere impedito di fatto l'acquisto di taluni prodotti finanziari o potrebbe essere offerto un servizio inadeguato[104], andando così incontro a forme di esclusione finanziaria[105]. Nel secondo caso, l'investitore potrebbe essere erroneamente ricondotto a uno specifico mercato di riferimento non adeguato, con la conseguente offerta di prodotti finanziari con rischio superiore rispetto al suo reale profilo.

Alla luce di quanto sopra, diviene evidente la necessità per il gestore di implementare i corretti presidi che garantiscano l'acquisizione e l'analisi di informazioni di qualità elevata per un puntuale adempimento delle regole di condotta, nonché quale

99 Come è stato correttamente evidenziato, «*an algorithm is only as good as the data it works with*», così, Barocas e Selbst, *Big data's disparate impact*, in Cal. L. Rev., 104, 2016, 671; per la dottrina italiana «(p)resupposto necessario ed imprescindibile diviene però la correttezza delle informazioni confluite nell'algoritmo (...), che condizionano inevitabilmente la qualità del servizio offerto», così, Paracampo, M.T. (2017), *La consulenza finanziaria automatizzata*, in *Fintech. Introduzione ai profili giuridici di un mercato unico tecnologico dei servizi finanziari*, Torino, 142; sulla rilevanza con specifico riferimento alle attività del robo advisor, si veda CONSOB, *op. cit.* (2019), 14 e ss..

100 Per una generale comprensione del fenomeno, si rinvia a Cai e Zhu, *op. cit.* (2015).

101 Ci si riferisce all'insufficienza di rappresentatività dei dati rispetto alla popolazione di riferimento; così, IIF (2019), *Bias and ethical implications in machine learning*, in *Machine Learning thematic series* (Part II), 7.

102 Con tale espressione si fa riferimento a una attività strutturata in più fasi. La prima fase consiste nella raccolta dei dati, svolta in prima persona o servendosi di dati raccolti da appositi soggetti; la seconda, ancor più importante, prevede il cosiddetto *data cleaning*, cioè la fase in cui si tenta di eliminare quei dati che immediatamente sono riconosciuti come insufficienti. Solo a questo punto è possibile effettuare il processamento delle informazioni così ottenute.

103 Si veda Schermer, B.W. (2011), *The limits of privacy in automated profiling and data mining*, in Comp. L. & security Rev., 27, 48.

104 Va invero precisato che un tale problema si pone in realtà solo in astratto. Di fatto, risulta estremamente raro che un soggetto possa essere escluso dal perimetro di un servizio come la gestione di portafoglio, vista la varietà di differenti profili di rischio possibili. Ne deriva che difficilmente, il solo l'utilizzo dei *big data* possa condurre a una forma di esclusione, potendo questa esserci solo in compresenza di altri fattori.

105 Sul punto, si rinvia a, CONSOB, *op. cit.* (2019),14-15. Analoghe considerazioni si ritrovano in Paracampo, *op. cit.* (2017), 142 ss.; in ambito internazionale, si veda Financial Services User Group (FSUG, 2016), *Assessment of current and future impact of Big Data on Financial Services*, in www.ec.europa.eu, 14.

presupposto di una corretta costruzione del portafoglio con un livello di rischio coerente alle caratteristiche del cliente. In tal modo, l'utilità di implementare le TA di analisi dei *big data* per il rispetto delle regole di condotta si rifletterà direttamente su una più efficiente *asset allocation*.

Quanto fin qui esposto su benefici e rischi connessi all'utilizzo di TA, che aumentano le risorse in possesso del gestore di portafoglio da utilizzare per la valutazione di adeguatezza e la revisione periodica del prodotto previste dalla disciplina MiFID II, potrebbe chiedere di ridiscutere il contenuto dell'obbligo di diligenza dovuta dall'intermediario.

Si potrebbe infatti immaginare in futuro di ritenere opportuna/necessaria la valutazione degli obblighi di diligenza del professionista alla luce di parametri che tengano conto anche dell'utilizzo o meno dei *big data* nella logica di tenere costantemente conto di tutte le informazioni disponibili sul cliente[106].

Allo stato attuale, invero, il carattere ancora sperimentale delle TA porta a escludere che il loro mancato impiego sia configurabile come negligenza nella prestazione del servizio di gestione e ad attribuire, per converso, mero carattere facoltativo al loro utilizzo.

Per la medesima ragione, in tale ultimo caso sembra valere la regola prevista dalla disciplina del contratto d'opera per le prestazioni che implicano la «*soluzione di problemi tecnici di particolare difficoltà*»: la responsabilità non sussiste «*se non in caso di dolo o colpa grave*» (art. 2236 c.c.).

1.4 Rischi connessi all'uso di *big data* e IA

La qualità dei dati è una questione centrale per l'intermediazione finanziaria. Tuttavia, l'uso della tecnologia dei *big data* e dei suoi derivati (quali, ad esempio, *data analytics*, *machine learning* e *deep learning*) porta con sé problemi sia per l'economia in generale sia per gli investitori in particolare[107].

Innanzitutto, è stato sostenuto che «*l'adozione di modelli di calcolo basati sulla deep learning può rendere il sistema finanziario meno resiliente a scenari di mercato avversi*»[108]. Allo stesso tempo, è stato sottolineato che le caratteristiche dell'intelligenza artificiale potrebbero condurre a rischi di standardizzazione dei comportamenti e di discriminazione nei confronti dei soggetti più deboli[109].

106 In tal senso, un possibile spunto può essere tratto dalle linee guida ESMA, *Guidelines on certain aspects of the MiFID II suitability requirements*, 2018, www.esma.europa.eu, 50; in particolare: «*In order to match clients with suitable investments, firms should establish policies and procedures to ensure that they consistently take into account: all available information about the client necessary to assess whether an investment is suitable, including the client's current portfolio of investments (and asset allocation within that portfolio)*».

107 Per una generale comprensione del fenomeno, si rinvia a Mengoni, M. (2021), *La nuova strategia della Commissione Europea in tema di finanza digitale: quid iuris per i (futuri) servizi finanziari offerti dalle società Tech*, in Paper di Diritto Europeo.

108 Mengoni, *op. cit.* (2021), 120.

109 Invero, l'utilizzazione dell'intelligenza artificiale può condurre ad altri rischi, precisamente di mancanza di trasparenza nel funzionamento degli algoritmi e di violazione della *privacy*.

Per quanto riguarda la standardizzazione, è stato sostenuto che «*l'utilizzazione generalizzata di algoritmi che favoriscono l'assunzione di best practices e/o modelli predittivi che si assomigliano tra loro possono dare luogo ad una sorta di «monocultura»* tra gli intermediari finanziari»[110].

Questa idea di monocultura, che potrebbe essere interpretata come omogeneità di pensiero dei partecipanti al sistema finanziario, potrebbe creare un ambiente favorevole alla pro-ciclicità. La crescente omogeneità di visione, obiettivi e azioni amplifica il potenziale rischio sistemico[111]. Un altro fattore che potrebbe portare a risultati simili a quelli menzionati sarebbe legato ai dati forniti agli algoritmi. In un contesto in cui aziende BigTech, TechFin, fornitori di *cloud computing* sono incentivati a creare un unico data-base, questo ridurrebbe le fonti di dati a cui gli algoritmi avrebbero accesso. La condivisione dei dati porterebbe all'uniformità degli *out-put* e alla probabile omogeneità nel processo decisionale[112].

Inoltre, l'uso di intelligenza artificiale, specialmente nei processi di *profiling* con sistemi di *machine learning*, può determinare varie forme di discriminazione. A prima vista, questo processo potrebbe sembrare efficiente poiché l'obiettivo degli algoritmi di *profiling* è separare le persone in gruppi. Tuttavia, in alcune ipotesi i dati potrebbero fenomeni di esclusione di gruppi sociali ed episodi individuali di discriminazione (ad esempio, per ragioni di sesso, religione o colore della pelle)[113].

Più nello specifico, la discriminazione mediante l'utilizzazione di sistemi di apprendimento automatico potrebbe anche tradursi in esclusione finanziaria. Si può considerare l'ipotesi di un sistema di *machine learning* che non consenta l'apertura di un conto bancario, la concessione di un prestito di denaro o l'accesso a un mutuo per scopi abitativi personali, solo per il fatto che la persona appartiene a un certo gruppo etnico, in violazione dei principi costituzionali comuni degli Stati membri dell'UE[114]. Per ovviare a questi inconvenienti la Direttiva 2008/48/CE[115] ha imposto agli Stati Membri l'obbligo di prevedere nel proprio ordinamento il diritto del consumatore di ottenere informazioni sul rifiuto opposto da un intermediario creditizio a una domanda di credito, specificando gli estremi della banca dati eventualmente consultata da quest'ultimo per la valutazione del merito creditizio.

110 Sul tema si rinvia a Mengoni, *op. cit.* (2021), 121.

111 Danielson, J., Macre, R. e Utheman, A. (2017), *Artificial Intelligence, financial risk management and systemic risk, Systemic Risk Center*, SRC Special Paper n. 13, novembre.

112 Sul tema si rinvia a Mengoni, *op. cit.* (2021), 121-122.

113 Haijan, S. e Domingo Ferrer, J. (2012), *A Methodology for Direct and Indirect Discrimination Prevention in Data Mining*, in *IEEE Transactions on Knowledge and Data Engineering*.

114 Questo sarebbe contrario ai principi di buona fede, non discriminazione ed equità. Al riguardo l'art. 21 della Carta dei diritti fondamentali dell'Unione europea stabilisce il divieto di «*qualsiasi forma di discriminazione fondata, in particolare, sul sesso, la razza, il colore della pelle o l'origine etnica o sociale, le caratteristiche genetiche, la lingua, la religione o le convinzioni personali, le opinioni politiche o di qualsiasi altra natura, l'appartenenza ad una minoranza nazionale, il patrimonio, la nascita, la disabilità, l'età o l'orientamento sessuale*».

115 Direttiva 2008/48/CE del Parlamento europeo e del Consiglio del 23 aprile 2008, relativa ai contratti di credito ai consumatori che abroga la direttiva 87/102/CEE.

Tuttavia, le tutele introdotte dal legislatore italiano, come in altri Stati membri UE, prevedono solo la possibilità per le imprese finanziarie di valutare il merito creditizio degli utenti di servizi finanziari consultando banche dati pubbliche o private che memorizzano situazioni pregresse di individui insolventi. Questa legislazione non prende in considerazione i criteri operativi di prestito attualmente utilizzati dagli algoritmi di *deep learning*. In questo contesto, la sfida normativa è quella di trovare un equilibrio tra la necessità di risultati predittivi efficaci e la necessità di evitare la discriminazione finanziaria contro alcune categorie di individui[116].

Infine, è opportuno rilevare che il continuo flusso di informazioni rende ancora più complesso il fenomeno delle asimmetrie informative. La possibilità per il gestore di analizzare una grande quantità di dati sul cliente aumenta il *gap* informativo tra le due parti del rapporto, problema poi aggravato dalla scarsa trasparenza e conoscibilità della logica sottesa alla attività di analisi dei dati[117].

2 La *governance* degli algoritmi

Il ricorso alle tecniche di automazione per la gestione del portafoglio comporta il rischio di esternalità negative, in particolare nelle forme di rischio sistemico conseguente a errori o *bias* nel *design* o nel funzionamento dei relativi algoritmi[118]. Ne consegue pertanto, la ragionevolezza di un intervento funzionale a prevenire tali rischi e gli eventuali conseguenti danni per l'investitore attraverso una efficiente *governance* degli algoritmi.

2.1 La trasparenza come requisito per un sistema di IA affidabile

La realizzazione di una *governance* efficiente degli algoritmi solleva un'esigenza ineliminabile di trasparenza. In particolare, il corretto funzionamento dei mercati finanziari alimenta la pretesa della conoscenza e della comprensibilità delle caratteristiche degli algoritmi, dei dati utilizzati per il loro funzionamento e dei modelli di *business* in forza dei quali sono stati progettati.

Infatti, la predisposizione di algoritmi trasparenti consentirebbe di perseguire una maggiore tutela dell'investitore mediante una riduzione del rischio connesso all'utilizzazione dei sistemi di IA, e la relativa eliminazione delle difficoltà di individuazione del responsabile per i danni provocati.

116 Sul tema si rinvia a Mengoni, *op. cit.* (2021), 123 – 124.

117 Bogroff e Guégan, *op. cit.* (2019).

118 Un ulteriore rischio derivante da una eccessiva oggettivizzazione che possa mettere in discussione non solo l'autonomia dell'investitore, ma anche la sua stessa personalità, cosiddetta *dignitary rationale* è evidenziato da Kaminsky, M.E. (2019), *Binary governance: lessons from the GDPR's approach to algorithmic accountability*, in S. Cal. L. Rev., 92, 1537-1552. Ad avviso dell'autrice, il principale motivo che giustifica l'intervento sugli algoritmi rimane, però, il rischio di un loro errore (cosiddetto '*instrumental rationale*'). L'autrice sottolinea inoltre come le diverse ragioni sottese non siano indifferenti alle modalità di intervento del legislatore.

2.2 Un possibile approccio: la trasparenza generalizzata...

Un controllo *ex ante* presuppone un approccio di trasparenza alle TA, con la possibilità quindi di 'guardare da dentro' le modalità di funzionamento dell'algoritmo[119]. In particolare, risulta decisivo avere accesso agli algoritmi[120] e al loro *modus operandi* nell'elaborazione dei dati, che costituirà il vero oggetto della *disclosure*[121]. Tale accesso deve essere garantito anche quando i gestori forniscono i propri servizi sulla base di algoritmi acquisiti da società terze, onde evitare un facile aggiramento della disciplina di *disclosure* attraverso il ricorso a *provider* esterni. Sia in quest'ultima ipotesi che in caso di algoritmi elaborati *in house*, infatti, appare decisivo comunicare non solo il nome ma anche, e soprattutto, il codice matematico, le premesse su cui è costruito e i limiti di cui sia eventualmente a conoscenza come, ad esempio, il grado di accuratezza alla luce dei falsi positivi e/o negativi già accertati[122].

Sarà fondamentale altresì poter accedere anche ai dati che vengono utilizzati per far funzionare l'algoritmo[123]. La trasparenza dei dati, infatti, è funzionale a garantire che la loro qualità sia massima, affinché l'*output* prodotto sia altrettanto di valore. In tal senso, dovrebbero prevedersi doveri di comunicazione circa i *dataset* usati e le modalità di aggregazione delle informazioni utilizzate[124]. Inoltre, potendo risultare decisivo l'intervento umano ai fini della aggregazione in un'unica fonte di dati appartenenti a *dataset* differenti, sembra opportuno prevedere protocolli che regolino tali attività al fine di garantire *standard* qualitativi di assoluto livello[125].

Definito il perimetro oggettivo, la trasparenza degli algoritmi pone talune considerazioni sul versante dei soggetti ai quali è destinata tale *disclosure*. Un approccio classico di trasparenza[126] verso tutto il mercato mal si concilierebbe con la nota circostanza che i risparmiatori non sono tipicamente in grado di comprendere informazioni circa gli algoritmi, essendo caratterizzate da una naturale opacità, estrema

119 In tal senso, su tutti si veda Desai, D.R. e Kroll, J.A. (2017), *Trust but verify: a guide to algorithms and the law*, in Harv. J. L. & Tech, Vol. 31, No. 1, 8.

120 In generale, sul tema di ripensare la *disclosure* nel caso in cui i servizi di investimento siano prestati da un consulente automatizzato e non umano, si veda Iannarone, N.G. (2019), *Rethinking automated investment adviser disclosure*, in U. of Tol. L. Rev., 50, 433-445.

121 A tal proposito si veda Sandvig, C., Hamilton, K., Karahalios, K. e Langbort, C. (2014), *Auditing algorithms: research methods for detecting discrimination on internet platforms*, in *Annual Meeting of the International Communication Association*, 9 e ss..

122 Zarsky, T. (2013), *Transparent Predictions*, in University of Illinois Law Review, Vol. 13, No. 4, 1524; Colaert, V. (2017), *RegTech as a response to regulatory expansion in the financial sector*, www.ssrn.com, 18-19. L'autrice tratta nello specifico del RegTech, tuttavia, il discorso può essere applicato anche nel caso del FinTech poiché si riferisce alla base tecnica e algoritmica che è indifferente al tipo di utilizzo che ne viene fatto.

123 In tal senso Pasquale, F. (2015), *The black-box society*, Harvard, 141, sostiene che «*(d)ata is the fuel of the information economy, and the more data a company already has, the better it can monetize it*».

124 Zarsky, *op. cit.* (2013), 1523. Tutto questo risulta a maggior ragione rilevante con riferimento ai dati utilizzati per consentire l'apprendimento al sistema di intelligenza artificiale, specie nelle forme del *machine learning*. In tal caso infatti, i dati sono ancor più importanti poiché servono all'algoritmo per migliorare le proprie performances e ne deriva che la loro eventuale scarsa qualità inciderebbe in modo negativo sull'efficienza di quest'ultimo.

125 Sulla rilevanza dell'intervento dell'uomo nella fase di raccolta dei dati e conseguenti proposte di protocolli si veda Zarsky, *op. cit.* (2013), 1524.

126 Per un inquadramento generale sulla funzione della trasparenza nel mercato e, sull'oggettività dei limiti connessi alla efficient capital markets hypothesis, si veda Perrone, *op. cit.* (2018), 67 e ss..

tecnicità e notevole complessità. Di fatto, alla luce delle necessarie competenze tecniche richieste, si risolverebbe in una forma di tutela 'vuota di contenuto' per il cliente.

Analoghe perplessità su una *disclosure* generalizzata deriverebbero dal conflitto con il diritto di proprietà intellettuale dello sviluppatore dell'algoritmo e dal rischio di utilizzo di tali informazioni in maniera distorta o anticoncorrenziale (per ulteriori dettagli si rimanda alla Sezione '*Profili di* investor protection').

Ne consegue un necessario ripensamento del perimetro soggettivo di *disclosure* al fine di contenere costi e limiti dell'approccio classico. In quest'ottica, appare più opportuno un meccanismo di *enforcement* di natura pubblica, in forza del quale le informazioni siano destinate a uno specifico e qualificato soggetto e non all'intero mercato.

2.3 ...(segue) La trasparenza selettiva

La soluzione più equilibrata sembra essere una trasparenza qualificata, vale a dire una trasparenza selettiva[127] da parte dei gestori e dei *provider* nei confronti dell'Autorità di vigilanza. Solo un soggetto pubblico, infatti, potrebbe contemperare al meglio le opposte esigenze di *disclosure* e di segretezza, così risolvendo anche i limiti relativi a entrambi gli oggetti della trasparenza (ossia, i dati e l'algoritmo).

Una generalizzata possibilità di accesso ai *dataset* porterebbe a rischi connessi alla tutela della *privacy*[128]. Se, infatti, le informazioni dovessero essere rivelate ai *competitor* e a tutti gli investitori, il diritto alla riservatezza dei dati, specie quelli personali, sarebbe del tutto ignorato e privato di qualsiasi garanzia. La soluzione proposta, invece, avrebbe il pregio di dare tale accesso solo all'Autorità pubblica, così ridimensionando il rischio in questione poiché il soggetto pubblico non ha specifici interessi commerciali sui dati degli investitori[129].

Sotto il diverso profilo dell'algoritmo, invece, un approccio di trasparenza selettiva sembra risolvere, da un lato, il problema del cosiddetto '*gaming the system*' e, dall'altro, quello relativo alla opacità connaturata agli algoritmi.

Con l'espressione '*gaming the system*' si intende quella situazione per cui le regole originariamente pensate per regolare un meccanismo vengono sfruttate al fine di evadere il meccanismo stesso[130]. Con riferimento, ad esempio, alla distribuzione di

127 Si veda Pasquale, *op. cit.* (2015), 142, laddove afferma che «*I call this general trend 'qualified transparency'—limiting revelations in order to respect all the interests involved in a given piece of information*»; Pasquale, F. (2010), *Beyond innovation: the need for qualified transparency in internet intermediaries*, in NW U. L. Rev., 104, 160 e ss.; Colaert, *op. cit.* (2017), 19; Desai e Kroll, *op. cit.* (2017), 40-41.

128 «*A fully transparent society would be a nightmare of privacy invasion*», così Pasquale, *op. cit.* (2015), 142.

129 Così De Laat, P.B. (2018), *Algorithmic Decision-Making Based on Machine Learning from Big Data: Can Transparency Restore Accountability?*, in *Philosophy & Technology*, 31, 533-534. L'autore rimarca che, in una società, in cui già si verificano fenomeni di cessione e acquisto dei dati, è fondamentale, al fine di garantire responsabilità, restringere la platea di chi ne ha accesso, a chi può avere un potere di supervisione in merito.

130 Una possibile ipotesi si avrebbe nel caso in cui venissero rivelate a tutti i contribuenti le principali red flags utilizzate dalle agenzie di riscossione per individuare i soggetti con un profilo a rischio evasione. Tale *disclosure* permetterebbe a questi ultimi di adattare la propria condotta alla luce di tali indicatori eludendo i controlli.

prodotti finanziari, in caso di *disclosure* generalizzata, gli investitori conoscendo il meccanismo di funzionamento di un algoritmo potrebbero sfruttare le informazioni pubbliche per 'ingannare' l'intelligenza artificiale, al fine di ottenere ad esempio strumenti finanziari a cui non potrebbero normalmente avere accesso alla luce delle regole in tema di adeguatezza. Tale rischio è solo apparentemente secondario perché, alla luce di quello che è di fatto un falso positivo, strumenti come il *machine learning* potrebbero standardizzare quella soluzione con gravi rischi in chiave sistemica per tutti gli altri investitori. Al contrario, una *disclosure* selettiva riduce la possibilità di ingannare l'intelligenza artificiale[131], poiché non viene fornito all'investitore il *set* di regole da sfruttare a suo vantaggio[132].

Con riferimento all'opacità e complessità tecnica, si evidenzia che l'Autorità di vigilanza possiede le competenze sufficienti, quantomeno in astratto[133], per poter comprendere da un punto di vista squisitamente tecnico le informazioni circa il *modus operandi* degli algoritmi. In tal senso, la trasparenza sarebbe solo un primo passo per verificare in concreto il rispetto degli obblighi normativi, con il cosiddetto *algorithms auditing*[134], al fine di garantire una migliore gestione dell'algoritmo e, quindi, di conseguenza, una più efficace tutela del risparmiatore.

La trasparenza selettiva inoltre consentirebbe di risolvere anche il problema relativo alla tutela dell'algoritmo come proprietà intellettuale[135], poiché l'Autorità di vigilanza, destinataria delle informazioni, agendo nel suo ruolo di vigilante e non di *competitor*, non ha alcun diretto interesse a uno sfruttamento opportunistico e concorrenziale[136].

Per le ragioni sopra esposte, la trasparenza selettiva rappresenta probabilmente il miglior punto di partenza possibile nell'ottica di costruire un sistema di *governance* degli algoritmi. In particolare, se da un lato bisognerà tener conto dei potenziali limiti attuali, dall'altro è lecito aspettarsi che, in futuro, possano subentrare nuove

131 «*(...) another approach to curb gaming the system can be implemented any time: preventing the disclosure of said proxies to potential gamers. The public at large is no longer to be a beneficiary of transparency concerning the algorithm in use. Of course, individuals subjected to a decision from one algorithm or another should keep the option to receive a transparency report about that particular decision. That might give him/her some clues for future gaming, but only to a very limited extent. Obviously, intermediate oversight authorities are to retain full rights of transparency as far as the model and its proxies are concerned, otherwise all accountability is gone*» così De Laat, *op. cit.* (2018), 535.

132 Nel caso del servizio di consulenza, invero, la soluzione proposte non esclude in radice il rischio che l'investitore possa comunque fornire informazioni non veritiere al fine di ottenere un certo tipo di portafoglio.

133 La precondizione affinché ciò si verifichi è una integrazione di competenze informatiche e ingegneristiche nel personale dell'Autorità di vigilanza.

134 Tale attività consiste proprio in una valutazione di conformità dell'algoritmo alla eventuale normativa esistente o agli standard approvati. Ad oggi si ritiene però che un tale meccanismo non sia ancora sufficiente per garantire la piena comprensione, si veda Kroll, J.A., Huey, J., Barocas, S., Felten, W., Reidenberg, J.R., Robinson, D.G. e Yu, H. (2017), *Accountable algorithms*, in U. of Pa. L. Rev., 165, 660 e ss.

135 Circa il riconoscimento dell'algoritmo come proprietà intellettuale si veda, Pasquale, F. (2011), *Restoring Transparency to Automated Authority*, in *J. on Telecomm. and High Tech. L.*, 237; Sandvig et al., *op. cit.* (2014), 9.

136 In tal senso, tra tutti De Laat, *op. cit.* (2018), 536, sostiene l'utilità di prevedere un vero e proprio 'deposito in garanzia' (cosiddetto *escrow*), al quale avrebbero accesso solo le autorità competenti al fine di garantire la segretezza, in cui vengono depositati gli algoritmi, si veda Prosperetti, E. (2018), *Algoritmi dei big data: temi regolamentari, responsabilità, concorrenza*, in Falce, V., Ghidini, G. e Olivieri, G. (a cura di), *Informazione e Big Data tra Innovazione e Concorrenza*, Milano, 317.

modalità associate alla *computer science* in grado di garantire una gestione ancora migliore dell'intelligenza artificiale e dei suoi continui sviluppi[137].

2.4 La vigilanza pubblica...

Dalle considerazioni esposte sull'approccio di cosiddetta trasparenza selettiva, appare opportuno individuare nell'Autorità di vigilanza il fulcro di tale regime, potendo questa coniugare le esigenze di trasparenza con quelle di segretezza[138]. Allo stesso tempo, all'Autorità sarebbe affidato anche il compito più generale di sostenere lo sviluppo tecnologico tenendo adeguatamente in considerazione i relativi benefici e costi.

Per conseguire tali obiettivi sono ipotizzabili diversi approcci di intervento dell'Autorità pubblica, con un maggiore o minore grado di pervasività. Si potrebbe ipotizzare una cosiddetta *hard regulation*, in forza della quale l'Autorità esercita il potere di autorizzazione degli algoritmi[139]. Tale soluzione sembra prestare il fianco a diversi rilievi. Da un lato si rischierebbe di ostacolare l'innovazione tecnologica, poiché gli attori del mercato difficilmente investirebbero in maniera cospicua per realizzare tecnologie sempre più sofisticate, non avendo *ex ante* la certezza dell'autorizzazione dell'algoritmo[140]. D'altra parte, l'Autorità non sarebbe esente dal rischio di errori in fase di autorizzazione e approvazione di algoritmi caratterizzati da *bug* o *bias* tecnologici. Inoltre, vi sarebbe il rischio di standardizzare l'eventuale errore non individuato una volta che l'algoritmo sia stato approvato. Non da ultimo, infine l'Autorità sarebbe direttamente responsabile degli eventuali danni generati da un algoritmo autorizzato nonostante i *bug*, anche in considerazione di una eventuale portata sistemica.

137 Si veda Kroll et al., *op. cit.* (2017), 659-671. Gli autori, in particolare, forniscono alcuni esempi sia dei limiti tipici dell'approccio di trasparenza (evidenziando, ad esempio, la circostanza che forme come il *machine learning* cambiando nel tempo impediscono una valutazione definitiva), che i possibili vantaggi derivanti dalle nuove tecniche della computer science. Tra queste ultime, su tutte, la cosiddetta z*ero-knowledge proof*, la quale consente di provare che è stata seguita una specifica politica decisionale ma senza dover rivelare quale sia quella stessa politica. Circa le potenzialità della *computer science*, si veda Desai e Kroll, *op. cit.* (2017), 35-42; sulla necessità di andare oltre la trasparenza, si veda Perel, M. ed Elkin-Koren, N. (2017), *Black box tinkering: beyond disclosure in algorithmic enforcement*, in Fl. L. Rev., 69, 186-197.

138 In tal senso, si veda De Laat: «Obviously, intermediate oversight authorities are to retain full rights of transparency as far as the model and its proxies are concerned, otherwise all accountability is gone», De Laat, *op. cit.* (2018), 535. Dello stesso tenore, si veda Pasquale: «Sometimes the route to orderly and productive investigation is to entrust the job to a small group of experts (...). I call this general trend 'qualified transparency'—limiting revelations in order to respect all the interests involved in a given piece of information», Pasquale, *op. cit.* (2015), 142. Nello stesso senso, si veda anche Pasquale, *op. cit.* (2010), 161; Kroll et al., *op. cit.* (2017), 641.

139 Così Tutt, A. (2017), *An FDA for algorithms*, in Admin. L. Rev., 69, 111; Posner, E.A. e Weyl, E.G. (2013), *An FDA for financial innovation: applying the insurable interest doctrine to twenty-first-century financial markets*, in *NW U. L. Rev.*, 107, 1348 -1351. I prodotti finanziari potrebbero essere trattati come i farmaci, pertanto, al pari di questi ultimi, devono essere approvati da parte di chi sia in grado di ragionare secondo criteri di esperienza e di tutela della collettività, senza basarsi sulle preferenze di ciascuno. Sulla possibilità per l'autorità di sviluppare essa stessa un algoritmo o prodotti RegTech e per esteso anche prodotti FinTech sul presupposto che essa disponga di risorse competenti in materia, si rinvia a Enriques, L. (2017), *Financial supervisors and RegTech: four roles and four challenges*, in Revue Trimestrielle de Droit Financier, n. 53, 5.

140 Vi sarebbe infatti, il forte rischio che gli investimenti, funzionali allo sviluppo di servizi automatizzati sempre più evoluti, si traducano in un nulla di fatto a causa della mancata approvazione da parte dell'Autorità, qualora questi non soddisfino i requisiti previsti, si veda Posner e Weyl, *op. cit.* (2013), 1354-1355.

Passate in rassegna le criticità connesse a un approccio di *hard regulation*, sul versante opposto, si potrebbe far ricorso alla cosiddetta *self regulation*, caratterizzata dalla totale assenza di interventi pubblicistici. Anche tale soluzione, che presenta indubbi vantaggi in termini di stimolo all'innovazione tecnologica ed efficienza operativa, non sembra essere esente da potenziali criticità. Il principale rischio consiste nel fatto che la corsa allo sviluppo tecnologico potrebbe verificarsi a scapito dei fondamentali valori e principi che regolano il mercato, come la tutela del consumatore e la stabilità dei mercati[141]. Ulteriore problema sarebbe poi l'assenza del ruolo di un'Autorità con il compito di rendere effettivo il sistema di *enforcement* delle regole che gli stessi soggetti privati si sono dati[142]. A ciò si aggiunge, infine, il rischio che un sistema di *laissez faire* possa facilitare il consolidamento del potere di un soggetto che già gode di una posizione dominante all'interno del mercato di riferimento considerato[143].

I limiti esposti per una *hard* o una *self regulation* portano a ritenere preferibile un approccio mediano, che possa consentire di coniugare i benefici insiti a tali ipotesi limitando al contempo i suddetti rischi. L'Autorità agirebbe, quindi, in qualità di *soft regulator*[144], esercitando poteri di vigilanza prudenziale con possibilità di introdurre forme, appunto, di *soft regulation*[145], come ad esempio i cosiddetti *safe harbor* nelle forme di codici di condotta[146]. Tali strumenti, prevedendo una regolamentazione flessibile, in quanto non fondata su prescrizioni puntuali, consentono all'Autorità di indirizzare i vigilati al rispetto dei principi essenziali e di prevenire i principali rischi delle condotte operative.

2.5 ... (segue) e i limiti nelle risorse umane ed economiche

La bontà della soluzione proposta si fonda sull'assunto che l'Autorità disponga di risorse in grado di comprendere al meglio le informazioni circa l'algoritmo e sul suo funzionamento. Al riguardo, va rilevato che la complessità e la tecnicità intrinseca agli algoritmi e all'intelligenza artificiale postulano un ripensamento sul tipo di competenze richieste all'interno delle Autorità di vigilanza[147].

141 Un privato infatti normalmente agisce, in primis per un proprio interesse e, solo eventualmente, per un più ampio interesse pubblicistico. In tal senso, dunque, si metterebbe a rischio non solo la tutela dell'investitore ma anche la stessa stabilità dei mercati.

142 Si veda Kaminsky, *op. cit.* (2019), 1561.

143 In questo modo, Canepa, A. (2017), *L'era delle piattaforme fra opportunità e rischi*, in Paracampo, M.T. (a cura di), *op. cit.*, Torino, 61. Ad avviso dell'autrice un tale problema sarebbe tale da annullare totalmente i vantaggi propri della self-regulation.

144 Per capire il ruolo, si consideri la seguente definizione di soft regulation: «regulations that are low enough cost that they preserve freedom of choice and do not substantively limit the kinds of algorithms that can be developed or when or how they can be released», così Tutt, *op. cit.* (2017), 109.

145 Sull'opportunità di un tale approccio, anche per i *robo advisors*, si rinvia a CONSOB, *op. cit.* (2019).

146 Sull'ipotesi di definire anticipatamente una vera e propria carta di principi che indichi le tipologie di algoritmi o le loro alterazioni non desiderabili, si veda Prosperetti, *op. cit.* (2018), 317.

147 In merito, si rinvia al Discorso del Presidente al mercato finanziario, Incontro annuale con il mercato finanziario, 8 maggio 2017, Milano, www.consob.it), 21, nel quale è stato sostenuto che «(*l*)*e «Le stesse Autorità di controllo dovranno adattare al nuovo contesto strutture operative e metodi di lavoro. Nella CONSOB di domani ci sarà bisogno di più ingegneri e meno avvocati*».

In tal senso, sarà quindi fondamentale investire nel fattore umano perché è il miglior 'strumento' per gestire le innovazioni tecnologiche[148], affiancando a funzionari esperti nelle materie economiche e giuridiche anche personale qualificato in ambito informatico e ingegneristico[149].

Il percorso da intraprendere consisterebbe quindi nel formare professionisti all'interno e, ove necessario, poterli assumere dall'esterno. In tale ultima ipotesi però, la vera sfida consiste nella capacità di attrarre qualificati professionisti delle realtà private nel mondo delle Autorità governative, alla luce delle differenti capacità e/o possibilità di spesa tra soggetti privati e pubblici, ad appannaggio dei primi, circostanza che rappresenta un importante ostacolo per il conseguimento di tale obiettivo[150]. Il rischio è, quindi, quello di dover svolgere una complessa Attività di vigilanza ad ampio spettro senza avere i necessari mezzi per farlo in modo adeguato[151].

2.6 La collaborative governance

Una possibile soluzione con cui l'Autorità può mitigare gli effetti della limitatezza di risorse umane ed economiche consiste nella instaurazione di un regime di cosiddetta '*collaborative governance*'[152]. Tale regime si fonda su tre elementi essenziali: la compresenza di autorità governative e realtà del mondo privato; il ruolo attivo, e non solo di consulenza, riconosciuto ai soggetti privati e infine il funzionamento per consenso alla realizzazione di obiettivi e/o programmi pubblicistici.

Una siffatta collaborazione avrebbe importanti benefici a cominciare da una percezione di maggiore legittimità dei futuri obblighi normativi da parte dei soggetti vigilati, avendo questi ultimi collaborato alla loro redazione. L'assenza di imposizione, in questo senso, gioca un ruolo fondamentale per la realizzazione dell'obiettivo della *compliance* da parte dei singoli attori del mercato, agevolando, quindi, anche i meccanismi di *enforcement*[153].

148 Si veda Panisi, F. e Perrone, A. (2018), *Systems So Perfect That No One Wil Need to Be Good? RegTech and the 'Human Factor'*, in Orizz. dir. comm., n. 2, 9.

149 Si veda sul punto Panisi e Perrone, *op. cit.* (2018), 9; Enriques, *op. cit.* (2017), 6.

150 In tal senso Panisi e Perrone, *op. cit.* (2018), 10; Enriques, *op. cit.* (2017), 6; in senso contrario, sulla capacità delle autorità governative di attrarre esperti, Posner e Weyl, *op. cit.* (2013), 1353.

151 Per utilizzare le parole di Bréhier «*we have the feeling that market regulators are akin to police or customs officers driving an Austin Mini in pursuit of drug traffickers driving German sports cars*» così Bréhier, B. (2013), *High frequency trading: should technological developments be considered a potential threat to financial markets and be subject to specific regulation?,* in *ERA forum*, 78.

152 Con tale espressione si intende: «*A governing arrangement where one or more public agencies directly engage non-state stakeholders in a collective decision-making process that is formal, consensus-oriented, and deliberative and that aims to make or implement public policy or manage public programs or assets*». Così Ansell, C. e Gash, A. (2012), *Collaborative governance in theory and practice*, in *J. of Publ. Adm. Res. and Th.*, 544. Un'altra possibile definizione è quella fornita da Kaminsky: «*Collaborative governance, or 'new governance' deploys private- public partnerships towards public governance goals*», Kaminsky, *op. cit.* (2019),1559.

153 In questo senso, Kaminsky, *op. cit.* (2019), 1562; in riferimento ad altri ambiti, si veda, tra tutti, Rasche, A. (2010), *Collaborative governance 2.0*, www.ssrn.com, 4; Albareda, L. (2008), *Corporate responsibility, governance and accountability: from self-regulation to coregulation*, in Corporate governance International Journal of Business in Society, 8, 437.

In secondo luogo, la collaborazione porterebbe un beneficio importante anche per la stessa Autorità che, attraverso la partecipazione di un maggior numero di esperti del mondo privato nella attività regolamentare e di governo, potrebbe consolidare e perfezionare le proprie competenze[154]. Tale considerazione di principio è valida *a fortiori* nel campo dell'intelligenza artificiale, atteso che gli algoritmi sono caratterizzati sia dal continuo impulso al cambiamento e allo sviluppo sia da un alto livello di tecnicità[155].

Come si evince da quanto illustrato, l'approccio di *collaborative governance*, instaurando di fatto una *partnership* tra privati e autorità, può essere un valido strumento con cui mitigare i limiti di risorse umane ed economiche tipici di una autorità di vigilanza pubblica e realizzare così un meccanismo virtuoso che giochi a favore di tutti i partecipanti[156]. La collaborazione e lo scambio di competenze infatti consentirebbero non solo di stare al passo con gli sviluppi tecnologici ma anche di trovare le soluzioni più efficienti e adeguate ai nuovi problemi che, di volta in volta, potrebbero emergere nella prassi operativa[157].

2.7 La trasparenza nella proposta UE in materia di intelligenza artificiale

La trasparenza è altresì uno dei valori fondamentali promossi dall'UE per lo sviluppo, la diffusione e l'uso dei sistemi di IA. Dall'inizio del processo politico per la regolamentazione dell'IA, tutti i documenti ufficiali delle istituzioni dell'Unione europea hanno promosso la trasparenza quale principio guida nella disciplina dell'utilizzazione dei sistemi di IA[158]. Anche la proposta di Regolamento UE in materia di intelligenza artificiale ruota intorno al principio di trasparenza[159]. In particolare, la proposta definisce concetti chiave in materia di intelligenza artificiale e suddivide i sistemi di IA

154 Si veda Hirsch, D.D. (2011), *The Law and Policy of Online Privacy: Regulation Self-Regulation, or Co-Regulation?*, in Seattle U. L. Rev., vol. 34, n. 2, 467; Kaminsky, *op. cit.* (2019),1562. Tra gli altri vantaggi entrambi gli autori segnalano il miglioramento nelle tecniche di controllo e di vigilanza sulla compliance e il più facile adeguamento a una realtà, come quella delle tecnologie, in continua evoluzione.

155 Sul punto si veda, Kaminsky, *op. cit.* (2019), 1570.

156 Il funzionamento di un tale meccanismo si fonda sulla necessità di sviluppare strumenti di dialogo tra l'autorità e i soggetti vigilati. Tra questi possono segnalarsi, i codici di condotta condivisi, figure professionali e indipendenti che facciano da intermediari rispetto alla compliance degli algoritmi, report periodici sulle *performance* da indirizzare all'autorità, tutela del *whistle-blower*. In merito, si veda Perrone, *op. cit.* (2021), 718-719.

157 Così, Rasche, *op. cit.* (2010), 4 ss.; per altri ambiti quale quello ambientale, si veda, tra tutti, Börzel, T.A. e Risse, T. (2005), *Public-Private Partnerships: Effective and Legitimate Tools of International Governance?*, www.researchgate.com, 14. In senso parzialmente contrario, c'è chi ritiene che il continuo sviluppo tecnologico potrebbe invece mettere in discussione questo vantaggio perché occorrerebbe continuamente rideterminare gli obiettivi dell'intervento e, qualora cambiassero anche i soggetti della collaborazione, sarebbe necessario rimettere in discussione ogni aspetto fino a quel momento concordato, così Huxhman, C. e Vangen, S. (2000), *The Challenge of Collaborative Governance: Public Management an International Journal of Research and Theory*, in Public Management Rev., 345.

158 Tra questi possono essere citati i seguenti documenti: le linee guida etiche per l'IA degna di fiducia emesse dal gruppo di esperti di alto livello sull'IA nel dicembre 2018, il Libro bianco sull'IA pubblicato dalla Commissione europea del 19 febbraio 2020 e la Risoluzione del Parlamento europeo recante raccomandazioni alla Commissione concernenti il quadro sugli aspetti etici dell'IA, della robotica e delle tecnologie correlate dal 20 ottobre 2020.

159 Le norme sulla trasparenza sono espressamente incluse in due articoli della proposta: l'art. 13 in tema di trasparenza e fornitura di informazioni agli utenti e l'art. 52 in tema di obblighi di trasparenza per determinati sistemi di IA.

sulla base del rischio potenziale per le persone (*risk-based approach*), dal quale derivano divieti di operatività dei sistemi di IA e differenti obblighi di trasparenza.

Innanzitutto, i sistemi a rischio inaccettabile sono espressamente vietati. Gli altri due sistemi (rischio elevato e rischio limitato), invece, sono sottoposti a regole in tema di trasparenza generalizzata e trasparenza selettiva.

Più nello specifico, con riferimento ai sistemi ad alto rischio, la proposta imputa la responsabilità in capo agli sviluppatori dei sistemi di IA (art. 13), i quali devono predisporre un livello 'adeguato' di trasparenza, ancorché la medesima proposta non precisi cosa debba intendersi per 'adeguato'. A tale riguardo, è predisposta una disciplina relativa ai dati e alla *governance* dei dati per i sistemi ad alto rischio che usano banche dati di informazione, con l'elencazione delle pratiche da seguire per l'addestramento, la convalida e la prova dei *set* di dati (art. 10, par. 2) e l'indicazione di criteri di pertinenza, rappresentatività, completezza e correttezza dei dati (art. 10, par. 3).

I sistemi di IA ad alto rischio devono contenere le informazioni tecniche prima che siano immessi sul mercato o messi in servizio: le informazioni devono essere indicate in modo tale che il sistema sia conforme al regolamento (art. 11) e consenta la registrazione automatica di tutti gli eventi una volta entrato in funzione (art. 12). Allo stesso tempo, questi sistemi devono essere previamente approvati e registrati da parte dell'autorità di vigilanza prima dell'immissione sul mercato (art. 40-51) e devono essere progettati e sviluppati in modo tale da garantire la supervisione e il monitoraggio umano durante la sua utilizzazione (art. 14).

Viceversa, i sistemi a basso rischio sono disciplinati con una regolamentazione meno puntuale e incisiva. È previsto unicamente un obbligo di comunicazione qualora le persone si relazionino con i sistemi di IA e le loro emozioni o caratteristiche vengono riconosciute attraverso processi automatizzati (art. 52).

Oltre agli ampi obblighi imposti per lo sviluppo, la distribuzione e l'uso di sistemi di IA, la proposta contiene una serie di misure volte a sostenere l'innovazione, proponendo la creazione un quadro normativo uniforme in materia di intelligenza artificiale[160]. Con riferimento al settore finanziario, occorre chiedersi quali siano le ricadute per gli intermediari finanziari. I sistemi di IA, qualificati come sistemi ad alto rischio (art. 6, par. 2), potrebbero essere utilizzati per la valutazione del merito creditizio nei confronti delle persone fisiche (All. III, punto 5, lett. b). Dal canto loro, gli intermediari finanziari potrebbero essere chiamati a rispondere in solido con gli sviluppatori in quanto eserciterebbero una funzione di controllo e di monitoraggio su quest'ultimi (art. 29), anche mediante una valutazione di pertinenza dei dati inseriti rispetto alla finalità prevista (art. 29, par. 3). Saranno quindi obbligati a monitorare il funzionamento di

160 La proposta prevede l'attivazione delle *regulatory sandbox* per incentivare nuove modalità di interazione tra imprese digitali e autorità competenti e sviluppare e convalidare sistemi di IA innovativi. Tale strategia non è del tutto inedita in ambito UE. Un precedente è la sperimentazione avente per oggetto il regime pilota per la negoziazione degli strumenti finanziari nell'attività di infrastrutture di mercato basate su registri elettronici distribuiti. In particolare, l'art. 53 prevede che una o più autorità competenti degli Stati membri o il Garante europeo della protezione dei dati possano istituire un ambiente controllato che facilita lo sviluppo, le prove e la convalida di sistemi di IA innovativi per un periodo di tempo limitato prima della loro immissione sul mercato o della loro messa in servizio conformemente a un piano specifico.

questi sistemi, così come a informare lo sviluppatore o il distributore se sospettano di un malfunzionamento del sistema, nonché obbligati a conservare i *log* generati automaticamente dal sistema (art. 29, parr. 4 e 5). Invero, a norma della proposta di Regolamento questi obblighi dovrebbero considerarsi soddisfatti se è rispettata la disposizione dell'art. 74 della Direttiva 2013/36/UE per quanto riguarda le norme sui dispositivi, sui processi e sui meccanismi di *governance* interna[161].

In definitiva, la proposta ambisce a realizzare sistemi di IA più trasparenti. Indubbiamente, solleva numerose critiche, in particolare in materia di tutela della proprietà intellettuale; ciononostante, sembrerebbe che costituisca un equilibrato compromesso tra le restrizioni al diritto alla libertà d'impresa, alla libertà artistica e scientifica e le esigenze di salvaguardia della salute, della sicurezza e della protezione dei consumatori.

2.8 Risultati della *survey*

Nonostante la proposta di Regolamento in materia di intelligenza artificiale sia ancora in fase di approvazione, la *governance* degli algoritmi costituisce un argomento in agenda degli intermediari finanziari.

La *survey* ha dimostrato che l'utilizzazione dei sistemi di IA non avviene fuori da qualsiasi coinvolgimento umano. Quattro delle società coinvolte nell'indagine hanno rappresentato che la supervisione umana è intensa in quanto il sistema di IA non è autonomo nel processo decisionale. Due hanno riconosciuto che il sistema di IA utilizzato è parzialmente autonomo con attribuzione della decisione finale all'uomo; un'altra società, invece, ha rilevato che vi è solo un controllo umano sul sistema di IA nella decisione finale[162].

Per quanto riguarda il trattamento dei sistemi di IA i risultati dell'indagine hanno dato riscontri diametralmente opposti. Da un lato, tre delle otto società di gestione hanno sostenuto che i processi decisionali basati su sistemi di IA non dovrebbero essere trattati diversamente da quelli che utilizzano il solo fattore umano. D'altra parte, altre cinque società hanno manifestato la necessità che il controllo umano sia sempre necessario e il monitoraggio sia continuo.

161 In particolare, la proposta di Regolamento dispone che gli enti creditizi possano adottare le misure previste dalla disciplina di settore (e, precisamente, la direttiva 2013/36/UE, Capital Requirements Directive 4 – CRD4) per adempiere agli obblighi imposti a loro carico quali fornitori in ordine al sistema di gestione di rischi e di qualità, alla conservazione di documenti e registri, al monitoraggio e alla notifica di incidenti relativi ai sistemi di intelligenza artificiale ad alto rischio. Sennonché, la proposta non fa riferimento agli altri intermediari autorizzati a concedere previsti (in Italia, gli intermediari finanziari iscritti all'albo ex art. 106 TUB), che non rientrano nella definizione di ente creditizio bensì in quella di ente finanziario. Tale limitazione dell'ambito di applicazione soggettivo potrebbe alterare la concorrenza fra intermediari finanziari e intermediari creditizi e aprire alla possibilità di arbitraggi regolamentari. Si veda Sepe, M. (2021), *Innovazione tecnologica, algoritmi e intelligenza artificiale nella prestazione dei servizi finanziari*, in Riv. trim. dir. ec., 3, Suppl., 186 ss.

162 Per quanto riguarda il trattamento del processo decisionale dei sistemi di IA, tre su otto società consultate hanno ritenuto che non dovrebbe essere diverso da quello dato ai processi che coinvolgono solo l'intervento umano. Tuttavia, altre cinque società consultate hanno indicato motivazioni legate alla necessità che un controllo umano sia sempre necessario o che il monitoraggio continuo e la validazione dei modelli alla base dei sistemi di IA, come parte integrante del processo decisionale basato su tali sistemi, richiede processi decisionali o dedicati e/o processi di natura diversa.

Ciò che più rivela la *survey* è la circostanza che le società consultate abbiano già considerato la questione dell'*algo-governance*. Cinque delle otto società hanno già utilizzato procedure specifiche di *governance* dei dati, mentre le altre hanno dichiarato di non averle ancora attivate.

3 Ruolo dei *provider*s: profili di responsabilità dello sviluppatore

3.1 I possibili fatti dannosi

Le tecniche di automazione utilizzate nella gestione di portafoglio possono presentare taluni vizi estrinseci e intrinseci in grado di cagionare un danno nei confronti degli investitori. Gli eventuali rimedi civilistici *ex post* richiedono dunque di identificare i fatti generatori della responsabilità e il relativo criterio di imputazione

Una prima categoria di eventi dannosi in grado di inficiare il corretto funzionamento delle TA riguarda l'utilizzo di dati di cattiva qualità[163] nelle diverse fasi di operatività dei sistemi di intelligenza artificiale. Innanzitutto, qualora lo sviluppatore utilizzi dati qualitativamente scarsi durante la fase di addestramento[164] ne deriverebbe un algoritmo incapace di elaborare un modello idoneo a fornire gli *output* desiderati. Analogamente, se il gestore fornisse a un sistema intelligente, perfettamente addestrato, informazioni di scarsa qualità, l'*output* ottenuto dall'elaborazione di tali *input* sarebbe erroneo in quanto si fonderebbe su presupposti (i dati) inaffidabili. Diversa da tali ipotesi, è la possibilità che l'algoritmo elabori una scelta di investimento erronea per una inadeguata attività di auto-apprendimento svolta attraverso una fallace acquisizione di dati dall'esperienza. In quest'ultimo caso, è stato rilevato che lo sviluppatore potrebbe, in fase di ideazione dell'algoritmo, gestire il rischio operativo attraverso specifiche cautele, tra cui la possibilità di inserire nel sistema «blocchi di sicurezza idonei a 'impedire' al prodotto intelligente di porre in essere determinate condotte»[165].

Come si evince dalle ipotesi descritte, l'affidabilità dei dati analizzati è un fattore decisivo per una efficiente gestione. Ne consegue che i soggetti coinvolti nella fattispecie saranno tenuti a implementare sistemi informatici in grado di prevenire e mitigare gli effetti di possibili attacchi informatici capaci di pregiudicare la qualità dei dati.

La seconda categoria di fatti dannosi riguarda l'erronea ideazione e/o progettazione dell'algoritmo. In tal senso, una prima ipotesi desumibile dalla definizione

163 Sulla qualità dei dati, si vedano a Cai e Zhu, *op. cit.* (2015), 2 e ss. e CONSOB, *op. cit.* (2018), 14 e ss..

164 L'addestramento (*training*) del *machine learning* è l'attività con cui lo sviluppatore fornisce all'algoritmo un set di regole al fine di svolgere l'attività di apprendimento autonomo.

165 Sul punto, Amidei, A. (2019), *Intelligenza artificiale e product liability: sviluppi del diritto dell'Unione Europea*, in Giur. it., 1722. Secondo tale impostazione, quindi, lo sviluppatore sarebbe in grado di limitare le possibili condotte danno derivanti da un'erronea attività di apprendimento autonomo dell'algoritmo. Alcuni autori, addirittura, suggeriscono l'ipotesi di codificare delle regole di condotta etiche leggibili dall'algoritmo. Così, Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. e Floridi, L. (2016), *The Etichs of Algorithms: Mapping the Debate*, in Big Data & Society, 3, 11.

stessa di algoritmo quale «*a set of computational rules to be followed to solve a mathematical problem*»[166], si avrebbe nel caso in cui una delle istruzioni all'interno della sequenza algoritmica fosse erronea. Ne deriverebbe un *output* non conforme alla finalità prevista e potenzialmente in grado di cagionare una perdita economica per il cliente.

Una seconda ipotesi, relativa alla fase di ideazione dell'algoritmo può derivare dai *bias*[167] cognitivi e statistici. In particolare, i '*cognitive biases*' si verificano quando vengono utilizzate inaccurate metodologie di raccolta dei dati che portano a una falsa rappresentazione della realtà. Essi possono essere causati da dati di scarsa rappresentatività oppure da una erronea percezione della realtà degli sviluppatori dell'algoritmo[168]. Gli '*statistical biases*'[169], invece, si riferiscono alla tendenza dell'algoritmo a ottenere risultati discriminatori laddove i dati che vengono analizzati in un certo settore sono sbilanciati[170].

Quanto alla dimensione discriminatoria, i *bias* possono essere classificati secondo l'intento soggettivo del programmatore come 'volontari' o 'involontari'. Nel primo caso, l'algoritmo viene volutamente elaborato preventivamente per selezionare specifiche categorie di dati sensibili (quali razza, etnia e orientamento sessuale) che portano a una fisiologica segmentazione basata su tali specifiche informazioni. Nella seconda ipotesi si riferisce tipicamente al caso degli *statistical biases* e la discriminazione è effetto non di una decisione in tal senso, bensì dalla scarsa qualità dei dati utilizzati dallo sviluppatore o dal gestore di portafoglio (sul tema dei *bias* statistici e della discriminazione algoritmica si tornerà nell'Appendice '*Intelligenza artificiale e tutela della persona*').

3.2 La responsabilità dell'intermediario

Alla luce della descritta casistica di eventi potenzialmente dannosi per l'investitore, va rilevata la complessità della fattispecie concreta che deriva anche dalla molteplicità dei soggetti potenzialmente coinvolti nella prestazione del servizio di gestione del portafoglio automatizzata. Come sottolineato da alcuni approfondimenti[171], infatti,

166 Così Financial Stability Board, *op. cit.* (2017), 35.

167 Il fenomeno può essere visto sotto molteplici aspetti: in primo luogo, la nozione di *bias* si potrebbe riferire all'inadeguatezza dei dati o del modello di analisi a raggiungere a raggiungere l'obiettivo statistico desiderato; in secondo luogo, il concetto sarebbe sinonimo di discriminazione; infine, esso potrebbe identificare un meccanismo di pensiero che causa una deviazione del giudizio introdotto nell'algoritmo. Si veda Bogroff e Guégan, *op. cit.* (2019), 19 e ss.. Sul punto, si veda Barocas, S. e Selbst, A.D. (2016), *Big Data's Disparate Impact*, in Cal. L. Rev., vol. 104, 677 e ss..

168 Si pensi, ad esempio, all'utilizzo di categorie irrilevanti nello specifico settore in cui esse dovrebbero essere utilizzate.

169 La distinzione proposta nella trattazione viene effettuata in Malgieri, G. e Comandè, G. (2017), *Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation*, in International Data Privacy Law, 9 e ss.. Sul tema, Bamberger, K. (2010), *Technologies of Compliance*, in Texas L. Rev., 88, 711 e ss.. Per un'esaustiva trattazione dei *biases* unitamente a un'indicazione della dottrina maggiormente autorevole sul tema si veda Mittelstadt et al., *op. cit.* (2016), 7 e ss..

170 Si pensi, ad esempio, ai casi di discriminazione basati sul sesso o fasce di età.

171 Si rinvia a CONSOB, *op. cit.* (2019) e BaFin, *op. cit.* (2018), 55.

l'intermediario potrebbe sviluppare in proprio il sistema di intelligenza artificiale, eventualmente dopo la fusione o acquisizione di una *start up*[172], oppure sottoscrivere con società specializzate degli accordi esternalizzando così alcune fasi ed elementi del servizio prestato.

Le eventuali perdite economiche subite dall'investitore durante lo svolgimento del servizio devono qualificarsi secondo il paradigma della responsabilità contrattuale e il gestore deve considerarsi inadempiente alle prestazioni convenute contrattualmente. Ne consegue che, qualora uno dei rischi sopra indicati si realizzi a causa della mancata adozione di cautele adeguate da parte dell'intermediario, l'investitore potrà tutelarsi agendo in via contrattuale chiedendo il risarcimento per inadempimento, in considerazione del rapporto giuridico esistente fra gestore e investitore costituito dal contratto di gestione di portafoglio[173].

Un interessante tema in merito potrebbe porsi sulla eventuale prova liberatoria fornita dall'intermediario nel caso in cui il malfunzionamento del sistema di intelligenza artificiale fosse dovuto a una condotta erronea dello sviluppatore (ad esempio, in fase di realizzazione dell'algoritmo). In particolare, bisognerebbe chiedersi fino a che punto deve estendersi il potere di controllo del gestore, finalizzato a evitare il danno, sulle tecniche di automazione esternalizzate allo sviluppatore.

Appare preferibile, anche in tale ipotesi, imputare contrattualmente la responsabilità in capo all'intermediario nei confronti dell'investitore in quanto l'utilizzo di algoritmi forniti da terzi soggetti rientrano nel perimetro della disciplina della esternalizzazione di funzioni. La *ratio* sottesa a tale normativa impone infatti al gestore del portafoglio di adottare le misure ragionevoli per prevenire i rischi[174]. Al riguardo, si può notare peraltro il parallelismo con la medesima soluzione prospettata dall'art. 1228 c.c. secondo cui il debitore che si avvale dell'opera di terzi per adempiere puntualmente alle proprie obbligazioni risponde dei fatti dolosi e colposi da questi cagionati.

Nessun dubbio sussiste ovviamente nel caso di TA sviluppate *in house* dal gestore del portafoglio, essendo questi tenuto al risarcimento secondo le regole della responsabilità contrattuale per inadempimento, alla luce dell'esistenza di un unico rapporto giuridico con l'investitore danneggiato.

172 Si rileva spesso la tendenza degli *incumbent* di legarsi a dei soggetti, molto spesso *start up*, il cui scopo è creare nuove tecnologie da applicare nel campo della finanza.

173 Si rinvia per analoghe considerazioni in tema di consulenza automatizzata a CONSOB, *op. cit.* (2019), 83 ss.

174 A tale riguardo, l'art. 19 del Regolamento in materia di organizzazione e procedure degli intermediari che prestano servizi di investimento o di gestione collettiva del risparmio (aggiornato con delibera CONSOB nel febbraio 2018) stabilisce che «quando, nella prestazione dei servizi e delle attività di investimento, gli intermediari affidano a un terzo l'esecuzione di funzioni operative essenziali o importanti o di servizi o di attività di investimento, adottano misure ragionevoli per mitigare i connessi rischi». Inoltre, secondo l'art. 31, par. 1, Regolamento delegato MiFID II «*le imprese di investimento che esternalizzano funzioni operative essenziali o importanti restano pienamente responsabili di tutti gli obblighi imposti loro dalla direttiva 2014/65 UE*».

3.3 La possibile soluzione efficiente della fattispecie: la responsabilità dello sviluppatore

Qualora il sistema di intelligenza artificiale venga acquistato o esternalizzato dal gestore bisogna individuare il modello di responsabilità applicabile allo sviluppatore, fatta salva la responsabilità contrattuale parallela dell'intermediario verso l'investitore qualora si verifichino i presupposti sopra indicati.

Per individuare il soggetto chiamato a rispondere del costo del danno prodotto, è fondamentale individuare colui che si può qualificare nella fattispecie in esame come il cosiddetto '*cheapest cost avoider*'[175]. Secondo questa teoria, la responsabilità deve essere allocata nei confronti di colui che è in grado di evitare il danno, prima della sua verificazione, nella maniera più economica possibile. La fattispecie *de qua* suggerisce di considerare proprio lo sviluppatore come il soggetto più idoneo a valutare il grado di rischio dell'attività svolta e ad adottare le cautele necessarie sostenendo costi contenuti. A titolo esemplificativo, si prenda il caso di errori nella configurazione o nella disposizione della serie di operazioni che costituiscono l'intero algoritmo. In tale ipotesi, chi può evitare in maniera più efficiente la perdita patrimoniale derivante da tale errore è sicuramente colui che lo ha progettato, potendo lo sviluppatore, per esempio, svolgere test[176], con costi contenuti, sul funzionamento delle tecniche di automazione anche in considerazione della sua completa conoscenza dei meccanismi sottesi a queste.

La giustificazione dogmatica della soluzione economica raggiunta richiede uno sforzo interpretativo maggiore nella fattispecie analizzata, nella quale non esiste alcun rapporto giuridico tra l'investitore e lo sviluppatore, essendo mediato dall'intermediario che utilizza le tecnologie per la prestazione del servizio di investimento. A tale riguardo, una possibile risoluzione del problema interpretativo è riconducibile alla teoria che identifica nel rischio di impresa il criterio di imputazione della responsabilità[177].

175 Una delle opere a cui si fa risalire l'origine dell'analisi economica del diritto è il saggio di Calabresi, G. (1961), *Some Thoughts on Risk Distribution and the Law of Torts*, in Yale L. J., vol. 70, n. 4, 499 e ss.. L'espressione «*cheapest cost avoider*» viene utilizzata da Calabresi per identificare il soggetto che può evitare la causazione del danno nella maniera più efficiente così da ridurre la ricaduta del costo del danno sulla società. Tale concetto viene formulato da Calabresi nella sua celebre opera *The costs of accident: a Legal and Economic Analysis*, New Haven (1970), trad. it De Vita, A., Varano, V. e Vigoriti, V. (a cura di, 1975), *Costo degli incidenti e responsabilità civile* Milano, 183 e ss.. In un saggio successivo, Calabresi afferma che il «*cheapest cost avoider*» è il soggetto «*best suited to make the cost-benefit analysis between the accident costs and the avoidance costs*», Calabresi, G. (1975), *Optimal Deterrence and Accidents: to Flaming James Jr.*, il miglior fabbro, in Yale L. J., vol. 84, n. 4, 666.

176 Segnala l'importanza di testare e monitorare continuamente le tecnologie impiegate per presidiare costantemente eventuali imprecisioni delle TA, BaFin, *op. cit.* (2018), 26 e ss..

177 La prima teorizzazione sistematica della responsabilità per rischio di impresa risale a Trimarchi, P. (1961), *Rischio e responsabilità oggettiva*, Milano. Inoltre, Castronovo svolge una puntuale ricognizione storica e comparatistica sull'annosa questione del superamento della colpa come principale criterio di ascrizione della responsabilità. A tale riguardo, si veda Castronovo, C. (2006), *La nuova responsabilità civile*, Milano, 275 e ss.. Si veda in tal senso anche Alpa, G. e Bessone, M. (1975), *La responsabilità civile*, Genova, 121 e ss.. L'Autore, tuttavia, ritiene che le fattispecie di responsabilità oggettive del codice civile siano troppo diverse per ritenere che possano essere accomunate in un unico criterio di imputazione della responsabilità civile. Il tema viene approfondito nell'ambito della responsabilità ambientale da Degli'innocenti, F. (2013), *Rischio di impresa e responsabilità civile*, Firenze, 27 e ss..

Il principio fondamentale di questa responsabilità *sine culpa* è rappresentato dal brocardo latino '*cuius commoda eius et incommoda*' secondo cui chi ottiene dei benefici da un bene o da una attività deve assumersene i relativi oneri, tra cui sopportare i costi dei danni cagionati[178].

Sulla scorta della dottrina civilistica in materia di responsabilità da prodotto, la ricostruzione del rischio di impresa come criterio di imputazione della responsabilità si fonda sul principio di cui è espressione l'art. 2049 c.c. secondo cui l'imputazione dei fatti illeciti dei dipendenti all'imprenditore costituisce l'ipotesi paradigmatica di «fatti dannosi imputabili all'impresa, alla stessa struttura organizzata»[179]. Analogicamente si potrebbe ritenere applicabile alla fattispecie *de qua*, la responsabilità da prodotto in quanto riconducibile al rischio di impresa poiché il danno cagionato dall'algoritmo è frutto dell'inadeguata organizzazione dell'impresa di cui il rapporto di preposizione e il prodotto difettoso sono le manifestazioni originarie.

Alla luce di tale interpretazione, lo sviluppatore sarebbe tenuto a risarcire in via oggettiva i danni subiti dagli investitori perché attraverso la realizzazione di dispositivi sofisticati egli produce un rischio per il mercato dei capitali ottenendo dei benefici i cui costi, qualora non fosse ritenuto responsabile, sarebbero sostenuti dagli investitori.

La soluzione interpretativa proposta soddisfa anche gli elementi strutturali del rischio identificati dalla dottrina. Oltre che a essere cagionato nell'ambito di una attività economica[180], il danno infatti deve essere la concretizzazione di un rischio tipico[181] rispetto all'attività di impresa svolta. Nel nostro caso, non sembrano esserci dubbi che l'eventualità di realizzare un sistema di intelligenza artificiale difettoso per

178 Questa è la tesi di Forchielli, P. (1983), *Responsabilità civile*, Cedam, 78 e ss.. Sul tema, si vedano anche Frigida, F. (2010), *Responsabilità del sorvegliante dell'incapace, dei genitori e tutori, dei padroni e committenti*, in Fava, P. (a cura di), *La responsabilità civile*, Milano, 1809 e ss. e Franzoni, M. (2010), *L'illecito*, Franzoni, M. (a cura di), *Trattato della responsabilità civile*, Milano, 767 e ss..

179 Castronovo, C. (1964), *Problema e sistema nel danno da prodotti*, Milano. Nonostante, Tuttavia, parte della dottrina ritiene che l'art. 2049 c.c. non possa essere applicato al caso di danni causati dall'uso dell'intelligenza artificiale, come sarebbe il caso della responsabilità del formatore dell'algoritmo. Questo perché si ritiene che al di là dell'interpretazione analogica che si può dare all'articolo, esso deve essere circoscritto all'esistenza del comportamento umano. Questo perché sarebbe veramente difficile identificare il 'danno causato dall'agente' con il danno causato da un sistema di intelligenza artificiale, dato che la norma rende il mandante responsabile di un'ipotesi di fallibilità del suo dipendente, essendo l'intelligenza umana l'elemento soggettivo inevitabile della norma. Ruffolo, U. (a cura di, 2019), *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in Giur. it.

180 Nella fattispecie esaminata, sicuramente il primo requisito è soddisfatto: lo sviluppatore è un soggetto perlopiù organizzato in forma societaria che svolge un'attività di programmazione e realizzazione di sistemi intelligenti in maniera economica, professionale, organizzata e lecita. Per un'analisi approfondita sui singoli requisiti dell'attività di impresa si veda Presti, G. e Rescigno, M. (2015), *Corso di diritto commerciale*, Bologna, 17 e ss.. Riguardo al requisito dell'organizzazione, l'unico problema può sorgere nel caso in cui il programmatore di software che svolge singolarmente la propria attività. In questo caso, lo sviluppo di software può ritenersi un'attività di impresa solo nel caso in cui si ritenga sufficiente l'auto-organizzazione come requisito richiesto per la qualificazione. Per un'ulteriore ed esaustiva illustrazione dei requisiti per qualificare un soggetto come imprenditore, si veda Graziani, A., Minervini, G. e Belviso, U. (2011), *Manuale di diritto commerciale*, Padova, 39 e ss.. Proprio Trimarchi afferma l'applicabilità della responsabilità per rischio di impresa come paradigma generale per i danni cagionati nell'ambito dell'attività di impresa. Il regime di responsabilità oggettiva è, pertanto, eccezionale rispetto alla colpa nell'ambito delle attività biologiche che hanno infatti «*il carattere della necessità e della non economicità: non ha senso [...] parlare di un conto attivo e passivo, sul quale deve gravare il costo del rischio creato*», Trimarchi, *op. cit.* (1961), 43 e ss..

181 Il rischio deve essere tipico poiché «*alcuna pressione economica può essere esercitata rispetto ai singoli eventi imprevedibili e atipici che non sono manifestazione di un rischio costante, ma che determinano un danno occasionale non pertinente ad un'attività programmata*». Sul punto, si veda Degl'innocenti, *op. cit.* (2013), 17.

la presenza di un *bug* o per l'utilizzo di dati scarsamente qualitativi nella fase di addestramento, deve considerarsi aspetto patologico connaturato all'attività economica di realizzazione di sistemi tecnologicamente avanzati.

Va sottolineato peraltro che i rischi derivanti dall'automazione attuata mediante tecnologie di scarsa qualità potrebbero assumere un'entità notevole in considerazione del folto numero di investitori che potrebbero aver sottoscritto lo stesso contratto di gestione automatizzata rivelatosi errato[182]. Inoltre, qualora vi siano più algoritmi aventi la stessa funzione collegati tra loro, il rischio di errori nella fase di sviluppo può riflettersi su altri sistemi di intelligenza artificiale e assumere così una dimensione sistemica[183]. Tali circostanze rilevano nella traducibilità del rischio, dopo la sua concretizzazione, in un costo effettivo che possa essere calcolato dall'imprenditore affinché possa realizzarsi la funzione utile del paradigma[184].

Nella ipotesi descritta il pregiudizio subito dal cliente può inoltre configurarsi come 'danno ingiusto' ai sensi dell'art. 2043 c.c. poiché l'ordinamento prevede la tutela degli investitori attraverso una corretta prestazione del servizio di investimento[185] che sarebbe frustrata da malfunzionamenti nelle tecniche di automazione fornite all'intermediario[186].

Il rispetto delle regole di condotta da parte del gestore, infatti, potrebbe dipendere in maniera significativa dalla qualità delle tecnologie implementate e quindi anche da eventuali errori commessi dallo sviluppatore. Al riguardo, si ricordano le norme a protezione del risparmiatore previste dalla disciplina della vigilanza ispettiva

182 Per un caso concreto, si veda Ferrari, V. e Lusardi, G. (2019), *Fintech: Chi è responsabile se l'intelligenza artificiale sbaglia ad investire*, www.agendadigitale.eu. In sintesi, l'investitore ha chiesto un risarcimento di 23 milioni di dollari, adducendo una rappresentazione erronea delle performance del sistema K1 da parte degli investitori del fondo. Per maggiori dettagli sugli sviluppi del procedimento, si vedano Tre, L. (2019), *Borse, chi paga se è l'algoritmo a perdere i soldi?*, www.ilsole24ore.com e Beardsworth, T. e Kumar, N. (2019), *Who to Sue When a Robot Loses your Fortune*, www.bloomberg.com.

183 Sul tema dei possibili rischi sistemici finanziari derivanti dall'utilizzo dell'IA si veda Financial Stability Board, *op. cit.* (2017), 25. Lo stesso problema viene affrontato in BaFin, *op. cit.* (2018), 165.

184 Trimarchi affronta esaustivamente la funzione del paradigma di responsabilità per rischio di impresa. Tale modello si applica anche alle imprese che non svolgono attività intrinsecamente pericolose, poiché «*qualsiasi impresa, implicando organizzazione e continuità, se causa danni, grandi o piccoli, frequenti o infrequenti, li causa con una certa regolarità, calcolabile per lunghi periodi, talché il rischio relativo deve essere tradotto in costo*», così Trimarchi, *op. cit.* (1961), 36.

185 Sul punto, si richiama la clausola generale espressa dall'art. 21, Tuf, co. 1, lett. a).

186 In merito, si rileva che l'approccio adottato da parte delle autorità di vigilanza europee si basa sul principio della neutralità tecnologica per cui «*le regole di comportamento dettate dall'ordinamento per la prestazione dei servizi di investimento non vengono meno se l'impresa si avvale di internet per lo svolgimento di intermediazione*». Sul punto, CONSOB, Comunicazione n. 30396, (2000), www.consob.it, 1. Tuttavia, la Commissione Europea, sebbene sostenga che «*la neutralità tecnologica è uno dei principi guida delle politiche della Commissione*», riconosce che «*i rapidi progressi delle tecnologie finanziarie stanno determinando cambiamenti strutturali nel settore finanziario. In un ambiente in così rapida evoluzione una regolamentazione eccessivamente prescrittiva e precipitosa rischia di produrre effetti indesiderati. Tuttavia, il mancato aggiornamento delle politiche e dei quadri normativi potrebbe porre in posizione di svantaggio i prestatori di servizi finanziari dell'UE in un mercato sempre più globale. Esiste inoltre la possibilità, ad esempio nel caso della cyber-sicurezza che alcuni rischi importanti non siano affrontati*». Sul punto, Commissione Europea, *Piano d'azione per le tecnologie finanziarie: per un settore finanziario europeo più competitivo ed innovativo*, (2018), www.eur-lex.europa.eu, 11, 20.

e delle sanzioni amministrative[187], che costituiscono indici di una tutela rafforzata per l'investitore in caso di esternalizzazione di funzioni essenziali o importanti.

Alla luce delle argomentazioni proposte, si può ritenere che la specifica applicazione del rischio di impresa alla fattispecie considerata qualifichi una ipotesi di 'responsabilità da algoritmo'[188]. In considerazione della fallibilità dell'intelligenza artificiale, infatti, appare corretta l'allocazione di tale responsabilità non in capo all'intermediario, bensì al soggetto che progetta e fornisce l'algoritmo su cui si basa la prestazione del servizio di gestione di portafoglio[189]. L'applicabilità della responsabilità per rischio di impresa avrebbe inoltre il pregio di favorire la tutela dell'investitore anche da un punto di vista processuale, posto che sarebbe molto difficoltoso per quest'ultimo provare la colpa dello sviluppatore a causa della elevata complessità dei meccanismi sottesi alle tecniche di automazione.

3.4 Considerazioni attuali sulla responsabilità civile dei sistemi di IA

La regolamentazione della responsabilità civile per l'utilizzazione di sistemi di intelligenza artificiale è oggetto di dibattito anche nell'Unione europea[190]. Sebbene non sia stata ancora adottata una regolamentazione in materia, il Parlamento europeo ha pubblicato una risoluzione sul '*Regime di responsabilità civile per l'intelligenza artificiale*'[191].

La proposta di Regolamento sulla responsabilità per il funzionamento dei sistemi di intelligenza artificiale, allegata alla risoluzione, poi seguita dalla proposta di Regolamento sull'intelligenza artificiale, fatte salve le ipotesi di responsabilità oggettiva per i sistemi di intelligenza ad alto rischio (art. 4), riconduce la responsabilità in capo all'operatore dei sistemi di IA in caso di colpa[192] (per dettagli in merito alla pro-

187 L'art. 6-*ter*, co. 1 e l'art. 190, co. 1, Tuf introducono nel perimetro della vigilanza ispettiva e delle sanzioni amministrative i soggetti a cui gli intermediari abbiano esternalizzato funzioni essenziali o importanti.

188 Tale espressione, seppure in una diversa accezione, viene utilizzata da Ruffolo, U. (2019), *Intelligenza Artificiale, machine learning e responsabilità da algoritmo*, in *Giur. it.*, 1693; Ruffolo, U. (2017), *Per i fondamenti di un diritto della robotica self-learning; dalla machinery produttiva all'auto driverless: verso una 'responsabilità da algoritmo?'*, in Ruffolo, U. (a cura di, 2018), *Intelligenza artificiale e responsabilità*, Milano, 16.

189 Sul punto, alcuni studiosi statunitensi si sono spinti a prospettare l'introduzione di un regime di responsabilità da prodotto che prescinda dall'esistenza di un difetto quale soluzione al tema della 'responsabilità da intelligenza artificiale', si veda Vladeck, S. (2014), *Machines Without Principals: Liability Rules and Artificial Intelligence*, in Wash. L. R., 89, 146.

190 Con l'obiettivo generale di creare un 'mercato digitale europeo', l'Unione europea ha affrontato una serie di questioni compresa la responsabilità civile per l'uso di sistemi di intelligenza artificiale. Uno dei primi documenti è stata la 'Risoluzione di raccomandazione del Parlamento Europeo alla Commissione del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica', seguita dalla Comunicazione del 25 aprile 2018 'L'intelligenza artificiale per l'Europa' e dal 'Libro bianco dell'intelligenza artificiale' del 19 febbraio 2020. Il documento più ambizioso è la Risoluzione sul 'Regime di responsabilità civile per l'intelligenza artificiale', adottata il 20 ottobre 2020.

191 Risoluzione del Parlamento europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale in www.europarl.europa.eu

192 L'art. 8 della risoluzione distingue il regime di responsabilità a seconda del rischio dei sistemi di IA: per i sistemi di IA ad alto rischio imputa oggettivamente la responsabilità in capo all'operatore; per gli altri sistemi di IA stabilisce un criterio basato sulla colpa.

posta di Regolamento sull'intelligenza artificiale si veda l'Appendice '*Intelligenza artificiale e tutela della persona*'). Attribuisce altresì a quest'ultimo, per entrambi i sistemi di responsabilità, la possibilità di esercitare un'azione di regresso contro il produttore in caso di 'sistema IA difettoso', in applicazione della direttiva 85/374/CEE[193] e delle disposizioni nazionali che disciplinano la responsabilità dei prodotti difettosi (art. 12, par. 3). Tuttavia, questo sistema non è stato ancora recepito in una regolamentazione UE applicabile all'interno degli Stati membri, per cui l'unica normativa attualmente vigente è la direttiva in materia di responsabilità da prodotti difettosi. Soltanto qualora i sistemi di IA siano considerati un prodotto difettoso secondo quanto definito dalla medesima direttiva, il danno derivante dalla loro utilizzazione potrebbe essere imputato oggettivamente a carico del produttore del sistema: il soggetto leso dovrebbe dimostrare la difettosità del prodotto, il pregiudizio patito e il nesso di causalità tra quest'ultimo e il difetto del prodotto indipendentemente dalla sussistenza dell'elemento soggettivo del dolo o della colpa[194].

Ciò premesso, la direttiva fa salva la possibilità da parte del produttore di liberarsi dall'imputazione della responsabilità mediante la prova della sussistenza di una delle condizioni previste dall'articolo 7[195]. Tra queste, a livello esemplificativo, il produttore potrebbe dimostrare l'esistenza del cosiddetto 'rischio da sviluppo', con il quale si intende il difetto, non prevedibile al momento della messa in circolazione del prodotto o sorto successivamente. Si intende che la prova liberatoria dello sviluppatore non è ammessa quando era prevedibile, già prima della messa in circolazione, che il prodotto avrebbe potuto manifestare un comportamento imprevisto[196], ragionamento dal quale dovrebbe dedursi che, anche quando siano utilizzati sistemi di IA, il produttore non potrebbe invocare il 'rischio da sviluppo' qualora insorga un comportamento deviante[197].

Oltre alla responsabilità del produttore, vi può essere un addebito di responsabilità in capo al programmatore dell'algoritmo qualora il malfunzionamento del *software* sia riconducibile all'algoritmo, dal quale scaturiscono i danni prodotti, configurandosi così una responsabilità solidale nei confronti dell'utente[198].

193 Direttiva 85/374/CEE del Consiglio del 25 luglio 1985 relativa al ravvicinamento delle disposizioni legislative, regolamentari e amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi in www.eur-lex.europa.eu

194 Leanza, C. (2021), *Intelligenza artificiale e diritto: ipotesi di responsabilità civile nel terzo millennio*, in Resp. civ. prev., 3, 1011.

195 L'art 7, Dir. 85/374/CEE stabilisce: «Il produttore non è responsabile ai sensi della presente direttiva se prova: a) che non ha messo il prodotto in circolazione; b) che, tenuto conto delle circostanze, è lecito ritenere che il difetto che ha causato il danno non esistesse quando l'aveva messo in circolazione o sia sorto successivamente; c) che non ha fabbricato il prodotto per la vendita o qualsiasi altra forma di distribuzione a scopo economico, né l'ha fabbricato o distribuito nel quadro della sua attività professionale; d) che il difetto è dovuto alla conformità del prodotto a regole imperative emanate dai poteri pubblici; e) che lo stato delle conoscenze scientifiche e tecniche al momento in cui ha messo in circolazione il prodotto non permetteva di scoprire l'esistenza del difetto; f) nel caso del produttore di una parte componente, che il difetto è dovuto alla concezione del prodotto in cui è stata incorporata la parte o alle istruzioni date dal produttore del prodotto».

196 High-Level Expert Group on Artificial Intelligence (2019), *Ethics Guidelines for Trustworthy Artificial Intelligence*, https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

197 Burgio, E. e De Simone, L. (2021), *Intelligenza Artificiale e responsabilità civile*, in Media Laws, 15 aprile.

198 *Ibidem*.

Questa ricostruzione, consentita nella misura in cui l'algoritmo è inteso quale componente del prodotto finale, mediante l'estensione della definizione di 'prodotto' dell'art. 2 dalla Direttiva 85/374/CEE[199], attribuisce la responsabilità in capo al prestatore di servizi nei confronti del cliente, eventualmente in solido con il produttore del sistema di IA e il terzo a cui è stato esternalizzato il servizio di *big data analytics*. Si tratta di un orientamento in linea con i principi di diritto comune; in particolare, nella prospettiva di conciliare esigenze di efficienza con istanze di giustizia commutativa, l'intento è quello di imputare la responsabilità al soggetto (produttore, programmatore, intermediario) che è nella migliore posizione per evitare il danno.

Tuttavia, l'incertezza degli elementi non consente di pervenire alla definizione di un sistema di responsabilità univoco. Più in particolare, la natura della responsabilità civile, la difficoltà di dimostrazione del danno subito a causa della complessità degli algoritmi, l'effetto sistematico dell'evento dannoso e l'eventuale incapacità finanziaria del responsabile indicano inequivocabilmente la sussistenza di una sproporzione tra il rimedio privato e la natura del problema[200]. Sembra quindi preferibile adottare un approccio che, senza escludere necessariamente il ricorso al risarcimento, si concentri su approcci alternativi: strategie di regolamentazione preventiva di supervisione pubblica, con la previsione di sistemi di indennizzo per proteggere le parti lese. Sarebbe definito un modello di responsabilità oggettiva, compendiato da un sistema di assicurazione obbligatoria, analogamente a quanto previsto nella risoluzione del Parlamento europeo in materia di sistemi di intelligenza artificiale ad alto rischio[201].

3.5 Considerazioni dalla *survey*

Le implicazioni in tema di responsabilità civile del produttore, del programmatore e dell'intermediario di sistemi di IA possono avere ripercussioni nel mercato finanziario.

Come emerge dalla *survey*, la maggior parte delle società di gestione consultate ritiene che l'utilizzazione e il perfezionamento dei sistemi di IA costituiscano una priorità strategica; a dimostrazione di ciò, alcune società di gestione già fanno ricorso a sistemi di IA.

Con particolare riferimento al monitoraggio di questi sistemi, nessuna della società ha dichiarato che i sistemi di IA utilizzati sono in grado di garantire la coerenza tra le decisioni/scelte e le politiche di investimento del fondo. Ciò ha ripercussioni indubbiamente in termini di responsabilità e postula l'individuazione di un intervento umano ai fini della riconducibilità del danno prodotto in capo a un agente (produttore, programmatore, intermediario).

199 L'art. 2 Dir 85/374/CEE stabilisce: «i fini della presente direttiva, per «*prodotto*» si intende ogni bene mobile, ad eccezione dei prodotti agricoli naturali e dei prodotti della caccia, anche se forma parte di un altro bene mobile o immobile. Per «*prodotti agricoli naturali*» si intendono i prodotti del suolo, dell'allevamento e della pesca, a esclusione dei prodotti che hanno subito una prima trasformazione. Per «*prodotto*» si intende anche l'elettricità».

200 Alpa, G. (2021), *Quale modello normativo europeo per l'intelligenza artificiale?*, in Contr. impr., n. 4, 1003.

201 In questi termini si esprime Perrone, *op. cit.* (2021), 720.

Allo stesso tempo, i risultati dell'indagine hanno evidenziato l'utilizzazione di sistemi di IA di origine mista, cioè in parte sviluppati internamente e in parte sviluppati o procurati esternamente.

La ragione principale legata allo sviluppo tecnologico in *outsourcing* o in *partnership* risiede nei lunghi tempi di realizzazione e implementazione, rilevati da sei società e dalla mancanza di professionalità interne competenti in materia di nuove tecnologie.

In conseguenza, la dinamica di funzionamento dei sistemi di IA evidenzia che la causazione di un eventuale danno può coinvolgere gli interessi dell'investitore, dell'intermediario e dello sviluppatore. In particolare, qualora il sistema di IA sia stato sviluppato soltanto dall'intermediario finanziario, a quest'ultimo sarà unicamente imputato il danno prodotto, a meno che non dimostri di avere posto in essere tutti i comportamenti necessari per evitarlo.

Viceversa, la situazione diventa più complessa quando il sistema IA coinvolto è stato sviluppato da un terzo, che non è coinvolto nella sottoscrizione del contratto tra investitore e intermediario.

La *survey* ha dimostrato che l'utilizzazione dei sistemi di IA non avviene fuori da qualsiasi coinvolgimento umano. Quattro società coinvolte nell'indagine hanno evidenziato che la supervisione umana è intensa in quanto il sistema di IA non è autonomo nel processo decisionale. Due delle società hanno riconosciuto che il sistema di IA utilizzato è parzialmente autonomo con attribuzione della decisione finale all'uomo; un'altra società, invece, ha rilevato che vi è solo un controllo umano sul sistema di IA nella decisione finale[202].

Il mantenimento della capacità decisionale e la sussistenza di un residuo controllo può suscitare perplessità sulla riconducibilità della responsabilità in capo allo sviluppatore. Nello specifico, è discusso se lo sviluppatore possa rispondere in solido quando la decisione finale non è stata assunta dal sistema di IA ma dall'uomo. Parimenti, la stessa questione riguarda la responsabilità dello sviluppatore qualora l'intermediario eserciti un controllo sul funzionamento del sistema di IA.

Attualmente non è possibile fornire soluzioni univoche a queste questioni che potranno essere affrontate soltanto da scelte di politica legislativa.

202 Per quanto riguarda il trattamento del processo decisionale dei sistemi di IA, tre delle società consultate hanno ritenuto che non dovrebbe essere diverso da quello dato ai processi che coinvolgono solo l'intervento umano. Tuttavia, altre cinque società consultate hanno indicato motivazioni legate alla necessità che un controllo umano sia sempre necessario o che il monitoraggio continuo e la validazione dei modelli alla base dei sistemi di IA, come parte integrante del processo decisionale basato su tali sistemi, richiede processi decisionali o dedicati e/o processi di natura diversa.

4 Tutela della privacy

4.1 Introduzione

L'utilizzo di TA aumenta le risorse che il gestore di portafoglio può utilizzare per la valutazione di adeguatezza e la revisione periodica del prodotto previste dalla disciplina MiFID II[203]. Tale possibilità richiede di discutere la conformità alla disciplina sulla tutela dei dati personali di un *customer relationship management* realizzato tramite *data analytics*.

Il tema del trattamento dei dati personali è del resto sempre più centrale nella disciplina dei mercati finanziari, ponendosi, come è stato correttamente osservato «*al crocevia tra diversi ambiti operativi e regolamentari e per il fatto di mettere in evidenza la sempre maggiore rilevanza che assumono i big data e le relative modalità di trattamento ed elaborazione come 'materia prima' nel processo di creazione del valore nei diversi comparti del settore finanziario*»[204].

In tale ambito, come noto, il Regolamento generale sulla protezione dei dati personali n. 679/2016/UE (GDPR)[205] costituisce la *sedes materiae* della normativa europea in materia di protezione dei dati personali.

Nel contemperare l'interesse all'innovazione tecnologica e la tutela dei soggetti sottoposti al trattamento, l'art. 6 del GDPR ha previsto le condizioni necessarie della liceità del trattamento[206].

Tali condizioni, oltre al caso di consenso prestato dal cliente, ricorrono, per quanto qui interessa, nel caso della profilatura prevista da MiFID II, essendo questa finalizzata all'«*esecuzione di un contratto*» di cui l'investitore è parte e il relativo trattamento è funzionale per «*adempiere un obbligo legale al quale è soggetto il titolare del trattamento*»[207].

203 Direttiva 2014/65/UE del Parlamento europeo e del Consiglio, del 15 maggio 2014, relativa ai mercati degli strumenti finanziari e che modifica la direttiva 2002/92/CE e la direttiva 2011/61/UE.

204 Così CONSOB (2019), *Financial Data Aggregation e Account Information Services. Questioni regolamentari e profili di business. Questioni regolamentari e profili di business*, in Quaderni FinTech, 4, www.consob.it, 4.

205 Approvato con Regolamento UE 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016, pubblicato nella Gazzetta Ufficiale europea il 4 maggio 2016 ed entrato in vigore il 24 maggio 2016, è applicabile a decorrere dal 25 maggio 2018. Con la sua entrata in vigore, il GDPR ha sostituito la direttiva sulla protezione dei dati (95/46/CE) e, con riferimento al nostro ordinamento, ha abrogato gli articoli del codice per la protezione dei dati personali (d.lgs. 196/2003) con esso incompatibili.

206 L'art. 6, rubricato 'Liceità del trattamento', al par. 1), stabilisce che «*(i)l trattamento è lecito solo se e nella misura in cui ricorre almeno una delle seguenti condizioni: a) l'interessato ha espresso il consenso al trattamento dei propri dati personali per una o più specifiche finalità; b) il trattamento è necessario all'esecuzione di un contratto di cui l'interessato è parte o all'esecuzione di misure precontrattuali adottate su richiesta dello stesso; c) il trattamento è necessario per adempiere un obbligo legale al quale è soggetto il titolare del trattamento; d) il trattamento è necessario per la salvaguardia degli interessi vitali dell'interessato o di un'altra persona fisica; e) il trattamento è necessario per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento; f) il trattamento 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. La lettera f) del primo comma non si applica al trattamento di dati effettuato dalle autorità pubbliche nell'esecuzione dei loro compiti*».

207 Così, l'art. 6, par. 1, lett. b e c, GDPR.

Alla medesima soluzione si giunge con riferimento all'attività di ribilanciamento del portafoglio, anch'essa basata sul profilo di rischio del cliente, poiché come la valutazione di adeguatezza costituisce una prestazione tipica del contratto di gestione.

Qualora invece, l'acquisizione e la successiva memorizzazione di informazioni riguardanti il cliente è svolta attraverso soggetti terzi che, svolgendo differenti attività, sono in grado di acquisire con maggiore facilità dati relativi all'intera esperienza di vita degli investitori (i *data providers*)[208], la base giuridica del trattamento va rinvenuta nel consenso prestato dal cliente a questi per il trattamento ed eventuale successivo trasferimento di dati a favore dell'intermediario[209].

Con riferimento alla più puntuale fattispecie della comunicazione dei dati dovuta dal *distributor* al *manufacturer* ai sensi delle regole MiFID II sulla revisione periodica del prodotto, una trasmissione di dati anonimi[210] al *manufacturer* in forma aggregata consente il rispetto delle regole di *product governance* anche in assenza di uno specifico consenso del cliente alla comunicazione. Sembra muovere in tal senso, il Considerando 26 del GDPR, che prevede come «i principi di protezione dei dati non dovrebbero potersi applicare a informazioni anonime», cioè «informazioni che non si riferiscono a una persona fisica identificata o identificabile o a dati personali resi sufficientemente anonimi da impedire o da non consentire più l'identificazione dell'interessato».

Giova qui ricordare, che tanto con riguardo al *target market assessment* che al *suitability assessment*, il principio cardine che guida il processo di raccolta ed elaborazione delle informazioni del cliente consiste nella 'proporzionalità' rispetto alla natura del prodotto di investimento, servizio di investimento e mercato di riferimento[211].

4.2 La gestione automatizzata del portafoglio

Una specifica riflessione sembra opportuna con riferimento al diritto di opposizione del soggetto interessato (ossia l'investitore) a «una decisione basata unicamente sul trattamento automatizzato» previsto dall'art. 22 del GDPR[212]. In astratto, l'attività

208 Come noto tra i *Data provider* vi sono le *BigTech* (come Facebook, Twitter, Google) aziende in grado di acquisire una grandissima quantità dati di varia natura nelle loro attività tipiche utilizzabili nella prestazione di servizi finanziari.

209 In tal senso, l'art. 6, par. 1, lett. a, GDPR. Evidenzia la sovranità dei dati in capo al soggetto interessato, si veda BaFin, *op. cit.* (2018), 38 e ss..

210 Diversa dall'anomizzazione è la pseudonimizzazione intesa come «il trattamento dei dati personali in modo personale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive, a condizione che tali informazioni aggiuntive siano conservate separatamente e soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile» (art. 4, par. 1, n. 5, GDPR). La differenza con l'anonimizzazione si basa sul fatto che in quest'ultimo caso non è più in alcun modo possibile ricostruire la provenienza del dato.

211 Si rinvia a CONSOB, *op. cit.* (2019), 99.

212 L'art. 22 GDPR stabilisce un divieto di decisioni basate unicamente sul trattamento automatizzato, ad eccezione delle seguenti ipotesi: a) se necessario per la conclusione e l'esecuzione di un contratto tra l'interessato e il titolare del trattamento; b) se autorizzato dal diritto UE o dallo Stato membro cui è soggetto il titolare del trattamento; c) se è il consenso esplicito dell'interessato.

di gestione automatizzata del portafoglio potrebbe essere riconducibile alla fattispecie prevista dal par. 1, di processo decisionale automatizzato, così derivandone la possibilità per il cliente di opporsi alla propria profilatura. Invero, in concreto appare più corretto ritenere operante l'esenzione prevista nel par. 2, che in maniera simile all'art. 6 GDPR, prevede la non applicabilità del diritto di opposizione quando l'attività di analisi dei dati automatizzata sia «*necessaria per la conclusione o l'esecuzione di un contratto tra l'interessato e il titolare del trattamento*». La prestazione gestoria *de qua*, infatti, necessita dell'utilizzo dei *big data* e delle relative tecniche di analisi, essendo queste imprescindibili per il corretto svolgimento del relativo servizio[213]. Va evidenziato però che tale fattispecie non si applicherebbe in tutti i casi di presenza (anche minima) di intervento umano[214] che porti a escludere una decisione basata 'unicamente' sul trattamento automatizzato.

In ogni caso, restano fuori da tale esenzione, le ipotesi di utilizzo dei dati sensibili. L'art. 22, par. 4, GDPR, infatti, configura in capo al soggetto interessato il diritto di opposizione per alcune delle eccezioni previste giustificando la deroga alla luce delle specifiche esigenze di tutela della *privacy* in relazione a informazioni particolarmente sensibili[215].

Nella fattispecie esaminata *ex* art. 22 GDPR, l'intermediario dovrà predisporre «*misure appropriate per tutelare i diritti*» dei soggetti interessati, tra cui la possibilità di «*ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione*»[216]. In merito, come è stato opportunamente osservato[217], le tutele previste non garantiscono una sufficiente e adeguata tutela dell'investitore se non vi è una comprensione del processo decisionale, in particolare laddove vengono usati algoritmi con '*black box*' (sul tema si veda anche l'Appendice '*Intelligenza artificiale e tutela della persona*').

Le disposizioni previste dall'art. 22 GDPR, interpretate alla luce dell'art. 13, par. 2, lett. f. GDPR[218] – secondo cui il titolare del trattamento dovrebbe fornire al soggetto interessato «informazioni significative sulla logica utilizzata» del processo automatizzato, nonché «l'importanza e le conseguenze previste di tale trattamento per l'interessato» – sembrano muovere verso una effettiva trasparenza degli algoritmi.

213 Anche in caso di esternalizzazione di tale funzione, l'implementazione di tecniche di automazione può considerarsi un'attività aziendale essenziale, poiché «*un'anomalia nella sua esecuzione o la sua mancata esecuzione*» inficerebbe la capacità dell'intermediario di continuare «*a garantire la conformità delle condizioni e degli obblighi della sua autorizzazione o agli altri obblighi della MiFID II*» e di mantenere «*la solidità o la continuità*» dei servizi o delle attività di investimento prestate (art. 30, par. 1, Regolamento delegato MiFID II).

214 Sul rischio che l'autonomia di giudizio umana possa essere influenzata negativamente dall'eccesiva confidenza nella correttezza delle decisioni proposte dagli algoritmi (cosiddetti *automation biases*), si rinvia a Panisi e Perrone, *op. cit.* (2018), 5 e ss..

215 Sulla disciplina dei dati sensibili si rinvia all'art. 9 del GDPR.

216 Così, l'art. 22, par. 3, GDPR.

217 Sul punto, si veda Expert group on regulatory obstacles to financial innovation ROFIEG (2019), 30 reccomendations on regulation, innovation and finance, www.ec.europa.eu, 38 e ss..

218 Nonché dei successivi artt. 14, par. 2, lett. g, e 15, par. 1, lett. h), che ne riprendono pedissequamente la sua formulazione.

L'attuale formulazione di tali norme però non consente di individuare con precisione la portata delle informazioni sulla logica degli algoritmi, potendosi queste ridursi a una mera indicazione delle tecnologie specificamente implementate. Inoltre, non viene previsto uno specifico rimedio per rispondere a una eventuale violazione dell'obbligo sancito dalla disciplina europea.

È opportuno sottolineare del resto, che la previsione di un diritto alla trasparenza, che pur con qualche difficoltà può configurarsi in astratto, rischia di trasformarsi in una finzione, se non garantisce in concreto un diritto alla comprensione[219]. Gli investitori, infatti, non possedendo tendenzialmente specifiche conoscenze in materia, difficilmente potranno giungere a una significativa comprensione dei meccanismi riguardanti le tecniche di automazione anche quando queste venissero illustrate minuziosamente[220].

4.3 Il diritto all'oblio e l'obbligo di trasparenza finanziaria

Un'altra importante disposizione contenuta nel GDPR è il diritto alla cancellazione (o diritto all'oblio), secondo il quale l'interessato (ossia, l'investitore) «*ha il diritto di ottenere dal titolare del trattamento la cancellazione dei dati personali che lo riguardano senza ingiustificato ritardo*», quando i suoi dati non sono più necessari, non vengono più trattati o ha ritirato il consenso prestato (art. 17 del GDPR)[221].

Da una lettura sommaria del regolamento, sembrerebbe che questa disposizione crei una potenziale incoerenza normativa tra il GDPR e la MiFID II. L'art. 16, par. 6, della direttiva MIFID prescrive l'obbligo per le imprese di investimento (ossia gli intermediari finanziari) di tenere una registrazione sufficiente di «*tutti i servizi prestati e tutte le attività e le operazioni effettuate*», mentre il par. 7 del medesimo articolo stabilisce che le registrazioni delle operazioni concluse devono essere «*conservate per un periodo di cinque anni*».

219 Più che un diritto alla spiegabilità degli algoritmi, auspicano un diritto di 'legibility by design'. Malgieri e Comandè, *op. cit.* (2017), 3 e ss.. Una mera spiegazione dei meccanismi sottesi alla decisione algoritmica non garantisce infatti che l'investitore comprenda a pieno le logiche utilizzate dall'algoritmo.

220 Sul tema, propongono il passaggio da un diritto di spiegabilità a un diritto alla miglior decisione, Edwards, L. e Veale, M. (2018), *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions?'*, in IEEE Security & privacy, 7.

221 L'art 17 rubricato 'Diritto alla cancellazione («diritto all'oblio»)' al par. 1), stabilisce che «*(l)'interessato ha il diritto di ottenere dal titolare del trattamento la cancellazione dei dati personali che lo riguardano senza ingiustificato ritardo e il titolare del trattamento ha l'obbligo di cancellare senza ingiustificato ritardo i dati personali, se sussiste uno dei motivi seguenti: a) i dati personali non sono più necessari rispetto alle finalità per le quali sono stati raccolti o altrimenti trattati; b) l'interessato revoca il consenso su cui si basa il trattamento conformemente all'articolo 6, paragrafo 1, lettera a), o all'articolo 9, paragrafo 2, lettera a), e se non sussiste altro fondamento giuridico per il trattamento; c) l'interessato si oppone al trattamento ai sensi dell'articolo 21, paragrafo 1, e non sussiste alcun motivo legittimo prevalente per procedere al trattamento, oppure si oppone al trattamento ai sensi dell'articolo 21, paragrafo 2; d) i dati personali sono stati trattati illecitamente; e) i dati personali devono essere cancellati per adempiere un obbligo legale previsto dal diritto dell'Unione o dello Stato membro cui è soggetto il titolare del trattamento; f) i dati personali sono stati raccolti relativamente all'offerta di servizi della società dell'informazione di cui all'articolo 8, paragrafo 1*».

Tuttavia, il diritto all'oblio non dovrebbe essere considerato come un diritto assoluto[222], ma dovrebbe essere considerato all'interno dell'ecosistema finanziario ed essere operativo nei limiti della relativa compatibilità. Un'interpretazione sistematica del GDPR e della MiFID II garantirebbe che la raccolta di dati da parte degli intermediari sia fondata innanzitutto su un obbligo legale[223]. Essa permetterebbe agli intermediari di raccogliere, elaborare e conservare le informazioni, non solo per fornire correttamente i loro servizi, ma anche per garantire gli strumenti necessari per la trasparenza finanziaria.

In effetti, le persone avranno il diritto di richiedere la cancellazione dei loro dati non appena questa sarà giustificata da uno dei motivi di cui all'art. 17 GDPR. Comunque, gli intermediari finanziari avrebbero il diritto di rifiutare la richiesta qualora venisse proposta nel termine di cinque anni previsto dalla MiFID II.

Allo stesso tempo, il rifiuto potrebbe essere giustificato da altri motivi previsti dall'art. 6 GDPR e, per esempio: a.- nel consenso dato dall'investitore nella misura delle transazioni finanziarie (art. 6, par. 1, lett. a GDPR); b.- nella esecuzione di contratti con i clienti che portano a transazioni finanziarie (art. 6, par. 1, lett. b GDPR); c.- in un ordine del tribunale o un'autorità competente per mantenere quelli basati sulla difesa dell'interesse pubblico (art. 6, par. 1, lett. e GDPR); o d.- sul legittimo interesse di protezione che può essere tenuto dall'investitore per proteggere il pubblico interesse (art. 6, par. 1, lett. e GDPR); oppure e.- nel legittimo interesse di protezione che gli intermediari finanziari possono avere contro eventuali azioni di responsabilità civile e nella protezione anche dei propri clienti (art. 6, par. 1, lett. f GDPR).

Inoltre, sarebbe opportuno che gli intermediari mantengano una comunicazione fluida con gli investitori, affrontando congiuntamente i diritti concessi loro dal GDPR e dalla MiFID II. Questo potrebbe aiutare a evitare o almeno a mitigare i possibili processi burocratici coinvolti nell'*e-discovery*[224] a cui gli intermediari finanziari potrebbero essere vincolati dalle suddette regole e anche migliorare la tutela dei dati personali di un *customer relationship management* realizzato tramite *data analytics*.

In conclusione, le distinte regolamentazioni non sono in contraddizione ma piuttosto in rapporto di reciproco coordinamento. La possibilità per gli individui di richiedere la cancellazione, oltre all'accesso e alla rettifica dei dati, aumenterebbe le possibilità di rilevare le mancanze nel rispetto degli obblighi di conservazione dei dati stabiliti dalla MiFID II, incoraggiando gli intermediari a migliorare gli strumenti, i processi e i sistemi utilizzati a tal fine.

222 Il carattere non assoluto di questo diritto è stata riconosciuta nella sentenza della Corte di giustizia UE del 24 settembre 2019, 'GC e a. (Deindicizzazione di dati sensibili)', causa C-136/17, EU:C:2019:773, punti 55, 56 e 57, nonché dalle 'Linee guida 5/2019 sui criteri per l'esercizio del diritto all'oblio nel caso dei motori di ricerca, ai sensi del RGPD (parte 1)' adottate dallo European Data Protection Board il 7 luglio 2020. Si veda sul tema Assonime, *Diritto all'oblio e deindicizzazione dei motori di ricerca: la giurisprudenza della Corte di giustizia*, in Note e Studi, 1/2020, 1 e ss..

223 Dobrauz, G. (2018), *MIFID2 and GDPR – When Opposites Attract*, PwC.

224 E-discovery can be defined as «*the exchange of data between parties in civil or criminal litigation. The process is largely controlled by attorney who determine what data should be produced based on relevance or withheld based on claims of privilege. Forensic examiners, however, play crucial roles as technical advisors, hand-on collectors and analysts*», Holley, J.O., Luehr, P. , Reust Smith, J. e Schwerha, J. (2010), *Chapter 3 – Electronic Discovery*, in Casey et al. (a cura di), *Handbook of Digital Forensics and Investigation*, 63-133, Academic Press.

4.4 Conclusioni tratte dalla *survey*

Per concludere questa sezione, è opportuno fare riferimento alla *survey* condotta. Anche se non affronta direttamente la questione della *privacy*, i risultati assumono un certo rilievo anche su questo profilo.

Come emerge dall'indagine, sette delle otto società utilizzano sistemi di IA. Questi sono utilizzati soprattutto nei processi di investimento, nell'efficienza operativa e nella gestione e analisi dei dati. In particolare, l'indagine evidenzia che sei delle otto aziende applicano o prevedono di applicare sistemi di IA nella raccolta, gestione e analisi dei dati. Di queste sei società, due hanno sistemi in piena implementazione, altre due sono in fase di sperimentazione e di finalizzazione del progetto e le ultime due stanno valutando possibili applicazioni e opzioni esistenti.

Per quanto riguarda la fonte dei dati utilizzati da queste società, sei delle otto società hanno dichiarato che i dati utilizzati sono in parte interni e in parte esterni, e le altre due hanno dichiarato che utilizzano solo dati da fonti esterne.

Infine, va notato che nessuna di queste società cadrebbe sotto il divieto dell'art. 22 GDPR, poiché non ci sono decisioni completamente automatizzate. Secondo l'indagine, i sistemi di IA utilizzati sono sempre riconducibili all'uomo. Quattro delle otto società consultate hanno dichiarato che il controllo umano è molto forte perché il sistema non viene utilizzato direttamente per prendere decisioni rilevanti per il *business*. Due su otto hanno dichiarato che il sistema è solo parzialmente autonomo con *input* umano che influenza le decisioni in modo predeterminato e solo una delle aziende il sistema IA è autonomo ma con il controllo umano della decisione finale.

Appendice

Intelligenza artificiale e tutela della persona

1 Alcune questioni aperte

Dall'opera pionieristica di Alan Turing del 1950[225] ad oggi, ci sono stati grandi progressi nell'utilizzo di sistemi di intelligenza artificiale (IA). Solo a titolo esemplificativo, attualmente l'intelligenza artificiale permette di generare sistemi capaci di guida automatica (cosiddetti *self driving car, car driverless* o *autonomous car*), identificazione biometrica, diagnosi medica e interazione vocale.

Tale sviluppo è riconducibile alla capacità di modelli statistico-matematici (algoritmi di *machine learning* e di *deep leerning*) di raccogliere e analizzare una mole indefinita di informazioni (i *big data*)[226], di stabilire relazioni e connessioni, di elaborare predizioni sulla base dell'esperienza acquisita, di sviluppare interazioni e assistenza con l'intelligenza umana e, perfino, di fornire decisioni tendenzialmente o totalmente autonome dal controllo dell'uomo[227].

I sistemi di auto-apprendimento e decisione hanno il vantaggio di svolgere operazioni e fornire prognosi molto più accurate di quelle umane, in temi rapidi e con maggiore precisione. Sollevano però questioni di natura interdisciplinare, per le implicazioni dell'innovazione tecnologica in ambito economico, sociale e giuridico, e pongono problematiche anche sul piano etico[228], nell'ottica della tutela della dignità umana, dei diritti e delle libertà fondamentali, per le potenziali conseguenze discriminatorie e per la lesione della sfera personale[229].

225 Turing, A.M. (1950), *Computing machinery and intelligence*, in Mind. A Quarterly Review of Psychology and Philosophy, Volume LIX, Issue 236, October, 433 e ss..

226 Sui *big data* si rinvia a Autorità garante della concorrenza del mercato (AGCM) – Autorità garante per le comunicazioni (AGCOM) – Garante per la protezione dei dati personali (GPDP), *Indagine conoscitiva sui Big Data*.

227 Abriani, N. e Schneider, G. (2021), *Diritto delle imprese e intelligenza artificiale*, Bologna, 21 e ss., distinguono i sistemi di intelligenza artificiale sulla base di due differenti approcci. Secondo un primo approccio i differenti sistemi di intelligenza artificiale sono suddivisi in ragione dei differenti modelli statistico-matematici di elaborazione delle informazioni e di apprendimento automatico (*machine learning, supervised learning, reinforcement learning, unsupervised learning* e *deep learning*). Secondo un secondo approccio, i sistemi di intelligenza artificiale sono identificati sulla base della loro capacità di interazione con l'intelligenza umana, per cui sono distinti sistemi di intelligenza assistita, sistemi di intelligenza aumentata, sistemi di intelligenza amplificata e sistemi di intelligenza autonoma.

228 Si veda Casonato, C. (2022), *L'intelligenza artificiale e il diritto pubblico comparato ed europeo*, in DPCE online, 1, 169; D'Aloia, A. (2019), *Il diritto verso 'il mondo nuovo'. Le sfide dell'Intelligenza Artificiale'*, in BioLaw Journal – Rivista di BioDiritto, n. 1, 3 e ss.; Italiano, G.F. (2019), *Le sfide interdisciplinari dell'intelligenza artificiale*, in AGDE, 1, 9 e ss..

229 Celotto, A. (2019), *Come regolare gli algoritmi. Il difficile bilanciamento fra scienza, etica e diritto*, in AGDE, 1, 47 e ss..

1.1 L'intelligenza algoritmica

In via preliminare, è opportuno precisare che utilizzare il termine 'intelligenza' per questi sistemi è fuorviante in quanto sottintende un perdurante antropocentrismo, estraneo alle applicazioni artificiali[230]. Più esattamente, qualificare 'intelligenti' i sistemi artificiali può essere utile per indagare la capacità delle 'macchine' di sostituirsi all'uomo nell'assolvimento di compiti e nel raggiungimento di risultati predefiniti, a prescindere dal percorso generativo che conduce a un determinato processo. Come è stato sottolineato, ciò non equivale però ad affermare che i sistemi di intelligenza artificiale siano capaci di pensare e di ragionare ma unicamente a sostenere che un determinato comportamento possa essere eseguito senza fare ricorso all'uomo e alle sue capacità cognitive, cioè facendo a meno della sua intelligenza[231].

Tutt'al più, la sinonimia di intelligenza umana e intelligenza artificiale[232] può essere circoscritta esclusivamente all'ambito comunicativo, ovvero agli eventi che i sistemi di intelligenza artificiale producono sul piano dell'effettività sociale, economica e giuridica[233], soltanto qualora un medesimo risultato possa essere raggiunto indifferentemente dall'uomo o da un'applicazione di intelligenza artificiale. In queste ipotesi, per deduzione logica, poiché un determinato comportamento, se compiuto dall'uomo, è qualificato intelligente, allora il medesimo comportamento, eseguito dal sistema artificiale, potrà essere parimenti qualificato intelligente[234].

Un altro aspetto evidenzia la difficile sovrapposizione di intelligenza umana e intelligenza artificiale. I sistemi di intelligenza artificiale mancano del cosiddetto «senso comune» dell'intelligenza umana, cioè della coscienza che consente all'uomo di meditare una decisione sulla base dei valori e degli ideali nei confronti del mondo e, in generale, della collettività[235]. Le decisioni delle applicazioni artificiali sono desoggettivizzate e, quindi, disumanizzate nella misura in cui non sono tendenzialmente sempre riconducibili a una coscienza in forza di un ragionamento per principi e valori, tipicamente umano, ma sono espressione di meccanismi computazionali che definiscono *output* sulla base di classificazioni per omologazioni e inferenze con quanto precedentemente accaduto, difficilmente capaci di plasmare una risposta sempre adeguata ed efficace alle imprevedibili fattispecie concrete[236].

230 Finocchiaro, G. (2020), *Intelligenza artificiale e responsabilità*, in Contr. impr., n. 2, 724. Si veda altresì le notazioni introduttive critiche sull'espressione 'intelligenza artificiale' e, appunto, sul relativo connubio dei termini 'intelligenza' e 'artificiale' di Di Rosa, G. (2021), *Quali regole per i sistemi automatizzati 'intelligenti'?*, in Riv. dir. civ., n. 5, 824-825.

231 Floridi, L. (2022), *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide* (trad. it. a cura di Durante, M.), Milano, 34 e 52.

232 Sulle differenze sui meccanismi di funzionamento del cervello umano e del cervello artificiale si veda Maira, G. (2021), *Intelligenza umana e intelligenza artificiale*, in federalismi.it, n. 7, 1 e ss..

233 Teubner, G. (2019), *Soggetti giuridici digitali? Sullo status privatistico degli agenti software autonomi* (trad. it. a cura di Femia, P.), Napoli, passim, in part. 41-42 e 51-52.

234 Si fa riferimento al cosiddetto test di Turing, basato su un giudizio di natura controfattuale, che consente di qualificare 'intelligente' il processo computazionale di sistemi artificiali, guidati da algoritmi, qualora producano il medesimo risultato raggiungibile dall'uomo.

235 Cirillo, G.P. (2020), *I soggetti giuridici digitali*, in Contr. impr., n. 2, 573, 589.

236 Asaro, P.M. (2011), *A Body to Kick, but Still No Soul to Damn: Legal Perspectives on Robotics*, in Lin, P., Abney, K. e Bekey G. (eds.), *Robot Ethics: The Ethical and Social Implications of Robotics*, Cambridge, 169 e ss.. Si rinvia altresì in senso conforme nella dottrina italiana a Crisci, S. (2018), *Intelligenza artificiale ed etica dell'algoritmo*, in Foro amm., 10,

Più correttamente, l'intelligenza artificiale sembrerebbe realizzare una dissociazione tra agire e intelligenza in quanto una serie di compiti prima svolti dalle persone fisiche sono progressivamente delegati ed eseguiti da sistemi computazionali automatizzati in grado di apprendere dall'esperienza e di fornire decisioni, anche totalmente autonome dall'uomo. Queste innovative modalità di azione degli apparecchi artificiali non sono dotate della capacità di produzione cognitiva umana, ma producono comunque i medesimi effetti e risultati sulla realtà concreta e sulla collettività senza utilizzare alcuna forma di intelligenza[237].

Tantomeno, i sistemi più sofisticati, basati su reti neurali, cioè di capacità di auto-apprendimento, sono in grado di essere assimilati ai processi cognitivi dell'uomo. Seppure forniti di capacità di immagazzinamento di una mole indefinita di informazioni, di meccanismi computazionali statistico-matematici in grado di operazioni fuori dalla portata umana, il loro funzionamento è meramente oggettivo e non guidato da coinvolgimento emotivo tipico del genere umano[238]. Queste modalità di azione divengono ancora più rilevanti poiché evocano rischi e scenari inediti di associazione, derivanti dalla stretta cooperazione tra uomo e sistemi artificiali, e di interconnessione quando più sistemi artificiali agiscono in stretta interdipendenza[239].

1.2 La decisione algoritmica

Gli innovativi meccanismi artificiali stanno progressivamente producendo una «mutazione diretta» e una «mutazione indiretta» della tecnologia che finisce per compromettere in generale i diritti fondamentali e pregiudicare, in particolare, la capacità di autodeterminazione della persona[240].

La «mutazione diretta» avviene per il tramite di un processo di progressivo affidamento delle decisioni in capo a sistemi di intelligenza artificiale tendenzialmente o totalmente autonomi dall'uomo. Questi sistemi divengono 'membri non umani' della

1787. Sulla difficoltà di individuare univocamente solo vantaggi o solo svantaggi dall'utilizzo della decisione robotica e sull'opportunità di alimentare una «feconda interazione» tra uomo e macchina si esprime Mattera, R. (2019), *Decisione negoziale e giudiziale: quale spazio per la robotica?*, in NGCC, n. 1, 198 e ss..

237 L. Floridi, *Etica dell'intelligenza artificiale. Sviluppi, opportunità, sfide*, *op, cit.*, passim.

238 Sull'impossibilità di una sovrapposizione di intelligenza umana e intelligenza artificiale è paradigmatica la posizione del Parlamento europeo nella proposta di Regolamento sulla responsabilità dei sistemi di intelligenza artificiale, allegata alla Risoluzione del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale. Il sistema di intelligenza artificiale è espressamente qualificato quale «sistema basato su software o integrato in dispositivi hardware che mostra un comportamento che simula l'intelligenza» (art. 3, lett. a). Tuttavia, nella proposta successiva del Regolamento sull'intelligenza artificiale del 21 aprile 2021, il riferimento alla simulazione dell'intelligenza da parte dei sistemi artificiali scompare per abbracciare una nozione basata sulle tecniche e sugli approcci alla base del loro funzionamento (art. 3, par. 1).

239 Teubner, *op. cit.* (2019), passim.

240 Si esprime in termini di «mutazione diretta» e di «mutazione indiretta» Simoncini, A. (2019), *L'algoritmo incostituzionale: intelligenza artificiale e il futuro delle libertà*, in BioLaw Journal – Rivista di BioDiritto, 1, 69-76. Si veda anche Papa, A. (2020), *La problematica tutela del diritto all'autoderminazione informativa nella big data society*, in *Liber Amicorum per Pasquale Costanzo* (reperibile su *Consulta Online*), 17 aprile.

società e pongono la questione, non più meramente astratta, del riconoscimento, quantomeno parziale, di soggettività giuridica[241] o, comunque, dell'individuazione di criteri di collegamento dei comportamenti 'artificiali' alla sfera decisionale dell'uomo[242]. In mancanza dell'adesione alternativa a una delle prospettate ricostruzioni giuridiche, anche per il tramite di una certa *fictio iuris*, sarebbe pertanto concreto il pericolo di lasciare pregiudicati beni giuridicamente protetti e di non reintegrare la posizione soggettiva lesa in caso di causazione di effetti dannosi da parte di sistemi automatici o semiautomatici[243].

La dispersione e la proliferazione delle informazioni, lasciate dalla 'scia digitale' di ogni individuo nella rete, sollevano l'altra questione di pari rilievo, ovvero la «mutazione indiretta» delle decisioni dell'uomo[244]. Ad esempio, l'accesso indiscriminato e incontrollato alle informazioni veicolate dalle piattaforme tecnologiche determina un pericolo di condizionamento dell'individuo, ormai «*non più libero, ma profilato, prigioniero di meccanismi che non sa e che non può controllare*»[245]. In linea generale, le informazioni digitali potrebbero anche determinare esternalità positive, qualora consentano alla persona/utente di selezionare beni e servizi migliori sulla base delle

241 La Risoluzione del Parlamento UE del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica ha proposto l'attribuzione di una soggettività giuridica piena almeno per i sistemi di intelligenza artificiale più sofisticati in modo tale da consentire l'applicazione di meccanismi di riparazione per equivalente a causa del danno cagionato dal funzionamento dei medesimi. Tuttavia, questa prospettiva ha ricevuto parere negativo da parte del Comitato economico e sociale europeo (CESE), nel documento pubblicato il 31 agosto 2017, in quanto avrebbe compromesso il principio di prevenzione in materia di responsabilità giuridica. È stata anche successivamente abbandonata sia dalla Risoluzione del Parlamento UE del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale sia dalla proposta di Regolamento (UE) sull'intelligenza artificiale. In dottrina a favore del riconoscimento di soggettività giuridica in capo ai sistemi di intelligenza artificiale Teubner, *op. cit.* (2019), 55-60 e 70-78, e Cirillo, *op. cit.* (2020), 580-581, i quali postulano il riconoscimento di una capacità giuridica parziale, ovvero la capacità di essere rappresentante, in quanto possono assumere decisioni autonome e per ciò stesso possono causare conseguenze in punto di responsabilità.

242 Si veda in tema di imputazione della responsabilità civile derivante dall'utilizzo di sistemi di intelligenza artificiale Leanza, C. (2021), *Intelligenza artificiale e diritto: ipotesi di responsabilità civile nel Terzo Millennio*, in Resp. civ. prev., n. 3, 1011 e ss.; Palmerini, E. (2020), *Soggettività e agenti artificiali: una soluzione in cerca di un problema?*, in Oss. dir. civ. comm., n. 2, 445 e ss.; Ratti, M. (2020), *Riflessioni in materia di responsabilità civile e danno cagionato da dispositivo intelligente alla luce dell'attuale scenario normativo*, in Contr. impr., n. 3, 1174 e ss.; Frattari, N.F. (2020), *Robotica e responsabilità da algoritmo. Il processo di produzione dell'intelligenza artificiale*, in Contr. impr., n. 1, 435 e ss., Ruffolo, U. (2019), *Intelligenza artificiale, machine learning e responsabilità da algoritmo*, in Giur. it., n. 7, luglio, 1689 e ss.; Capilli, G. (2019), *Responsabilità e robot*, in NGCC, n. 3, 621 e ss.; Coppini, L. (2018), *Robotica e intelligenza artificiale: questioni di responsabilità civile*, in Pol. dir., n. 4, 713 e ss.. Sulle indicazioni provenienti dalla Risoluzione del Parlamento europeo del 20 ottobre 2020 su 'Raccomandazioni alla Commissione sul regime di responsabilità civile e intelligenza artificiale' si veda Salanitro, U. (2020), *Intelligenza artificiale e responsabilità: la strategia della Commissione europea*, in Riv. dir. civ., n. 6, 1246 e ss.; Fusaro, A. (2020), *Quale modello di responsabilità per la robotica avanzata? Riflessioni a margine del percorso europeo*, in NGCC, n. 6, 1344 e ss.; Serrao d'Aquino, P. (2021), *La responsabilità civile per l'uso di sistemi di intelligenza nella Risoluzione del Parlamento europeo 20 ottobre 2020: 'Raccomandazioni alla Commissione sul regime di responsabilità civile e intelligenza artificiale'*, in DPER online, n. 1, 248 e ss..

243 È stato rilevato in dottrina (Di Rosa, *op. cit.* (2021), 839) che il riconoscimento di una soggettività giuridica in capo ai sistemi di intelligenza artificiale «*rischia di trasporre in maniera pericolosa e subdola l'utilizzata metafora giuridica (assunta la essenziale esclusività umana) sul piano della realtà*». Si veda altresì Perlingieri, C., *L'incidenza dell'utilizzazione della tecnologia robotica nei rapporti civilistici*, in Rass. dir. civ., 2015, 1246, per il quale il riconoscimento di soggettività in capo ai sistemi di intelligenza artificiale avrebbe l'effetto di «*reputare reale la trasformazione della similitudine in identità e, con riguardo al tema in esame, assimilare l'azione dei robot in grado di svolgere compiti cognitivi complessi a quella di essere viventi*».

244 Punzi, A., *Il diritto e i nuovi orizzonti dell'intelligenza umana*, in AGDE, n. 1, 2019, 28.

245 Rodotà, S., *Il mondo nella rete. Quali i diritti, quali i vincoli*, Roma-Bari, 2014, 39.

informazioni acquisite, incidere, in particolare, sulle imprese produttive per migliorarne le offerte e, in generale, per favorire la competitività del sistema produttivo[246].

Invero, questo potenziale carattere virtuoso è messo in discussione dalla corrispondente profilazione della persona/utente, sviluppata dagli algoritmi, in grado di influenzarne le decisioni individuali. In ambito economico, la personalizzazione algoritmica, se allarga le opportunità e le scelte del consumatore, ne influenza i comportamenti, i quali sono poi sviluppati, anche indirettamente, in forza di suggerimenti di acquisto sulla base della precedente cronologia di navigazione o sulla base della emulazione di comportamenti diffusi nella collettività[247].

La combinazione di queste strategie di elaborazione dei dati e delle informazioni raccolte, sviluppate da meccanismi di indicizzazione dei motori di ricerca, realizza un'invasiva intromissione nella sfera privata e crea *target* specifici e nuove forme oggettivate di segmentazione sociale, basate sulle preferenze manifestate, sulla capacità economica o sulla provenienza territoriale[248].

1.3 La trasparenza algoritmica

Alla questione della mediazione della decisione da parte dei sistemi di intelligenza artificiale è associato un profilo di non minore rilievo, quello della qualità delle informazioni, della loro intrinseca opacità e del carattere fuorviante del funzionamento degli algoritmi nei processi decisionali pubblici e privati in violazione delle garanzie partecipative di ciascun procedimento[249].

Anzitutto, la prestazione del consenso all'utilizzo dei dati personali, in particolare nei confronti delle piattaforme tecnologiche di comunicazione, prevista dalla normativa a tutela della *privacy* dal Regolamento (UE) 2016/679 (GDPR)[250], non assicura una protezione adeguata della sfera personale. L'individuo/utente non è nelle condizioni di conoscere anticipatamente il futuro impiego dei *big data* né, tantomeno, di fornire una preventiva legittimazione all'utilizzo di dati di cui non è effettivo titolare

246 Punzi, *op. cit.*, 35.

247 Si veda Thaler, R.H. e Sunstein, C.R. (2019), *La spinta gentile. La nuova strategia per migliorare le nostre decisioni su denaro, salute e felicità* (trad. it. a cura di Olivieri, A.), Milano, i quali hanno elaborato la definizione di «spinta gentile» (*nudge*), quale metodo per migliorare le decisioni umane nei campi più disparati.

248 Si veda Pezzoli, A. e Tonazzi, A. (2019), *Discriminazione e collusione tacita tra lessico, intelligenza artificiale e algoritmi*, in AGDE, n. 1, 203.

249 Sulla questione della qualità delle informazioni si rinvia a Mittelstadt, B.D., Allo, P., Taddeo, M., Watcher, S., Floridi, L., *The ethics of algorithms: Mapping the debate*, in Big Data & Society, July - December, 2016, 1-21, e, più recentemente, Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Taddeo, M., Floridi, L., *The ethics of algorithms: key problems and solutions*, in AI & Society, January – June, 2021, 1-16. Gli Autori individuano sei questioni etiche, suddivise tra tre questioni epistemiche, due questioni normative e, infine, una rilevante sia a fini epistemici sia a fini normativi. In particolare, i fattori epistemici (prove inconcludenti, prove imperscrutabili e prove fuorvianti) evidenziano la rilevanza della qualità e dell'accuratezza dei dati in mancanza dei quali i risultati del funzionamento degli algoritmi possono portare a conclusioni ingiuste. Le questioni normative riguardano esiti ingiusti ed effetti trasformativi di azioni e decisioni guidate da algoritmi. Da ultimo, le questioni epistemiche e le questioni normative possono rendere difficile la tracciabilità della catena di eventi e fattori che conducono a un determinato risultato.

250 Regolamento (UE) 2016/679, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (Regolamento generale sulla protezione dei dati).

per la loro caratterizzazione in forma massiva[251]. A essere pregiudicato è il diritto dell'individuo a conoscere l'esistenza del procedimento automatizzato e dei processi di profilazione organizzati sulla base delle informazioni personali rilasciate e diffuse a seguito delle attività di navigazione digitali[252].

In secondo luogo, la mancanza di trasparenza dei meccanismi statistico-matematici (i già richiamati algoritmi di *machine learning* e di *deep learning*) potrebbe rendere imprevedibile il comportamento dei sistemi di intelligenza artificiale, specie di quelli più sofisticati basati su capacità di auto-apprendimento e decisione. Gli individui non sarebbero in grado di misurare i loro comportamenti preventivamente sia per l'impossibilità di accedere ai codici sorgente sia per la difficoltà di decifrare le stesse stringhe alfanumeriche di composizione dell'algoritmo[253].

Questa presumibile indecifrabilità degli algoritmi di auto-apprendimento porterebbe a escludere una univoca comprensione delle relative manifestazioni esteriori e della logica a esse sottese non soltanto da parte dei destinatari delle decisioni ma altresì da parte dell'operatore e dello stesso programmatore. Ciò accade in quanto gli algoritmi sviluppano i loro processi sulla base di relazioni statistiche e correlazioni automatiche delle informazioni iniziali, ma anche di quelle successivamente acquisite ed elaborate autonomamente senza alcuna assistenza di una persona fisica. Il tutto si potrebbe tradurre nella violazione del principio del divieto di trattamento algoritmico automatizzato, sancito dall'art. 22 del GDPR, delle garanzie di partecipazione dell'individuo al procedimento di assunzione delle decisioni e di tutela giuridica delle posizioni soggettive eventualmente lese dall'azione di questi sistemi di intelligenza artificiale.

Ipotizzare poi un'effettiva e generalizzata *disclosure* delle cosiddette *black boxes*[254] e degli algoritmi aprirebbe una questione di violazione del divieto di divulgazione di informazioni coperte da privativa industriale o segreto commerciale e sarebbe parzialmente efficace in mancanza di una conoscenza diffusa di meccanismi statistico-matematici da parte della collettività. Indagare e sindacare *a posteriori* la logica e la correttezza di un determinato processo informatico, mediante un giudizio di natura

251 In questo senso De Minico, G. (2019), *Big Data e la debole resistenza delle categorie giuridiche. Privacy e lex mercatoria*, in Dir. pubbl., 1, 89 ss., in part. 92. Si veda altresì De Mari, M., *La profilatura finanziaria algoritmica*, in Orizz. dir. comm., 2021, n. 1, 168, e Mattassoglio, F., *La profilazione dell'investitore nell'era dei big data*, in Riv. trim. dir. ec., 2016, Suppl. n. 1 suppl., 248 e ss..

252 Non sembra cogliere la necessità di un cambio di registro nemmeno la proposta di Regolamento (UE) in materia di mercati digitali che disarticola il consenso alla protezione dei dati personali in tanti specifici consensi senza coniugare il tutto con la previsione di obblighi di protezione a carico delle piattaforme digitali. Si intende fare riferimento agli obblighi dei gatekeeper, previsti dall'art. 5 della proposta di Regolamento (UE) relativo a mercati equi e contendibili nel settore digitale (legge sui mercati digitali). In particolare, la disposizione citata stabilisce che il gatekeeper «si astiene dal combinare dati personali ricavati da tali servizi di piattaforma di base con dati personali provenienti da qualsiasi altro servizio offerto dal gatekeeper o con dati personali provenienti da terzi e dall'accesso con registrazione degli utenti finali ad altri servizi del gatekeeper al fine di combinare dati personali, a meno che sia stata presentata all'utente finale la scelta specifica e che quest'ultimo abbia prestato il proprio consenso ai sensi del regolamento (UE) 2016/679». Su questa proposta di regolamentazione si rinvia al commento di Contaldi, G., *Il DMA (Digital Markets Act) tra tutela della concorrenza e protezione dei dati personali*, in Ordine internazionale e diritti umani, 2021, n. 2, 292 ss.

253 Mula, D. (2019), *Big Data vs. Data Privacy*, in Finocchiaro, G. e Falce, V. (a cura di), *Fintech: diritti, concorrenza, regole. Le operazioni di finanziamento tecnologico*, Bologna, 363 e ss..

254 Questa espressione è stata coniata da F. Pasquale (2015), *The black-box society: The secret algorithms that control money and information*, Cambridge-London.

controfattuale, basato su un paradigma logico-deduttivo, per valutarne la relativa affidabilità, sarebbe quantomeno impossibile a causa della difficoltà di analizzare e processare tutte le variabili e i criteri statistico-matematici in virtù dei quali operano gli algoritmi[255].

1.4 La discriminazione algoritmica

I meccanismi di selezione, immagazzinamento, elaborazione e sfruttamento delle informazioni, basati su algoritmi predittivi di profilazione, non assicurano la realizzazione di un processo e di un risultato finale libero da pregiudizi e malfunzionamenti. Anzi, la tecnologia degli algoritmi dei sistemi di intelligenza artificiale non è per nulla neutrale, rappresenta la realtà incorporata in quelle informazioni e ne costituisce una 'cassa di risonanza', per cui le relative operazioni statistico-computazionali, alla base delle quali sono sviluppati le decisioni dei sistemi di intelligenza artificiale, possono finire per amplificare le distorsioni già esistenti in mancanza di meccanismi preventivi di programmazione e correzione[256].

A ciò si aggiunge la mancanza di un senso comune e di un patrimonio valoriale autonomo dei sistemi di intelligenza artificiale, rispetto a quello esistente nel tessuto sociale, che può portare a ulteriori pregiudizi e produrre fenomeni di esclusione di interi gruppi sociali ed episodi individuali di discriminazione. Gli algoritmi di profilazione, più degli algoritmi di gestione quantitativa, possono generare potenziali discriminazioni qualora siano utilizzati dati di natura personale (come etnia, orientamento sessuale e religioso, opinioni politiche, appartenenza sindacale)[257].

Volendo esemplificare, se le informazioni raccolte contengono già in fase di partenza errori di valutazione o pregiudizi (*bias*), il fenomeno algoritmico, basato su analisi computazionali, non potrà che ripetere e amplificare la realtà esistente, riproducendolo anche un'infinità di volte[258]. Più esattamente, l'algoritmo già in fase di creazione potrebbe contenere '*in re ipsa*' i pregiudizi cognitivi dello sviluppatore potenzialmente discriminatori. Ciò può avvenire infatti sia qualora vengano selezionate *ex ante* specifiche categorie di dati sensibili che producano scelte inclusive/esclusive basate (solo o soprattutto) su tali informazioni, sia nel caso di *bias* statistici, in virtù dei quali, nonostante l'involontarietà dello sviluppatore, permarrebbe una scelta discriminatoria

255 Sui fattori che rendono oscuri i modelli algoritmici si veda Lo Sapio, G. (2021), *La black box: l'esplicabilità delle scelte algoritmiche quale garanzia di buona amministrazione*, in federalismi.it, n. 16, 121. In senso contrario si veda Donati, F., *Intelligenza artificiale e giustizia*, in Rivista AIC, 2020, n. 1, 428.

256 Si veda Mobilio, G. (2020), *L'intelligenza artificiale e i rischi di una 'disruption' della regolamentazione giuridica*, in BioLaw Journal – Rivista di BioDiritto, 2, 401 ss., in part. 406-407, il quale afferma in generale che la presunta neutralità della tecnologia è in realtà «un equivoco» dal momento che la tecnologia è utilizzata per raggiungere un qualsiasi scopo.

257 Sul punto si rinvia a Paracampo, M.T. (2019), *FinTech tra algoritmi, trasparenza e algo-governance*, in Diritto della banca e del mercato finanziario, n. 2, 12 e ss., la quale ha rilevato che gli algoritmi di profilazione «destano maggiore preoccupazione ai fini di eventuali discriminazioni, soprattutto in funzione dei processi decisionali automatizzati e della tutela dei dati personali».

258 L'ipotesi di avveramento di un siffatto rischio viene solitamente definito dai data scientist come 'GIGO', ovvero «*garbage in garbage out*». Si veda Carcaterra, A. (2019), *Macchine autonome e decisione robotica*, in Carleo, A. (a cura di), *Decisione robotica*, Bologna, 38 e ss.

figlia di una errata distribuzione statistica dei dati utilizzati durante la fase di addestramento, che non rappresenta perfettamente tutto il campione[259].

Con riferimento all'ambito economico e finanziario, nel quale l'intelligenza artificiale ha già ricevuto una vasta applicazione, già possono essere registrati alcuni episodi di discriminazione. La definizione delle strategie di investimento e la valutazione del merito creditizio (cosiddetto *credit scoring*) sono sempre più affidate ad algoritmi di profilazione nei confronti di utenti che richiedono raccomandazioni di investimento o finanziamenti a istituti di credito. Tuttavia, se questi meccanismi di natura statistica consentono di elaborare e processare rapidamente una grande massa di dati e a velocizzarne il relativo processo di erogazione, possono altresì realizzare discriminazioni. Categorie di persone o singoli possono essere escluse dall'erogazione del credito o ricevere finanziamenti a condizioni maggiormente onerose rispetto a quelle ottenute da altri. Ciò può avvenire perché l'intelligenza artificiale può amplificare stereotipi pregiudizievoli a causa degli effetti di una relazione statistica tra l'appartenenza a una minoranza razziale, alle classi meno abbienti o ad altre condizioni materiali o sociali[260].

Come illustrato, le potenziali discriminazioni sono alimentate qualora si utilizzino algoritmi di *deep learning*, non previamente addestrati e allenati dall'attività dell'uomo, ma non è escluso che le medesime discriminazioni siano il portato della programmazione iniziale e, quindi, di ideali e valori di progettatori o sviluppatori di *software*[261].

Queste ultime forme di discriminazione potrebbero essere eliminate direttamente nei confronti di operatori, fornitori e sviluppatori, responsabili della immissione sul mercato e messa in servizio di sistemi di intelligenza artificiale, mediante le disposizioni attualmente in vigore. Viceversa, le discriminazioni prodotte autonomamente dagli algoritmi richiedono invece la predisposizione di un complesso regolamentare che renda la programmazione automatica algoritmica incapace di produrre pregiudizi fondati su differenze di trattamento e che rispecchiano parzialmente la realtà. In altri termini, l'algoritmo dovrebbe essere addestrato in modo tale che determinati fattori siano ignorati o incorporati con la costruzione di meccanismi di carattere dissuasivo capaci di prevenire effetti discriminatori[262].

259 Stradella, E., *Stereotipi e discriminazioni: dall'intelligenza umana all'intelligenza artificiale*, in *Liber Amicorum per Pasquale Costanzo* (reperibile su *Consulta Online*), 30 marzo 2020, 2-3. Si rinvia altresì a Molaschi, V., *Algoritmi e nuove schiavitù*, in federalismi.it, 2021, n. 18, 205 e ss., in part. 226-228; P. Zuddas, *Intelligenza artificiale e discriminazioni*, in *Liber Amicorum per Pasquale Costanzo* (reperibile su *Consulta Online*), 16 marzo 2020, 1 ss.; Giorgini Pignatiello, G., *Il contrasto alle discriminazioni algoritmiche: dall'anarchia giuridica alle Digital Authorities?*, in federalismi.it, 2021, n. 16, 164 e ss..

260 Mostacci, E. (2022), *L'intelligenza artificiale in ambito economico e finanziario: rischi e prospettive*, in DPCE online, n. 1, 361 e ss., in part. 367-368. Si veda, con particolare riguardo all'accesso al credito degli immigrati e alle relative discriminazioni, Mattarella, G. (2020), *Big Data e accesso al credito degli immigrati: discriminazioni algoritmiche e tutela del consumatore*, in Giur. comm., n. 4, 696 e ss..

261 Si veda Colapietro, C., e Moretti, A. (2020), *L'Intelligenza Artificiale nel dettato costituzionale: opportunità, incertezze e tutela dei dati personali*, in BioLaw Journal – Rivista di BioDiritto, n. 3, 373-374.

262 Un'indicazione potrebbe essere ricavata dal Regolamento GDPR. Il Considerando n. 71 stabilisce l'opportunità che «*il titolare del trattamento utilizzi procedure matematiche o statistiche appropriate per la profilazione, metta in atto misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificate i fattori che comportano inesattezze dei dati*».

2 I principi etici sull'intelligenza artificiale nell'Unione europea

L'utilizzo dell'intelligenza artificiale rischia di sottoporre a revisione critica o, addirittura, scardinare alcune categorie concettuali elaborate in molti campi del diritto[263]. Da qui discende l'interesse dell'Unione europea a definire regole per disciplinare i sistemi di intelligenza artificiale. Uno specifico riferimento ai principi etici è presente nella Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica[264]. In particolare, nel declinare i principi etici, il Parlamento europeo pone l'accento «*sul principio della trasparenza, nello specifico sul fatto che dovrebbe sempre essere possibile indicare la logica alla base di ogni decisione presa con l'ausilio dell'intelligenza artificiale che possa avere un impatto rilevante sulla vita di una o più persone*» e ritiene che «*debba sempre essere possibile ricondurre i calcoli di un sistema di intelligenza artificiale a una forma comprensibile per l'uomo e che i robot avanzati dovrebbero essere dotati di una 'scatola nera' che registri i dati su ogni operazione effettuata dalla macchina, compresi i passaggi logici che hanno contribuito alle sue decisioni*»[265].

Un tentativo più organico in materia è stato effettuato nel 2018 dal Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale istituto dalla Commissione Europea con la pubblicazione degli 'Orientamenti etici per un'IA affidabile'[266].

Sebbene non vincolanti, tali Orientamenti hanno il pregio di offrire un quadro d'insieme dei principi fondamentali per una intelligenza artificiale affidabile, riconoscendo la necessità che essa sia 'antropocentrica' e tesa allo sviluppo e promozione di una società equa.

I cardini essenziali prodromici a una IA affidabile sono riconosciuti in tre componenti che devono essere sempre presenti durante l'intero ciclo di vita del sistema, ovvero la legalità, l'eticità e la robustezza[267]. Per essere affidabile infatti l'IA deve garantire la compatibilità con le norme etiche ed essere robusta sia dal punto di vista tecnico che sociale. Questi presupposti di legalità, eticità e robustezza sono altresì definiti come correlati e complementari poiché si integrano a vicenda.

Tra i passaggi chiave degli Orientamenti, va sottolineata la declinazione del principio di uguaglianza, che in un contesto di IA, implica che «*il funzionamento del*

263 In questo senso si esprime la Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica ha condiviso questa opinione, precisando che «*sono palesi le carenze dell'attuale quadro normativo anche in materia di responsabilità contrattuale, dal momento che le macchine progettate per scegliere le loro controparti, negoziare termini contrattuali, concludere contratti e decidere se e come attuarli rendono inapplicabili le norme tradizionali; [...] ciò pone in evidenza la necessità di norme nuove, efficaci e al passo con i tempi che corrispondano alle innovazioni e agli sviluppi tecnologici che sono stati di recente introdotti e che sono attualmente utilizzati sul mercato*».

264 Parlamento Europeo, *Risoluzione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica*, www.europarl.europa.eu (2017), 10-14.

265 Ivi, p. 12

266 Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale, *Orientamenti etici per un'IA affidabile*, in www.ec.europa.eu, 2018.

267 Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale, *Orientamenti etici*, cit., 2.

sistema non possa generare risultati ingiustamente distorti (ad esempio, i dati utilizzati per istruire i sistemi di IA dovrebbero essere il più inclusivi possibile e rappresentare gruppi di popolazione diversi)»[268].

Il gruppo di esperti individua quattro principi etici, radicati nei diritti fondamentali ai quali occorre aderire per garantire che i sistemi di IA siano sviluppati, distribuiti e utilizzati in modo affidabile. Sono definiti come imperativi etici affinché gli operatori del settore dell'IA si adoperino sempre per aderirvi. Essi consistono nel rispetto dell'autonomia umana, nella prevenzione dei danni, nell'equità e nell'esplicabilità.

Appare molto rilevante rimarcare la centralità della esplicabilità, componente necessaria per creare e mantenere la fiducia degli utenti, presupposto quest'ultimo essenziale per la tenuta del sistema finanziario. Tale principio infatti implica che «*i processi devono essere trasparenti, le capacità e lo scopo dei sistemi di IA devono essere comunicati apertamente e le decisioni, per quanto possibile, devono poter essere spiegate a coloro che ne sono direttamente o indirettamente interessati*».

Tali informazioni sono essenziali perché una decisione non può essere debitamente impugnata senza conoscerli. Tuttavia, il gruppo di esperti è consapevole che non sempre è possibile spiegare perché un modello ha generato una particolare decisione, con particolare riferimento ai sistemi di 'scatola nera', pertanto possono essere necessarie altre misure per garantire l'esplicabilità, come «*la tracciabilità, la verificabilità e la comunicazione trasparente sulle capacità del sistema*».

Gli esperti sottolineano che gli Orientamenti costituiscono principi che devono essere resi effettivi attraverso la loro traduzione in requisiti concreti.

Il gruppo di esperti ha stilato un elenco, ancorché non esaustivo, di sette requisiti[269]. Al riguardo, diventano essenziali tutte le diverse categorie di portatori di interessi nel garantire che i requisiti siano soddisfatti: *in primis*, «*gli sviluppatori che attuano e applicano i requisiti ai processi di progettazione e sviluppo*»; poi i distributori nel garantire che i sistemi – che utilizzano e i prodotti e i servizi che offrono – soddisfino i requisiti e infine «*gli utenti finali e la società in generale che devono essere informati su tali requisiti e hanno la facoltà di domandarne il rispetto*»[270].

Il citato documento inoltre descrive i metodi tecnici e non tecnici funzionali a garantire un'IA affidabile che possono essere incorporati nelle fasi di progettazione, sviluppo e utilizzo di un sistema di IA. Tra questi, sono indicati i metodi per garantire i

268 Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale, *Orientamenti etici*, cit., 12.

269 Nel dettaglio sono: 1) Intervento e sorveglianza umani (inclusi i diritti fondamentali, l'intervento umano e la sorveglianza umana); 2) Robustezza tecnica e sicurezza (inclusi la resilienza agli attacchi e la sicurezza, il piano di emergenza e la sicurezza generale, la precisione, l'affidabilità e la riproducibilità); 3) Riservatezza e governance dei dati (inclusi il rispetto della riservatezza, la qualità e l'integrità dei dati e l'accesso ai dati; 4) Trasparenza (incluse la tracciabilità, la spiegabilità e la comunicazione); 5) Diversità, non discriminazione ed equità (incluse la prevenzione di distorsioni inique, l'accessibilità e la progettazione universale, e la partecipazione dei portatori di interessi); 6) Benessere sociale e ambientale (inclusi la sostenibilità e il rispetto ambientale, l'impatto sociale, la società e la democrazia); 7) Accountability (inclusi la verificabilità, la riduzione al minimo degli effetti negativi e la loro segnalazione, i compromessi e i ricorsi).

270 Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale, *Orientamenti etici*, cit., 16.

principi etici fin dalla progettazione (*by-design*) sul presupposto che la conformità alle norme può essere implementata nella progettazione del sistema di IA.

3 La proposta di Regolamento (UE) sull'intelligenza artificiale

Gli orientamenti etici degli esperti confluiscono nell'ambito di una strategia europea per l'intelligenza artificiale, insieme a una serie di documenti elaborati dalle istituzioni UE[271]. Di questa strategia fa parte la '*Proposta di Regolamento del Parlamento Europeo e del Consiglio, che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*', pubblicata il 21 aprile 2021.

Le istituzioni UE stanno lavorando all'introduzione di questa ambiziosa disciplina in materia di intelligenza artificiale, con il proposito di facilitare l'innovazione tecnologica e di limitarne le criticità per la costruzione di un quadro giuridico compatibile con il patrimonio costituzionale europeo. Gli obiettivi di fondo della proposta di Regolamento (UE) sono due: in primo luogo, l'armonizzazione della disciplina in materia di intelligenza artificiale tra gli Stati membri UE e la predisposizione di misure destinate ad assicurare l'instaurazione e il funzionamento del mercato interno (art. 114 Tfue); in secondo luogo, l'adozione di specifiche regole sulla protezione dei dati personali basate sull'art. 16 Tfue, come le particolari restrizioni all'utilizzo di sistemi di intelligenza artificiale per l'identificazione biometrica da remoto in spazi accessibili al pubblico ai fini di attività di contrasto[272].

Rispetto all'atteggiamento assunto nei confronti della regolazione dell'intelligenza artificiale negli Stati Uniti d'America e in Cina assumono particolare rilevanza gli approcci, alla base della proposta legislativa delle istituzioni UE, per il perseguimento e il mantenimento di quei principi e valori che definiscono l'impianto costituzionale dei Trattati (Tue e Tfue) e della Carta dei diritti fondamentali UE. In particolare, i primi commentatori hanno individuato sostanzialmente quattro approcci: un approccio orizzontale, un approccio basato sul rischio, un approccio antropocentrico e un approccio flessibile e collaborativo[273].

271 Tra gli altri documenti è possibile citare la 'Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni del 25 aprile 2018, L'intelligenza artificiale per l'Europa', il 'Libro bianco sull'intelligenza artificiale della Commissione europea del 19 febbraio 2020 – Un approccio europeo all'eccellenza e alla fiducia' e la 'Risoluzione del Parlamento UE del 20 ottobre 2020 recante raccomandazioni alla Commissione su un regime di responsabilità civile per l'intelligenza artificiale'.

272 Si veda la relazione alla proposta di Regolamento (UE) sull'intelligenza artificiale, pp. 6-7.

273 La proposta di Regolamento (UE) in materia di intelligenza artificiale ha sollecitato una serie di commenti in dottrina. Si veda ex multis Donati, F. (2021), *Diritti fondamentali e algoritmi nella proposta di regolamento sull'intelligenza artificiale*, in *Il Diritto dell'Unione europea*, nn. 3-4, 453 e ss.; Contaldi, G. (2021), *Intelligenza artificiale e dati personali*, in *Ordine internazionale e diritti umani*, n. 5, 1193 e ss.; Casonato, C. e Marchetti, B. (2021), *Prime osservazioni sulla proposta di regolamento della Commissione UE in materia di intelligenza artificiale*, in BioLaw Journal – Rivista di BioDiritto, n. 3, 415 e ss.; Proietti, G. (2021), *Intelligenza artificiale: una prima analisi della proposta di regolamento europeo*, in *dirittobancario.it*, maggio; Marchianò, G. (2021), *Proposta di Regolamento della Commissione europea del 21 aprile 2021 sull'intelligenza artificiale con particolare riferimento alle IA ad alto rischio*, in Riv. giur. Ambientediritto.it, n. 2, 1 e ss..

Con riferimento specifico alla dimensione antropocentrica della proposta, secondo un'analisi preliminare di una parte della dottrina questa sarebbe perseguita mediante la combinazione di diverse tecniche di protezione: il principio di prevenzione, il principio di controllo e il principio di collaborazione[274].

La combinazione di queste differenti modalità di tutela dell'individuo delinea una regolazione differenziata dei sistemi di intelligenza artificiale a seconda del livello di rischio che i medesimi presentano. Sono previsti alcuni divieti in materie di pratiche di intelligenza artificiale 'inaccettabili'[275] e alcune prescrizioni per i sistemi di intelligenza artificiale 'ad alto rischio'[276]. Più nel dettaglio, a seguito di un giudizio di bilanciamento, le istituzioni UE, con i divieti, in applicazione del principio di prevenzione, intendono inibire gli utilizzi dell'intelligenza artificiale immediatamente lesivi di diritti e valori costituzionali e quelli potenzialmente capaci di causare discriminazioni. Ammettono espressamente i cosiddetti sistemi a rischio limitato, i quali richiedono unicamente obblighi di trasparenza a carico di fornitori e utenti (art. 52)[277].

A metà strada si colloca la regolamentazione dei sistemi ad alto rischio, per i quali invece non sono escluse l'immissione sul mercato, la messa in servizio o l'uso ma, in applicazione del principio di precauzione, sono stabilite una serie di prescrizioni e cautele in quanto detti sistemi sono potenzialmente in grado di procurare un rischio per la salute, la sicurezza e i diritti fondamentali delle persone fisiche. Sono ammessi nel mercato unico a condizione di rispettare alcuni requisiti, previa valutazione di conformità, ai quali sono affiancati determinati obblighi a carico di fornitori e utenti[278].

274 Sull'individuazione di questi principi si veda Alpa, G. (2021), *Quale modello europeo per l'intelligenza artificiale*, in Contr. impr. n. 4, 1011.

275 In primo luogo, è vietato l'utilizzo di sistemi di IA che, mediante artifici, possono ingannare (inconsapevolmente) e pregiudicare un individuo (art. 5, par. 1, lett. a). In secondo luogo, è vietato l'utilizzo di sistemi di IA che, approfittando della situazione di vulnerabilità di un individuo, ne ledono la relativa sfera giuridica fisicamente o psicologicamente (art. 5, par. 1, lett. b). In terzo luogo, è vietato il cosiddetto *Social Scoring* (o *Social Credit System)*, mediante il quale è valutato il comportamento degli individui (art. 5, par. 1, lett. c). Infine, è vietato l'utilizzo di sistemi di identificazione biometrica remota in luoghi aperti al pubblico. Fanno eccezione le ipotesi nelle quali i sistemi di IA sono adoperati i) per la ricerca mirata di specifiche e potenziali vittime di reati, compresi bambini scomparsi; ii) per la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o per l'incolumità fisica delle persone o in caso di attacco terroristico; iii) per l'individuazione, la localizzazione, l'identificazione o il perseguimento di un autore o sospettato di alcuni reati in particolare punibile nello Stato membro interessato con una pena detentiva per un periodo massimo di almeno tre anni (art. 5, par. 1, lett. d).

276 L'utilizzo di questi sistemi di IA è subordinato a una serie di requisiti che devono essere verificati da un'apposita autorità. In particolare, l'art. 6, par. 1, della proposta definisce ad alto rischio il sistema di IA in presenza di alcune condizioni: «*a) il sistema di IA è destinato a essere utilizzato come componente di un prodotto, o è esso stesso un prodotto, disciplinato dalla normativa di armonizzazione dell'Unione elencata nell'allegato II; (b) il prodotto, il cui componente di sicurezza è il sistema di IA, o il sistema di IA stesso in quanto prodotto è soggetto a una valutazione della conformità da parte di terzi ai fini dell'immissione sul mercato o della messa in servizio di tale prodotto ai sensi della normativa di armonizzazione dell'Unione elencata nell'allegato II*».

277 A carico di fornitori, i quali devono garantire che i sistemi di IA destinati a interagire con persone fisiche siano progettati e sviluppati in modo tale che le persone fisiche siano informate del fatto di stare interagendo con un sistema di IA, a meno che non risulti evidente dalle circostanze e dal contesto di utilizzo. A carico di utenti per i sistemi di IA:- utilizzati per il riconoscimento emotivo o per la categorizzazione sulla base di dati biometrici, - generatori o manipolatori di immagini o contenuti audio o video artificiali tali da apparire falsamente autentici o veritieri per una persona (cosiddetti *deep fake*).

278 Nell'ambito di questi sistemi ad alto rischio possono essere ricompresi, per esempio, macchine, giocattoli, ascensori, apparecchi e sistemi di protezione destinati a essere utilizzati in atmosfera potenzialmente esplosiva, apparecchiature radio, attrezzature a pressione, attrezzature per imbarcazioni da diporto, impianti a fune, apparecchi che bruciano carburanti gassosi, dispositivi medici e dispositivi medico-diagnostici in vitro. Un'ulteriore indicazione dei sistemi ad

A seguito di queste riflessioni preliminari, si potrebbe concludere che la disciplina abbia unicamente carattere preventivo e oppositivo e non intenda disciplinare la fase di riparazione per eventuali danni provocati dai sistemi di IA. Vi sarebbe l'intenzione di salvaguardare le regole previgenti e imputare la responsabilità in capo all'uomo, rimanendo applicabili le regole in materia da responsabilità da prodotti difettosi qualora il sistema di IA sia considerato un prodotto 'difettoso'. Manca, infatti, qualsiasi riconoscimento di personalità giuridica in capo ai sistemi di intelligenza artificiale, con un tentativo di individuare un collegamento dei relativi effetti con i fornitori e gli utenti. Si colloca in questa direzione la previsione, per i sistemi ad alto rischio, del dovere di sorveglianza umana (cosiddetto *duty of human oversight*), in forza del quale debbono essere progettati e sviluppati in modo tale da potere essere supervisionati dall'uomo (art. 14) e della garanzia in capo agli utenti di interpretare l'*output* del sistema e di permettere un monitoraggio permanente (art. 13).

* * * * *

Il tutto appare in continuità con il divieto delle persone fisiche di essere sottoposte a un trattamento interamente automatizzato, previsto dall'art. 22 del GDPR (si veda anche il Capitolo 4.2 nella sezione '*Profili di investor protection*'). A dire il vero, il paradigma dell'umanizzazione di qualsiasi sistema di intelligenza artificiale è destinato ben presto a infrangersi. Difficilmente un supervisore assumerà il rischio e la responsabilità di disattendere una decisione oggettiva, imparziale e dotata di un elevato livello statistico di probabilità[279].

Secondo altra dottrina il mancato raggiungimento del proclamato antropocentrismo della proposta sarebbe percepibile anche in altri caratteri. I valori europei di dignità umana, libertà, uguaglianza, democrazia e diritti fondamentali (in particolare, il diritto alla non discriminazione, alla protezione dei dati, alla *privacy*) non sono oggetto di valorizzazione nel corpo della proposta di Regolamento: il testo utilizza solo una volta i termini «umano» e «dignità» a dispetto delle premesse[280]. Più in particolare, l'influenza dei diritti e delle libertà non copre tutta la catena di produzione degli effetti dei sistemi di intelligenza artificiale ma unicamente il momento della raccolta dei dati e non è estesa alla fase successiva di elaborazione dei dati e di produzione di *output* decisionali per la difficoltà di comprendere il funzionamento e i meccanismi che governano gli algoritmi di auto-apprendimento (la già citata *black box*). Una protezione nei confronti dell'individuo non potrebbe tantomeno essere assicurata dal GDPR per una duplice circostanza: il funzionamento dell'intelligenza artificiale non è basato

alto rischio avviene per il tramite del rinvio del par. 2 dell'art. 6 della proposta all'allegato III della proposta. L'allegato contiene un elenco di ambiti in cui potrebbe essere utilizzata la tecnologia in questione con la precisazione della finalità per la quale il sistema di IA potrebbe essere utilizzato.

279 Casonato, C., e Marchetti, B., *op. cit.* (2021), 429.

280 Il Considerando 15 della proposta indica il parametro di riferimento per valutare l'ammissibilità dei sistemi di intelligenza artificiale e, precisamente «*i valori dell'Unione relativi al rispetto della dignità umana, della libertà, dell'uguaglianza, della democrazia e dello Stato di diritto e dei diritti fondamentali dell'Unione, compresi il diritto alla non discriminazione, alla protezione dei dati e della vita privata e i diritti dei minori*».

esclusivamente su dati personali; i rischi di *bias* non sono manifestati necessariamente per effetto di dati personali ma a causa dei *big data*[281].

Complessivamente la proposta intende frenare l'espansione dei sistemi di intelligenza artificiale, ma non promuove meccanismi di protezione diretti, a favore degli individui vittime di decisioni automatizzate, discriminatorie o comunque errate, per ripristinare la posizione giuridica lesa. Com'è stato evidenziato dallo European Data Protection Board (EDPD) e dallo European Data Protection Supervisory (EDPS), sarebbe stato opportuna la promozione di modalità proattive e tempestive per informare gli utenti sullo stato decisionale del sistema, anche mediante la predisposizione di allarmi rapidi su risultati nocivi, per consentire alle persone destinatarie degli *output* di reagire e impugnare le relative decisioni[282].

4 Osservazioni conclusive

Come illustrato, progressivamente gli algoritmi hanno acquisito crescenti capacità e hanno assunto un ruolo decisionale, per cui diviene fondamentale comprendere le modalità con le quali essi agiscono, determinare le ragioni della loro decisione e individuare meccanismi di imputazione della responsabilità per il loro comportamento.

Le istituzioni UE, tanto nei principi enucleati dagli Orientamenti etici e nei metodi per implementarli quanto nella nuova proposta legislativa di regolamentazione dei sistemi di intelligenza artificiale, abbandonano un certo *laissez-faire* e propongono un ambizioso tentativo di conciliare le esigenze di innovazione e di tutela dei diritti e delle libertà fondamentali. Tuttavia, secondo alcuni commentatori, la proposta di regolamentazione, a livello UE, non sembrerebbe offrire complessivamente un disegno unitario per una molteplicità di ragioni.

Anzitutto, l'intelligenza artificiale è difficilmente comprensibile e la relativa disciplina diviene immediatamente obsoleta per i rapidi progressi della tecnica, di fronte ai quali le scienze sociali e il diritto faticano ad adeguarsi. Di questo la disciplina in corso di definizione ha pensato di farsi carico sia abbandonando il principio di neutralità tecnologica sia predisponendo meccanismi di aggiornamento della normativa mediante procedimenti di *soft-law* e di *self-regulation*, dei quali la prassi applicativa consentirà di accertare l'efficacia e l'efficienza.

In secondo luogo, sarebbe necessario interiorizzare i meccanismi di protezione di diritti e libertà fondamentali all'interno dei circuiti tecnologici perché la fase di elaborazione delle informazioni e di produzione degli *output* non produca pregiudizi nei confronti dei destinatari delle decisioni. Per la produzione di questo processo virtuoso,

281 Si veda Pollicino, O., De Gregorio, G. e Paolucci, F. (2021), *La proposta di Regolamento sull'intelligenza artificiale: Verso una nuova governance europea*, in Privacy & Data Protection Technology Cybersecurity, n. 3.

282 Si veda European Data Protection Board (EDPD) – European Data Protection Supervisory (EDPS), *Parere congiunto 5/2021 sulla proposta di Regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale)*, 18 giugno 2021, 22.

ovviamente, è opportuno rinunciare a sistemi immediatamente lesivi della dignità umana.

Allo stesso modo, come accade in ambito ambientale, dovrebbero essere implementati negli algoritmi meccanismi precauzionali. Ciò dovrebbe avere l'obiettivo di evitare che i sistemi di intelligenza artificiale producano potenzialmente effetti pregiudizievoli nei confronti dei singoli o della collettività. È necessaria però, dapprima, una reciproca integrazione di competenze tecnologiche e giuridiche, affinché possano essere plasmati sistemi artificiali e algoritmi costituzionalmente conformi[283] e, poi, come sottolineato in dottrina[284], immettere i valori alla base del patrimonio costituzionale in forma digitale negli algoritmi e integrare la razionalità algoritmica con l'etica tipicamente umana mediante una regolazione giuridica[285].

Un ultimo profilo da tenere in considerazione riguarda la responsabilità per i danni provocati dai sistemi di intelligenza artificiale. Sviluppare una disciplina di carattere unicamente preventivo e oppositivo senza una contestuale declinazione di profili di *enforcement* potrebbe determinare una cornice regolamentare non esattamente proporzionata ai rischi e alla rilevanza dei rischi per la tutela dei diritti e delle libertà fondamentali. In definitiva, sia che si riescano o meno a programmare 'comportamenti etici' dell'algoritmo, sia che si arrivi a riconoscere uno status giuridico specifico o una personalità elettronica ai robot più sofisticati[286], al momento il 'fattore umano', nel monitorare i processi algoritmici assumendosi la responsabilità finale dei processi di supervisione, sembra essere ancora il presidio più efficace ed etico per la protezione dei diritti, inclusi quelli dei risparmiatori[287].

283 Si veda Enriques, *op. cit.* (2017).

284 Celotto, *op. cit.* (2019), 59.

285 *Ibidem*.

286 Si veda il punto n. 59, lett. f) della Risoluzione recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica, www.europarl.europa.eu (2017).

287 Si veda Panisi e Perrone, *op. cit.* (2018).

Quaderni FinTech

9 – giugno 2022 **L'intelligenza artificiale nell'*asset* e nel *wealth management***
N. Linciano, V. Caivano, D. Costa, P. Soccorso,
T.N. Poli, G. Trovatore; in collaborazione con Assogestioni

8 – aprile 2021 **La portabilità dei dati in ambito finanziario**
A cura di A. Genovese e V. Falce

7 – settembre 2020 **Do investors rely on robots?**
Evidence from an experimental study
B. Alemanni, A. Angelovski, D. Di Cagno, A. Galliera,
N. Linciano, F. Marazzi, P. Soccorso

6 – dicembre 2019 **Valore della consulenza finanziaria e *robo advice* nella percezione degli investitori**
Evidenze da un'analisi qualitative
M. Caratelli, C. Giannotti, N. Linciano, P. Soccorso

5 – luglio 2019 **Marketplace lending**
Verso nuove forme di intermediazione finanziaria?
A. *Sciarrone Alibrandi, G. Borello, R. Ferretti, F. Lenoci,*
E. Macchiavello, F. Mattassoglio, F. Panisi

4 – marzo 2019 **Financial Data Aggregation e Account Information Services**
Questioni regolamentari e profili di business
A. Burchi, S. Mezzacapo, P. Musile Tanzi, V. Troiano

3 – gennaio 2019 **La digitalizzazione della consulenza in materia di investimenti finanziari**
Gruppo di lavoro CONSOB, Scuola Superiore Sant'Anna di Pisa, Università Bocconi,
Università di Pavia, Università di Roma 'Tor Vergata', Università di Verona

2 – dicembre 2018 **Il FinTech e l'economia dei dati**
Considerazioni su alcuni profili civilistici e penalistici
Le soluzioni del diritto vigente ai rischi per la clientela e gli operatori
E. Palmerini, G. Aiello, V. Cappelli
G. Morgante, N. Amore, G. Di Vetta, G. Fiorinelli, M. Galli

1 – marzo 2018 **Lo sviluppo del FinTech**
Opportunità e rischi per l'industria finanziaria nell'era digitale
C. Schena, A. Tanda, C. Arlotta, G. Potenza