GOVERNANCE DEI BIG DATA E ASPETTI GIURIDICI

R. Mugavero, E. Marvelli

UNIVERSITÀ DEGLI STUDI
DELLA REPUBBLICA DI SAN MARINO
DIPARTIMENTO DI STUDI STORICI
CENTRO DI FORMAZIONE
SULLA SICUREZZA

CEMEC
European Centre
for Disaster Medicine
EUR-OPA
EUR-OPA MAJOR HAZARDS AGREEMENT
ACCORD EUR-OPA RISQUES MAJEURS
COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

GOVERNANCE DEI BIG DATA E ASPETTI GIURIDICI

Premessa

I principali trends di sviluppo delle c.d. Tecnologie Emergenti e Dirompenti (EDTs) hanno evidenziato come il volume, la varietà e la velocità di generazione dei Dati (c.d. Big Data) rendano indispensabile l'implementazione dei processi di Data Governance per l'intero ciclo di vita degli stessi (c.d. Data Life Cycle), con specifico riferimento alla:

- capacità di raccolta, memorizzazione e trattazione dei dati;
- disponibilità di sistemi di analisi e predizione di supporto al processo decisionale (c.d. Data Analytics);
- applicazione di adeguate forme di protezione che impediscano l'uso indiscriminato e, potenzialmente, malevolo dei dati e del patrimonio informativo estratto da essi (c.d. Data Mining).

Tale processo richiede la definizione di nuove ed appropriate regole etico-giuridiche, in grado di bilanciare adeguatamente esigenze di privacy e disponibilità delle informazioni, riassumibili nel paradigma Privacy vs. Sicurezza Nazionale.

Il presente documento intende indagare potenzialità e criticità derivanti dal crescente impiego dei Big Data nel contesto internazionale, con particolare attenzione alle esigenze di armonizzazione ed interoperabilità di tecnologie e culture giuridiche applicate al campo della security e dei nuovi scenari di minaccia asimmetrici.

L'attività di indagine si estenderà alla definizione degli elementi per la governance e l'utilizzo dei Big Data nel campo della sicurezza anche mediante lo sviluppo, nonché l'impiego, di dimostratori in grado di valorizzare i prodotti di studio e ricerca nello specifico settore delle minacce asimmetriche globali, attualmente enucleate nel c.d. *CBRNe Risk* (rischio chimico, batteriologico, radiologico, nucleare ed esplosivo).

Obiettivi della ricerca

L'attività si è articolata in due momenti fondamentali e strettamente connessi sul piano della continuità logica e temporale, che possono essere così sintetizzati: una prima fase - finalizzata all'attività di ricerca, e una seconda fase, consecutiva alla prima, orientata alla produzione scientifica sull'argomento.

La fase di ricerca ha esplorato i riflessi applicativi dei Big Data nella sua duplice dimensione di massimizzazione del valore dei BD, con conseguente implementazione della national security, e di minimizzazione del rischio potenzialmente derivante dall'impiego massivo e non adeguatamente regolamentato dei medesimi identificato, dalla dottrina più accreditata, nella violazione della privacy dei cittadini. Non a caso, il nucleo attorno al quale si condensano le attuali direttrici di ricerca sul tema è costituito dall'annoso contemperamento tra le due istanze – entrambe da considerarsi di rango primario, come verrà evidenziato nel prosieguo della trattazione – della garanzia della sicurezza nazionale e della tutela della riservatezza dei cittadini, in un contesto marcatamente globalizzato e, per ciò stesso, foriero di multiformi e costanti minacce asimmetriche (Antares Fumagalli, 2018). Gli outcomes della fase esplorativa saranno riassunti nel presente documento, nel quale si descriverà lo stato dell'arte in materia di Big Data Governance, con particolare attenzione alle modalità di bilanciamento delle istanze sopra citate.

La successiva fase semestrale avrà ad oggetto la produzione scientifica individuale inerente ad una concreta applicazione della metodologia di ricerca OSINT (Open Source Intelligence) [1] : trattasi, nello specifico, dell'attività tesa alla digitalizzazione di *OSDIFE CBRNe Report*, la pubblicazione mensile a cura dell'Osservatorio Sicurezza e Difesa CBRNe – OSDIFE (Roma), i cui contenuti sono attualmente in corso di migrazione verso un sito web dedicato, che ospiterà un applicativo denominato *GAT Report – Global Asymmetric Threats Observatory*. L'applicativo in esame, appositamente sviluppato dal Research Team di OSDIFE, è principalmente finalizzato ad offrire, ai potenziali stakeholders, un efficace strumento di analisi degli attuali trend di minaccia chimica, biologica, radiologica, nucleare ed esplosiva a livello globale (CBRNe Risk).

Per meglio comprendere la stretta correlazione esistente tra la dimensione concettuale e quella operativa in tema di Big Data Governance, è sufficiente delineare l'attuale scenario in materia, che vede il massiccio ricorso di Agenzie pubbliche e private, Dipartimenti e Organizzazioni governative, ma anche Organizzazioni non governative e Società commerciali, alla metodologia OSINT per motivi di sicurezza e di mercato, complice la crescita improvvisa dell'impiego di strumenti digitali (smartphone, social network, dispositivi IoT, ecc.) e il conseguente enorme volume di dati

[1] L'espressione "Open Source Intelligence" è utilizzata da decenni per descrivere l'attività di raccolta di informazioni attraverso risorse disponibili al pubblico. Nella storia recente, l'attività OSINT è stata introdotta durante la Seconda guerra mondiale come strumento di intelligence da molte Agenzie di sicurezza di varie nazioni. Il *NATO Open Source Intelligence Handbook* definisce OSINT *"(…) information that has been deliberately discovered, discriminated, distilled, and disseminated to a select audience, generally the commander and his/her immediate staff, in order to address a specific question*". La principale differenza delle fonti OSINT rispetto a quelle impiegate per altre forme di intelligence sta nel fatto che quest'ultime devono essere legalmente accessibili al pubblico senza comportare alcuna violazione di copyright o di normativa sulla privacy.

generati dagli utenti della rete (Banu & Yakub, 2020).

L'indagine relativa alla Big Data Governance, nella sua duplice dimensione della privacy e della sicurezza nazionale, è stata condotta prendendo le mosse dall'evoluzione della c.d. Data Protection da diritto individuale ad interesse pubblico, secondo l'interpretazione proposta da Antares Fumagalli (2018, 134), che pone l'accento sul concetto di *"asimmetria informativa"* la quale, se *"(…) venisse meno in favore di pochi attori, gli equilibri potrebbero essere messi a repentaglio, i tradizionali approcci rischierebbero l'obsolescenza e ogni rapporto di forza potrebbe essere stravolto*". La tesi sostenuta dall'Autore pare essere condivisibile. Se, infatti, l'indisponibilità di dati sul conto dell'individuo rappresenta una sorta di "patrimonio negativo", di cui il medesimo beneficia - in quanto gli permette di preservarne la riservatezza, impedendone la conoscenza generalizzata da parte del prossimo - la cessazione di detto privilegio porrebbe il potenziale opponente in una posizione di ingiustificato vantaggio nei confronti del primo. Scarsità cognitiva e legittima reticenza, del resto, costituiscono da sempre un punto di forza di ciascuno dei contendenti, anche in fatto di analisi strategica, come peraltro descritto da Sun Tzu nel più antico trattato di strategia militare attualmente conosciuto (V sec. a.C.).

Il gap informativo tra contendenti produce conseguenze apprezzabili anche su istituti di diritto pubblico essenziali alla tutela degli interessi afferenti alla sicurezza nazionale: il riferimento è d'obbligo al segreto di Stato (Fratini, 1997). Per altri versi, tuttavia - ed è questo lo snodo che consente di raccordare le due dimensioni-chiave pubblico/privato cui si informa attualmente il paradigma dei Big Data e il conseguente dilemma Privacy vs. National Security – l'acquisizione di informazioni, anche in violazione della sfera individuale (e, in generale, delle libertà civili), potrebbe assicurare l'implementazione di un patrimonio informativo cruciale per garantire la sicurezza nazionale (ed internazionale) da aggressioni esterne, solo che si pensi all'attività di Web Intelligence (WEBINT) posta in essere per prevenire attacchi terroristici (Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and other National Goals, National Research Council, 2008). In ogni caso, l'odierno ragionamento strategico si rivela essere concettualmente condizionato dal fattore informativo come mai prima d'ora (Antares Fumagalli, 2018).

Metodologia di ricerca

Le precedenti considerazioni rendono ragione dell'attività conclusasi con la stesura del presente Report, prevalentemente condotta attraverso la metodologia OSINT, ossia mediante ricerca ed elaborazione di notizie settoriali tratte da fonti aperte, con specifico interesse per le pubblicazioni accademiche e gli articoli specialistici, integrati da articoli giornalistici provenienti da testate specializzate, ove di interesse per il tema trattato. La ricerca si è estesa anche alla consultazione di siti web dedicati, ospitati nel c.d. surface web, previa selezione delle fonti rilevanti mediante l'impiego dei più diffusi search engines *(Google, Yahoo, Bing*[2]). Tale ricerca è stata condotta anche avvalendosi dell'impiego di Summaries a cura di Organizzazioni pubbliche e private di settore (es. *Google Scholar)* La peculiarità delle fonti impiegate e la vastità del loro numero hanno, peraltro, richiesto particolare attenzione tanto nell'individuazione preventiva e nella selezione delle fonti

[2] Per ampliare l'area di ricerca, si è utilizzato il search engine con l'estensione .com, ovvero un gTLD (generic Top Level Domain), un dominio di primo livello generale che, in quanto tale, non ha un collegamento territoriale ben preciso (es. *Google.com*).

rilevanti, quanto nella determinazione del loro grado di attendibilità.

Nella ricerca dei testi online si è fatto largo uso degli operatori booleani, in particolare "AND" (es. Big Data AND national security), "OR" (es. Big Data OR Artificial Intelligence), oppure una specifica espressione (es. "Information and Communication Technologies"; "privacy-shield"); per le ricerche più complesse, invece, si è utilizzata una stringa di termini associati tra loro con il simbolo "+" (es. big+data+suicide+terrorism). La medesima modalità è stata utilizzata anche per setacciare i contenuti di siti specializzati in pubblicazioni accademiche (summary) come *Google Scholar*, impiegando il metodo delle c.d. citazioni a catena, a partire dai riferimenti bibliografici di singole pubblicazioni di settore. In caso di versioni plurime dello stesso articolo, si è selezionata la versione recante la data più recente.

Posto che la presente attività di ricerca non ha pretesa di completezza – né potrebbe ambire ad averne, in ragione della produzione sterminata sul tema "Big Data Governance" che si è accumulata nel corso dell'ultimo trentennio – la letteratura raccolta ed esaminata è stata suddivisa in quattro macroaree: la prima, relativa alle caratteristiche generali e all'evoluzione del paradigma dei Big Data; la seconda, attinente ai profili giuridici che investono la regolamentazione dei BD nelle varie fasi di acquisizione, archiviazione, manipolazione, gestione, analisi, estrazione di conoscenza, sicurezza, privacy e visualizzazione dei dati, con particolare attenzione alle criticità che ciascuna di dette fasi sottende, non ultime le conseguenze sul piano etico; la terza, dedicata alle potenzialità dei Big Data negli innumerevoli campi di applicazione; la quarta, inerente ai rischi e alle sfide che questa tecnologia emergente presenta in un contesto fortemente globalizzato.

Per quanto attiene alla struttura del Report, il presente documento riflette in maniera coerente le quattro macroaree sopra delineate e risulta, pertanto, suddiviso in altrettante sezioni, con l'aggiunta di una quinta sezione conclusiva. In ciascuna delle sezioni verrà presentato sommariamente lo stato dell'arte relativo allo specifico aspetto della tematica generale preso in esame, evidenziando elementi comuni, dissonanze e criticità emergenti dall'analisi sistematica delle fonti raccolte. Esso include, inoltre, una bibliografia ragionata per ciascuna delle macroaree (e relative sezioni) esaminate, nella quale verranno riportate le fonti consultate durante l'attività di ricerca, senza (giova ripeterlo) alcuna pretesa di completezza. In ragione dei riferimenti bibliografici contenuti nelle fonti citate, ognuna di esse potrà essere validamente impiegata per ampliare l'indagine in materia di Big Data, selezionando percorsi specifici nell'ambito di ciascuna delle macroaree di cui si compone il presente documento.

Per quanto attiene la ricerca e la valutazione delle fonti, merita precisare che l'intera attività è stata condotta conformemente al Processo OSINT[3], con specifico riguardo alle prime tre fasi, rispettivamente della *Discovery, Discrimination* e *Distillation* delle fonti e del patrimonio informativo ad esse riconducibile: la peculiarità delle fonti impiegate e la vastità del loro novero richiedono, infatti, specifica attenzione tanto nell'individuazione preventiva e selezione di quelle rilevanti, quanto nella determinazione del loro grado di attendibilità, qui inteso come livello di aderenza alla realtà dell'informazione che da tali fonti può essere

[3] Come riportato da Riservato e Scaini (2018), cui si rinvia per una dettagliata analisi del Ciclo OSINT, *"L'origine dell'Intelligence da fonti aperte si può far risalire all'utilizzo di informazioni disponibili, verbali o scritte, per pianificare investimenti di tipo sia militare che economico/commerciale; un esempio classico è quello dato dai Lloyds di Londra dove, già nel 1688, la "Coffee house" di Edward Lloyd era un luogo molto frequentato da marinai, mercanti e imprenditori navali, i quali si ritrovavano in questo luogo per discutere dei propri affari e dal quale nacque la compagnia di assicurazioni più famosa al mondo".*

ottenuta, mediante processo di estrazione (Riservato & Scaini, 2018). Com'è noto, *"con l'acronimo OSINT, secondo la definizione generalmente accettata in dottrina, si intende descrivere l'informazione presente sulle c.d. "*fonti aperte*"- quindi, liberamente disponibile per essere fruita dall'utenza della rete – che è stata "setacciata" con applicativi in grado di filtrare la ricerca, così da soddisfare il bisogno informativo espresso attraverso un intelligence requirement"* (Di Stefano, 2020).

L'individuazione delle fonti (*Discovery*) è stata compiuta mediante interrogazione, con l'ausilio dei più diffusi search engines (*Google, Bing, Yahoo*), di una data key word o di stringhe di caratteri utilizzati contestualmente ai c.d. operatori booleani (AND, NOT e OR), prediligendo, tra i risultati ottenuti, le c.d. fonti primarie, ossia quelle aventi attinenza diretta con il tema o il fenomeno di interesse. Sono state, pertanto, privilegiate le fonti maggiormente esperte o che hanno maturato una maggiore esperienza nello scenario oggetto di analisi, soprattutto in ragione del ruolo istituzionale e/o della mission universalmente riconosciuta dell'Autore: in questo caso, le pubblicazioni accademiche nonché i Report rilasciati da Istituti di Ricerca nazionali o internazionali e da Agenzie governative, purché attinenti alla tematica della governance dei Big Data e riflessi applicativi. Successivamente, sono state indagate le fonti c.d. secondarie – quelle, cioè, incidentalmente riconducibili alle fonti primarie – nella misura in cui manifestavano un coinvolgimento anche solo parzialmente diretto con lo scenario da analizzare: in questo caso, i contenuti prodotti da Autori e/o Organizzazioni note presenti su siti web o portali di informazione accreditati. Coerentemente con tale impostazione, infatti, sono stati selezionati i soli articoli giornalistici e le ricerche imprenditoriali citati nelle pubblicazioni accademiche e nei Report di ricerca rilasciati da Enti istituzionali o Agenzie governative.

Il principale ambiente di raccolta delle fonti (sebbene, non esclusivo) è stata, dunque, la c.d. Cybersfera, quella dimensione virtuale che ospita i contenuti digitali prodotti dall'essere umano in quanto utente della rete Internet: primi fra tutti, materiali elaborati da esperti di settore (ad es. ricerche accademiche, analisi specialistiche rilasciate da Data Scientist, ecc.), ma anche documenti come ricerche imprenditoriali o articoli giornalistici tratti da periodici online in tema di Big Data Governance, unitamente a materiale scientifico/culturale di varia origine e formato (es. tesi di laurea/dottorato; registrazioni di convegni, ecc.) ottenuti esplorando Siti Web, Forum, Blog, Riviste specializzate online afferenti ad Università, Istituti di Ricerca e Biblioteche, Social Networks e portali di informazione tecnologica dedicati. Nell'attività di ricerca si è, talvolta, incorsi nella c.d. letteratura grigia, con ciò intendendosi quel materiale non pubblicato né distribuito, catalogato o diffuso se non in ambienti ristretti, inclusivo di documenti interni, bozze, reports tecnici e altro materiale non accessibile (Riservato & Scaini, 2018), solo successivamente reso noto perché, ad esempio, declassificato o desecretato dall'Ente/Agenzia che ne vantava la paternità.

Poiché, non di rado, la fonte identificata restituiva un'informazione parziale e limitata rispetto alla totalità e alla complessità dell'evento osservato, si è reso necessario compiere un esame incrociato delle fonti, per ottenere maggiori dettagli sugli aspetti salienti del fenomeno in esame e delle dinamiche ad esso sottese, soprattutto in presenza di fonti c.d. *de relato* o di seconda mano. A tal proposito, la gerarchizzazione delle fonti si è rivelata di fondamentale importanza al fine di identificarne la provenienza (soprattutto in presenza di una fonte c.d. di seconda mano) e valutarne l'attendibilità, ossia l'aderenza alla realtà. Tale processo ha manifestato le sue potenzialità anche in relazione alla seconda fase del ciclo OSINT, ossia alla *Discrimination*, intesa come selezione delle fonti rilevanti rispetto a quelle ininfluenti ai

fini dell'oggetto della ricerca (Riservato & Scaini, 2018). Posto che la valutazione di una fonte attiene al livello di affidabilità della stessa, espressa nella misura in cui essa è in grado di produrre informazioni utili al processo OSINT, sono state privilegiate fonti considerate "Completamente affidabili" – "History of complete reliability", "Storia di completa affidabilità", secondo la definizione fornita dal manuale *Admirality Code NATO System*, attraverso la c.d. *griglia 6X6 NATO* – ossia, nessun dubbio sull'autenticità, sull'affidabilità o sulla competenza della fonte.

Il codice di grado massimo nella griglia di validazione di attendibilità delle fonti ("A Completely reliable") è stato agevolmente conseguito rispetto alle pubblicazioni accademiche o ad altri contenuti provenienti da fonti istituzionali, che hanno peraltro costituito la parte principale della web query, in ragione del ruolo istituzionale e/o della mission universalmente riconosciuta all'Autore del documento - o, in subordine, dell'appartenenza o affiliazione dell'Autore/i ad un Dipartimento universitario o ad un Ente di ricerca accreditato presso la Comunità Scientifica, ecc. Per ciò che concerne, invece, gli articoli giornalistici o la stampa specializzata, *"(…) spesso la "fonte" rimanda a due soggetti diversi: l'articolista/autore e la testata/sito/editore che ha pubblicato o rilanciato l'informazione; molte notizie sono, poi, pubblicate su siti web registrati come periodici a stampa e online, con contenuti caratterizzati da identificazione ISSN* [4] *o rimandanti a pubblicazioni singole catalogate con un codice univoco ISBN*[5]*. I contenuti collazionati da articolista/autore/testata/sito/editore traggono, in genere, spunto da fonti ed archivi attendibili, sintetizzandone i contenuti con la citazione della fonte originaria o richiamando fedelmente frammenti degli stessi documenti".* Ciò che consentirebbe di considerare la fonte in questione di qualificata e tracciata attendibilità, sostanzialmente pari a quella assegnata alla precedente categoria di fonti. In un simile contesto, infatti, l'assoggettamento di articolista/autore/sito/pubblicazione/editore alla disciplina penale sulla stampa sembra ragionevolmente costituire un efficace deterrente alla messa in rete di notizie di particolare enfasi ma prive di genuinità (Di Stefano, 2020).

Al pari delle fonti, anche le informazioni da esse estratte *(Distillation)* hanno necessitato di essere organizzate gerarchicamente. A tal fine, le informazioni distillate dai documenti in esame sono state distinte in certe e attendibili. Certe sono state considerate le informazioni sulle quali esistono prove inequivocabili circa la loro fondatezza: in ragione di detta peculiarità, sono state poste a fondamento della trattazione oggetto del Report (si pensi ai dati forniti dai Ministeri o dalle Agenzie governative, ritenuti certi in quanto provenienti da soggetti istituzionali ufficialmente autorizzati ad esprimersi ed agire nel contesto indicato). Successivamente, sono state considerate le informazioni attendibili, quelle cioè prodotte da autori o pubblicazioni considerati come affidabili o autorevoli (l'esempio è quello di notizie tratte da testate giornalistiche specializzate rilanciate da Enti istituzionali): in ogni caso, affidabilità e attendibilità sono parametri da porsi in relazione al soggetto in esame.

Le informazioni certe, ritenute rilevanti ai fini delle analisi compiute nel Report, sono

[4] L'International Standard Serial Number (ISSN) è il codice numerico, di otto cifre, che identifica l'edizione periodica a stampa o elettronica. I periodici pubblicati sia a stampa che in formato elettronico possono avere due codici ISSN: un ISSN stampa (p-ISSN) e un ISSN elettronico (e-ISSN o e-ISSN).

[5] Il codice ISBN (International Standard Book Number) è una sequenza numerica di 13 cifre usata internazionalmente per la classificazione dei libri. Ogni codice ISBN identifica in modo univoco ogni specifica edizione di un libro (non però le semplici ristampe, che mantengono lo stesso codice dell'edizione cui si riferiscono) e, una volta assegnato, non può più essere riutilizzato.

state sottoposte a verifica mediante l'impiego di una pluralità di fonti, qualora la provenienza delle informazioni fosse multipla (es. il valore del traffico-dati in un determinato contesto) mentre le informazioni attendibili sono state incrociate con quelle certe, ed eventualmente scartate se in contrasto con queste ultime. A tal fine, si è impiegato un metodo di ricerca orizzontale - utilizzando un vasto numero di fonti e di informazioni per evitare di escludere a priori elementi utili – integrato da una ricerca di natura verticale, che ha consentito di approfondire le informazioni raccolte attraverso un'indagine più accurata e dettagliata degli elementi selezionati come fondamentali per la formulazione di analisi, inferenze ecc.

Introduzione

Nel linguaggio ICT (Information and Communication Technologies), la locuzione generica e onnicomprensiva "Big Data"[6] indica il flusso di dati generato quotidianamente dagli utenti della rete Internet, in modo massiccio e continuo, nell'epoca dei social media e del crescente sviluppo tecnologico (Diebold, 2012; Laney, 2001). L'espressione, talvolta tradotta con il vocabolo "megadati", indica *"(…) una quantità di dati ed informazioni così estesa in termini di volume, velocità e varietà da richiedere tecnologie e metodi analitici specifici per l'estrazione di valore o conoscenza. L'espressione viene utilizzata, cioè, per indicare la capacità di analizzare ovvero di estrapolare e mettere in relazione un'enorme mole di dati eterogenei, strutturati e non strutturati – grazie a sofisticati metodi statistici e informatici di elaborazione – al fine di scoprire i legami tra fenomeni diversi – ad esempio, correlazioni – e prevedere quelli futuri"* (Rolli & D'Ambrosio, 2022, 784). Circostanze contingenti come la recentissima pandemia di COVID-19, esponendo la collettività a un maggior numero di interazioni online e ad una accresciuta dipendenza sociale dalla tecnologia, hanno contribuito ad aumentare in modo significativo il bacino globale di dati immessi in Rete (Hammond-Errey, 2022).

L'espressione "Big Data" è stata adottata anche dalle aziende IT per descrivere il problema della gestione di un'enorme mole di dati (Big Data Governance), così alludendo ad una delle principali criticità derivanti dall'impiego di questa Tecnologia Emergente, con importanti ripercussioni, soprattutto nel contesto della sicurezza nazionale. Il progresso e l'innovazione tecnologica offrono grandi opportunità ma, al contempo, comportano molteplici sfide per la società, l'economia, la politica ed il mondo militare, la cui entità sfugge ancora ad una reale comprensione (Stato Maggiore della Difesa, 2022). Certamente, i Big Data stanno esacerbando le attuali minacce alla sicurezza nazionale, generandone di nuove e imprevedibili (Hammond-Errey, 2022).

Vero è che, sottesa al predetto flusso di dati, vi è la promessa di estrapolare preziose intuizioni per meglio attagliare campagne di marketing, rilevare sintomi di patologie specifiche nei pazienti, mappare focolai di malattie infettive (Mugavero & Thorossian, 2021), anticipare eventi imminenti pregiudizievoli per la sicurezza nazionale, fino a gestire maxiemergenze di natura sanitaria (XXI Congresso Nazionale AIMC, 2023), per citare solo alcune delle numerose potenzialità di impiego dei Big Data. Il valore di essi, dunque, si fonda sull'ipotesi che, qualora i dati disponibili siano correttamente raccolti, conservati e sottoposti a metodologie analitiche adeguate, gli stessi saranno in grado di generare intuizioni precedentemente occulte, fornendo ai fruitori consapevolezza situazionale in tempo reale.

[6] A tal proposito, si richiama la nota definizione delle "3 V", coniata da Doug Laney nel 2001, che caratterizza i ""Big Data *" (…) as a circumstance where the volume, velocity and variety of data of an organization's storage go beyond the computation capacity for precise and well-timed decision making"*. Per una evoluzione ragionata del concetto, si rinvia a Banu & Yakub (2020).

Altrimenti detto, il valore dei BD si riferisce alla probabilità che un set di dati, se sottoposto ad analisi appropriate, produrrà nuove tendenze o altri prodotti analitici di valore per l'adozione di una decisione finale (Chi, 2017).

Ciò posto, i Big Data impongono l'individuazione di innovativi strumenti analitici e gestionali per estrapolare informazioni da essi, principalmente attraverso l'impiego delle tecniche di Data Mining più avanzate, denominate Big Data Analytics (BDA)[7]. Del resto, l'espressione stessa è il risultato della fusione di "Big Data" e "Data Analytics", ciò che ha generato una delle tendenze maggiormente radicate nell'attuale Business Intelligence (BI) (Banu & Yakub, 2020; Russom, 2011). Questi nuovi metodi di analisi, tuttavia, impiegano una serie di tecnologie complesse e all'avanguardia, che attraversano diverse discipline, dall'informatica alla statistica, dalla matematica applicata all'economia, fino al Machine Learning (ML) e all'Intelligenza Artificiale (AI) (Nurkin & Konaev, 2022). L'ampiezza e complessità di dette tecnologie sono tali da legare indissolubilmente il concetto di BD a molteplici fenomeni emergenti, non ultimo la crescita esponenziale di impiego dell'Intelligenza Artificiale in svariati contesti applicativi (Banu & Yakub, 2020).

Secondo Fourcade & Healy (2017), le organizzazioni moderne seguirebbero una sorta di *"imperativo di dati istituzionali"* che imporrebbe loro di raccogliere il maggior numero di dati possibile: certo è che l'analisi di dati massivi richiede un impegno multilivello per estrarre la conoscenza necessaria ad implementare il processo decisionale (Acharjya & Ahmed, 2016). La stessa comunità scientifica concorda sulla necessità di massimizzare i benefici analitici dei Big Data attraverso il trattamento dei dati come un asset strategico, investendo sulle persone oltre che sui processi e sulle BD Technologies (Henke, Bughin, Chui et al., 2016). Il mantenimento di un processo di analisi del tipo *human-in-the-loop* è stato, inoltre, proposto come un'ipotetica soluzione al problema della c.d. opacità algoritmica, intesa come il deficit di comprensione del funzionamento dell'algoritmo da parte degli utenti umani, a causa della complessità della sua struttura (Burrel, 2016). Una opacità che, peraltro, rischia di compromettere il metodo scientifico, fondato sulla riproducibilità degli esperimenti e il nesso razionale tra causa ed effetto (Facchini & Termine, 2022). L'analisi dei Big Data appare, perciò, come un'area attuale di ricerca e sviluppo, che pone innumerevoli questioni aperte, anche rispetto ai costi di impiego del c.d. capitale umano nell'attività analitica stessa (Acharjya & Ahmed, 2016).

Dunque, parallelamente allo sviluppo di adeguate tecnologie digitali, si rende necessario l'utilizzo di risorse umane, ovvero di una comunità di analisti, preferibilmente esperti in Open Source Intelligence (OSINT)[8] che, soprattutto per le Agenzie di sicurezza nazionale, sia in grado di analizzare i trends di sviluppo di eventi critici elaborati dalle ITs, così da costruire modelli predittivi di minaccia e di rischio. Posto che l'intento dell'analisi OSINT è quello di collezionare elementi di conoscenza al fine di soddisfare una precisa necessità informativa, il recupero di informazioni, soprattutto quelle attinenti a persone fisiche, potrebbe facilmente sconfinare in attività intrusive della sfera individuale, non di rado tecnicamente complesse e raffinate, quando non addirittura illecite (hacking, social engineering o attività di spionaggio vere e proprie, solo per citarne alcune). L'impiego di tali tecniche di indagine, infatti, può consentire l'acquisizione di dati altamente sensibili, quali nomi di persone fisiche e loro

[7] Sono attualmente disponibili un gran numero di strumenti per elaborare i BD: tra quelli emergenti, spiccano *Hadoop, Map Reduce, Apache Spark e Storm* (Acharjya & Ahmed, 2016).

[8] A tal proposito, sarebbe più corretto parlare di WEBINT (Web Intelligence), ossia quella sottocategoria di OSINT limitata all'impiego delle sole fonti aperte disponibili online.

riferimenti (dati anagrafici, numeri di telefono ed occupazione, indirizzo di residenza, ecc.) nonché dei relativi contatti della rete sociale; username, immagini, e-mail personali; geolocalizzazione; informazioni personali collegate a domini Internet; daily routine, come presenza sui social media, ma anche dati sanitari e molto altro.

A tal proposito, merita citare, tra le fonti prevalentemente utilizzate nell'attività OSINT [9], una molteplicità di risorse provenienti dal dominio digitale, che spaziano dall'accesso alla rete Internet – forum, blog, social network, condivisione video, wiki, record Whois di nomi di dominio registrati, metadati e file digitali, risorse web scure, dati di geolocalizzazione, indirizzi IP, motori di ricerca delle persone e qualsiasi altro dato reperibile online - alle foto e video, inclusi metadati e informazioni geospaziali, unitamente ad altre informazioni di pubblico dominio, come rapporti governativi o ministeriali, piani finanziari, dati demografici, conferenze e comunicati stampa, discorsi, avvisi di istituzioni e forze di polizia, avvisi dell'Aeronautica e della Marina, dibattiti legislativi e atti politici, purché disponibili online. Con l'avvertenza che l'attività esplorativa summenzionata non si limita all'utilizzo dei principali motori di ricerca, posto che la quasi totalità dei dati presenti in rete - secondo taluni esperti di settore, addirittura oltre il 95% - non sarebbe indicizzata da *Google, Bing, Yahoo* o qualsiasi altro search engine. A ciò si aggiunga che almeno l'80% dei Big Data (secondo alcuni Autori, addirittura il 90%) risulta, attualmente, non strutturato e richiede ancora un'analisi (Malick, 2023; Sashidharan, 2023).

Nel settore commerciale, dove l'uso principale dei Big Data si pone a fondamento dell'analisi predittiva basata sul comportamento del consumatore per trarne profitto (Zuboff, 2019) – emblematico è il caso di applicativi come *Google AdSense* che creano 'profili' di singoli utenti sulla base delle loro storie di ricerca per predirne condotte future, come dimostrato negli annunci pubblicitari di *Netflix* e *Amazon*, in cui i suggerimenti personalizzati sono la funzione di un profilo comportamentale elaborato a partire dai dati relativi alla visualizzazione e all'acquisto precedenti (Chi, 2017) – una potenziale intrusione nella sfera privata dell'utente può essere ragionevolmente stigmatizzata, nella misura in cui appare assoggettata a discutibili ragioni di lucro (Fourcade & Healy, 2016).

Diverso, ed eticamente più complesso sul piano delle possibili conseguenze, il caso in cui l'acquisizione di informazioni personali consenta la profilazione di soggetti le cui attività potrebbero risultare potenzialmente pericolose per la collettività, funzione ritenuta di estremo interesse dagli analisti impegnati nella mappatura di una rete globale di minacce asimmetriche (Chi, 2017). Profilazione, in questo senso, significa impiego delle informazioni disponibili per prevedere condotte illecite dai riflessi di portata transnazionale, in quanto amplificate dal processo di globalizzazione. Si badi che le stesse organizzazioni terroristiche utilizzano fonti OSINT per pianificare gli attacchi, raccogliere informazioni sui targets selezionati attraverso l'analisi di social media, acquisire informazioni militari rivelate accidentalmente dai Governi e diffondere la loro propaganda in tutto il mondo impiegando canali social (Ricci, 2018).

Posto che, anche nel settore della sicurezza nazionale, esiste la possibilità di impiegare indicatori automatici in grado di analizzare i profili comportamentali, finanziari e delle precedenti attività lesive di attori

[9] Il lungo elenco comprende anche fonti tradizionali - mass media quali televisione, radio, giornali, libri e riviste – oltre a fonti altrimenti qualificate, come riviste specializzate, pubblicazioni accademiche, tesi di laurea, atti di convegni, profili aziendali, relazioni annuali, notizie aziendali, profili dei dipendenti/curricula, et similia.

pregiudizievoli, con l'intento di identificare potenziali minacce imminenti – quelli che Hossain, Harutyunyan, Ning, et al. (2022) definiscono *"set of precursors"* - il *New York State Intelligence Center's 'Terrorism Indicators Reference Card* già da tempo elencava indicatori, presenti nei profili dei singoli viaggiatori, che correlavano con precedenti condotte di stampo terroristico (New York State Intelligence Center, 2008), tali da costituire un'efficace strumento di analisi predittiva di futuri attacchi di natura simile[10]. In casi come questi, le ragioni di national security sembrano prevalere rispetto a possibili condotte lesive della sfera individuale, purché legittimamente motivate: la questione, tuttavia, è fortemente dibattuta e continua a presentare profili di criticità, soprattutto negli ordinamenti giuridici marcatamente garantisti e informati al modello democratico occidentale (Antares Fumagalli, 2018; Fourcade & Healy, 2016).

In ogni caso, l'uso dei Big Data nel settore della sicurezza nazionale deve superare sfide di complessità, qualità e quantità dei dati, risultati potenzialmente discriminatori e problemi di privacy (Chi, 2017). Il binomio esistente tra informazione e potere, conseguente alla presa d'atto delle enormi potenzialità informative scaturenti dal progresso tecnologico in campo informatico, impone, nondimeno, una rinnovata lettura in chiave strategica del concetto di Data Protection ai fini della sicurezza nazionale. La riflessione dottrinale sul punto suggerisce - già da un quinquennio ormai, precisamente all'indomani dell'emanazione del Regolamento UE 2016/679, denominato Regolamento Generale sulla Protezione dei Dati (GDPR, utilizzando l'acronimo anglosassone) - un focus sui rischi strutturali, per il sistema Paese, connessi alla profonda rivoluzione concettuale e operativa avviata da tecnologie come Big Data e Deep Learning[11] (Antares Fumagalli, 2018).

Certo è che *"(…) Il dato, oggi, diventa una materia prima di grande valore e vede esplodere il proprio potenziale informativo grazie alla capacità dei moderni algoritmi di estrarne descrittori e metterli in relazione tra loro, individuando correlazioni con altri dati e moltiplicandone così i contenuti"* (Antares Fumagalli, 2018, 133). Ciò pone una serie di riflessioni su quella che è stata felicemente definita la rivoluzione avviata dalla *"quarta dimensione"* (Pansa, 2017), la quale ha rapidamente investito alcuni degli assiomi posti alla base del nostro ordinamento giuridico ed economico, oltre che ai rapporti di forza tra attori geopolitici. Com'è facilmente intuibile, infatti, le nuove tecnologie richiedono la disponibilità di ingenti somme di denaro, con il conseguente accentramento di un'enorme mole di dati nelle mani di pochi - ad avvalersene principalmente sono i grandi gruppi societari, i proprietari delle grandi piattaforme social o commerciali e quelli d'infrastrutture per il web, gli hosting provider o gli apparati pubblici legittimati ad acquisire dati sulla popolazione – da cui deriva una distribuzione asimmetrica di capacità informativa e, com'è stato acutamente osservato *"(…) la diretta conseguenza dell'equazione 'informazione=potere' è che a un salto di*

[10] Tra gli indicatori listati, alcuni appaiono particolarmente suggestivi: 'Recente viaggio d'oltremare'; "Possesso di un visto di studio, ma non competente in inglese"; 'Rifiuto del servizio di pulizia (del proprio alloggio)"; "Possesso di un'unità GPS" e "Comportamento insolitamente calmo e distaccato". Indicatori alternativi hanno incluso anche gruppi jihadisti con corsi di paintball in Australia (Radden Keefe, 2006). Va, tuttavia, precisato che, sul piano criminologico, alcuni di tali indicatori potrebbero avere valenza aspecifica, necessitando di essere inquadrati in un contesto adeguato.

[11] Con l'avvertenza, tuttavia, che *"(…) La disciplina europea, occorre sottolinearlo, non si estende alle politiche di difesa e sicurezza, appannaggio esclusivo degli stati membri, rivolgendosi piuttosto al bilanciamento tra diritti fondamentali dei cittadini ed esigenze di tutela del mercato comune"* (Antares Fumagalli, 2018, 143).

qualità nella capacità informativa corrisponde un accrescimento del potere" stesso (Antares Fumagalli, 2018, 134). Gli attuali scenari, nondimeno, confermano come le capacità tecniche e analitiche, essenziali per il funzionamento delle società, siano sempre più concentrate nelle mani di un piccolo numero di entità commerciali (Hammond-Errey, 2022).

Quanto esposto in precedenza restituisce solo in minima parte l'ampiezza e la complessità delle implicazioni inerenti alla gestione dell'enorme mole di dati attualmente disponibili (Big Data Governance) e, in generale, alla definizione del nuovo rapporto tra essere umano ed Intelligenza Artificiale (Human Automomy Teaming), in termini di opportunità e di potenzialità, così come di criticità e di continue sfide. L'obiettivo (ambizioso) della presente attività di ricerca è quello di contribuire ad ampliare la expertise in merito al c.d. paradigma dei Big Data, anche mediante la concreta applicazione delle attuali conoscenze in materia ad uno sviluppatore dedicato al monitoraggio del rischio CBRNe nel mondo, ossia il GAT Report – Global Asymmetric Threats Observatory di OSDIFE.

1 Evoluzione e Caratteristiche Generali dei Big Data

Rispetto alla prima macroarea di ricerca, parte della letteratura esaminata in tema di evoluzione del concetto di "Big Data" [12] riconduce le origini del fenomeno ad alcune decadi fa, quando invalse l'uso di tecniche di analisi dei dati per supportare il processo decisionale, noto come *Data-Driven Decision Making Process* (Banu & Yakub, 2020; Picciano, 2012; Provost & Fawcett, 2013 e bibliografia ivi citata). Il lavoro di Halevi & Moed (2012), esplorando l'impiego del termine "Big Data" nella produzione scientifica recente, ne fa risalire la prima apparizione ad un articolo del 1970 sui rilevamenti atmosferici e oceanici condotti nell'Arcipelago delle Barbados. Altri Autori posticipano l'origine dell'evento collocandola nel contesto informatico statunitense intorno alla metà degli anni '90, attribuendo i primi riferimenti accademici significativi a Weiss & Indurkhya (1998) in informatica e Diebold in statistica/econometria[13] (Diebold, 2012).

Siti online specializzati – esplorati sommariamente mediante il motore di ricerca *Google.com* inserendo stringhe di operatori booleani del tipo "big+data+history" - propongono analisi del fenomeno in chiave storica, individuando nella Grande Biblioteca di Alessandria d'Egitto (fondata tra il 285 e il 246 a.C.), uno dei primi esempi di memorizzazione di dati agglomerati su ampia scala. Tale affermazione sarebbe, peraltro, suffragata anche da fonti accademiche (Wiegand & Donald Jr., 2015), che documentano altresì l'uso degli Antichi Romani di consultare le statistiche militari per determinare lo schieramento ottimale degli eserciti, grazie all'impiego di efficaci modalità di archiviazione di dati di massa utilizzanti per formulare ipotesi predittive. I numerosi progressi tecnologici compiuti per scopi militari durante la Seconda Guerra

[12] Per restringere l'area di ricerca, sono state considerate le sole pubblicazioni contenenti le key-words "big data+evolution", "big data+origins" o "big+data+history" nel titolo o nell'abstract. La ricerca è stata condotta sia attraverso i search engines più diffusi (*Google, Yahoo, Bing*) sia attraverso il summary di *Google Scholar*, con l'impiego degli operatori booleani. I risultati ottenuti sono stati ulteriormente filtrati selezionando quelle pubblicazioni che trattavano l'argomento in maniera specifica e non in via meramente incidentale. In caso di review della letteratura, è stata riportata la sola opera principale, con rinvio ai titoli in essa citati.

[13] Benché siano riportati alcuni riferimenti ai Big Data antecedenti al 2000, sia accademici che non accademici, si tratterebbe di episodi scarsamente rilevanti, poiché il termine sarebbe stato utilizzato senza una reale consapevolezza del fenomeno (Diebold, 2012).

Mondiale[14] e la loro estensione, dapprima al settore commerciale e, successivamente, al vasto pubblico con la diffusione del personal computing, rappresenterebbero ulteriori pietre miliari di detto percorso evolutivo (Jackson-Barnes, 2023).

George Firican, sul sito web *lightsondata.com* si spinge addirittura agli albori della civiltà citando, come primo esempio di memorizzazione e Data Analytics, niente di meno che i bastoni di conteggio rinvenuti in Uganda negli anni Sessanta e risalenti al 18.000 a.C.[15]. La dettagliata (e ben documentata) timeline riportata da Phillips (2021) sul sito *www.techtarget.com* colloca, invece, le origini dei Big Data nel moderno significato del termine – ossia come momento in cui gli uomini di scienza hanno iniziato a comprendere il valore delle statistiche per attribuire significato alla realtà circostante - nel XVII° secolo a Londra, e precisamente nel 1663, quando John Graunt introdusse l'analisi dei dati statistici per lo studio dell'andamento dei contagi della peste bubbonica, pubblicando la prima raccolta di registri di salute pubblica recanti i tassi di mortalità e le loro variazioni sul territorio inglese.

Sebbene la sommaria indagine compiuta in rete abbia evidenziato disaccordi circa le origini del termine "Big Data"[16] - talune fonti accreditano lo scienziato informatico americano Mashey come il "padre dei Big Data", talaltre ritengono che la locuzione sia stata coniata nel 2005 da Mougalas, altre fonti ancora sostengono che il concetto in questione non sia davvero decollato fino al 2010, solo per citare alcune delle open sources riportate da Firican (2023), cui si rinvia – tutti gli Autori attribuiscono rilevanza dirimente alla "Internet Age" per ciò che concerne quella che è stata definita "l'alba dei big data" (Phillips, 2021), ossia il momento in cui i personal computers hanno avviato la condivisione delle informazioni a tassi esponenzialmente maggiori, in seguito alla creazione della rete Internet. Secondo uno studio condotto da IBM, infatti, il 90% di tutti i dati del mondo sarebbe stato generato solo negli ultimi due anni (Sashidharan, 2023).

L'analisi circostanziata riportata dal sito *www.bigdataframework.org* scorpora il processo evolutivo dei BD in tre momenti fondamentali: una prima fase, quella dei contenuti strutturati, basata sulle tecniche di archiviazione, estrazione e ottimizzazione dei dati comunemente impiegati nei sistemi di gestione dei database tradizionali (RDBMS); una seconda fase, quella dei contenuti non strutturati originati dalle applicazioni Web, che hanno implementato una nuova forma di conoscenza a favore delle Organizzazioni aziendali, come quella relativa alle strategie comportamentali degli utenti della rete; la terza e attuale fase, guidata dall'adozione della tecnologia dei dispositivi mobili e dei dati da essi generati (IoT) [17] . Posto che tali dispositivi sono connessi a Internet senza soluzione di continuità, i dati da essi generati forniscono un quadro in tempo reale e senza precedenti del comportamento degli utenti. Ciascuna delle tre fasi, dunque, appare dominata da

[14] Secondo Jackson-Barnes (2023), le origini del c.d. electronic storage possono essere fatte risalire allo sviluppo del primo computer programmabile al mondo, l'*Electronic Numerical Integrator and Computer* (ENIAC), progettato dall'esercito americano durante la Seconda Guerra Mondiale per risolvere problemi numerici, come il calcolo della gittata dell'artiglieria.

[15] Gli uomini paleolitici, infatti, erano soliti contrassegnare bastoni o ossa di animali con apposite tacche, per conservare memoria dell'attività commerciale o dei rifornimenti – ciò che consentiva loro di eseguire calcoli per formulare previsioni relative, ad esempio, alla durata delle scorte di cibo.

[16] Lo stato dell'arte sul tema è stato efficacemente tratteggiato da Diebold (2012), secondo cui, testualmente *"(…) : the origins of the term are intriguing and a bit murky, involving both industry and academics, computer science and statistics/econometrics"*.

[17] Secondo il sito citato, solo per fornire una vaga stima del fenomeno, si pensi che nel 2020 circa 10 miliardi di dispositivi risultavano connessi a Internet, ciascuno di essi in grado di generare dati ogni secondo della giornata.

specifici progressi tecnologici, vantando proprie caratteristiche e rispettivi outcomes.

Quale che sia la prospettiva di analisi adottata, la totalità delle fonti esaminate concorda nell'attribuire rilevanza decisiva all'impiego massiccio di Internet e dei social network nel corso degli ultimi due decenni (Lee, 2017; Chae, 2019 e bibliografia ivi citata). L'emergere del World Wide Web, nell'ultima decade del secolo scorso (1989-1999); il controllo del volume dei dati, social media e cloud computing da parte di attori Big Tech, agli albori del Nuovo Millennio - aziende come *Amazon, eBay* e *Google* hanno contribuito a generare enormi quantità di traffico web, con conseguente combinazione di dati strutturati e non strutturati, tra il 2000 e il 2010[18] – fino all'approdo alle attuali tecniche di ottimizzazione, ai dispositivi mobili e allo IoT (dal 2010 ai giorni nostri), hanno rappresentato le principali tappe evolutive del paradigma dei BD e, al contempo, le maggiori sfide in tema di Big Data Governance [19] (Jackson-Barnes, 2023). È, peraltro, opinione comune nelle fonti esaminate che i Big Data costituiscano un dominio di conoscenza ormai consolidato, sia in ambito accademico che industriale.

L'ampiezza del numero di lavori in tema di evoluzione dei BD, così come il crescente interesse accademico per l'argomento, sono suffragati dalle meta-analisi compiute nell'ultimo quinquennio: primo tra tutti, il contributo di Gupta & Rani (2018), che include un approfondito studio bibliometrico delle pubblicazioni accademiche e industriali relative al periodo 2000-2017, cui si rinvia per un esame più dettagliato delle fonti. Appare improntato ad una metodologia simile anche lo studio di Raban & Gordon (2020), che indaga le pubblicazioni accademiche su Big Data e Data Science in chiave evolutiva, analizzando la relazione tra i due flussi di letteratura mediante l'impiego di vari indicatori bibliometrici, tra cui le aree di ricerca e la loro origine, le riviste centrali, i paesi che producono e finanziano la ricerca e le organizzazioni di startup, le dinamiche di citazione, la dispersione e l'impegno dell'autore. Sul metodo biometrico si incentra anche la meta-analisi di Batistič & Van der Laken (2019), giungendo, sostanzialmente, ai medesimi risultati dei colleghi, ossia il costante aumento dell'attenzione accademica per il tema dell'evoluzione dei Big Data in chiave diacronica [20]. Anche l'ampia review condotta da Sardi, Sorano et al., (2020), che ha esaminato ben 873 articoli relativi a trends, evoluzione e future opportunità dei Big Data, impiegando l'analisi biometrica, concorda nell'indicare un incremento significativo del numero di pubblicazioni sul tema, evidenziando, tuttavia, una carenza di studi nelle aree aziendali, gestionali e contabili.

È importante osservare come, sebbene la locuzione "Big Data" sia comunemente associata all'informatica, la produzione scientifica ne documenti l'applicazione a molteplici discipline. Segmentando la linea temporale ed esaminando le aree tematiche trattate nei diversi periodi, infatti, si può notare che nei primi lavori (cioè, quelli prodotti fino al 2000) ad essere interessati dal fenomeno sono l'ingegneria - in particolare le aree dell'ingegneria informatica (reti neurali, Intelligenza Artificiale, simulazione informatica, gestione, estrazione e archiviazione dei dati) - ma anche aree come materiali da costruzione, generatori elettrici, ingegneria elettrica ed elettrotecnica, apparecchiature per le telecomunicazioni e il trasporto pubblico, nonché sistemi di telefonia cellulare ed elettronica. Dal 2000 in poi, il

[18] Circostanza, peraltro, indirettamente suffragata dalla rassegna di Halevi & Moed, che indica nel 2008 l'anno dell'esplosione di pubblicazioni contenenti il termine "Big Data"

[19] Secondo un rapporto del 2017 di *Data Never Sleeps* di Domo, 2,5 quintilioni di byte di dati vengono generati quotidianamente a livello globale (Jackson-Barnes, 2023).

[20] "*Academic research on the topic also skyrocketed. Searching for the term 'big data', the Web of Science Core Collection yields 3347 hits in 2015, and over 4000 in both 2016 and 2017*" (Batistič & Van der Laken, 2019, 229).

settore risulta dominato dall'informatica, seguito da ingegneria e matematica. I dati, oltre a documentare l'estensione del termine alle discipline mediche e umanistiche, alle arti e alle scienze ambientali, sottolineano la predominanza delle conferenze scientifiche nella progressione delle pubblicazioni di settore, seguite dagli articoli su riviste specializzate. Dal punto di vista geografico, la ricerca evidenzia la leadership degli Stati Uniti, seguiti dalla Cina e da alcuni Paesi europei (Halevi & Moed, 2012).

Gli outcomes sopra riportati appaiono perfettamente in linea con gli attuali approdi del percorso evolutivo dei Big Data, sia in merito alla natura trasversale e "ubiquitaria" dei medesimi – Hammond-Errey afferma, testualmente, nel suo Report del 2022: *"Data abundance, digital connectivity, and ubiquitous technology now enable near complete coverage of human lives across the planet, often in real-time"* - sia rispetto agli attuali trends di sviluppo e alla leadership globale in materia di ICTs, a tutt'oggi attribuite a Stati Uniti e Cina (Knight, 2023). Posto che taluni Autori legano indissolubilmente l'evoluzione dei BD allo sviluppo dell'Intelligenza Artificiale (AI), secondo un meccanismo di empowerment reciproco (Duan, Edwards & Dwivedi, 2019 e bibliografia ivi citata), non sarebbe un caso se il primato tecnologico statunitense venisse attualmente insidiato dal competitor cinese. Il nuovo rapporto pubblicato dal *Massachusetts Institute of Technology (MIT)*, in collaborazione con il think tank *Council on Competitiveness* e la società di investimenti *Silicon Catalyst,* mostra, infatti, come, nell'ultimo quinquennio, la quota americana dei c.d. supercomputer più potenti al mondo sia calata sensibilmente, al pari del vantaggio degli Stati Uniti nell'informatica avanzata, che si andrebbe progressivamente assottigliando, soprattutto rispetto alla Cina (Thompson, Evans & Armbrust, 2023).

Il trend a livello globale sarebbe, peraltro, supportato anche da recenti indagini condotte da Società specializzate in BD Analytics come *Gartner*, la quale ha rilevato che l'89% delle aziende starebbe investendo nei Big Data per ottenere un vantaggio competitivo, e ciò varrebbe non solo per le Big Tech: la National Small Business Association avrebbe, infatti, riportato un utilizzo sistematico dei BD da parte del 63% delle piccole imprese, con l'intento esplicito di migliorare le prestazioni di mercato. I risultati riferiti non dovrebbero affatto stupire, soprattutto alla luce delle recentissime stime compiute da *McKinsey Co.*, secondo cui i Big Data potrebbero garantire un aumento della produttività del 2-3%, conducendo ad una riduzione dei costi del 20-25% (Sashidharan, 2023 e fonti ivi citate).

Per quanto concerne le caratteristiche dei Big Data, è necessario prendere le mosse dalla tradizionale (e oltremodo citata) definizione di Laney che, in una nota di ricerca non pubblicata del META Group del 2001 (Diebold, 2012 e bibliografia ivi citata), evidenzia il c.d. "Modello delle 3 V" dei BD, ossia Volume, Varietà e Velocità. La totalità delle fonti esaminate riporta i suddetti parametri come caratteri distintivi del paradigma dei Big Data: disponibili in enormi volumi, si presentano con formati destrutturati e caratteristiche eterogenee e sono sistematicamente generati ad elevatissime velocità. Secondo l'analisi compiuta da Rezzani. (2018), mole di dati come transazioni bancarie e movimenti sui mercati finanziari, solo per citare un esempio, assumono valori di un'ampiezza tale da non poter essere gestiti con i tradizionali tools tecnologici [21] ; mentre la diversità dei formati e l'assenza di una struttura rappresentabile attraverso una

[21] Secondo Idc, i dati sono in perenne crescita: nel 2019 la loro ecosfera, a livello globale, raggiungeva i 40 Zettabytes. Entro il 2025 questo valore è destinato a quintuplicarsi (Castigli, 2022).

tabella in un database relazionale sarebbero peculiarità dei BD principalmente riconducibili alla loro mancata strutturazione. Varie sono anche le fonti da cui essi vengono prodotti: alcuni sono generati automaticamente da macchine - come i dati provenienti da sensori o i log di accesso a un sito web o quelli del traffico su un router - altri, invece, sono generati dagli utenti del web.

E' la velocità con cui i nuovi dati si rendono disponibili il terzo fattore di identificazione dei Big Data e, proprio detto parametro renderebbe necessario l'impiego di strumenti in grado di garantirne il corretto immagazzinamento. Tra le tecnologie capaci di gestire i dati "ad alta velocità" siti specializzati come *www.dataskills.it* elencano i database *Historian* (per l'automazione industriale) e quelli denominati *Streaming Data* o *Complex Event Processing* (CEP), come *Microsoft StreamInsight*, un framework per lo sviluppo di applicazioni CEP che consente il monitoraggio di più fonti di dati, analizzando questi ultimi in modo incrementale con una bassissima latenza. Le applicazioni CEP trovano impiego con successo anche negli ambiti industriale, scientifico, finanziario e in quello relativo all'analisi degli eventi generati sul web.

Mentre rispetto ai tre caratteri distintivi originariamente enucleati da Laney non emergono difformità nell'ambito della letteratura esaminata, posizioni divergenti si registrano per quanto concerne i parametri aggiuntivi, da talune fonti identificabili nella Veridicità, ossia la capacità di determinare la certezza e la coerenza dei dati, e nel Valore, inteso come la capacità di ottenere approfondimenti sui dati e dai dati (Hammond-Errey, 2022 e bibliografia ivi citata). Secondo altri Autori, sarebbero tre le ulteriori discriminanti ad assumere rilievo nell'odierno scenario dei BD: Variabilità**,** una caratteristica riferita alla possibile inconsistenza dei dati analizzati, ossia la soggezione dei dati stessi a cambiamenti del loro significato, in ragione della loro origine da differenti contesti; Complessità, che aumenta in maniera direttamente proporzionale alla dimensione del dataset e Veridicità**,** relativa al valore informativo che è possibile estrarre dai dati, intesa come utilità informativa e/o predittiva[22] (Rezzani, 2018) - caratteristica, quest'ultima, rispetto alla quale parrebbe esserci una sostanziale convergenza in seno alla letteratura specialistica presa in esame, sebbene indicata con una terminologia differente (Ravera, 2023). In ogni caso, i parametri identificativi sono frutto dell'incremento delle fonti informative così come dell'evoluzione tecnologica e il progressivo ampliamento del catalogo degli stessi lo dimostrerebbe. L'upgrade in atto si traduce in quello che gli esperti definiscono, rispettivamente, il "Modello delle 5 V" (Castigli, 2022), delle "10 V" (Soomro, 2023), delle "17 V" (Arockia, Varnekha & Veneshia, 2017 e bibliografia ivi citata), in ragione del paradigma dei Big Data cui si aderisce. Tuttavia, la progressiva complessità dello scenario tecnologico promette ulteriori e significativi upgrade in materia.

2 Profili Giuridici ed Implicazioni Etiche dei Big Data

Come la precedente, anche la presente macroarea è stata indagata con metodologia di ricerca e analisi delle fonti OSINT, utilizzando operatori booleani generici e key words impostate sul linguaggio giuridico-normativo. In particolare, la web query è stata condotta mediante inserimento, nella barra di ricerca, di stringhe di caratteri del tipo "big data AND

[22] I dati rappresentano un'entità, un fenomeno o un avvenimento espressi in maniera codificata. L'informazione non è però una rappresentazione codificata, bensì il risultato di un processo di Data Analytics (Castigli, 2022).

profili giuridici" [23] oppure espressioni generiche come "diritto delle nuove tecnologie" "diritto dell'informatica", con progressivo affinamento della ricerca mediante l'impiego di key words-filtro (es. "diritto PENALE dell'informatica"), in ragione degli aspetti emersi nel corso dell'attività esplorativa delle fonti ritenuti rilevanti, pertinenti o, quantomeno, di interesse per l'attività in essere (Discovery). La selezione delle fonti (Discrimination) è stata compiuta avvalendosi della expertise in materia maturata principalmente durante il "Master di Secondo Livello in International Security Studies" recentemente conseguito presso l'Ateneo sammarinese. Ciò ha, peraltro, facilitato l'individuazione di peculiari criticità in tema di Big Data Governance, prima fra tutte quella inerente al binomio privacy & sicurezza dei dati - con un focus specifico sulla tutela della riservatezza dell'utente titolare dei dati personali oggetto di trattamento da parte di Organizzazioni pubbliche o private, non ultime le c.d. Big Tech – così da orientare rapidamente la ricerca, condotta prevalentemente su siti specializzati in materie giuridiche e avvalendosi di Summaries curati da Agenzie/Enti di settore.

Posto che le moderne capacità di raccolta ed elaborazione dei dati, conseguenti al rapido sviluppo delle ICT, consentono di utilizzare gli stessi come fonte di alimentazione per varie tipologie di business, oltre che per la soddisfazione di esigenze governative – marketing, profilazione, fidelizzazione e definizione del sentiment delle persone sono soltanto alcune delle finalità cui assolvono i Big Data - l'evoluzione tecnologica ha condotto alla nascita del concetto di "Data Protection" (DP) inteso come tutela delle informazioni personali: da ciò, la recente previsione di strumenti e procedure volte a limitare l'ingerenza delle tecnologie nella sfera individuale dei cittadini [24] . Nella protezione dei dati personali rientra, altresì, l'utilizzo trasparente degli stessi, che si inserisce nella prospettiva della relazione tra la raccolta di informazioni e dati e i dispositivi tecnologici che ad essa possono procedere per veicolarle. Nella maggior parte dei casi, lo scopo delle metodologie di DP è quello di assicurare il giusto compromesso tra la tutela del diritto alla privacy della persona fisica e l'utilizzo dei dati a scopi commerciali (Cerrone, 2022). Altrimenti detto, chi acquisisce dati "sensibili" ha il dovere, giuridicamente sanzionabile, di tutelare la persona identificata utilizzando le informazioni in proprio possesso solo per gli scopi consentiti dalla legge e per i quali è stato esplicitamente autorizzato al trattamento dal soggetto interessato (mediante consenso), in quanto titolare dei dati medesimi.

Data l'ampiezza e la complessità della tematica, la letteratura in materia è sterminata: la pregevole monografia di Nicola Fabiano (2020), dal titolo *"GDPR & Privacy: consapevolezza e opportunità. L'approccio con il Data Protection and Privacy Relationships Model (DAPPREMO)"*, e la ricchissima bibliografia proposta (cui si rinvia), rappresenta, senza dubbio, una delle opere fondamentali nel panorama italiano e sammarinese, dalla quale prendere le mosse per ulteriori ricerche. Altrettanto apprezzabile è il contributo di Emilio Tosi nell'ambito del volume *"Privacy digitale. Riservatezza e protezione dei dati personali tra GDPR e nuovo Codice Privacy"* che, pur se risalente al 2019, affronta argomenti-chiave

[23] Poiché larga parte delle fonti è prodotta in lingua inglese, nella web query sono state impiegate anche espressioni equivalenti, come "big data AND legal aspects", "data protection AND law" et similia.

[24] Il principale riferimento normativo in materia di protezione dei dati personali è il Regolamento Generale sulla Protezione dei Dati, meglio noto come GDPR, entrato in vigore il 27 aprile 2016. Il regolamento stabilisce "*norme relative alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché norme relative alla libera circolazione di tali dati*" e si propone di proteggere "*i diritti e le libertà fondamentali delle persone fisiche, in particolare il diritto alla protezione dei dati personali*".

di stringente attualità, primo fra tutti il controverso rapporto tra persona e mercato digitale all'indomani dell'entrata in vigore del Regolamento UE del 2016 e successive modifiche (GRPR). Di estremo interesse si è rivelata l'esplorazione della rassegna ragionata di fonti proposta dall'Autore: trattasi di voci di carattere dottrinale e giurisprudenziale, provenienti dalla produzione scientifica nazionale e comunitaria, cui è d'obbligo il rinvio per la ricchezza dei titoli raccolti e la sistematicità dell'opera. Il catalogo delle voci riportate consente, infatti, l'analisi del fenomeno a partire dai primi anni Duemila, così da rendere edotto il ricercatore, che a quelle voci si approccia, delle dinamiche evolutive del paradigma dei Big Data e delle relative implicazioni sul piano della tutela dei diritti individuali.

Dello stesso tenore appaiono i contributi di Autori italiani che forniscono dettagliate analisi dell'impatto del GDPR sul diritto interno, muovendo dall'esame del razionale sotteso alla riforma europea per giungere all'approfondimento dei principi generali da essa sanciti e al loro recepimento nell'ambito del c.d. Codice Privacy italiano. Tra questi, si segnalano gli scritti di De Stefani (2018), Lucchini Guastalla (2019), Ottonelli (2020), Calzolaio (2017) e le bibliografie ivi citate, anche con riferimento a profili squisitamente tecnologici: il richiamo è d'obbligo agli istituti cc.dd. della Privacy by Design e Privacy by Default, che hanno consentito l'aggiornamento della disciplina europea in materia di protezione dei dati personali rispetto all'avvento della società digitale, introducendo un modello di Data Protection fondato sulla rischiosità del trattamento, sulla responsabilità del titolare del trattamento e sulla protezione dei dati sin dal momento della progettazione del trattamento medesimo e per impostazione predefinita (Saetta, 2018).

Nel novero dei titoli bibliografici appaiono anche Tesi di Laurea o Dissertazioni post-dottorali aventi ad oggetto aspetti peculiari della normativa europea, come il c.d. Data Protection Impact Assessment (DPIA) - un processo finalizzato alla valutazione del rispetto dei principi privacy (i principi di necessità e proporzionalità, in primis) – e alla gestione dei rischi inerenti al trattamento dei dati personali (Chihai, 2019; Donato, 2020) nonché l'impatto del GDPR su aziende locali esaminato attraverso la lente esperta del Data Protection Officer (DPO), figura deputata alle procedure di risk assessment concernenti il trattamento dei dati nella realtà aziendale (Magnanelli, 2022). Anche in questi casi, sia consentito rinviare alle voci citate nelle ampie rassegne bibliografiche, un'analisi sommaria delle quali evidenzia uno spiccato (e comprensibile) propensione all'impiego di Case Studies per testare i risultati concreti dell'applicazione della Data Protection Law (DPL) nel contesto comunitario.

Un'area di applicazione meno nota della DPL (di certo meno indagata dalle consuete traiettorie di analisi), è quella relativa all'impiego dei dati individuali a fini di studio e ricerca, come si evince dal contributo di Bougleux (2021), in cui l'Autrice mette in luce limitatezza e rischi di un approccio – quello sotteso all'applicazione del GDPR (General Data Protection Regulation) nel settore specifico della ricerca etnografica - che propone una gestione unica di dati agglomerati per una molteplicità di attività accademiche differenti e dalle esigenze non riducibili ad una categoria unitaria. Il saggio, composto a latere di un webinar dal titolo *"General Data Protection Regulation in Antropologia"* promosso dalla SIAC (Società Italiana di Antropologia Culturale) nel febbraio 2021, è meritevole di attenzione sia in ragione dell'originalità dell'approccio che dell'ampia bibliografia citata, cui si rinvia per eventuali approfondimenti. Il nodo problematico affrontato in seno al dibattito è, in ogni caso, quello relativo alla differenza strutturale tra dato grezzo (raw data) e dato processato, con ripercussioni tanto sul versante dell'anonimato della fonte quanto su quello dell'interpretazione del dato e dell'attribuzione di significato ad esso (Leonelli, 2018).

Lo European Data Protection Supervisor (EDPS)[25] aveva, infatti, recentemente varato un'interessante "Opinion" sull'inedito binomio ricerca scientifica-protezione dei dati personali, ribadendo l'applicazione di ciascuno dei principi di cui all'art. 5 del GDPR – ossia, liceità-correttezza-trasparenza, limitazione della finalità, minimizzazione dei dati, esattezza, limitazione della conservazione, nonché integrità e riservatezza, cui si aggiunge il principio dell'*accountability* (o responsabilizzazione), che sancisce l'obbligo, per il titolare del trattamento, di dimostrare il rispetto dei sei principi precedenti - a tutti i trattamenti di dati, compresi quelli a fini di ricerca. Da notare come, anteriormente all'entrata in vigore del GDPR, la Direttiva 95/46/CE riconoscesse la ricerca come un importante settore di interesse pubblico, tale da giustificare deroghe alle "norme generali". Tanto premesso, detta Direttiva delegava, in massima parte, la tematica della protezione dei dati nei settori della sanità e della ricerca medica agli Stati Membri UE. (Mischitelli, 2020).

In una siffatta cornice normativa, dottrinale e giurisprudenziale si inscrive l'Indagine conoscitiva denominata *IC53 – Big Data*, avviata congiuntamente dalle tre Autorità Garanti italiane AGCOM (Autorità per le Garanzie nelle Comunicazioni), AGCM (Autorità Garante della Concorrenza e del Mercato) e Garante per la Protezione dei Dati Personali (c.d. Garante Privacy), a far data dal 30 maggio 2017. Al predetto documento - volto ad approfondire la conoscenza delle conseguenze del fenomeno dei Big Data nel contesto economico, politico e sociale del Paese, con specifico riferimento al quadro normativo vigente – ha fatto seguito un elaborato finale, licenziato nel mese di febbraio 2020, in cui le tre Autorità coinvolte hanno riportato i risultati dell'attività esplorativa. Le tappe dei lavori della Commissione d'indagine sono riportate in dettaglio da Piretti (2020) - sul portale giuridico *Diritto.it*, cui si rinvia per gli opportuni approfondimenti – così come da Cerrone (2022), sulla rivista "*Data Protection Law"*.

Orbene, tre sono le questioni principali affrontate dall'indagine, incentrata sull'analisi della propensione degli utenti web a consentire l'uso dei propri dati a fronte dell'erogazione di servizi online di varia natura: la verifica del grado di consapevolezza degli utenti delle piattaforme digitali in merito alla cessione e all'utilizzo dei dati individuali (Donato, 2020); la disponibilità degli utenti a cedere i dati personali come forma corrispettiva dei servizi online (Cerrone, 2022); infine, il controllo della c.d. portabilità dei dati, con ciò intendendosi la migrazione dei medesimi tra piattaforme digitali diverse (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2019). In ordine al citati profili, è stato opportunamente osservato che *"(…) gli agglomerati di dati ottenuti a seguito di elaborazione sono frutto del trattamento di dati personali le cui finalità non vengono specificate all'interno dell'informativa che dovrebbe essere consegnata preventivamente ex art. 13 GDPR. In aggiunta, anche il diritto alla portabilità ex art. 20 GDPR riesce difficoltosamente ad operare in virtù dei limiti di interoperabilità derivanti dall'utilizzo di diverse piattaforme da parte degli operatori economici. Un altro aspetto da non sottovalutare è quello relativo alla fruizione di servizi ed applicazioni offerti da provider OTT (Over The Top – Facebook, Amazon) poiché spesso, in questi casi, l'utente non ha possibilità di esprimersi sul trattamento dei dati personali effettuato dal provider; in tal caso, infatti, si presuppone che l'utente dia il consenso a tutte le condizioni di utilizzo, comprendenti anche quelle relative al trattamento dei propri dati personali con la conseguenza che l'utente sia ignaro non solo di dove risiedano i propri*

[25] Si tratta dell'autorità europea che vigila sul rispetto della protezione dei dati presso gli organi e le istituzioni dell'Unione Europea.

dati, ma anche dell'uso che ne fa l'OTT (…)" (Piretti, 2020).

L'approccio multidisciplinare adottato dalle tre Autorithies ha consentito di estrapolare quelle che Cerrone (2022) definisce *"le caratteristiche economiche delle piattaforme digitali"*, espresse principalmente in termini di importanza che per esse assume l'acquisizione dei Big Data, i quali fungerebbero da moneta di scambio per la fruizione di prestazioni offerte senza un corrispettivo monetario strettamente inteso. Proprio tale ultimo aspetto inciderebbe in maniera significativa sul comportamento dei consumatori, indotti a compiere scelte irrazionali e inconsapevoli. Secondo l'Autrice, infatti, l'uso ormai invalso nella pratica di *"(…) acconsentire al trattamento dei dati personali – attraverso strumenti rapidi ma problematici come quello delle "terms and conditions" - può rivelarsi rischioso non soltanto dal punto di vista della privacy, ma altresì da quello – da molti non preso in considerazione – della tutela del consumatore, il quale diviene preda facilmente catturabile attraverso l'utilizzo di pratiche commerciali scorrette, e dunque mediante l'inserimento – da parte dell'imprenditore – di cunei di ingannevolezza o aggressività nei rapporti commerciali"* (Cerrone, 2022, 5). Si tratta di un considerevole cambio di paradigma, posto che, abitualmente, la materia dei dati personali viene analizzata secondo una lente di tipo squisitamente personalistico. Approccio, quest'ultimo, riconducibile alla definizione stessa di "dato personale", inteso come informazione che identifica (o rende identificabile) una determinata persona fisica: trattasi di un dato personale poiché, appunto, riferito "alla persona", e perciò tutelabile alla stregua dell'art. 2 Cost., che garantisce *"i diritti inviolabili dell'uomo, sia come singolo sia nelle formazioni sociali ove si svolge la sua personalità"*.

Una riflessione a margine in tema di protezione della sfera individuale, in primo luogo mediante la salvaguardia e il controllo delle informazioni personali condivise, attiene all'atteggiamento dei consumatori sui mercati digitali, che si traduce in quello che è stato felicemente battezzato il *"paradosso della privacy"*. Posto che *"(…) la comprensione degli atteggiamenti e dei comportamenti dei consumatori in relazione alla privacy e alla protezione dei dati è una questione chiave nell'analisi delle dinamiche competitive nei mercati che coinvolgono i dati degli utenti"* (Donato, 2020, 70), la letteratura specialistica documenta una sorta di atteggiamento "schizofrenico" da parte degli utenti della rete, oscillante tra (comprensibili) preoccupazioni attinenti alla violazione della sfera personale e condivisione diffusa e generalizzata (non di rado, incauta) di dati sensibili.

Un'indagine condotta nel 2015 dalla Commissione Europea su oltre 15.000 internauti ha rilevato, infatti, come l'80% degli intervistati ritenga di non essere in grado di esercitare un pieno controllo sui dati personali condivisi online e come tale circostanza costituisca fonte di preoccupazione per almeno due terzi di essi (European Commission, 2015). Ciò sembrerebbe, tuttavia, ampiamente smentito dall'effettivo comportamento degli utenti in rete, che tradirebbe una certa "disinvoltura" nella divulgazione di informazioni private, soprattutto sulle c.d. piattaforme di attenzione[26], social network in modo particolare (Prat & Valletti, 2021). Le evidenze empiriche dimostrano come detta discrepanza risulti influenzata da molteplici fattori di contesto (Acquisti, Taylor &

[26] Vengono definite piattaforme (o broker) di attenzione quelle aziende capaci di ottenere informazioni sulle preferenze individuali degli utenti e indirizzare singolarmente, sulla base di queste, gli annunci pubblicitari, mediante attività di profilazione. Dette piattaforme, cioè, non chiedano ai consumatori un prezzo in termini monetari per accedere ai loro servizi, quanto piuttosto la loro "attenzione" (intesa come il tempo di permanenza sulla piattaforma aziendale), con conseguente trasferimento dei dati personali, che diviene perciò il "corrispettivo" della prestazione digitale (Purra & Carlsson, 2016 e bibliografia ivi citata).

Wagman, 2016 e bibliografia ivi citata), unitamente ad una gamma di pregiudizi comportamentali in materia di privacy, non ultima la miopia dei consumatori, propensi alla condivisione dei propri dati in vista di un beneficio immediato, senza considerare le implicazioni sulla privacy a lungo termine - quello che Donato (2020) definisce, provocatoriamente, «il costo del "gratuito"». La produzione scientifica, dottrinale e giurisprudenziale sul tema è ampia e variegata, seppur concorde nel rilevare la criticità di un siffatto approccio da parte degli internauti, come minuziosamente documentato dal *"Consumer Data Rights and Competition - Background note"* licenziato dal Directorate For Financial And Enterprise Affairs Competition Committee Secretariat dell'OECD nel giugno del 2020.

Big Data, Privacy & Concorrenza (e il loro reciproco embricarsi) si rivela una delle tematiche più complesse ed indagate nell'attuale panorama di studio e di ricerca a livello internazionale, come peraltro testimoniano l'elaborato di Donato (2020) e la fiorente rassegna bibliografica ivi riportata, cui si rinvia per gli approfondimenti di settore. Da una disamina della produzione sul tema emergono, principalmente, i molteplici rischi connessi al c.d. duopolio dell'attenzione, costituito da Big Tech del calibro di *Google* – il principale search engine utilizzato dagli utenti nel mondo occidentale per la ricerca di contenuti online – e *Facebook* che, con oltre 2,5 miliardi di utenti, si attesta come la piattaforma di incontro virtuale maggiormente diffusa a livello globale. Ciò che ne ha consentito il consolidamento delle rispettive leadership sui mercati digitali, in veste di gatekeeper tra le imprese e i potenziali clienti – tematica, questa, diffusamente affrontata sia sul piano scientifico (Binns, Lyngs, Van Kleek, et al., 2018; Prat & Valletti, 2021; Robertson, 2019 e bibliografie ivi citate) che istituzionale (Federal Trade Commission, 2014; OECD, 2020).

Per comprendere la complessità della questione – e la presenza di aree di sovrapposizione concettuale non irrilevanti sul piano giuridico - è sufficiente richiamare l'accurata analisi attinente all'ampiezza semantica di espressioni quali "consumer data" e "personal data", riportata nella summenzionata nota ufficiale dell'Organizzazione per la Sicurezza e la Cooperazione in Europa (OECD, 2020), a mente della quale, per "consumer data" devono intendersi i dati relativi agli utenti in quanto oggetto di raccolta, scambio e/o utilizzo nell'ambito di un rapporto commerciale. Concetto, quest'ultimo, sensibilmente più ristretto di quello di "personal data" – definito come *"(…) qualsiasi informazione relativa a un individuo identificato o identificabile (soggetto dei dati)"* (OECD, 2013, 13). Altrimenti detto, il concetto di "personal data" risulta essere ben più ampio di quello di "consumer data", includendo la totalità dei dati relativi all'individuo, nella duplice veste di cittadino e consumatore.

Il concetto di "consumer data" deve, inoltre, ritenersi riferibile ai soli dati relativi all'individuo "in quanto consumatore", posto che la disciplina applicabile congiuntamente in materia di privacy e concorrenza (almeno nell'alveo di operatività della citata nota OECD del 2020) investe le sole transazioni commerciali. Pertanto, dai "dati dei consumatori" esulano quelli raccolti, scambiati e impiegati da Governi o altri Agenti/Organizzazioni non commerciali, rispetto ai quali potrebbero sorgere, nondimeno, problematiche di differente natura (OECD, 2020). Per altri versi, tuttavia, l'espressione "consumer data" si rivela più ampia di "personal data", in quanto potenzialmente inclusiva di dati relativi ai consumatori anche qualora i predetti non siano riconducibili all'individuo. Sebbene tali dataset non sollevino le stesse criticità in termini di Data Protection Law, gli stessi possono assumere rilevanza sotto il diverso (ma non secondario) profilo della disciplina antitrust (OECD, 2020).

Si consideri, inoltre che, l'impossibilità di ricondurre un set di "personal data" a un

individuo identificato (o identificabile) secondo la categorizzazione ISO/IEC 19441[27], riduce la probabilità di assoggettamento degli stessi alla normativa sulla privacy, erodendo significativamente i margini di tutela dell'utente/fruitore dei mercati digitali. Ciò significa che il trattamento di tali dati sarà in grado di generare vantaggi economici e competitivi a favore delle organizzazioni commerciali (leggi, profitto) senza integrare alcuna violazione, giuridicamente perseguibile, della privacy del loro legittimo titolare (OECD, 2019).

A corollario delle precedenti considerazioni, il citato documento del 2020 presenta un'ampia panoramica degli istituti giuridici di carattere generale posti a tutela dei c.d. consumer data – rispettivamente applicabili nei Paesi del circuito OECD e in quelli della comunità internazionale - oltre a delineare specifici diritti riconosciuti ai titolari di personal data dalla normativa interna dei singoli Stati. Di tali discipline viene di seguito riportata una sintetica rassegna (OECD, 2020; 2013 e bibliografia ivi citata), aggiornata con i risultati di un'indagine condotta su siti web specializzati.

Per quanto concerne i c.d. Consumer Data Rights, larga parte dei Paesi membri del circuito comunitario dispone di una forma di legislazione in materia di Data Protection, coerente con il quadro definito nelle "Privacy Guidelines" e con gli otto principi enucleati dall'OECD nel 2013[28]. Le singole legislazioni tendono a fornire una protezione basilare della privacy degli utenti online, riconoscendo agli interessati il diritto di esercitare il controllo sui propri dati in maniera concreta ed efficace. La maggior parte delle legislazioni nazionali prevede,

[27] *"Personal data can also be classified according to the extent to which it is personally identifiable. In particular, ISO/IEC 19441 distinguishes between five categories including: 1. Identified data, which is unambiguously associated with a specific person. 2. Pseudonymised data, in which aliases are used in place of personal identifiers; aliases can only be reversed by the party that assigned them. 3. Unlinked pseudonymised data, in which all identifiers are erased or substituted by aliases for which the assignment function is erased or irreversible, so that the linkage cannot be re-established by anyone. 4. Anonymised data, which is unlinked and altered (e.g., attributes' values are randomised or generalised) in such a way that there is a reasonable level of confidence that a person cannot be identified. 5. Aggregated data, which does not contain individual-level entries and is combined from information about enough different persons that individual-level attributes are not identifiable*" (Technical Committee ISO/IEC JTC 1/SC 27, 2018).

[28] 1. Limitazione della raccolta: la raccolta dei dati personali deve essere limitata, i dati devono essere ottenuti con mezzi leciti e previa conoscenza o con il consenso dell'interessato.
2. Qualità dei dati: i dati personali devono essere pertinenti alle finalità per le quali vengono raccolti e per cui devono essere utilizzati, devono essere accurati, completi e aggiornati.
3. Specificazione delle finalità: le finalità per cui vengono raccolti i dati personali devono essere specificati (al più tardi al momento della loro raccolta) e l'uso successivo deve essere limitato a tali finalità.
4. Limitazione dell'uso: i dati personali non devono essere divulgati, resi disponibili o altrimenti utilizzati per scopi diversi da quelli specificati, salvo: a) con il consenso dell'interessato; o b) in base alla legge.
5. Salvaguardie di sicurezza: i dati personali devono essere protetti da ragionevoli garanzie di sicurezza contro rischi quali la perdita o l'accesso non autorizzato, la distruzione, l'uso, la modifica o la divulgazione dei dati.
6. Apertura: le pratiche di raccolta e utilizzo dei dati devono essere aperte, in modo da consentire alle persone di stabilire l'esistenza e la natura dei dati personali, l'uso dei dati, nonché l'identità e la residenza del responsabile del trattamento.
7. Partecipazione individuale: l'individuo deve avere il diritto di: a) sapere se un responsabile del trattamento dei dati possiede dati che lo riguardano; b) ricevere i propri dati (i) entro un termine ragionevole; (ii) ad un costo, se del caso, non eccessivo; (iii) in modo ragionevole e (iv) in una forma che sia prontamente intelligibile; c) essere motivato in caso di rifiuto della richiesta e di poter impugnare il rifiuto; d) contestare i dati che lo riguardano e, se la contestazione ha esito positivo, ottenerne la cancellazione, la rettifica, il completamento o la modifica.
8. Responsabilità: il responsabile del trattamento dei dati deve essere responsabile del rispetto delle misure che danno attuazione ai principi sopra enunciati (OECD, 2013).

infatti, un regime basato sul principio del consenso, che dovrebbe offrire ai consumatori la possibilità di verificare il modo in cui i loro dati vengono raccolti e utilizzati - fornendo o negando il proprio assenso, appunto, rispetto alla pratica suddetta. In talune legislazioni, il sistema di Data Protection Law conferisce ulteriori diritti, tra cui: il diritto di rettifica delle informazioni inesatte, che riconosce agli interessati il diritto di chiedere ed ottenere la rettifica, da parte del titolare del trattamento, delle informazioni personali errate; il diritto all'oblio, ossia alla cancellazione dei dati personali quando non più necessari/funzionali al trattamento; il diritto alla portabilità dei dati, che consente agli interessati di trasferire i dati personali da un "responsabile/titolare del trattamento" a un altro.

Com'è noto, tali diritti sono inclusi nella principale normativa europea sulla protezione dei dati - il Regolamento Generale sulla Protezione dei Dati (GDPR) del 2016 e successive modifiche e integrazioni - che prevede un catalogo di diritti a tutela dei cittadini europei, tra cui i diritti di accesso ai dati personali (art. 15) e rettifica dei dati personali inesatti (art. 16); il diritto all'oblio (art. 17) e alla portabilità dei dati (art. 20); il diritto a non essere assoggettati a processo decisionale automatizzato, compresa la profilazione personale, salvo eccezioni (art. 22). Ulteriori aspetti chiave del GDPR sono costituiti da una definizione rivista e ampliata della nozione di "personal data" (art. 4); dal rafforzamento del regime di consenso applicato alla raccolta e all'elaborazione dei dati personali (artt. 6 e 7); dall'obbligo di trasparenza e facile accessibilità alle politiche sulla privacy (artt. 12, 13 e 14); dall'ampliamento dell'ambito di applicazione territoriale della normativa europea (art. 3); dalla previsione di requisiti specifici per la protezione dei dati in fase di progettazione e per impostazione predefinita (art. 25); infine, un inasprimento del trattamento sanzionatorio in caso di mancata applicazione della normativa europea (art. 83). Diritti dai contenuti similari sono contemplati, già da tempo, anche dalle legislazioni di alcuni Paesi degli Stati Uniti, come la California, con il *California Consumer Privacy Act (CCPA)* (CMA, 2015; Moazed, 2019).

Poliedrica e assai diversificata è la normativa di settore nel panorama internazionale. A differenza dell'Unione Europea, infatti, gli Stati Uniti difettano di un quadro normativo completo sulla privacy che stabilisca uno standard universale per il trattamento dei dati personali. In assenza di una normativa di carattere generale di livello federale, diversi Stati americani hanno emanato (o stanno emanando) leggi interne sulla privacy, al fine di tutelare le informazioni personali dei propri residenti. Il 2023 si preannunciava un anno significativo per la privacy dei dati negli USA, con diversi sviluppi degni di nota. Il 1° gennaio, infatti, sono entrati in vigore il *California Privacy Rights Act* e il *Virginia Consumer Data Protection Act*, inaugurando così una nuova era di normative sulla privacy. Il Connecticut, il Colorado e lo Utah sono pronti a implementare le proprie leggi in materia nel corso dell'anno, aggiungendosi al panorama delle misure di protezione dei dati attualmente in vigore. Anche Indiana, Montana, Iowa e Tennessee hanno adottato leggi sulla privacy, che entreranno in vigore tra il 2024 e il 2026. Simili sviluppi dimostrano la crescente importanza attribuita alla salvaguardia delle informazioni personali nel contesto statunitense, riflettendo l'impegno condiviso per migliorare le pratiche di tutela della privacy in tutto il Paese (Frazao & Strachan, 2023).

Negli Stati Uniti - dove la disciplina sui dati è di competenza statale, sebbene la Federal Trade Commission (FTC) abbia l'autorità per garantire che le Aziende non adottino atti o pratiche sleali o ingannevoli, anche per ciò che concerne le c.d. Personal Data Practices – l'anno in corso si sta rivelando il più prolifico sul piano della produzione normativa in materia. Benché la principale economia su scala globale abbia resistito (almeno per il momento) alla

tendenza di emanare una legislazione completa sulla protezione dei dati a livello federale[29], i singoli Stati hanno adottato, o stanno adottando, misure a tutela dei dati personali dei consumatori. Il nutrito catalogo delle leggi entrate in vigore nel 2023 (e di quelle che dovrebbero essere approvate entro l'anno in corso) è riportato dai maggiori siti statunitensi specializzati in materia di Data Protection Law (*secureprivacy.ai*; *verasafe.com; loginradius.com*, ecc.).

Di particolare interesse tra quelli visitati, si è rivelato il sito specializzato in DPL *iapp.org* [30] , che ospita un documento contenente un'accurata sintesi degli *Opt-Out Rights* e degli *Opt-Out Preference Signal Requirements* imposti dalle normative statali alle Aziende statunitensi operanti sui mercati digitali. In sostanza, detto documento, rilasciato dalla Thompson Hine LLP, illustra la gamma di diritti e obblighi in materia di protezione dei consumer data[31], previsti dalle singole leggi statali in vigore dal 1° gennaio 2023. I "Legal Requirements" riportati nell'apposita tabella indicano le modalità per mezzo delle quali i consumatori potranno opporsi all'utilizzo dei loro dati a fini di pubblicità mirata o di cessione a terze parti. Gli Stati federali attualmente interessati sono California, Colorado, Connecticut, Utah e Virginia: *California Privacy Rights Act 2020 (CPRA); Colorado Privacy Act (CPA); Connecticut Data Privacy Act (CDPA); Utah Consumer Privacy Act (UCPA) e Virginia Consumer Data Protection Act (VCDPA)* sono le rispettive normative interne di riferimento. In breve, ciascuno Stato ha elaborato un differente quadro di conformità (framework) indicante le modalità con cui le Aziende coinvolte dovranno consentire ai consumatori l'esercizio di tali diritti, mediante l'uso di "link" di opt-out pubblicati sui siti web aziendali e/o attraverso l'implementazione di una tecnologia per i siti web che si adatti alle configurazioni di privacy di un dispositivo digitale [32] . Gli Stati federali interessati stanno attualmente elaborando e perfezionando le proprie norme in materia[33],

[29] La legislazione federale in fase di studio già dal 2020 prevedeva l'approvazione di un progetto di legge denominato DASHBOARD, che avrebbe dovuto imporre alle piattaforme digitali di migliorare la trasparenza sulla raccolta e l'utilizzo dei dati degli utenti, nonché la legge *"Own Your Own Data"*, che avrebbe dovuto conferire agli utenti un diritto di proprietà esclusivo sui propri dati online (Chakrovorti, 2020). Era in esame anche il progetto di legge ACCESS, che avrebbe dovuto prevedere la portabilità dei dati per le piattaforme di social media con oltre 100 milioni di utenti negli Stati Uniti, secondo quanto riportato nella Background Note OECD, 2020.

[30] Come si legge nella homepage del sito, *"The International Association of Privacy Professionals (IAPP) is a resource for professionals who want to develop and advance their careers by helping their organizations successfully manage these risks and protect their data. In fact, we're the world's largest and most comprehensive global information privacy community"*.

[31] Come precisato a titolo di disclaimer, la tabella non contempla i casi in cui il consenso sia richiesto nell'ambito della pubblicità mirata o della cessione di dati personali, dei diritti di opt-out per la profilazione o di obblighi simili in materia di privacy.

[32] Secondo il CPRA 2020, ad es. un'azienda che ceda, condivida, utilizzi o divulghi informazioni personali sensibili per determinati scopi dovrà installare un link *"Do Not Sell or Share My Personal Information"* sulla sua pagina web per consentire ai consumatori di opporsi alla vendita o alla condivisione delle informazioni personali. L'impiego di un c.d. Segnale di Preferenza Opt-Out consente, invece, ai consumatori di opporsi alla cessione o alla condivisione dei dati personali sensibili (o di limitarne l'uso) mediante un apposito messaggio, di carattere generale, rivolto a tutti gli operatori con cui interagiscono online, senza la necessità di inviare richieste personalizzate a ciascuna Azienda (IAPP, 2023).

[33] A mero titolo esemplificativo, il *Colorado Privacy Act (CPA)* conferisce al Procuratore Generale del Colorado l'autorità di adottare norme in materia di privacy e richiede che, entro il 1° luglio 2023, il PG adotti norme indicanti le tecniche di un meccanismo di opt-out universale, che consenta di comunicare chiaramente la scelta affermativa, libera e inequivocabile del consumatore di opporsi al trattamento dei dati personali a fini di pubblicità mirata o di vendita

come peraltro desumibile dall'esplorazione dei websites governativi, cui si rinvia per un esame aggiornato dello stato dei lavori (*cppa.ca.gov; coag.gov,* ecc.).

Degno di nota anche il sito web *verasafe.com*, che propone una disamina ragionata delle normative statali attualmente in vigore negli USA, evidenziandone similitudini e divergenze nel blog post a firma congiunta di Frazao & Strachan, datato 26 giugno 2023 (cui si rinvia per un'analisi dettagliata della materia). Meritevole di interesse appare lo scenario prefigurato dai due Autori in tema di Data Protection Law negli Stati Uniti. Nel prossimo triennio, è prevista l'entrata in vigore di leggi sulla privacy in Montana, Indiana, Iowa, Tennessee e Texas. Diverse proposte di legge di portata generale sulla privacy sono in fase di elaborazione nelle legislature statali: attualmente, sono in corso proposte di legge in Delaware, Massachusetts, New Hampshire, New Jersey, North Carolina, Oregon, Pennsylvania e Rhode Island. A livello federale, *l'American Data Privacy Protection Act (ADPPA)* è all'esame del Congresso e mira a eliminare parte della complessità introdotta dalle diverse leggi statali sulla privacy. L'ADPPA, tuttavia, non crea uno standard di privacy uniforme, posto che continueranno ad essere parzialmente applicate specifiche discipline settoriali.

Particolare attenzione merita altresì il monitoraggio dei progressi del *Data Privacy Framework UE-USA* [34] un nuovo quadro transatlantico per il trasferimento dei dati che sostituisce il Privacy Shield Framework[35] La Commissione Europea, infatti, potrebbe adottare una decisione in merito durante l'anno in corso. Vero è che, mentre la maggior parte delle leggi sulla privacy degli Stati Uniti si fonda sugli stessi principi di base - allineandosi, in qualche misura, ai principi del Regolamento Generale sulla Protezione dei Dati (GDPR) dell'Unione Europea - i requisiti e gli obblighi specifici differiscono significativamente da uno Stato all'altro e richiedono un'attenta considerazione e analisi per garantirne un'efficace conformità (Frazao & Strachan, 2023).

Il sito specializzato statunitense *longinradius.com* ospita un blog post di Alok Patidar - responsabile della sicurezza informatica presso la LoginRadius, azienda leader nel settore della gestione dell'identità digitale e dei consumer data - che presenta una rassegna commentata delle normative essenziali in vigore, a livello internazionale, per l'anno 2023. L'assunto di base dal quale prende le mosse è che, nell'era digitale, le organizzazioni devono necessariamente navigare in un complesso panorama di leggi e regolamenti per salvaguardare le informazioni personali dei propri clienti - ciò che implica una conoscenza delle discipline di settore a livello globale, posto che, dall'entrata in vigore del GDPR, sono emerse in tutto il mondo numerose leggi sulla privacy dei dati. Ad ampliare il novero delle normative vigenti, la *Health Insurance Portability and Accountability Act (HIPAA)*,

di dati personali (Stransky, Knight, & Zych, 2023).

[34] Il 13 dicembre 2022 la Commissione Europea ha pubblicato la bozza di decisione di adeguatezza per il DPF UE-USA, (https://commission.europa.eu/document/e5a39b3c-6e7c-4c89-9dc7-016d719e3d12_en). Successivamente, la maggioranza dei membri del Parlamento Europeo ha votato a favore di una Risoluzione che si oppone all'adozione di una decisione di adeguatezza nell'ambito del *Data Protection Framework*.

[35] Com'è noto, il *Privacy Shield Framework* era un meccanismo approvato per il trasferimento di dati personali dall'UE e dalla Svizzera agli Stati Uniti, grazie al quale le aziende avrebbero avuto protezioni "adeguate" nel trasferire dati personali a società autocertificate, come richiesto dal GDPR. Pur rimanendo operativo, il *Privacy Shield* non può essere utilizzato come meccanismo legittimo per trasferire dati personali negli Stati Uniti. La questione di fondo sottesa al trasferimento dati UE-USA risiede nella diversità di approccio tra i due sistemi giuridici, che assume la forma del conflitto tra la legislazione statunitense, centrata sulla sorveglianza, e quella continentale, che impone il rispetto della privacy. Le tappe salienti del processo tutt'ora in atto sono riportate dal sito specializzato *noyb.eu,* cui si rinvia.

una legge statunitense concepita per la tutela delle informazioni sanitarie protette (PHI-Personal Health Information) degli individui. Applicabile ad operatori specializzati - fornitori di servizi sanitari, centri di clearing sanitario e partners d'affari – impone rigorosi standard di privacy e sicurezza per le PHI, tra cui limitazioni al loro uso e divulgazione, requisiti per la loro archiviazione e trasmissione sicura oltre all'implementazione di salvaguardie amministrative, fisiche e tecnologiche per proteggere le informazioni sensibili da accessi o divulgazioni non autorizzati.

Dello stesso tenore, il *New York Stop Hacks and Improve Electronic Data Security (SHIELD) Act* che potenzia i requisiti di privacy dei dati e di sicurezza informatica per le aziende che trattano le informazioni private dei residenti dello Stato di New York. Tale risultato viene conseguito anche mediante l'ampliamento della definizione di "informazioni private", che include ora dati biometrici, indirizzi e-mail e nomi-utente combinati con le password. La legge, inoltre, rafforza gli obblighi di notifica delle violazioni, imponendo alle aziende di informare tempestivamente le persone interessate e le Autorità competenti in caso di violazione dei dati, in maniera sostanzialmente simile all'istituto del "data breach notification" previsto dall'art. 33 GDPR (Patidar, 2023).

Nel panorama normativo internazionale, un cenno a parte merita il *Consumer Data Right* (CDR) - un articolato quadro legislativo e di standard per la portabilità dei dati dei consumatori, introdotto dal governo australiano nel 2019 – che consente ai clienti di accedere ai dati relativi alle loro transazioni bancarie, energetiche, telefoniche e internet (Barbaschow. 2017; Easton, 2017; Philipson, 2017). La normativa, attualmente in vigore[36], fornisce alle persone fisiche (consumatori) il diritto di accedere in modo efficiente ai dati specifici che li riguardano, detenuti dalle imprese (titolari dei dati); autorizza l'accesso sicuro ai dataset da parte di terze parti accreditate (destinatari accreditati dei dati), oltre ad imporre alle aziende di fornire al pubblico l'accesso alle informazioni su determinati prodotti offerti online. La legislazione sul CDR stabilisce un quadro di riferimento per consentire la progressiva applicazione della normativa ai vari settori dell'economia (Jones, 2022). La nuova disciplina del 2019 entrava in scia del *Privacy Act* australiano del 1988, che già conferiva una serie di diritti agli utenti della rete, tra cui il diritto di conoscere quali informazioni personali sarebbero state raccolte, il loro effettivo utilizzo e i soggetti a cui verranno divulgate; il diritto alla c.d. anonimizzazione dei dati; il diritto di richiedere l'accesso alle informazioni personali e di ottenere la correzione delle informazioni personali errate. Altri esempi di legislazione internazionale in materia sono rappresentati dalla legge sulla privacy emanata in Nuova Zelanda (*Privacy Act 1993*) e la legge sulla protezione delle informazioni personali e dei documenti elettronici (*Personal Information Protection and Electronic Documents Act* - PIPEDA) nonché la legge sulla privacy emanata in Canada nel 1985 e successive modifiche e integrazioni (*Privacy Act R.S.C., 1985, c. P-21*).

Alcuni Paesi del circuito OCSE hanno definito una normativa ad hoc per disciplinare specifiche pratiche a tutela degli utenti/consumatori operanti sui Digital Markets, prima tra tutte, quella attinente ai c.d. diritti di portabilità dei dati – istituto presentato come strumento di incremento della "'autodeterminazione informativa" e di garanzia dei risultati della concorrenza, poiché atto ad agevolare il cambio di fornitore di beni e/o servizi online (European Commission Data Protection, 2017). In una simile ottica, la portabilità e mobilità di dati tra diverse piattaforme digitali dovrà essere favorita anche mediante l'adozione di modelli

[36] Nel settembre 2022, il Governo australiano ha pubblicato una revisione statutaria indipendente sul quadro normativo del CDR e sulla sua attuazione negli ultimi anni (Kelly, 2022).

aperti e interoperabili, così da superare l'attuale impasse tecnologico, o quantomeno, mitigarne la portata (Camera dei Deputati, Parlamento Italiano, 2020). Nei Paesi membri dell'OECD esistono, allo stato attuale, numerose iniziative legislative e regolamentari in materia di portabilità dei dati, promosse dai governi nazionali, un'analisi comparata delle quali è compiuta da Takatsuki (2019), sul blog del sito specializzato *fieldfisher.com*, cui si rinvia per ulteriori approfondimenti.

Nel contesto europeo, l'articolo 20 del GDPR, entrato in vigore nel 2018, riconosce ai cittadini il diritto di ricevere e trasferire i propri dati personali in un *"formato strutturato, di uso comune e leggibile da dispositivo automatico"*. Tale disposizione si applica a tutti i dati personali forniti volontariamente dall'utente e trattati con mezzi automatizzati, in relazione all'esecuzione di una transazione commerciale (es. contratto). L'esercizio di questo diritto, tuttavia, non deve pregiudicare i diritti e le libertà di altri soggetti (ad esempio, la privacy di terzi). Sempre in Europa, la *Payment Service Directive for Payment Businesses (PSD2)* impone agli istituti di credito di consentire l'accesso a terzi ai dati dei conti di pagamento dei clienti, previo consenso esplicito di questi ultimi (Takatsuki, 2019).

Più articolata appare la disciplina statunitense, nella quale il *California Consumer Privacy Act (CCPA)*, entrato in vigore all'inizio del 2020, "fonde" i due istituti del diritto di accesso e diritto alla portabilità dei dati. In sostanza, la normativa riconosce ai consumatori il diritto di ottenere dalle aziende la rivelazione delle categorie di informazioni personali raccolte dall'Azienda in questione sul loro conto, unitamente ad informazioni circa le finalità aziendali o commerciali della raccolta o della cessione a terzi di informazioni personali; sui dettagli delle categorie di terzi con cui l'azienda condivide o a cui cede le informazioni personali e sulle categorie di informazioni personali altrimenti cedute o divulgate per scopi commerciali (Takatsuki, 2019).

Come in precedenza descritto, il Governo australiano, nel 2017, ha annunciato l'introduzione di un istituto di *Consumer Data Right - Treasury Laws Amendment Bill*, approvato nel 2019 e ufficialmente in vigore dal luglio 2020 - per garantire ai consumatori e alle piccole imprese un maggiore accesso e controllo sui propri dati. La finalità dichiarata di detto istituto è quella di promuovere la concorrenza, facilitando la migrazione degli utenti tra fornitori concorrenti mediante la portabilità dei propri dati (Australian Competition & Consumer Commission, 2023). Il CDR, che si concentra principalmente sull'esperienza dei consumatori, mira a modificare nel tempo il panorama economico, agevolando l'uso dei dati a favore dei nuovi operatori sul mercato, in un ampio numero di settori, primo tra tutti, quello bancario, seguito da quello dell'energia e delle telecomunicazioni [37] . L'introduzione della nuova disciplina sarà graduale - così si legge nella sezione dedicata del sito governativo *cdr.gov.au* - nel quale si definisce Il diritto ai dati dei consumatori concepito come un diritto di portata economica[38] così denotando l'accettazione di un'evidenza ormai acquisita nella recente letteratura specialistica (Zuboff, 2019).

Orbene, nell'era digitale, in cui il trattamento dei dati personali riveste un ruolo cruciale per la competitività tra le imprese e le dimensioni della privacy e della concorrenza appaiono strettamente

[37] Testualmente, *"The Consumer Data Right allows consumers to safely share the data that businesses hold about them. It can help consumers to compare products and services to find offers that best match their needs (...). Giving consumers the power to share their data also encourages innovation in new products and services"* (ACCC website, https://www.accc.gov.au/by-industry/banking-and-finance/the-consumer-data-right).

[38] Testualmente, *"(…) The Consumer Data Right is designed to be an economy-wide right"* (Australian Government website, Consumer Data Right, https://www.cdr.gov.au/rollout).

correlate, gli aggregati di dati individuali assumono rilevanza anche in termini patrimoniali – quella che Cerrone (2022) definisce *"la dimensione economica dei dati personali",* con ciò intendendosi, appunto, l'aspetto (anche) patrimoniale del set informativo del singolo utente (Rodotà, 1999)[39] - posto che la disponibilità di un asset di dati (personali e non) consente ai possessori delle informazioni da essi estrapolate di maturare una posizione di decisivo vantaggio competitivo nei mercati online, mediante la configurazione di specifici profili di abitudini di consumo (c.d. profilazione) funzionali alla proposizione di contenuti graditi all'utente[40]. Si perfeziona, cioè, quel processo di patrimonializzazione - ben delineato dalla dottrina recente (Senigaglia, 2020; Ricciuto, 2020, 2019; Thobani, 2018; De Franceschi, 2017) e richiamato nella nota sentenza del Consiglio di Stato 2631/2021[41], sull'obbligo di chiarezza gravante sul professionista nel fornire le informazioni al consumatore, nella

[39] La recente sentenza del Consiglio di Stato (Cons. St. 29 marzo 2021 n. 2631) ha evidenziato la "commerciabilità" del dato personale, così consentendo il riconoscimento di tutele di natura patrimoniale anche qualora siano coinvolte questioni (solo) apparentemente estranee alla logica commerciale dello scambio economicamente inteso - ossia bene a fronte di un corrispettivo economico, c.d. prezzo. Secondo Cerrone (2022, 9), infatti, *"(…) l'errore in cui solitamente si incorre è credere che la mancanza di un "prezzo", così come comunemente viene inteso, ossia come esborso monetario reso a fronte di un bene o di un servizio, permette di imprimere carattere gratuito alla prestazione che ci viene resa, per la quale, a ben vedere, il costo è rappresentato da noi stessi e dal nostro universo di informazioni private"*. La citata pronuncia evidenzia, in particolare, il processo di patrimonializzazione che tali dati subiscono all'atto della sottoscrizione di un accordo con i c.d. social network.

[40] La pubblicità mirata o target advertising ha luogo quando le aziende pubblicano annunci destinati ad utenti/consumatori specifici in base alle loro caratteristiche e ai loro interessi personali stimati. Essa comporta due importanti vantaggi per gli inserzionisti, rispetto alla pubblicità non mirata: le pubblicità target possono essere potenzialmente visualizzate solo dai consumatori interessati al prodotto o al servizio (in questo modo, si riduce la pubblicità dispendiosa) e il contenuto dell'annuncio può essere customizzato. Poiché si basa sugli interessi di un consumatore, un annuncio mirato comporta un'alta probabilità di una buona corrispondenza e tende quindi ad aumentare l'efficacia della pubblicità stessa in termini di acquisto del bene o servizio (Favretto, 2020).

[41] La materia del contendere è riassunta nella nota di Casalini (2021), che ne riporta gli elementi essenziali. Con sentenza n. 2631 del 29 marzo 2021, il Consiglio di Stato respingeva il ricorso presentato dalla società *Facebook Ireland Limited* avverso il provvedimento dell'Autorità Garante della Concorrenza e del Mercato (AGCM), adottato il 29 novembre 2018 (AGCM, Provv. n. 27432 del 29 novembre 2018), già impugnato di fronte al TAR Lazio (TAR Lazio, sez. I, 10 gennaio 2020, n. 260). Nello specifico, l'AGCM contestava alle società *Facebook Inc. e Facebook Ireland Limited* due distinte pratiche commerciali scorrette, in violazione degli artt. 20, 21, 22, 24 e 25 del Codice del consumo: per un verso, pratiche qualificabili come ingannevoli, in quanto la piattaforma digitale non informerebbe adeguatamente e immediatamente l'utente, in fase di attivazione dell'account, dell'attività di raccolta e utilizzo, per finalità informative e/o commerciali, dei dati che egli cede, rendendolo edotto della sola gratuità della fruizione del servizio, così da indurlo ad assumere una decisione che non avrebbe altrimenti preso (registrazione e permanenza sulla piattaforma)**;** per altro verso, pratiche qualificabili come aggressive, consistenti nell'esercizio di un indebito condizionamento nei confronti dei consumatori registrati, i quali, in cambio dell'utilizzo del social network verrebbero indotti ad autorizzare Facebook/terzi alla raccolta e all'utilizzo, per finalità informative e/o commerciali, dei dati che li riguardano in modo inconsapevole e automatico, tramite un sistema di preselezione del consenso alla cessione e utilizzo dei dati. La piattaforma, tuttavia, a seguito della pronuncia del TAR Lazio, confermativa del provvedimento dell'AGCM, si adeguava solo parzialmente al provvedimento, sì eliminando il *claim* sul carattere gratuito del servizio offerto dalla *homepage* ("Iscriviti. È gratis e lo sarà per sempre"), ma senza fornire indicazioni chiare circa l'uso commerciale dei dati degli utenti. L'organo giudicante rileva come la patrimonializzazione del dato personale, che nel caso di specie avviene inconsapevolmente, costituisca il risultato dell'intervento delle società attraverso la messa a disposizione/cessione dei dati degli utenti (e della relativa profilazione) a fini commerciali.

fattispecie riguardante il social network *Facebook* e l'uso di dati personali (Casalini, 2021). Del resto, la stessa Corte di Giustizia dell'Unione Europea [42] ha espressamente riconosciuto la natura economica dell'attività posta in essere da un soggetto digitale che offre un bene o un servizio a prezzo apparentemente nullo. Invero, la presunta mancanza del prezzo viene sopperita con gli introiti ottenuti tramite la pubblicità mirata, per la quale un ruolo cruciale è assolto proprio dai dati personali degli utenti stessi (Cerrone, 2020).

Con specifico riferimento al caso di *Facebook*, l'obbligo di chiarezza sancito dal c.d. Codice del consumo[43] non risultava rispettato dalla suddetta piattaforma, atteso che le informazioni rese all'utente al primo contatto - lungi dal contenere gli elementi essenziali per comprendere condizioni e limiti delle conseguenze che, a fronte della gratuità dei servizi offerti, sarebbero derivati dalla profilazione, in termini di indefinibilità dei soggetti che utilizzeranno i dati personali messi a disposizione e del tipo di utilizzo commerciale connesso - lasciavano supporre che fosse possibile ottenere immediatamente e facilmente, ma soprattutto "gratuitamente" (e per tutto il periodo in cui l'utente avrebbe mantenuto l'iscrizione in piattaforma), il vantaggio derivante dalla fruizione dei servizi tipici di un social network senza oneri economici, omettendo di comunicare che, invece, ciò sarebbe avvenuto (e si sarebbe mantenuto) solo se (e fino a quando) i dati sarebbero stati resi disponibili a soggetti commerciali non definibili anticipatamente ed operanti in settori anch'essi non precedentemente indicati, per finalità di uso commerciale e di diffusione pubblicitaria. Tanto basta a integrare gli estremi della pratica ingannevole, in quanto nel contesto del messaggio iniziale non si attribuisce adeguato risalto alle suindicate conseguenze (cfr. Cons. St. 29 marzo 2021, n. 2631). Altrimenti detto, la condotta della Big Tech va inquadrata nell'ambito delle pratiche c.d. ingannevoli, giacché *Facebook* avrebbe indotto in errore gli utenti/consumatori, alterandone il processo decisionale. Nello specifico, il Garante ha spiegato che, inconsapevolmente e in maniera automatica, tramite un sistema di preselezione del consenso alla cessione[44] e utilizzo dei dati, questi vengono trasferiti a terzi operatori. L'Antitrust ha altresì rilevato che, per evitare di subire limitazioni nell'utilizzo del servizio, gli utenti sarebbero stati indotti a mantenere attivo il trasferimento e l'uso dei propri dati (Cerrone, 2022; Casalini, 2021).

L'agglomerato di dati raccolti dalla piattaforma digitale può divenire oggetto di ulteriore monetizzazione attraverso la cessione a soggetti terzi - in alcuni casi, in palese violazione delle norme in materia di Data Protection, in altri, facendo un uso dei dati personali in maniera eccedente il consenso acquisito, tale da configurare illeciti in tema di diritti e libertà individuali (Franca, 2021). A tal proposito, potrebbe

---

[42] Corte di Giustizia dell'Unione Europea, 26 aprile 1988, Bond van Adverteerders e altri c. Paesi Bassi, C-352/85, punti 16-17.

[43] Il c.d. Codice del consumo (D.Lgs. n. 206/2005 e successive modifiche ed integrazioni) riconosce ai consumatori ed agli utenti, come fondamentali, i diritti *"(…) ad una adeguata informazione e ad una corretta pubblicità"*, *"(…) all'esercizio delle pratiche commerciali secondo principi di buona fede, correttezza e lealtà"* e *"(…) alla correttezza, alla trasparenza ed all'equità nei rapporti contrattuali"*(art. 1). Stabilisce, inoltre, che *"(…) le informazioni al consumatore, da chiunque provengano, devono essere adeguate alla tecnica di comunicazione impiegata ed espresse in modo chiaro e comprensibile, tenuto anche conto delle modalità di conclusione del contratto o delle caratteristiche del settore, tali da assicurare la consapevolezza del consumatore"* (art. 5). L'obbligo di chiarezza gravante sul professionista deve essere da costui assolto sin dal primo contatto, attraverso il quale debbono essere messi a disposizione del consumatore gli elementi essenziali per un'immediata percezione della offerta pubblicizzata.

[44] Merita sottolineare la qualificazione che l'AGCM ha dato della cessione dei dati, perché intesa come *"(…) contro-prestazione del servizio offerto dal social network, in quanto dotati di valore commerciale"* (Provvedimento n. 27432/2018 § 18).

rappresentare un valido correttivo quello delineato dal Garante Privacy a carico dei titolari del trattamento, consistente nell'adozione di misure preventive e processi interni volti a commisurare il rischio gravante sui diritti degli interessati: altrimenti detto, la valutazione d'impatto del trattamento sulla protezione dei dati personali, così come prevista dall'art. 35 GDPR. Ciò con particolare riguardo ai casi in cui il trattamento comporti *"(…) una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche"*[45] (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020).

Sebbene in ambito europeo il riconoscimento della dimensione (anche) patrimoniale del dato personale risulti fortemente osteggiato - se quest'ultimo venisse inteso come "corrispettivo non pecuniario", verrebbe svilito del proprio senso, del proprio essere diritto fondamentale della persona[46], una dottrina ancora minoritaria propende per il superamento della tradizionale concezione del dato personale (Cuffaro, D'Orazio & Ricciuto, 2019) – inteso quale attributo della persona, appunto - che si incentra esclusivamente sulla personalità del soggetto a cui il dato stesso si riferisce (Parenzo, 2021). Senonché, dal riconoscimento della dimensione economica dei dati potrebbe facilmente discendere il rispetto (anche nell'ambito di dette transazioni commerciali) degli obblighi di chiarezza, completezza, e non ingannevolezza delle informazioni, previsti dalla legislazione a tutela del consumatore/fruitore di un qualsivoglia servizio, anche digitale, con conseguente ampliamento della sfera applicativa della Data Protection Law. Ciò che potrebbe, peraltro, contribuire a mitigare la presenza di asimmetrie informative a danno dei fruitori dei servizi online – il *"soggetto debole" che merita protezione in questa nuova realtà digitale"* (Cerrone, 2022, 13) - stante il ruolo svolto dalle piattaforme online, intermediari che creano interazioni commerciali tra due categorie diverse di soggetti economici, utenti e fornitori (es. providers), collocati, questi ultimi, in una posizione di indiscutibile superiorità.

D'altronde, la dimensione economica dei dati personali è stata recentemente riconosciuta anche dalla giurisprudenza amministrativa italiana - che ha evidenziato come tali informazioni costituiscano *"(…) un asset disponibile in senso negoziale, suscettibile di sfruttamento economico e, quindi, idoneo ad assurgere alla funzione di controprestazione in senso tecnico di un contratto"* (TAR Lazio 10 gennaio 2020 n. 260) - e, prima ancora, dall'Autorità Garante della Concorrenza e del Mercato la quale, riconoscendo il valore economico dei dati degli utenti dei social media, ha ritenuto configurabile un rapporto di consumo tra il professionista e l'utente (Provvedimento n. 26596, 11 maggio 2017). Non ultima, la Commissione Europea negli "Orientamenti per l'attuazione/applicazione della direttiva 2005/29/CE relativa alle pratiche commerciali sleali" del 25 maggio

[45] Cfr. art. 35 GDPR, par. 3, lett. a)

[46] In questi termini si è espresso il Garante Europeo per la Protezione dei Dati Personali - Opinion n. 4/2017 "On the Proposal for a Directive on certain aspects concerning contracts for the supply of digital content", nella quale si dissuadono gli attori istituzionali dal considerare i dati personali come controprestazione: *"(…) personal data cannot be compared to a price, or money. Personal information is related to a fundamental right and cannot be considered as a commodity"*. Nello stesso senso, anche il Comitato Europeo per la Protezione dei Dati (European Data Protection Board, "EDPB") che, nelle Linee Guida n. 2/2019 concernenti il trattamento dei dati personali svolto ex art. 6, par. 1 lett b), GDPR, ha preliminarmente ricordato che il diritto alla protezione dei dati personali è un diritto fondamentale garantito dall'art. 8 della Carta Europea dei Diritti Fondamentali, sottolineando che i dati personali non possono essere considerati merci commerciabili; pertanto, gli interessati possono acconsentire al trattamento dei propri dati, senza tuttavia poterne disporre, in quanto loro diritti fondamentali, come riporta Cerrone (2022).

2016, ha espressamente asserito che *"(…) i dati personali, le preferenze dei consumatori e altri contenuti generati dagli utenti hanno un valore economico de facto e vengono venduti a terzi"*. Siffatto orientamento giurisprudenziale si allinea all'indirizzo recentemente condiviso dalle altre Autorità europee a tutela della concorrenza nell'ambito dei mercati digitali: è noto il massiccio ricorso, da parte delle imprese, a colossi digitali, i c.d. GAFAM20 (definiti anche "guardiani" o Gatekeepers della rete), per "agganciare" i consumatori, nei cui confronti i Garanti della concorrenza hanno intrapreso un'intensa attività di indagine, sfociata nell'irrogazione di sanzioni piuttosto pesanti (Cerrone, 2022).

Gli outcomes dell'attività di indagine *IC53 – Big Data* avviata dalle tre Autorità Garanti nazionali , sono stati tradotti in apposite Linee Guida, articolate in 11 punti, in cui si auspica la promozione di una policy unica e trasparente in materia di estrazione, accessibilità e utilizzo dei dati per la creazione di un mercato unico digitale; la riduzione delle asimmetrie informative tra utenti e operatori digitali nella fase di raccolta dei dati, nonché tra le piattaforme digitali e gli operatori che si avvalgono di queste ultime; il potenziamento delle tutele a favore dell'utente, mediante l'incremento della trasparenza delle informazioni a cui ha accesso e la qualità dei servizi, al fine di preservare il principio della libera concorrenza sul mercato (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2019). Non vi è dubbio che la natura massiva delle operazioni di trattamento automatizzato di ingenti quantità di dati - tra i quali possono essere ricompresi anche quelli di natura personale, nell'accezione fornita dall'art. 4 del Regolamento (UE) 2016/679 e successive modifiche[47] – ponga quesiti giuridicamente rilevanti, soprattutto per quanto attiene al processo di acquisizione dei dati stessi[48]. Com'è noto, infatti, la fase di raccolta dei Big Data prende avvio con la generazione di essi, che si realizza principalmente nell'ambito di attività compiute dagli utenti in un contesto informatizzato (IoT - Internet of Things). Nello IoT, infatti, in cui tutti i contenuti media sono resi disponibili in formato digitale e gran parte delle attività economiche e sociali vengono migrate in rete, sono le attività degli utenti (sia di tipo online che offline) a rappresentare i principali generatori degli enormi volumi di dati[49].

Il documento congiunto del 2019 è confluito in un elaborato finale, rilasciato nel mese di febbraio 2020, nel quale le tre Autorità coinvolte hanno riportato i risultati e le conclusioni definitive dell'indagine triennale (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020). Prendendo le mosse dalle considerazioni formulate dal Garante Privacy italiano, le attività connesse all'utilizzo dei Big Data sembrano evidenziare chiari profili di contrasto con aspetti fondamentali della

[47] Sul piano giuridico, assume rilevanza dirimente il carattere dei dati oggetto di elaborazione che, se di natura personale, risultano assoggettati ad uno specifico regime di protezione nell'ambito del quadro normativo recentemente definito a livello europeo, a cui concorrono sia il GDPR (General Data Protection Regulation), sia regole speciali per le attività *online*, individuate nella direttiva 2002/58/CE, relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche, e nella direttiva 2009/136/CE17. Qualora si tratti di dati aventi natura non personale, troverà invece applicazione il Regolamento (UE) 2018/1807 del 14 novembre 2018, inerente al quadro normativo applicabile alla libera circolazione dei dati non personali nell'Unione Europea (Piretti, 2020).

[48] Così la Risoluzione del Parlamento Europeo del 14 marzo 2017 sulle implicazioni dei *Big Data* per i diritti fondamentali: privacy, protezione dei dati, non discriminazione, sicurezza e attività di contrasto [2016/2225(INI)].

[49] Nel processo di acquisizione dei dati possono intervenire anche i c.d. data broker, ossia soggetti che aggregano dati da diverse fonti (principalmente siti internet) e li organizzano per metterli a disposizione di soggetti terzi. Tali intermediari, operando contemporaneamente su molteplici siti, consentono di aumentare significativamente l'ampiezza della raccolta dati (Pennasilico, 2018).

disciplina di protezione dei dati. Ciò principalmente con riferimento ai principi di liceità e correttezza nel trattamento – aspetto, quest'ultimo, implicante un'effettiva consapevolezza degli interessati circa le operazioni connesse all'utilizzo dei dati personali[50] – nonché al principio di finalità, posto che l'istituto del Data Protection si sostanzia, anzitutto, nel potere dell'interessato di esercitare il controllo in merito all'uso dei dati a sé riferiti, anche in relazione ai fini per i quali i medesimi vengono trattati. Com'è noto, qualora si tratti di dati personali, il GDPR ne assoggetta le attività di raccolta e utilizzo alla richiesta di consenso dell'interessato (o al ricorrere di una delle condizioni previste dall'art. 6); il trattamento dovrà informarsi ai principi di liceità, correttezza e trasparenza, essere compiuto per finalità determinate, esplicite e legittime. I dati raccolti, inoltre, dovranno essere adeguati, pertinenti e limitati a quanto necessario, avuto riguardo alle finalità per le quali gli stessi sono stati raccolti, in ossequio al c.d. principio di minimizzazione[51]. Tale operazione, tuttavia, per quanto apparentemente lineare, difficilmente si concilia con l'acquisizione di enormi quantitativi di dati, dovendosi concretamente determinare, di volta in volta, il concreto utilizzo che dei dati verrà effettuato, al fine di limitarne la raccolta a ciò che si renda necessario per fornire il servizio richiesto (Piretti, 2020).

Un plausibile approccio normativo – questa la proposta avanzata dal Garante Privacy nazionale nelle "Conclusioni" contenute nell'elaborato finale del 2020 - potrebbe essere quello di "attagliare" le privacy policies alle caratteristiche ed esigenze degli utenti/fruitori, anche valutandone la proporzionalità rispetto alla tipologia di servizi richiesti, così da implementarne la consapevolezza in materia di consenso al trattamento dei proprio corredo di dati [52]. Inoltre, stante la progressiva incertezza del confine tra dati personali e non personali, una regolamentazione orientata ad una protezione dei dati di carattere generale, piuttosto che settoriale e/o per categorie, potrebbe rivelarsi maggiormente efficace (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020). Studiosi di settore (Piretti, 2020) ritengono che uno strumentario normativo così impostato potrebbe offrire un valido contributo alla riduzione dell'asimmetria informativa esistente tra titolare dei dati e titolare del trattamento, garantendo un'adeguata informazione circa le finalità della raccolta e dell'utilizzo dei medesimi, così da consentire un esercizio consapevole dei propri diritti in materia di Data Protection e, in ultima analisi, l'adozione di oculate scelte di consumo da parte dell'utente finale (Alvisi, 2019).

L'analisi del fenomeno dei Big Data nella cornice italiana (e l'emergere delle relative criticità) impone una cooperazione rafforzata e un'interlocuzione delle Agenzie Garanti con altri soggetti istituzionali – primi tra tutti, le Autorità indipendenti di settore, cui sono rimessi poteri di vigilanza e di regolamentazione di specifici settori (assicurativo, bancario, finanziario, energetico ecc.) – così come la previsione di

[50] Studi empirici hanno evidenziato come gli utenti non prestino diligente attenzione a lunghe informative sulla privacy, con conseguenti forti asimmetrie informative riguardo alle transazioni economiche cui le informative suddette accedono (Piretti, 2020).

[51] Nello specifico, detto principio impone che, qualora il titolare intenda raccogliere dati ulteriori rispetto a quelli in suo possesso, o trattare i dati per una finalità diversa rispetto a quella comunicata, dovrà richiedere il relativo consenso all'interessato.

[52] A titolo esemplificativo, l'interim report dell'Indagine curato dall'AGCOM evidenzia come, in molteplici occasioni, l'utente non sia in grado di comprendere appieno congruità, coerenza e proporzionalità dei "permessi" di accesso ai suoi dati, richiesti dall'applicazione in corso di installazione sul proprio smartphone, rispetto agli usi cui è destinata l'applicazione stessa (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020).

profili professionali (c.d. data scientist) in grado di operare nel contesto dei Big Data, soprattutto presso le Autorità di controllo, per assicurare la qualità dell'attività di trattamento. In ogni caso, le competenze di tali figure non possono prescindere da un'adeguata considerazione dei profili etici e giuridici - con specifico riguardo all'attuale disciplina di protezione dei dati personali - che tali trattamenti implicano (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020).

Appare, dunque, di tutta evidenza come l'utilizzo dei Big Data investa ambiti di competenza e di criticità di diversa natura e come le sfide, poste dallo sviluppo dell'economia digitale, richiedano l'impiego sinergico degli strumenti a tutela della privacy, del consumatore e della concorrenza, in un'ottica di necessaria interdisciplinarità. In aggiunta, è necessario considerare la duplice dimensione, personale ed economica, dei dati personali - quali informazioni afferenti alla sfera intima delle persone, ma al contempo, moneta di scambio nell'ambito del mercato digitale (Cerrone, 2022) – per adottare soluzioni legislative e regolamentari all'interno di un contesto in rapida e costante evoluzione. Le conclusioni e gli interventi auspicati da parte della tre Autorità Garanti nazionali, contenute nella Relazione finale (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020), sono di seguito riportate, così come sistematizzate da Piretti (2020).

In primo luogo, è opportuno che Governo e Parlamento nazionali si interroghino circa la necessità di promuovere un appropriato quadro normativo in grado di affrontare la questione della piena ed effettiva trasparenza nell'uso delle informazioni personali, tanto nei confronti dei singoli utenti/fruitori di servizi quanto riguardo alla collettività. In secondo luogo, la transnazionalità del fenomeno in esame impone un rafforzamento della cooperazione internazionale in materia di Policies condivise per il governo dei Big Data. Ciò presuppone – ed è questo il terzo punto saliente - la promozione di una Policy nazionale unica e trasparente in tema di estrazione, accessibilità e utilizzo dei dati, al fine di determinare politiche pubbliche a vantaggio di imprese e cittadini. A tale impegno istituzionale dovrà fare seguito un opportuno coordinamento tra detta Policy e le strategie europee già delineate, in vista della costituzione di un mercato unico digitale di ampiezza comunitaria.

Ulteriore obiettivo da perseguire sarà quello della riduzione delle asimmetrie informative tra utenti/fruitori di servizi e operatori digitali, già nella fase di raccolta dei dati (Prestipino, 2017, Zuboff, 2019), nonché tra le grandi piattaforme digitali e gli altri operatori che di tali piattaforme si avvalgono a scopi commerciali – pratica che Bellini (2019) definisce "Data Monetization", con evidente richiamo al noto fenomeno dell'Information brokerage (Favretto, 2020; Federal Trade Commission, 2014; Binns R., Lyngs U., Van Kleek et al., 2018; Pennasilico, 2018). Dalle enormi potenzialità informative scaturenti dal progresso tecnologico in campo informatico discende la nota equazione tra informazione e potere (D'Acquisto & Naldi, 2017; Sun Tzu, 1997; Zanasi, 2003), strettamente connessa alla *"(…) disponibilità di grandi masse di dati concentrate nelle mani di pochi soggetti che, a vario titolo, li raccolgono ed elaborano o che possiedono le risorse per acquisirli in forma grezza o elaborata. Ad avvalersene principalmente sono, infatti, i grandi gruppi societari, i proprietari delle grandi piattaforme social o commerciali e quelli d'infrastrutture per il web, gli hosting provider o gli apparati pubblici legittimati ad acquisire dati sulla popolazione",* come la recente dottrina non ha mancato di sottolineare (Antares Fumagalli, 2018, 134).

Superfluo precisare come la distribuzione asimmetrica di capacità informativa si manifesti sia sul versante delle Aziende operanti sui Digital Markets – qui derivante dalla inaccessibilità ai più delle tecnologie e somme di denaro necessarie alla concreta attuazione di pratiche di raccolta e gestione

dei BD, con conseguente creazione di un regime monopolistico o, quantomeno oligopolistico, sul mercato globale (Di Porto, 2016 e bibliografia ivi citata; Favretto, 2020; Franzini, 2019; Maggiolino, 2016; Orefice, 2016; Pitruzzella, 2016; Vallone, 2022) – sia sul versante del consumatore, per giunta in settori tradizionalmente critici sul piano della tutela del c.d. contraente debole, primo fra tutti quello relativo al governo del rischio nei contratti bancari e assicurativi.

Nel rapporto contrattuale di natura bancaria o assicurativa, in verità, l'asimmetria informativa correlata all'uso dei Big Data presenta una matrice di criticità del tutto peculiare, poiché implicante fenomeni di selezione avversa[53], ovvero di sbilanciamento di informazione tra due controparti (in questo caso, cliente – Istituto bancario/assicurativo, a beneficio del primo e a svantaggio di quest'ultimo): la questione, piuttosto complessa, è affrontata da Prestipino (2017), in un saggio che delinea i nuovi scenari di rischio per la protezione dei dati personali in contesti tecnologici c.d. data intensive. L'originalità dell'approccio adottato dall'Autrice merita un approfondimento anche in questa sede.

Orbene, l'originario squilibrio informativo - che potrebbe indurre l'operatore economico ad un'assunzione di rischio non consapevole, perché falsata da un corredo di dati incompleto - può essere riequilibrato mediante la somministrazione di informazione c.d. derivata, ossia quella proveniente dall'attività di raccolta e di analisi dei Big Data, con conseguente ribaltamento delle posizioni. Tuttavia, quando ciò accade, l'alterazione dell'asimmetria informativa si rivela difficilmente ripristinabile per l'assenza o l'inadeguatezza di risorse da parte del cliente, con conseguente dispersione del controllo sui propri dati. Vero è che *"(…) in tale contesto il titolare del trattamento algoritmico è la parte più informata potendo sfruttare sia una originaria ridondanza di dati che l'esclusiva prerogativa di elaborarne e ampliarne il significato con mezzi e risorse di gran lunga superiori a quelle disponibili al soggetto interessato, il quale - in scenari informazionali data intensive e di trattamento basato sullo schema Big data spesso non si accorge nemmeno di questa prevalenza informativa del titolare; questi può assumere decisioni di cui le persone non sono consapevoli, oppure clusterizzare le deduzioni ignorandone le diverse valenze semantiche al variare dei soggetti interessati"* (Prestipino, 2017, 9-10). Nei settori della gestione del rischio, i Big Data hanno così trasformato il dato personale in asset strategico tramite il quale abbattere l'asimmetria informativa (originariamente a vantaggio del titolare dei dati/cliente) con evidenti ripercussioni sul piano delle tutele fornite ai soggetti contraenti e sulla possibilità di controllo dei dati sensibili da parte della compagnia bancaria o assicuratrice (Giorgi & De Masi, 2019; Greggio, 2022; Ioannoni Fiore, 2022).

Simili profili di criticità sono stati parimenti ravvisati nel settore giuslavoristico e previdenziale, in tema di correlazione tra infortuni/malattie professionali e contesto ambientale, comportamenti aziendali generatori di stress, stile di vita del lavoro ed azioni di prevenzione - tematiche nell'ambito delle quali si è sviluppata una vivace riflessione giuridica ormai quinquennale (Rota, 2017 e bibliografia ivi citata). Nel mirino dell'indagine dei giuslavoristi, l'incidenza delle informazioni sensibili nella gestione del rapporto individuale di lavoro, nell'ambito del quale assumono rilievo, per un verso, le analisi a campione combinate con i Big Data – tools forieri di informazioni

[53] La selezione avversa è una problematica connaturata al divario tra le informazioni possedute dall'azienda e quelle possedute dal cliente. Studiata inizialmente da Akerlof negli anni Settanta, trova attuale applicazione soprattutto nel mercato bancario e assicurativo - ambiti nei quali la maggiore informazione (di cui Ente assicuratore o Istituto di credito non dispone) consente al cliente una più conveniente operazione di stipula (Prestipino, 2017)

dettagliate su ciascun lavoratore (in particolare, sui cicli di performance), capaci di fornire, al contempo, elementi utili per il recruiting e la gestione delle risorse umane (Weiss, Khoshgoftaar, & Wang 2016) - e, per altro verso, i meccanismi di profilazione algoritmica, ritenuti potenziali responsabili di nuove prassi discriminatorie, attentati alle libertà fondamentali e all'autodeterminazione informativa della persona (Tullini, 2016 e bibliografia ivi citata, per quanto concerne il dibattito sul tema maturato all'estero).

Ancora, secondo le tre autorità amministrative indipendenti, l'introduzione di nuovi strumenti per la promozione del pluralismo online, la trasparenza nella selezione dei contenuti nonché la consapevolezza degli utenti circa le informazioni ottenute in rete dovrà considerarsi funzionale al perseguimento dell'obiettivo di tutela del benessere del consumatore, da conseguirsi mediante l'applicazione degli istituti giuridici propri della c.d. normativa antitrust, con estensione anche alla valutazione di obiettivi relativi alla qualità dei servizi, all'innovazione e all'equità dei servizi digitali. Orbene, tali istanze sono già da tempo presenti in dottrina - dove si è assistito ad una crescente richiesta di intervento antitrust, diretto a neutralizzare i rischi connessi alla raccolta e all'utilizzo dei Big Data relativi ad informazioni personali - tanto di ipotizzare un'unione tra la disciplina a tutela della concorrenza e quella a tutela del consumatore, così da assicurare che le eventuali violazioni della privacy siano affrontate con lo strumentario normativo proprio dell'antitrust (Colangelo, 2016 e bibliografia ivi citata).

Le Autorithies nazionali (2020) ribadiscono, inoltre, la necessità di tutela dei diritti personali dell'utente/fruitore di servizi online – primo fra tutti, il diritto alla riservatezza – la quale dovrà realizzarsi anche attraverso la verifica preventiva (ossia, anteriore al trattamento) di natura e proprietà dei dati, al fine di valutare la possibile re-identificazione della persona a partire da dati individuali anonimizzati, La dottrina italiana tende, innanzitutto, a fare chiarezza sul punto, posto che, da un corretto inquadramento della tematica discende un'appropriata applicazione della disciplina dettata dal GDPR (Martorana & Pinelli, 2021; Iaselli, 2018).

Il principale chiarimento attiene alla differenza tra "anonimizzazione" e "pseudonimizzazione" – concetti spesso confusi nella pratica – rispetto ai quali il discrimine è rappresentato dalla possibilità di re-identificazione del titolare dei dati. Così, i dati pseudonimizzati sono da considerarsi, in linea di principio, ancora dati personali, poiché è possibile collegarli ad una persona fisica con l'impiego di informazioni aggiuntive [54] (Iaselli, 2018). Al contrario, i dati anonimizzati non sono (più) suscettibili di associazione rispetto ad individui specifici, poiché sottoposti a processi di anonimizzazione, appunto. Tuttavia, la difficoltà nella previsione di tecniche idonee – il GDPR, invero, non prescrive alcuna misura tecnica per l'anonimizzazione, rimettendo quindi ai singoli responsabili del trattamento la scelta di un procedura sufficientemente sicura ed efficace - si è tradotta in processi di anonimizzazione incompleti, che hanno condotto, non di rado, alla re-identificazione dei titolari dei dati medesimi (Martorana & Pinelli, 2021).

Ciò ha reso necessario il recentissimo intervento dello l'European Data Protection Board (EPDB) che, di comune accordo con la Agencia Española de Protección de Datos

[54] Tipico strumento di pseudonimizzazione è rappresentato dalla crittografia, nella misura in cui rende i dati attribuibili a una persona identificata o identificabile solo con l'uso della chiave per la decrittografia: in tal caso, la conoscenza di tale chiave rappresenta l'informazione aggiuntiva necessaria per poter attribuire i dati alle persone a cui si riferiscono (Antonielli, 2021; Caputo & Ferorelli, 2023).

(AEPD), ha rilasciato un documento congiunto, contenente un elenco dei "malintesi" maggiormente diffusi, noto come *AEPD-EPDS Joint Paper – 10 Misunderstandings about Machine Learning* (2022), oltre ad una serie di pubblicazioni inerenti a Best Practices per l'implementazione di tecnologie conformi alle disposizioni in materia di Data Protection Law (Markopoulou, Papakonstantinou, & De Hert, 2019). La questione non è meramente speculativa, posto che, ai sensi del "Considerando 26" del GDPR, i dati interamente "anonimizzati" non soddisfano i criteri necessari per qualificarsi come dati personali e, pertanto, non sono assoggettati alle medesime restrizioni sul trattamento dei dati personali tout court (Martorana & Pinelli, 2021). Stabilire se un dato è qualificabile o meno come dato personale, infatti, è derimente ai fini della stessa (dis)applicabilità tanto del Regolamento (UE) 2018/1725 (EUDPR) [55] quanto del GDPR che, giova ribadirlo, non trovano applicazione al trattamento di informazioni anonime, mentre sono pienamente applicabili al trattamento di dati pseudonimizzati (Caputo & Ferorelli, 2023).

Sui concetti di "anonimizzazione" e "pseudonimizzazione" si è, peraltro, espresso anche il Tribunale della Corte di Giustizia UE, con sentenza del 26 aprile 2023, nell'ambito del trattamento di dati che vedono principalmente coinvolte due diverse organizzazioni: una in qualità di titolare e mittente di un insieme di dati, l'altra in qualità di soggetto ricevente tali dati. La recentissima pronuncia, che si è presentata, per taluni versi, innovativa, ha fornito un'interpretazione inedita del *thema decidendum*, suscitando perplessità e aprendo un vivace dibattito in seno alla comunità degli esperti di settore. Dal tenore letterale della sentenza, infatti, il Tribunale sembrerebbe introdurre un'accezione relativistica della "personalità" dei dati: altrimenti detto, *"(…) dati che per il mittente sono dati personali non per forza devono considerarsi tali anche per il soggetto che li riceve, con la conseguenza che, ove non lo fossero, il titolare del trattamento non sarebbe tenuto a informare gli interessati della circostanza che i loro dati saranno trattati dal soggetto ricevente in qualità di destinatario"* (Caputo & Ferorelli, 2023).

Ciò non pare essere perfettamente aderente ad un'interpretazione letterale della norma, posto che, dalla lettura del "Considerando 26" del GDPR, il carattere anonimo di un dato non sembra doversi valutare in relazione allo specifico soggetto coinvolto in una specifica fase del trattamento, bensì nel complesso del processo trattamentale che coinvolge il set di dati in esame. Del resto, la gestione dei Big Data in una prospettiva orientata alla tutela della privacy degli individui – anche e soprattutto mediante un corretto inquadramento della natura personale del dato - rappresenta il punto di equilibrio dei contrapposti interessi del diritto alla riservatezza, da un lato, e dello sviluppo dell'economia digitale dall'altro (Mastrelia, 2018).

Le considerazioni che precedono impongono, nell'analisi conclusiva del 2020, una riforma del controllo delle operazioni di concentrazione oligopolistica sul mercato online, al fine di incrementare l'efficacia dell'intervento delle Autorità Garanti teso a ristabilire il rispetto dei principi concorrenziali. A tal proposito, si reputa opportuno il rafforzamento dei poteri di acquisizione delle informazioni da parte di AGCM ed AGCOM al di fuori dei procedimenti istruttori, unitamente ad un incremento del massimo edittale per le sanzioni, così da garantire il rispetto della normativa a tutela del consumatore. Per il conseguimento di detto obiettivo, l'istituzione di un "coordinamento permanente" tra le tre Autorità Garanti nazionali potrebbe rivelarsi

[55] Regolamento sulla tutela delle persone fisiche in relazione al trattamento dei dati personali da parte delle istituzioni, degli organi e degli organismi dell'Unione e sulla libera circolazione di tali dati.

cruciale. La produzione dottrinale sul tema è ampia e variegata, come peraltro già evidenziato da Perrucci (2019), uno dei principali commentatori dei risultati preliminari della citata indagine conoscitiva, cui si rinvia per ulteriori approfondimenti.

Certo è che l'utilizzo di ingenti volumi di dati influisce, inevitabilmente, anche sulle dinamiche competitive e concorrenziali del mercato: i modelli di business fondati sui Big Data costituiscono, di fatto, un aspetto saliente di paradigmi economici e servizi digitali caratterizzati da elevati livelli di concentrazione e dalla presenza di operatori che detengono posizioni dominanti, le c.d. Big Tech (Favretto, 2020 e bibliografia ivi citata). Ne è un esempio concreto il potere di mercato che piattaforme digitali come *Google, Apple, Facebook, Amazon e Microsoft* hanno assunto, sia in ragione dei processi di globalizzazione, sia per il ruolo centrale che detti operatori rivestono nell'abilitazione delle interazioni e delle transazioni digitali (Piretti, 2020).

Le precedenti osservazioni contribuiscono a delineare uno scenario in cui il dato personale, lungi dal costituire oggetto esclusivo di un diritto fondamentale dell'individuo, diviene bene negoziabile e, in quanto tale, suscettibile di sfruttamento economico. Vero è che lo sviluppo esponenziale delle nuove tecnologie ha inciso significativamente sulla sfera dei diritti soggettivi, che si è espansa fino a prevedere la tutela di istanze pressoché inedite fino ad un decennio fa. Emblematica è l'estensione del diritto alla privacy, evolutosi nel diritto alla protezione dei dati personali, ovvero di quelle stesse entità che costituiscono il principale motore del sistema economico contemporaneo (Amato Mangiameli, 2022). La produzione scientifica internazionale sul tema è molto vasta e concorda sul ruolo prioritario che la gestione massiva dei personal data riveste nell'economia digitale 4.0, ossia quella della c.d. Quarta Rivoluzione Industriale[56] (Albergaria, & Jabbour, 2020; De Minico, 2019, Jayagopal & Basser, 2022; Kościelniak, & Puto, 2015; Mandelli, 2017; Novikov, 2020; Perrucci, 2019; Rolli & D'Ambrosio, 2022; Sawicki, 2016; Sedkaoui, & Khelfaoui, 2020; Tan, Ji, Lim, & Tseng, 2017; Vassakis, Petrakis, & Kopanakis, 2018; Wahyudi, Meilinda, & Khoirunisa, 2022 e bibliografie ivi citate).

Buchi, Cugno & Castagnoli (2019), in particolare, presentano un originale lavoro sul tema, che esplora la relazione causale tra Industry 4.0 e internazionalizzazione. I risultati della ricerca - che si avvale di una systematic literature review condotta sui data-base scientifici *WoS, Ebsco* e *Scopus* relativa a pubblicazioni scientifiche apparse tra il 2011 e il maggio 2019 - mostrano come l'Industry 4.0 riconfiguri l'ambiente dell'impresa, influenzando principalmente settori quali global value chain, global supply chain; localizzazione e fasi del processo produttivo; personalizzazione prodotto; relazioni con i principali stakeholder - fornitori, clienti e dipendenti, in primis. Il filone di ricerca è stato ulteriormente ampliato da altri Autori a livello internazionale, giunti a risultati del tutto sovrapponibili (Ahi, Sinkovics, Shildibekov, et al., 2022; Bettiol, Capestro, De Marchi, et al., 2020; Castagnoli, Büchi, Coeurderoy et al., 2022; Chiarvesio & Romanello, 2018; Ciaramitaro, 2022; Genovino, Caprino, & Salmista, 2020; Hoyer, Gunawan, & Reaiche, 2020; Morelli, Musso, Murmura, et al., 2022; Perna & Runfola, 2017; Strange, & Zucchella, 2017).

È doveroso precisare che i dati sono privi di valore intrinseco (ossia, quali raw data), dal momento che sono le informazioni in essi contenute ad assumere significato economicamente rilevante, e ciò nella misura

[56] È paradigmatico l'esempio riportato da Amato Mangiameli (2022), secondo cui il successo commerciale di un prodotto sulle piattaforme digitali può dipendere dall'ordine con cui *Google* (o un qualsiasi altro search engine diffuso tra gli internauti) posiziona i risultati delle ricerche online.

in cui costituiscano l'esito delle attività di organizzazione, gestione, filtraggio ed estrazione compiute per mezzo dell'Intelligenza Artificiale [57] (Rolli & D'Ambrosio, 2022). L'impiego del patrimonio informativo estrapolato da tale risorsa consente di *"(…) accrescere l'efficienza dei processi produttivi, migliorare le capacità decisionali, prevedere con maggiore accuratezza le tendenze attuali e future e, in conseguenza di ciò. rendere più mirate e precise le attività commerciali nell'individuazione del proprio* target *di mercato"* (Marcelli, 2021). È in ciò che si sostanzia il potenziale economico dei Big Data e, ovviamente, nella possibilità del loro sfruttamento a fini di lucro (AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali, 2020). Il crescente fenomeno del c.d. market for data - vero e proprio ecosistema digitale, all'interno del quale operano aziende di brokeraggio impegnate nella raccolta di grandi quantità di dati successivamente ceduti ad operatori interessati ad una migliore profilazione dei clienti – di detto potenziale rappresenta senz'altro la massima espressione (Abraham, Schneider, & vom Brocke, 2023; Bonneau, & Preibusch, 2010; Eichler, Gröger, Hoos, et al., 2022; Hayashi & Ohsawa, 2020; Javalgi, Martin, & Young, 2006; Ofori, Boakye, & Narteh, 2018; Schaub, 2018).

La valenza negoziale dei dati personali, e il suo conseguente tradursi nel corrispettivo per la fruizione di prestazioni solo apparentemente gratuite, costituisce il tema dominante della costruzione teorica di Zuboff, condensata nel suo saggio del 2019, dal titolo evocativo "The Age of Surveillance Capitalism", che guarda allo sviluppo delle aziende digitali, proprietarie di enormi volumi di dati, suggerendo come i loro modelli di affari rappresentino una nuova forma di accumulazione capitalista. La tesi centrale sostenuta dall'Autrice gravita attorno all'affermazione secondo la quale, nell'era dei Big Data, oggetto di scambio (e fonte di profitto) sui mercati digitali siano non solo le informazioni sulle persone – come già da tempo la comunità scientifica è concorde nel sostenere - bensì le persone stesse, con le loro esperienze di vita.

Secondo l'analisi critica della poderosa opera della Zuboff compiuta da Franzini (2019), nel c.d. capitalismo di sorveglianza le Aziende digitali che detengono l'oligopolio sui digital markets (Big Techs) si approprierebbero dei dati relativi ai comportamenti individuali – quelli online ma anche quelli offline – i quali, previa accurata elaborazione, sarebbero solo parzialmente impiegati per migliorare l'offerta di beni e servizi, dunque, a scopi socialmente utili. La parte residua, apparentemente inutilizzabile (c.d. metadati) confluirebbe nei "prodotti di previsione" per essere commerciati sui nuovi "mercati comportamentali a termine"[58]: ecco la fonte del "behavioural surplus", quella sorta di plusvalenza comportamentale che garantirebbe l'accumulo di immense ricchezze a beneficio dei principali operatori sul mercato, artefici indiscussi di questa nuova forma di capitalismo.

"The Surveillance Capitalism" sarebbe devastante soprattutto perché potenzialmente fautore dell'estinzione dell'umanità – qui da intendersi come modalità di pensiero, ragionamento e condotta propri dell'essere umano in quanto tale - di cui l'autonomia e la dignità sarebbero i tratti distintivi. Esso si nutrirebbe, dunque, dello sfruttamento non

[57] *"Intelligenza artificiale e big data, ovvero database che raccolgono enormi quantità di informazione di vario tipo (dalle immagini ai video, dai testi all'audio, dai like su Facebook alle transazioni monetarie) e che richiedono l'utilizzo di calcolatori di grande potenza per la raccolta di questi dati eterogenei e sterminati, come pure per l'individuazione di relazioni (collegamenti, connessioni) e per l'estrapolazione di previsioni"* (Amato Mangiameli, 2022, 97).

[58] Trattasi di mercati che forniscono incentivi alla raccolta di dati e profilazione degli utenti poiché a basso costo, di molto inferiore a quello degli studi psicologici alla base dei tentativi di persuasione dei consumatori risalenti a qualche decennio fa (Franzini, 2019).

solo (e non tanto) del lavoro umano, come nella tradizionale impostazione di Marx, bensì dell'esperienza umana nella sua complessità. Con esso si imporrebbe, giocoforza, una nuova forma di potere che la Zuboff definisce "strumentale" (*instrumentarian*) capace di influenzare il comportamento umano per generare profitto, la cui forza deriverebbe da un'architettura computazionale di dispositivi intelligenti, di cose (Internet of Things) e spazi tra loro connessi.

"Il capitalismo della sorveglianza va, dunque, combattuto non soltanto per ragioni antiche (è monopolistico, viola la privacy) ma anche, e soprattutto, perché riduce a merce i comportamenti umani e attraverso il loro commercio consente arricchimenti straordinari. Un capitalismo che non si accontenta " di automatizzare i flussi di informazioni su di noi, ma mira a automatizzare noi stessi" (Franzini, 2019). Si tratterebbe, insomma, di un meccanismo perverso di mercificazione dell'esperienza umana e, in ultima analisi, dell'essere umano nella sua interezza. Al netto di legittime osservazioni critiche in merito alla posizione "assolutista" della Zuboff (Franzini, 2019), vero è che lo scambio tra gratuità dell'accesso alla rete e appropriazione dei dati da parte delle Big Techs avviene con modalità spesso inidonee a garantire una scelta consapevole da parte dell'utente, come la citata letteratura ha evidenziato. In ogni caso, il costrutto teorico dell'Autrice ha stimolato una produzione scientifica particolarmente florida intorno alla sua opera (Anderson, 2020; Ball, 2019; Bongiovi, 2019; Borradaile & Reeves, 2020; Ellinger, 2020; Evangelista, 2019; Kapadia, 2020; Stanger, 2022; Zuboff, Möllers, Wood, et al., 2019), unanime nel sostenere che *"(…) un flusso costante di informazione possa divenire res in commercio"* (Rolli & D'Ambrosio, 2022, 785).

In sostanza, il dato personale acquista un valore sul mercato in quanto consente a società di servizi online e gestori di piattaforme digitali di compiere un'attività di profilazione, ovvero di tracciare un "profilo" della persona dell'utente/consumatore. Si tratta di un'operazione che richiede l'analisi di centinaia di fonti pubbliche, all'esito della quale i dati raccolti, e le informazioni da essi estrapolate, verranno utilizzate per pianificare campagne pubblicitarie mirate e, perciò, più efficienti. Per dimensionare il fenomeno, Pennasilico (2018) riporta le stime relative ai proventi pubblicitari del 2017 negli Stati Uniti (circa 83 miliardi di dollari) divisi per gli utenti Internet attivi in USA (circa 287 milioni): in tal caso il "valore di mercato" dell'utente medio si attesterebbe intorno ai 289.19 dollari annui. Superfluo aggiungere come i Data Brokers[59] si stiano confermando attori di primaria importanza nelle dinamiche della Digital Business Transformation (Di Falco 2022 e bibliografia ivi citata). Una conferma, a tal proposito, è stata fornita dalla stessa Commissione Europea, che ha recentemente dichiarato

[59] Benché la pratica del Data Brokerage non comporti problematiche sotto il profilo strettamente civilistico della disciplina di protezione dei dati personali – posto che i dati trattati sono "pubblici" o resi pubblicamente conoscibili dalle PA che li detengono nell'adempimento degli obblighi di trasparenza sulle stesse gravanti o, ancora, forniti direttamente dai titolari, ad esempio pubblicandoli sui social network – la circostanza che tali dati siano liberamente conoscibili non implica che siano anche liberamente riutilizzabili da chiunque e per qualsiasi scopo. Il riutilizzo dei dati, infatti, in attuazione del "principio di finalità" di cui all'art. 11 del d.lgs. 196/2003 ("Codice Privacy" italiano), non può essere consentito "in termini incompatibili" con gli scopi originari per i quali gli stessi sono resi accessibili pubblicamente. Le pratiche dei brokeraggio di dati possono, tuttavia, sollevare censure in termini di privacy: in genere, tali operatori raccolgono, manipolano e condividono informazioni sui consumatori senza interagire direttamente con loro. I consumatori, pertanto, sono in larga parte inconsapevoli del fatto che soggetti terzi sono impegnati in dette pratiche ed eventuali informative in merito all'uso dei dati personali non sempre risultano accessibili e/o comprensibili da parte del pubblico, nonostante gli obblighi di trasparenza imposti dalla normativa europea a partire dal 2010 (Favretto, 2020).

come il settore dei media e dei servizi di comunicazione digitali rappresenti un mercato innovativo ad elevato tasso di crescita, caratterizzato da frequenti ingressi da parte di nuovi operatori commerciali (Favretto, 2020). Il fenomeno dei data brokers rappresenta il chiaro esempio di come l'informazione sia denaro, Si comprende, allora, perché con sentenza del 10.1.2020, n. 261 il Tar Lazio abbia ritenuto ingannevole l'informativa fornita dal social network *Facebook*, con la quale si affermava la gratuità dell'iscrizione in fase di prima registrazione, posto che i dati personali sono suscettibili di sfruttamento economico e, conseguentemente, assumono valore commerciale.

Ebbene, una delle caratteristiche intrinseche dei dati, in quanto entità immateriali, è la c.d. non-rivalità: ciò significa che gli stessi possono essere oggetto di raccolta e impiego innumerevoli volte, senza che ciò ne pregiudichi il valore. Posto che la collazione di dati da parte di un'impresa non impedisce ad imprese concorrenti di fare altrettanto (Favretto, 2020), tale attività può facilmente tradursi in una raccolta sistematica di Personal Data a scopo commerciale. Ne ha rappresentato un valido esempio la profilazione di massa compiuta da *Cambridge Analytica*, la società inglese coinvolta in una procedura di collazionamento, analisi e cessione di dati personali a titolo oneroso, sorta a latere di un progetto di ricerca accademico successivamente degenerato in una pratica commerciale scorretta [60] . Inconsapevoli e senza possibilità di opporsi al trattamento dei loro dati, gli utenti erano divenuti, loro malgrado, merce in vendita, dotata di un valore di cui non condividevano in alcun modo il profitto ma che rischiava di minare alla base la loro weltanschauung (González, Yu, Figueroa, et al., 2019; Heawood, 2018; Hinds, Williams, & Joinson, 2020).

È noto, infatti, come i Big Data assumano un rilievo che è, al contempo, economico scientifico, politico e sociale, nascente dallo stretto legame tra gli stessi dati e la conoscenza che i soggetti si formano del mondo in cui vivono (Favretto, 2020). Nel caso specifico, la società inglese avrebbe creato un meccanismo di manipolazione e persuasione dell'elettorato (impegnato in una importante competizione elettorale oltreoceano), avvalendosi di una notissima piattaforma digitale come strumento di profilazione iniziale e influenza successiva. Interessante notare come *Cambridge Analytica* abbia affermato di aver sviluppato un sistema di "microtargeting comportamentale", ossia una pubblicità altamente personalizzata sul singolo utente, in grado di appellarsi alle sue emozioni per condizionarne la condotta. Non a caso, in seguito allo scalpore suscitato dalla vicenda, nella comunità scientifica è invalso l'uso dell'espressione "Data Ethics", un esplicito richiamo ai rischi di simili pratiche per i valori della libertà e della democrazia (Passani, 2018).

Dunque, oltre agli indiscutibili vantaggi – primo fra tutti, la possibilità di processare un'enorme mole di dati con un ridotto dispendio di tempo e di risorse (umane ed economiche) - l'odierno sviluppo tecnico-scientifico solleva molteplici interrogativi, soprattutto sul versante etico, mostrando matrici di criticità eterogenee, non ultime quelle attinenti alle minacce ai sistemi democratici e alle discriminazioni conseguenti alla (ormai acclarata) non-neutralità degli algoritmi impiegati nei processi di Big Data Analytics. L'interrogativo che anima gli studiosi nell'odierno contesto ipertecnologico e globalizzato è il seguente: *"(…) l'attività svolta dai sempre più sofisticati software può considerarsi neutrale, ovvero, quell'insieme di algoritmi e di big data, entro la cornice di una intelligenza artificiale capace di*

[60] Allo sterminato database di milioni di identità così ottenuto, la Società digitale ha applicato sofisticati modelli matematici per stilare un "profilo" degli utenti, mediante il quale era possibile non solamente ipotizzare (con estrema precisione) le idee politiche, ma anche individuare interessi, gusti, aderenza o meno a taluni valori e vicinanza (o lontananza) da temi importanti come ecologia, armi e integrazione.

riconoscere, classificare, ragionare, diagnosticare, agire, può essere ritenuto di per sé in ogni caso obiettivo?" (Amato Mangiameli, 2022, 96). La risposta a tale interrogativo fornita dalla comunità accademica sembra essere tutt'altro che rassicurante, tanto da avere sollecitato un recente intervento ad hoc da parte della Commissione Europea, tradottosi nella "Proposta di Regolamento del Parlamento Europeo e del Consiglio Europeo", che stabilisce regole armonizzate sull'Intelligenza Artificiale, peraltro modificando precedenti atti legislativi dell'Unione (COM/2021/206 final).

La questione, dai molteplici risvolti, affonda le sue radici nella (ben nota) pervasività delle tecnologie digitali, accompagnata dalla diffusione ubiquitaria dell'algoritmo, con effetti potenzialmente rivoluzionari tanto nell'ambito della vita quotidiana, quanto nella ricerca sociale. Non ci sarebbe da stupirsi, dunque, *"(…) se gli algoritmi siano diventati un argomento centrale tanto nel dibattito pubblico (Gilbert, 2018; Gillespie, 2014) che nella comunità accademica. Ne è una prova la recente pubblicazione di numeri di riviste scientifiche interamente dedicati alla tematica (Amoore, 2019; Beer, 2017; Boccia Artieri, Marinelli, 2018; Ziewitz, 2016) che, da punti di vista e con obiettivi differenti, hanno scandagliato i molteplici aspetti di quello che tecnicamente può essere definito come «un procedimento di calcolo esplicito e descrivibile con un numero finito di regole che conduce al risultato dopo un numero finito di operazioni, cioè di applicazioni delle regole» (Treccani, 2019).* La *semplicità di questa formula è però solo apparente (Barocas, Hood, Ziewitz, 2013; Seaver, 2017) e contrasta in maniera netta rispetto all'estrema complessità dell'impatto sociale degli algoritmi"* (Campo, Martella & Ciccarese, 2018, 8).

La stessa indagine circa le modalità attraverso le quali gli algoritmi operano appare tutt'altro che agevole, anche tra gli addetti ai lavori (Hargreaves, Agosti, Menasché, et al., 2019; Boccia Artieri, 2014), manifestandosi, per un verso, nell'intrinseca opacità – intesa come l'incapacità di comprendere le motivazioni sottese allo specifico risultato restituito dall'algoritmo, talvolta da parte degli stessi progettisti (Pedreschi, Giannotti, Guidotti et al., 2018 e bibliografia ivi citata) – e, per altro verso, con il sottrarsi delle dinamiche algoritmiche alla percezione diretta da parte dell'utente, tale per cui *"(…) è possibile osservarne il risultato finale, il prodotto del loro operare, che molto spesso non viene presentato come l'esito di un processo di selezione tra diverse possibilità, ma come un semplice dato di fatto"* (Campo, Martella & Ciccarese, 2018, 9). Fenomeni per descrivere i quali la dottrina specialistica ha coniato l'espressione "Black Box" - altamente suggestiva per assenza di trasparenza dei meccanismi di funzionamento degli algoritmi (Innerarity, 2021; Paßmann & Boersma, 2017; Seaver, 2017; von Eschenbach, 2021; Wischmeyer, 2020 e bibliografie ivi citate) - ossia "scatola nera" il cui funzionamento è difficile o impossibile da decifrare (Pasquale, 2015).

Da almeno un quinquennio, invero, molte sono le voci ad essersi levate a sostegno di una maggiore comprensione dei processi endogeni ed esogeni sottesi alle analisi automatizzate (Beer, 2017 e bibliografia ivi citata) che, sul piano tecnico, presentano almeno tre cause di opacità - rispettivamente riconducibili a segretezza, complessità e imperscrutabilità del linguaggio algoritmico (Burrell, 2016; Lampo, Mancarella & Piga, 2020) - le quali, operando in maniera congiunta, ostacolano l'interpretabilità degli algoritmi, così da pregiudicare la comprensione dei meccanismi di produzione della conoscenza, degli eventuali errori o distorsioni di sistema (Pasquale, 2015; Pedreschi, Giannotti, Guidotti et al., 2018).

Approfonditi studi di settore hanno, peraltro, dimostrato come le affermazioni in merito alla natura impersonale, astratta ed oggettiva dell'algoritmo si scontrino con scenari concreti del tutto difformi rispetto alla pretesa neutralità dell'Intelligenza Artificiale, che chiamano in causa gli

intricati processi sociali di cui gli algoritmi costituiscono il risultato (Campo, Martella & Ciccarese, 2018). Per meglio comprendere il cambio di paradigma in atto, si consideri la direzione che stanno percorrendo gli studi di tecnoscienze i quali, lungi dal considerare l'algoritmo come un mero strumento tecnico, si inseriscono piuttosto nell'ormai decennale filone "Critical Algorithm Studies" (Gillespie & Seaver, 2016; Seaver, 2013). Altrimenti detto, si tratta di riconoscere l'algoritmo come l'esito dell'azione congiunta di molteplici elementi che, interagendo tra loro, contribuiscono a creare un sistema ben più complesso e articolato (Aragona & Felaco, 2019) - quello che Kitchin (2017) definisce una sorta di assemblaggio socio-tecnico, con ciò intendendosi un sistema composto da diversi apparati di natura tecnica e sociale, inestricabilmente intrecciati, che convergono nella produzione dei dati. In una simile prospettiva, gli algoritmi rappresentano il prodotto combinato di diversi apparati, di molteplici tecniche analitiche e di varie comunità di esperti in competizione tra loro: una combinazione di fattori responsabile dell'opacità dell'origine e del funzionamento algoritmico (Burrell, 2016).

Un dibattito che negli ultimi anni ha dimostrato una spiccata vitalità attiene ad una (presunta) analogia tra Impersonalità, oggettività, razionalità, discrezionalità politica e resa tecnica – attributi che si pretenderebbero conferiti sic et simpliciter agli algoritmi – e meccanismi di funzionamento nonché criteri di legittimazione tipici delle burocrazie occidentali (Aneesh, 2009). La tematica è dettagliatamente affrontata nel contributo di Visentin (2019), dal suggestivo titolo "*Il potere razionale degli algoritmi tra burocrazia e nuovi idealtipi"*, in cui l'Autrice evidenzia come, benché alcune caratteristiche generali della burocrazia - estensione e intensificazione del controllo e inibizione della discrezione umana, in primis – possano essere certamente accentuate dall'uso di algoritmi, opinioni contrastanti emergano in merito alla capacità dei processi di algoritmizzazione e burocratizzazione di produrre effetti similari su campi e organizzazioni differenti. Certo è che *"(…) rimane (…) una differenza sostanziale nelle modalità di esercizio del potere che si mostra chiaramente quando si tratta di opporvisi: chi è soggetto al potere burocratico può, almeno in linea di principio, comprenderne il funzionamento ed eventualmente contestarlo. Gli algoritmi invece sono fondamentalmente imperscrutabili per gli attori sociali e in alcuni casi sono coperti da segretezza*" (Campo, Martella & Ciccarese, 2018, 14). Del resto, il difetto di trasparenza nel trattamento di dati posto in essere in ambiti applicativi c.d. data intensive era già stato segnalato dalla dottrina specialistica come pregiudizievole della qualità del consenso del titolare: oggetti IoT e app installate su smartphone, invero, raccolgono e scambiano dati in modalità discreta, continua, pervasiva, praticamente in assenza di partecipazione da parte dell'utente interessato (Prestipino, 2017).

Un corretto inquadramento della tematica, anche ai fini della comprensione delle implicazioni etiche derivanti dall'uso dei Big Data, necessita di una precisazione fondamentale: gli algoritmi devono essere considerati come *"(…) il risultato dell'incontro di complessi processi sociali e il cui esito non è mai scontato. Gli algoritmi (cioè) producono i loro effetti nel tessuto sociale e nelle vite degli individui poiché sono il prodotto di un contesto sociale animato da interessi molteplici – spesso in contrasto tra loro – e condizionato anche da limiti tecnici"* (Campo, Martella & Ciccarese, 2018, 11). La dimensione discrezionale [e dunque politica, nel senso di "attinente alla polis" (Burrell, 2016; Mazzotti, 2015)] è, perciò, una componente fisiologica dell'algoritmo, in quanto strumento condizionato dalle distorsioni e dai limiti tipici dell'azione umana – quella dei progettisti, in modo particolare, come emerge dal lavoro di Aragona & Felaco (2019). I due Autori, analizzando le fasi di progettazione di famiglie algoritmiche anche attraverso l'intervista ai soggetti coinvolti, hanno evidenziato come lo stesso processo di ricerca

sia frutto di un imprescindibile contemperamento tra interessi e necessità differenti. Dello stesso tenore appare il saggio di Airoldi & Gambetta (2019), che contribuisce a decostruire il mito della "neutralità algoritmica", a lungo considerata come la nozione centrale del discorso egemonico intorno agli algoritmi stessi (Ballatore & Natale, 2018). La non-neutralità degli algoritmi sarebbe, perciò, in aperto contrasto con la retorica sociale tradizionale, che li vorrebbe strumenti oggettivi, efficienti ed imparziali (Sandvig, Hamilton, Karahalios, et al., 2016).

L'algoritmo come strumento socio-tecnico – l'espressione è mutuata da Aragona & Felaco (2019, 11) – è una realtà ormai acquisita nella comunità accademica internazionale, che riconosce il ruolo cruciale da esso rivestito nelle istituzioni politiche ed economiche e, più in generale, la sua influenza nel tessuto sociale (Boyd & Crawford, 2012; Grosser, 2014; Nakamura, 2013; Pasquale, 2015; Tufekci, 2015). Per giunta, si registra un significativo ampliamento del novero degli studiosi che contestano l'obiettività e la conseguente pretesa di affidabilità degli algoritmi (De Rosa & Aragona, 2017; Dourish, 2016; Napoli, 2013), consapevoli di quanto un programmatore, per quanto si sforzi di mantenere la propria imparzialità nella creazione di un algoritmo, in esso farà convogliare immancabilmente il proprio background di conoscenze, così come i propri modelli, siano essi culturali che di pensiero (Gillespie, 2014). Fattori, quest'ultimi, destinati a confluire nei processi di calcolo, selezione, traduzione e categorizzazione implementati dallo strumento algoritmico (Aragona & Felaco, 2019). Il contributo di Pedreschi e colleghi conferma la presenza di *"(…) human biases and prejudices, as well as collection artifacts"* (2018, 1) nella costruzione del dato, posto che tali elementi si rifletteranno sia nella scelta degli indicatori che nei metodi di misurazione riprodotti dalle procedure automatizzate di analisi algoritmica (Lampo, Mancarella & Piga, 2019).

Problematiche di natura discriminatoria, conseguenti all'impiego dello strumento algoritmico nell'ambito della Big Data Analytics, sono state segnalate già da tempo dai cultori della materia a livello internazionale (Cardon, 2016; Christl & Spiekermann, 2016; Bruce, Malcom, & O'Neill, 2017[61]) concordi nel ritenere come formule e modelli matematici, diagrammi e procedimenti formali celino, in realtà, meccanismi di alterazione dell'informazione e condizionamento dell'azione di cittadini, utenti e consumatori. L'allarme è stato rilanciato finanche da Berners-Lee (già creatore del World Wide Web), il quale, in molteplici occasioni, ha ribadito come la rete Internet sia ormai popolata da guardiani digitali (gatekeepers) sempre più potenti, le cui armi sono algoritmi in grado di manipolare le persone e di limitarne la libertà [62] (Campo, Martella & Ciccarese, 2018).

Siffatti meccanismi di sbarramento minerebbero in radice il pluralismo informativo, prerequisito per il corretto funzionamento di qualsivoglia sistema democratico. Ad amplificare il citato effetto distorsivo, interverrebbe il c.d. pregiudizio di conferma, ulteriormente potenziato dal c.d. paradosso della Digital Society – quello che minimizza il costo di transazione nell'acquisizione del patrimonio informativo rispetto al passato ma che, al tempo stesso,

[61] Gli Autori mettono in guardia dall'affidabilità e oggettività degli algoritmi, potendo questi esprimere pregiudizi a causa di una programmazione superficiale o essere impiegati per mettere in atto comportamenti truffaldini (si pensi al software che ha consentito ad un noto marchio automobilistico tedesco di alterare le rilevazioni delle emissioni inquinanti).

[62] Secondo Berners Lee (citato da Favretto, 2020) esisterebbero alcun gatekeepers (guardiani), ossia un piccolo gruppo di piattaforme digitali molto potenti in grado "di controllare quali idee e quali opinioni sono viste e condivise. E queste piattaforme dominanti sono capaci di proteggere la loro posizione creando barriere all'entrata per i concorrenti". i

comprime l'attività di ricerca delle informazioni rilevanti (Nicita, 2019). Detta ricerca, a causa del confirmation bias, si ridurrebbe alla selezione delle sole informazioni in grado di avvalorare credenze e convinzioni dell'utente, posto che, con sempre maggiore frequenza, è lo stesso algoritmo dei search engines o dei social network ad alimentare il bias di conferma, suggerendo ciò che è probabile possa soddisfare l'interesse dell'utente, in base al profilo rivelato da scelte passate o da quelle operate dal gruppo dei pari. L'esperienza di gruppo fungerebbe, infine, da catalizzatore, cristallizzando le distorsioni cognitive dell'utente e polarizzandole verso una visione ancora più estrema. Il rapporto tra disinformazione e confirmation bias, come sopra delineato, è oggetto di dettagliata analisi da parte di Nicita & Delmasto, in un saggio dato alle stampe nel 2019, nel quale i due Autori documentano i pesanti effetti dell'incidenza dei Big Data sul tessuto economico e politico di un qualsivoglia Paese ispirato ai valori occidentali.

A favore del predetto costrutto, milita un rapporto dell'AGCOM, dal titolo "News vs Fake nel sistema dell'informazione" (2018) che misura empiricamente tanto la rilevanza del pregiudizio di conferma nei social media quanto la sua capacità di generare polarizzazione delle idee e del dibattito pubblico. Gli outcomes del Report dimostrano come i modelli di consumo informativo e l'interazione degli utenti con le notizie caricate sulle piattaforme online risultino caratterizzati da una forte tendenza alla polarizzazione e all'esposizione selettiva. Gli utenti analizzati tenderebbero, cioè, a selezionare le informazioni coerenti con il proprio sistema valoriale di preferenze e convinzioni, formando gruppi polarizzati di persone con idee simili su narrazioni condivise, in cui le informazioni discordanti verrebbero sistematicamente ignorate (Nicita, 2019 e bibliografia ivi citata).

Studiosi del diritto della privacy, della concorrenza, delle assicurazioni, della finanza e dei rapporti di lavoro si confrontano quotidianamente sui rischi BD Analytics-correlati, delineando scenari critici in relazione alla garanzia dei diritti fondamentali della persona (Lamardini, 2016; Maggiolino, 2017; Mantelero, 2015; Porrini, 2020; Rota, 2017 e bibliografie ivi citate). Trattasi di rischi principalmente connessi alle potenziali pratiche discriminatorie a danno del cittadino consumatore/utente, con specifico riferimento ai fenomeni di discriminazione indotti dai software di Big Data Analytics con finalità predittive (Bailo, 2015; Mantelero, 2015; Neslen, 2021; Rainie & Anderson, 2012).

Ricerche condotte sulla possibilità di risultati discriminatori derivanti dalla Big Data Analytics attribuiscono valenza causale a problemi di rappresentatività intrinsechi ai dataset impiegati per la previsione (Barocas & Selbst, 2016): in sostanza, l'algoritmo può essere influenzato dalla scarsità di dati disponibili per un determinato sottogruppo di individui, così da determinare l'assegnazione di una classe di rischio più elevata a causa della maggiore incertezza sul loro comportamento[63] (Chi, 2017; Osoba & Welser IV, 2017).

Un nutrito gruppo di ricercatori ha confermato la rilevanza di questioni di rappresentatività del campione a fini predittivi: una popolazione statisticamente meno rappresentata genera maggiore incertezza nella classificazione algoritmica, con conseguente produzione di outcomes svantaggiosi per detta minoranza (Goodman & Flaxman, 2016). A dataset più ampi corrisponderebbero, quindi, migliori prestazioni predittive da parte dell'algoritmo

[63] L'algoritmo, cioè, potrebbe valutare un rischio maggiore sulla base della sottorappresentazione del gruppo oggetto di analisi predittiva.

che si è dimostrato particolarmente vulnerabile rispetto a "bias di incertezza": altrimenti detto, a parità di altre condizioni, l'incertezza può tradursi in un maggiore rischio percepito. *"Potentially, algorithms could assess greater risk based on under-representation"* conclude Chi (2017). Podesta, Pritzker, Moniz, et al., (2014) hanno rilevato come alcune tendenze divengano visibili solo nei Big Data, citando un caso di ricerca genetica in cui i marcatori relativi alla schizofrenia non erano interamente rilevabili in piccoli campioni, diventando statisticamente significativi e identificabili solo in un dataset di 35.000 casi.

Nel contesto della sicurezza nazionale[64], al contrario, ciò potrebbe tradursi in un regime di targeting - arbitrario e, come tale, illegittimo - a danno di obiettivi sovrarappresentati nei dati, con conseguente focalizzazione ristretta, nell'attività di indagine, su determinate tipologie di bersaglio, favorendo il fenomeno noto come *streetlight effect* [65] . Ciò giustifica le preoccupazioni che animano gli studiosi di criminologia i quali, già da tempo, vanno segnalando i pericoli insiti nell'uso dei Big Data e del processo decisionale automatizzato (Chan & Bennet Moses, 2016; Joh, 2017; Selbst, 2017; Shapiro, 2019). Il timore è che i dati relativi ai c.d. precedenti penali possano rafforzare le disparità nell'attività di polizia preventiva (Berk, 2021; Ridgeway, 2018), con conseguenti risultati distorti nella prevenzione e applicazione della legge penale (Brayne & Christin, 2021). Detto pregiudizio potrebbe, inoltre, tradursi in un circolo vizioso in cui i dati storici divengono la base per misure di monitoraggio più aggressive, traducendosi in sistematici eccessi di attività di polizia, fenomeno noto agli esperti di settore come *overpolicing*[66] (Couchman, 2019;. Okidegbe, 2019; Southerland, 2020; Plesničar, Završnik., & Šarf, 2020).

Brayne (2017), in particolare, osserva che, se per un verso l'uso del Big Data Analytics può costituire un fattore razionalizzante – grazie al suo potenziale di incremento dell'efficienza mediante l'accuratezza dei risultati dell'analisi predittiva – per altri versi, l'impiego degli strumenti predittivi può reiterare tecnologicamente biases e consolidare modelli esistenti di disuguaglianza sociale. Vero è che l'applicazione del Big Data Analytics ha comportato severe trasformazioni in merito alle pratiche di sorveglianza adottate dalle Forze dell'ordine, amplificando la raccolta, registrazione e classificazione di informazioni su persone, processi e istituzioni (Ferguson, 2017). Di fatto, la crescente pervasività delle pratiche di *policing* ha concretizzato i modelli di "sorveglianza di massa" (Rule, 1974) e "società di sorveglianza" (Lyon, 1994), preconizzati alcuni decenni orsono.

Sebbene si assista ad un costante incremento dell'attività di sorveglianza in tutti i settori della società (Ball & Webster 2003; Lyon 1994, 2003; Marx 1988, 2016), la sua diffusione appariva, già da tempo, distribuita in modo non uniforme (Fiske 1998), posto che alcuni individui, aree e istituzioni risultavano essere oggetto di targeting più di

[64] La questione verrà affrontata dettagliatamente nel Cap. 4 del presente lavoro "Big Data, Sicurezza Nazionale e Minacce Asimmetriche".

[65] Trattasi di un bias di osservazione consistente nella limitazione dell'analisi ai dati disponibili piuttosto che ai dati necessari. Noam Chomsky (1928 - scienziato cognitivo e padre della linguistica moderna) richiama la barzelletta dell'ubriaco sotto il lampione come metafora del funzionamento della scienza: «*A policeman sees a drunk man searching for something under a streetlight and asks what the drunk has lost. He says he lost his keys and they both look under the streetlight together. After a few minutes the policeman asks if he is sure he lost them here, and the drunk replies, no, and that he lost them in the park. The policeman asks why he is searching here, and the drunk replies, "this is where the light is"*» (Barsky, 1998)

[66] Il termine "policing" è difficilmente traducibile in italiano. Può essere inteso con sfumature semantiche del tipo attività di polizia, controllo, vigilanza, mantenimento dell'ordine pubblico et similia (dictionary.cambridge.org).

altri. Inoltre, differenti gruppi di popolazione erano sorvegliati per scopi diversi (Lyon 2003). Lo sviluppo tecnologico ha senz'altro esacerbato queste pratiche, anche mediante l'uso di Data System precedentemente separati e attualmente uniti in sistemi relazionali inclusivi di dati originariamente raccolti in altre istituzioni, estranee ai circuiti della giustizia penale (Andrejevic & Gates, 2014). Secondo Brayne (2017, 979), *"Big data is an emerging modality of surveillance"*.

A suscitare maggiori perplessità sono, in ogni caso, le pratiche emergenti di sorveglianza "dragnet" ossia, a strascico (Angwin, 2014; Mason, 2012; Parker, 2011; Pavletic, 2018) – quelle, cioè, che raccolgono dati indistintamente, piuttosto che su individui sospettati – le quali si traducono in un monitoraggio invasivo di gruppi *"precedentemente esenti dalla sorveglianza di routine"* (Haggerty & Ericson 2000: 606; Angwin 2014; Lyon 2015). L'attuale sorveglianza condotta con tecnologie digitali appare, pertanto, sia più ampia che più invasiva, potendo coinvolgere una cerchia più estesa di persone, in una gamma più variegata di contesti istituzionali[67] (Graham & Wood, 2017). Inoltre, a differenza della sorveglianza tradizionale che è induttiva - la "stretta osservazione, specialmente di una persona sospettata" (Oxford American Dictionary of Current English 1999) - la nuova sorveglianza è suscettibile di essere applicata categoricamente, è deduttiva, remota, a bassa visibilità o invisibile, involontaria, automatizzata, preventiva e inclusa nell'attività di routine (Marx 2002, 2016).

La letteratura specialistica evidenzia un ulteriore profilo di criticità, attinente al carattere pervasivo delle pratiche di sorveglianza attuali: la proliferazione di record digitalizzati consente, invero, di accorpare dati provenienti da fonti istituzionali, originariamente separate, in un sistema unitario, costituito da una pluralità di banche dati interoperabili. Ebbene, l'integrazione dei sistemi di dati istituzionali comporta lo "slittamento" dell'attività di sorveglianza dal circuito (fisiologico) della giustizia penale verso altre istituzioni non penali, generando l'effetto distorsivo c.d. function creep [68] - il fenomeno dei dati originariamente raccolti per uno scopo e utilizzati per uno scopo diverso – e contribuendo, in tal modo, ad un incremento significativo dei dati cui le Forze dell'ordine hanno accesso (Chan, 2021; Brayne, 2020, 2017). Tale tendenza sarebbe, per giunta, perfettamente aderente a quello che Fourcade & Healy (2017) definiscono "Imperativo dei dati istituzionali", che imporrebbe alle organizzazioni moderne la raccolta del maggior numero di dati possibile, a scapito del diritto alla riservatezza del cittadino/utente (Watney, 2019).

A riflettere sull'importanza del principio di limitazione dello scopo già in fase di progettazione (By Design), al fine di mitigare l'impatto negativo dell'IA sui diritti umani e sulla sicurezza dei sistemi informativi, sono intervenuti Fantin & Vogiatzoglou (2020) i quali, nel loro contributo dal titolo *"Purpose Limitation By Design As A Counter To Function Creep And System Insecurity In Police Artificial Intelligence"* sostengono come il continuo sviluppo delle capacità di un sistema possa consentire impieghi che esorbitano l'ambito di applicazione e lo scopo originariamente previsti. Una simile evenienza può causare un'involuzione delle funzioni di sistema aggravando gli incidenti

[67] A riprova della natura cruciale della tematica, Sarah Brayne dedica alla *"Dragnet Surveillance"* un intero capitolo del suo ultimo saggio.*"Predict and Surveil: Data, Discretion, and the Future of Policing"* dato alle stampe nel 2020.

[68] Function creep *"is the gradual widening of the use of a technology or system beyond the purpose for which it was originally intended, esp when this leads to potential invasion of privacy* (Collins English Dictionary, https://www.collinsdictionary.com/dictionary/english/function-creep)

di sicurezza, oltre ad avere un impatto intrusivo sui diritti umani: nel caso in esame, sistemi di AI destinati a specifici obiettivi di prevenzione del crimine potrebbero essere gradualmente riutilizzati per attività di sorveglianza ingiustificate, poiché non considerate in origine (Brayne, 2017). Posto che la disciplina europea sul punto è cristallina - il principio della limitazione delle finalità impone che siano specificati i fini per i quali i dati sono trattati e che il loro successivo utilizzo sia limitato ad essi (art. 5(1)(b), GDPR) – il grado di evoluzione degli attuali sistemi di AI non consente, tuttavia, un'applicazione sistematica di detto principio, sollecitando gli Autori a proporre una valutazione più incisiva degli scopi previsti già in fase di progettazione del sistema.

Le *"representation issues"* (Chi, 2017, 16) precedentemente descritte non sono di poco conto sotto il profilo etico, posto che da esse possono derivare pregiudizi e discriminazioni a carico di ampie fasce di popolazione, non di rado già stigmatizzate: ciò che accade nel Credit Rating System, dove l'assegnazione del punteggio (score) di rischio statistico e la definizione del c.d. sistema di soglia assume valenza dirimente nelle politiche creditizie degli istituti bancari. Uno studio specifico su algoritmi di apprendimento automatico[69] e discriminazione ha dimostrato come diversi tipi di soglia possono avere esiti differenti nella concessione del credito, con conseguenti outcomes discriminatori per i sottogruppi di richiedenti (Wattenberg, Viegas & Hardt, 2016). Anche la più recente letteratura in tema di Credit Scoring e impiego dei Big Data, ampia e variegata nell'approccio, ha sostanzialmente confermato i risultati precedenti (Chern, Lei, Huang, et al., 2021; Das, 2016; Hurley & Adebayo, 2016; Onay & Öztürk, 2018; Wang, Ding, Yu, et al., 2020; Wang & Perkins, 2019; Wen, Yang, Gan et al., 2021; Zhao, 2020). Secondo Chi (2017), sarebbero gli algoritmi di apprendimento i principali responsabili dei risultati deleteri del processo decisionale, poiché riprodurrebbero modelli discriminatori o altri pregiudizi diffusi congeniti ai dati oggetto di analisi.

Falsi positivi, falsi negativi e cicli di feedback comporterebbero, anch'essi, conseguenze rilevanti sul piano etico, secondo la dettagliata analisi compiuta dall'Australian Strategic Policy Institute nel 2017. Tassi disuguali di inclusione nei database possono produrre conseguenze significative: gli afroamericani, infatti, hanno sette volte più probabilità dei bianchi di essere ingiustamente condannati per omicidio (Gross, Possley & Stephens 2017). Qualsiasi algoritmo di valutazione del rischio può produrre falsi positivi - non importa quanto sia accurato il modello, il profilo individuale o i dati processati (Brayne, 2017) – con la probabilità che persone innocenti siano soggette a intrusive e dirompenti indagini da parte delle agenzie di Law Enforcement. La necessità di rimuovere il maggior numero possibile di falsi negativi dai propri sistemi predittivi potrebbe indurre le predette agenzie ad indagare su un numero significativamente maggiore di falsi positivi, estendendo così l'attività di sorveglianza preventiva (Chi, 2017).

È opinione ormai consolidata che le nuove tecnologie comportino rischi per i diritti umani fondamentali – protezione dei dati personali e della privacy, in special modo – così come pregiudizi, discriminazioni e profilazione illegale (Gkougkoudis, Pissanidis & Demertzis, 2022). L'abbondanza di dati consente alle grandi aziende tecnologiche, che ne sono

[69] L'algoritmo di apprendimento automatico (anche detto Machine Learning – ML) ha il compito di apprendere il sistema di ponderazione per il modello, ossia quel sistema che valuta la probabilità che i pattern che il modello sta apprendendo riflettano le relazioni effettive nei dati. L'apprendimento automatico è una variante alla programmazione tradizionale, nella quale in una macchina si predispone l'abilità di apprendere informazioni dai dati in maniera autonoma, senza istruzioni esplicite (Bishop & Nasrabadi, 2006).

proprietarie (*Big Tech*) [70] , di formulare inferenze su credenze, valori, preferenze, stato psicofisico e dettagli intimi, compresi sentimenti e vulnerabilità delle persone (Zuboff, 2010). Inoltre, sebbene parte dei dati siano aggregati in c.d. dataset "anonimizzati" (Crawford & Schultz, 2014), la quasi totalità di essi può essere "scorporata" e i dati nuovamente identificati a livello individuale, con l'impiego di appositi software (Rocher, Hendrickx, & de Montjoye, 2019). Le stesse attività quotidiane producono costantemente tracce digitali (Unsworth, 2016), rendendo di fatto impossibile condurre un'esistenza "anonima" (Omand & Phythian, 2018).

Le suesposte considerazioni sembrano dimostrare *«(…) come l'idea di ottenere un algoritmo privo di "bias" (distorsioni) sia discutibile sul piano epistemologico perché è fondata sull'assunto implicito che l'algoritmo debba semplicemente avvicinarsi il più possibile a un contesto di selezione "giusto", che esiste "là fuori", indipendentemente dall'algoritmo stesso. La possibilità di risolvere discriminazioni e disparità di trattamento presenti nella società attraverso una programmazione il più possibile unbiased si rivela quindi insoddisfacente: in primo luogo perché gli assiomi iniziali di un modello condizionano necessariamente i passaggi successivi e, per di più, la realtà su cui gli algoritmi operano, e i dati che questi utilizzano, non esistono indipendentemente dall'algoritmo, che contribuisce in parte a creare e a mettere in forma quella stessa realtà. le conseguenze sociali di queste criticità non risolte (e probabilmente irrisolvibili se affrontate solo da un prospettiva tecnica) sono quelle note come "effetto San Matteo" (Merton, 1968; Mingo, Bracciale, 2016), che consistono, cioè nella riproduzione e amplificazione delle disparità sociali, in questo caso nascoste dal velo di imperscrutabilità imposto dall'automatizzazione dalle scelte di selezione, classificazione, identificazione delle relazioni operate da algoritmi sempre più complessi"* (Campo, Martella & Ciccarese, 2018, 12-13).

Big Data e nuove tecnologie associate rappresentano strumenti che permettono una sorveglianza dall'ampiezza e pervasività senza precedenti, poiché in grado di tracciare passivamente un gran numero di soggetti, anche estranei ai circuiti penali tradizionali, con implicazioni significative in termini di disuguaglianza sociale e applicazione della legge (Brayne, 2020). Benché Il ruolo del sistema di giustizia penale nella produzione della disuguaglianza sociale abbia ricevuto notevole attenzione nella letteratura dell'ultima decade – Laub (2014, 13) conclude il suo saggio affermando che *"(…) while there are important inequalities that exist prior to justice system involvement, the justice system itself is implicated in the exacerbation of inequality, especially for blacks and other minorities"* - l'impatto dell'uso dei Big Data rimane una questione empirica aperta. Posto che gli stereotipi rivestono utilità cognitiva a fronte di informazioni incomplete – le ricerche di psicologia sociale dimostrano come gli esseri umani, "avari cognitivi" (Fiske & Taylor, 1991), si affidino a scorciatoie come la fusione tra oscurità e criminalità (Quillian & Pager 2001) per attribuire significato al mondo che li circonda - qualora i Big Data vengano utilizzati per implementare la conoscenza di fenomeni ignoti, ciò potrà consentire alle Forze di polizia di limitare il ricorso compensatorio a stereotipi di razza e di classe, con conseguente riduzione dell'iper-sorveglianza delle minoranze etniche ed erosione della fiducia da parte della comunità (Sampson & Bartusch, 1998).

In ogni caso, la poderosa ricerca di Brayne (2017; 2020) ha evidenziato una triplice matrice di criticità, fonte di disuguaglianze sociali: rispettivamente, il perpetuarsi della

[70] Le stime suggeriscono che *Google, Amazon, Microsoft* e *Facebook* memorizzino almeno 1200 petabyte, o 1,2 milioni di terabyte, di dati ciascuno mentre, dei 4000 broker di dati a livello globale, uno dei più grandi, *Acxiom*, si stima abbia 3000 data points per ciascuno dei 500 milioni di consumatori in tutto il mondo (Hammond-Errey, 2022).

sorveglianza a carico di soggetti precedentemente sospettati, attraverso un ciclo di feedback (Pasquale, 2015); l'ampliamento del raggio d'azione delle agenzie di Law Enforcement, mediante l'impiego di tecniche digitalizzate come la c.d. dragnet surveillance (Renan, 2016); l'operatività di Data System integrati, con conseguente espansione della sorveglianza oltre i confini dei circuiti penali istituzionali. In tutte e tre le ipotesi, seppure in maniera diversa, si evoca lo stigma del contatto con la giustizia penale (Becker 1963; Brayne 2014; Goffman 2014, 1963; Kohler-Hausmann 2013; Pager 2007; Rios 2011; Stuart 2016; Wakefield & Wildeman, 2013; Western & Pettit, 2005).

Sebbene parte del fascino dei Big Data risieda nella loro promessa di un processo decisionale meno discrezionale e più obiettivo (Espeland e Vannebo 2007; Hacking 1990), le nuove piattaforme digitali e le tecniche analitiche sono impiegate in contesti organizzativi preesistenti (Barley 1986, 1996; Kling 1991) e incarnano gli scopi dei loro creatori (Boyd & Crawford 2012; Gitelman 2013; Kitchin 2014). Pertanto, rimane una domanda empirica aperta fino a che punto l'adozione di analisi avanzate ridurrà inefficienze e disuguaglianze organizzative, o al contrario, favorirà il radicamento delle dinamiche di potere all'interno delle Organizzazioni (Brayne, 2017).

3 Potenzialità e Applicazioni dei Big Data

Le c.d. tecnologie Data-intensive – Internet of Things (IoT), Apps per smart-devices e Big Data, solo per citare le più diffuse su scala globale (Chen & Zhang, 2014) – comportano una massiva produzione di informazioni e un pervasivo scambio di dati, le cui implicazioni, a livello individuale, possono tradursi in una personalizzazione indubbiamente foriera di vantaggi (si pensi, ad es. alla varie tipologie di targeted advertising) così come servizi a valore aggiunto, ossia più efficienti e meno costosi e, pertanto, migliorativi della user-experience (Prestipino, 2017). Attualmente, l'abbondanza di dati, la connettività digitale e la tecnologia onnipresente consentono una "copertura" pressoché completa dell'esistenza umana, sull'intero pianeta, spesso in tempo reale mentre la pandemia da COVID-19, imponendo maggiori interazioni online e un'aumentata dipendenza sociale dalla tecnologia, ha contribuito significativamente all'incremento del pool globale di dati (Hammond-Errey, 2022).

I Big Data, nello specifico, consentono l'accorpamento e l'analisi di milioni di informazioni: dai punti di geolocalizzazione, alle transazioni finanziarie, dai profili dei social media, ai file medici e ai flussi video (Boyd & Crawford, 2012), dando vita ad un nuovo, complesso panorama informativo e infrastrutturale (Hammond-Errey, 2022). Essi si riferiscono, dunque, alla circostanza secondo cui sempre più aspetti e artefatti della vita quotidiana risultano disponibili in formato digitale - profili personali o aziendali, post sui social network e sui blog, storie di acquisti, cartelle cliniche, solo per citarne alcuni – per essere oggetto di analisi mediante uno strumentario tecnologico specifico, teso all'estrazione di conoscenza dai dati medesimi (De Mauro, 2019; Vossen, 2014).

Questa abbondanza di dati permette la formulazione di inferenze su credenze, valori, preferenze, stato psicologico e dettagli sensibili di chi li produce, compresi sentiment e vulnerabilità delle persone, comportando, tuttavia, severi rischi per l'integrità della sfera dei diritti individuali (Mazarr, Bauer, Casey, et al., 2019; Zuboff, 2010). Orbene, al netto dei rischi, delle sfide e delle criticità, sia sul piano etico che giuridico (già parzialmente affrontate nel capitolo precedente), numerose sono le potenzialità che i Big Data assicurano nell'attuale scenario globalizzato, che trovano concreta applicazione in molteplici contesti operativi, tanto nel settore privato, quanto in quello pubblico (Di Porto, 2017; Rezzani, 2013).

Prerequisito infrastrutturale essenziale del *"paradigm shift"* (Kitchin, 2014) perfezionatosi nell'ultimo ventennio è la c.d. connettività digitale, intesa come la capacità di associare persone, luoghi e idee attraverso reti virtuali. Essa è consentita dalla compresenza di miliardi di sensori e dispositivi dislocati in tutto il mondo, costantemente connessi alla rete Internet, e dalla relazione tra le cose e le persone resa possibile da piattaforme ad hoc e dai costanti progressi nella tecnologia delle telecomunicazioni (Nasser & Tariq, 2015; Schwab, 2017) – ciò che genera un sistema digitalizzato su scala globale (Australian Government Productivity Commission, 2016).

Già da tempo, i Big Data hanno attirato un'enorme attenzione da parte dei ricercatori nel campo delle scienze sociali e dell'informazione, dei decision makers politici e delle imprese (Chen & Zhang, 2014), così come un crescente interesse è stato registrato da parte delle Pubbliche Amministrazioni nell'esercizio delle funzioni istituzionali – interesse, quest'ultimo, interpretabile come una evoluzione delle recenti politiche c.d. di open government e open data (Di Porto, 2017).

L'indiscusso punto di forza dei Big Data è, indubbiamente, ravvisabile nelle capacità analitiche e predittive ad essi associate, garantite dall'impiego di algoritmi in grado di estrarre velocemente informazioni rilevanti da enormi volumi di dati, di diversa origine e natura (Crawford, 2013). Ciò è reso possibile dall'uso di tecniche di machine-learning, che consentono di individuare correlazioni e modelli (patterns) in grado di predire comportamenti, fatti e processi su larga scala (Al-Jarrah, Yoo, Muhaidat et al., 2015; Injadat, Moubayed, Nassif, et al., 2021; L'Heureux, Grolinger, Elyamany, et al., 2017; Nti, Quarcoo, Aning, et al., 2022; Wang, Fu, He, et al., 2020; Zhou, Pan, Wang, et al., 2017). Trattasi di una tipologia informativa capace di generare conoscenza, in quanto permette di tracciare profili comportamentali, testare la robustezza di ipotesi di rischio – quello clinico e ospedaliero risultano tra i settori maggiormente esposti al fenomeno (Beam & Kohane, 2018; Hinton, 2018; Naylor, 2018; Peterson, 2019; Shah, Steyerberg & Kent, 2018) - e, appunto, formulare inferenze di natura previsionale (Di Porto, 2017).

Certamente, nell'enorme volume di dati si cela un potenziale valore di grande utilità in molteplici settori, che spaziano dalle attività economiche e commerciali alla pubblica amministrazione, dalla sicurezza nazionale alle ricerche scientifiche. In sostanza, i Big Data si stanno rivelando un asset estremamente prezioso per garantire produttività nelle imprese e progressi evolutivi nelle discipline scientifiche. Senza dubbio, le competizioni in atto nel campo della produttività aziendale e delle tecnologie emergenti si misurano costantemente sul piano esplorativo della dimensione c.d. Data-intensive, come preconizzato ormai da circa un decennio (Chen & Zhang, 2014). La letteratura sul tema sviluppatasi negli ultimi anni appare assai ricca e variegata, documentando una pluralità di contesti di utilizzo dei Big Data: aziendali, amministrativi e accademici – discipline scientifiche e sociali, in primis – ma anche governativi, all'interno dei quali si procede all'adozione di tecnologie e metodi analitici specifici per l'estrazione di valore dai dati collazionati, che prendono il nome di Big Data Analytics (BDA) (De Mauro, 2019).

Dalla complessità [71] del fenomeno – i Big Data, infatti, superano i limiti operativi degli strumenti di database tradizionali (Di Porto, 2017) – discende un triplice ordine di considerazioni, in merito, rispettivamente, al

[71] La complessità è la sfida fondamentale dei Big Data, secondo Chi (2017), poiché l'attività analitica richiede l'uso di una pipeline di tecnologie, lavoro da parte di un team di analisti multidisciplinari e una buona integrazione dei sistemi.

business, alla dimensione tecnologica e a quella finanziaria (Rezzani, 2013). In primo luogo, l'impiego dei Big Data permette l'implementazione di nuovi modelli di business aziendale, con sensibili vantaggi competitivi rispetto ai modelli tradizionali. (Ajah & Nweke, 2019; Gupta, Drave, Dwivedi et al., 2020; Maroufkhani, Wagner, Wan et al., 2019; Mikalef, Pappas, Krogstie, et al., 2019; Mishra, Luo, Hazen et al., 2019; Singh & Del Giudice, 2019). La letteratura è concorde nel sostenere che le aziende debbano sviluppare forti capacità di BDA per sfruttare l'analisi dei Big Data e ottenere guadagni in termini di efficienza imprenditoriale (Mikalef, Boura, Lekakos, et al., 2019).

Secondariamente, la natura complessa dei Big Data impone l'impiego di tecnologie adeguate all'estrazione di conoscenza e di valore dagli stessi (Athmaja, Hanumanthappa & Kavitha, 2017; Divya, Bhargavi & Jyothi, 2018; Wang & Alexander, 2016). L'evoluzione quali-quantitativa nel patrimonio dei dati in materia di Business Intelligence impone un costante upgrade nelle tecniche di analisi. La ricchezza e varietà della produzione scientifica sottolinea come la proliferazione dei dati richieda un ripensamento delle tecniche di acquisizione, archiviazione ed elaborazione degli stessi, suggerendo, inoltre, che una buona gestione e manipolazione dei Big Data – ossia, mediante l'impiego di tecniche e strumenti adeguati – possa fornire intuizioni praticabili sul piano concreto, in grado di generare valori aziendali (Ajah & Nweke, 2019. Chen, Li, & Wang, 2022; Saggi & Jain, 2018; Sun, Sun, & Strang, 2018; Sun, Zou, & Strang, 2015).

Benché la letteratura di settore si sia concentrata, in maniera prevalente, sugli aspetti tecnologici della questione (Rashid & Khurshid, 2022) - *"The most well-known tools in use today are business and information analysis, predictive analysis, cloud technology, portable business intelligence, Big Data consulting, and visual analytics"* (Gupta & Jiwani, 2021, 2) – allo stato attuale, pochi sono ancora gli studi ad avere realmente esplorato il nucleo della conoscenza della Business Intelligence. L'analisi dei fattori basata sulla matrice di co-citazione [72] ha rivelato sette key-factors della conoscenza di base della BI: big data analytics; benefici e successo della BI; capacità e performance organizzative; accettazione e misurazione delle tecnologie informatiche (IT); informazioni e business analytics; social media text analytics e sviluppo della BI. In aggiunta, l'analisi dei cluster ha rivelato sei categorie: accettazione e misurazione dell'IT; successo e misurazione della BI; capacità e prestazioni organizzative; valore aziendale abilitato dai big data; social media text analytics; sistema di BI e analytics. Detti risultati sembrano indicare l'attuale emersione di numerosi temi di ricerca legati all'impiego dei BD nel settore della Business Intelligence, oltre a suggerire possibili future traiettorie di studio (Shiau, Chen, Wang, et al., 2023).

In conclusione, i vantaggi economici aziendali, conseguenti all'adozione della Big Data Analytics, devono costituire oggetto di prudente bilanciamento, mediante un'approfondita valutazione a priori dei costi necessari per implementare simili soluzioni (Awotunde, Adeniyi, Ogundokun et al., 2021; Loebbecke, & Picot, 2015; Qi & Deng, 2019; Ren, 2022; Tan & Zhan, 2017; Vasarhelyi, Kogan, & Tuttle, 2015; Zhu & Yang, 2021). È proprio quest'ultima traiettoria di analisi a rivestire particolare criticità in termini di budget aziendale, posto che la misura ex ante del valore del pool di dati non è agevole (Chi, 2017). Ciò rende ragione del forte incentivo alla raccolta massiva di dati ritenuti potenzialmente idonei a generare conoscenza, che rende, tuttavia, piuttosto problematiche le stime dei costi e dei benefici, complicando la pianificazione e l'approvvigionamento e attenuando la

[72] L'analisi delle co-citazioni è un metodo di ricerca bibliometrica.

precisione delle iniziative per integrare i set di dati in questione (Bartosik-Purgat & Ratajczak-Mrożek, 2018; Côrte-Real, Ruivo, Oliveira, et al., 2019; Ferraris, Mazzoleni, Devalle et al., 2019). Emblematica la considerazione formulata da Hammouri e colleghi che, in una recentissima review della letteratura sul tema apparsa nel novembre 2022, affermano: *"Traditional business practices could be completely transformed by big data analytics (BDA). Nevertheless, it is still unclear how BDA capabilities affect a firm's performance"* (Hammouri, Atobishi,, Altememi, et al., 2022, 1090).

I Big Dati vengono frequentemente descritti come fattori evolutivi nel mondo della Business Intelligence (BI)[73], in qualità di sistema di supporto ai processi decisionali per migliorare la performance aziendale (Rezzani, 2013). Il ruolo dei Big Data nella BI è ben delineato da Balakrishnan & Rahul: *"The cardinal element of business intelligence is data. Big Data points on the volume of both structured and unstructured data collected from the sources. The size of data relies upon the sources of data considered, the company's establishment in the market, it's short and long-term goals to be achieved, knowing its customers' need, it's business model etc., It is simple to explain the big data in just three words, variety, velocity and volume of data. Big data involves in these main activities with data i.e.: collection, storage, integration"* (2018, 21). A riprova di ciò, uno studio bibliometrico condotto da Liang & Liu su pubblicazioni indicizzate nel periodo 1990-2017 documenta uniformità di vedute nella comunità accademica internazionale circa la veste che i BD hanno assunto nelle dinamiche aziendali ispirate alla BI[74] (Ardito, Scuotto, Del Giudice, et al., 2019; He, Wang & Akula, 2017; Ram, Zhang & Koronios, 2016).

Ebbene, se la letteratura consultata definisce i Big Data come *"(…) a core element of Business Intelligence research"* da oltre due decadi (Liang & Liu, 2018, 2), vero è che gli stessi sono stati universalmente riconosciuti come nuovi strumenti e metodologie per lo sviluppo di teorie nella ricerca aziendale da almeno un decennio (George, Haas & Pentland, 2014). Una revisione quantitativa e sistematica dei Big Data nell'ambito della ricerca economica, compiuta con analoga metodologia bibliometrica da Zhang e colleghi (2021), ha inoltre evidenziato come la Artificial Intelligence (AI) sia diventata una vera e propria "parola d'ordine" (*"a buzzword"*) nel settore commerciale, specialmente per le industrie high-tech. È interessante notare come il filone di ricerca bibliometrico – nella forma della. Systematic Literature Review - sia apparso particolarmente fiorente nell'ultimo triennio, documentando un vivace interesse soprattutto da parte di ricercatori e accademici asiatici, con outcomes peraltro sovrapponibili a quelli sopra riferiti (Khanra, Dhir & Mäntymäki, 2020; Ying, Sindakis, Aggarwal et al., 2021; Liu, Sun, Wang et al., 2020; Zhang, Srivastava, Sharma et al., 2021).

Sulla scorta delle precedenti osservazioni, pare lecito affermare che la panoramica delle opportunità potenzialmente derivanti dall'impiego dei Big Data sia davvero molto ampia e correttamente inquadrata dal saggio di Rezzani, dal titolo "Big Data: Architettura, tecnologie e metodi per l'utilizzo di grandi

[73] Con la locuzione "Business Intelligence" si fa riferimento ad un insieme di processi aziendali di raccolta dati ed analisi di informazioni strategiche, nonché la tecnologia utilizzata per implementare detti processi e le informazioni ottenute come risultato dei processi medesimi (Bergamaschi, Bianconi & Mattavelli, 2023).

[74] Sia "Big Data" che "Business Intelligence" sono parole chiave in rapida crescita nell'attuale ricerca accademica. Mentre "Big Data" è diventato popolare in epoca più recente, "Business Intelligence" è stato proposto molto prima, quando Luhn (1958) iniziò a usare la locuzione BI per descrivere un sistema automatico che diffonde informazioni e supporta il processo decisionale. Il concetto è stato successivamente assimilato all'area del supporto alle decisioni e dei sistemi informativi (Liang & Liu, 2018).

basi di dati" (2013), che spicca tra la produzione scientifica per chiarezza di impostazione. A partire dal riconoscimento del potenziale informativo di tale risorsa, l'Autore delinea un nutrito elenco di tipologie di Big Data e i corrispondenti esempi di uso. A mero titolo esemplificativo, si citano i dati fiscali, bancari e patrimoniali, utilizzabili dalle Agenzie governative competenti per l'identificazione di comportamenti anomali da parte dei contribuenti, suggestivi di pratiche di evasione fiscale; i dati sanitari dei cittadini, utili per ricerca, monitoraggio e diffusione di malattie; i dati meteorologici, per la predizione di eventi atmosferici estremi; i dati relativi a quotazioni e transazioni finanziarie, per la formulazione di analisi predittive in materia di insider trading e andamento dei mercati economici; i dati provenienti da social network, blog e forum, per l'analisi del sentiment di utenti e consumatori, nonché come serbatoio di informazioni per la comunità di Intelligence; i dati provenienti da web server log, per il tracciamento del traffico sui web server e l'identificazione dei comportamenti di navigazione degli internauti; i dati provenienti dai sistemi di sorveglianza, impiegati da Forze di polizia, enti di vigilanza e servizi di Intelligence per scopi istituzionali; i dati geografici provenienti da sistemi GIS, per geolocalizzazione di persone, eventi, ecc.; infine, i dati documentali generici, utilizzabili per controlli incrociati finalizzati alla Fraud Detection, soprattutto nell'ambito delle frodi assicurative (Rezzani, 2013).

Focalizzando l'attenzione sui dati provenienti da fonti web, il cui potenziale informativo appare elevatissimo, gli stessi possono tradursi in un valore aggiunto tanto per le aziende private quanto per il settore pubblico (Di Porto, 2017; Rezzani, 2013). Dati strutturati - anagrafiche clienti o prospect[75], acquisiti attraverso interviste, unitamente a dati di vendita - e non strutturati consentono, congiuntamente, di creare modelli di analisi dei comportamenti d'acquisto e di opinioni rispetto a prodotti, aziende e competitor sul mercato (Hofacker, Malthouse, & Sultan, 2016; Hoyer, MacInnis & Pieters, 2016; Matz, & Netzer, 2017; Smith, 2019). Trattasi di analisi finalizzate sia all'incremento della fidelizzazione del cliente che alla messa in atto di politiche di cross selling [76] sulla clientela già acquisita (Boustani, Emrouznejad, Gholami et al., 2023; Chen, Fan & Sun, 2023;. Fadillah, Yulita, Pradana et al., 2021; Haag, Hopf, Vasconcelos et al., 2022; Zhang, Priestley, DeMaio et al., 2021). Con riferimento ai prospect, le attività analitiche tendono a potenziare l'acquisizione di nuova clientela, mediante strategie di marketing da tempo consolidate, quali campagne pubblicitarie mirate (Appel & Matz, 2021; Hermann, 2023; Matz, Kosinski, Nave et al., 2017).

Taluni Autori hanno evidenziato come l'emergere di simili pratiche di micro-targeting basate sull'impiego della BDA – si pensi, ad esempio, agli algoritmi di raccomandazione di contenuti personalizzati - generino una forte tensione per operatori di marketing, consumatori e decisori politici. Se tali pratiche, per un verso, possono

[75] Il prospect è una persona che rientra nel target di potenziali clienti dell'azienda, ma con cui non è ancora stata creata una connessione diretta. Letteralmente, è un "candidato", possibile o probabile, a trasformarsi in cliente (https://www.extrasys.it/it/magnetblog/differenza-tra-lead-e-prospect-contact#:~:text=Cos'%C3%A8%20un%20prospect%3F,probabile%2C%20a%20trasformarsi%20in%20cliente).

[76] Si tratta di una strategia di vendita consistente nel proporre al cliente, che ha già acquistato un particolare prodotto o servizio, anche l'acquisto di altri prodotti o servizi complementari. La finalità è quella di consolidare la relazione con il cliente – che spesso acquista più prodotti nello stesso processo d'acquisto – e di accrescerne la profittabilità, aumentando la varietà dei prodotti o servizi acquistati dal cliente tra quelli presenti nel portafoglio prodotti (https://www.glossariomarketing.it/significato/cross-selling/).

contribuire al benessere dei consumatori facilitandone i comportamenti d'acquisto, per altro verso, le stesse possono pregiudicare la loro percezione di autonomia nella scelta, minando il welfare individuale (André, Carmon, Wertenbroch et al., 2018). Le risposte dei clienti in un ambiente di "targeted e-commerce" sono state oggetto di un'originale indagine condotta in Vietnam da Le & Liaw (2017), i cui risultati hanno evidenziato che items come la ricerca di informazioni, il sistema di raccomandazione, i prezzi dinamici e i servizi al cliente generano benefici significativi sulla clientela, mentre la privacy e la sicurezza, la dipendenza dallo shopping e le influenze del gruppo dei pari incidono negativamente sulle risposte della medesima. Studi come questo possono contribuire a migliorare la comprensione delle risposte della c.d. customer base nell'era dei Big Data – ciò che potrebbe svolgere un ruolo determinante per lo sviluppo di un mercato di consumatori sostenibile (Khatri, 2021; Kumar, 2020; Liang, Jiao & Liu, 2020; Moşescu, Chivu, Căescu et al., 2020; Puri & Mohan, 2020).

Blog, tweet, post e commenti sui social network forniscono un prezioso feedback alle aziende, posto che i Big Data costituiti dalle informazioni condivise sulle piattaforme social rivestono primaria importanza per la ricerca in settori complessi come il marketing, la politica, la salute o la gestione dei disastri. Siti del calibro di *Facebook* e *Twitter* sono ampiamente utilizzati per condurre affari, commercializzare prodotti e servizi e raccogliere opinioni e feedback sugli stessi. Posto che i dati raccolti dalle piattaforme social sono aggiornati in tempo reale e vengono, per lo più, forniti spontaneamente dagli utenti, essi tenderanno ad essere maggiormente "realistici" e a riflettere il sentiment generale rispetto ad uno specifico target. L'analisi della mole di dati così prodotti potrà, pertanto, condurre alla formulazione di intuizioni accurate, anche mediante l'implementazione di un sistema cloud per la pubblicità mirata, basata sull'analisi del sentiment dei tweet (Asghar, Ali, Ahmad et al., 2019; Das, Behera, & Rath, 2018; Farooqui & Ritika, 2019; Kumar, Koolwal & Mohbey, 2019; Nair & Shetty, 2017; Patil & Lokesha, 2022).

I dati provenienti dal Web, fornendo indicazioni suggestive dei trends di pensiero degli internauti, possono costituire una preziosa risorsa per ricerche di mercato funzionali al concepimento di nuovi prodotti di successo, ma anche per la formulazione di previsioni sull'andamento dei mercati finanziari online se non, addirittura, di eventi di natura politica, come dimostrato dalla campagna elettorale statunitense per le elezioni presidenziali del 2012 [77] (Kreiss, 2016; Kreiss & Jasinski, 2016; Kreiss & Saffer, 2017) Di certo, la dimensione politica ed elettorale degli ultimi anni è stata permeata dell'ascesa del fenomeno c.d. Technology-intensive Campaigning (Hendricks & Kaid, 2014; Kreiss, 2016) e dall'uso corrispondente dei Big Data da parte di numerose formazioni politiche (Bartlett, 2018; Chadwick & Stromer-Galley, 2016; Lachapelle & Maarek, 2015; Nickerson & Rogers, 2014).

L'articolo di Kreiss & McGregor (2018)[78] offre la prima indagine sistematica del ruolo che le aziende tecnologiche (in particolare *Facebook, Twitter, Microsoft e Google*) svolgono nel forgiare la comunicazione politica nelle competizioni elettorali degli Stati Uniti. Dall'analisi empirica dell'attività delle aziende tecnologiche nell'ambito della Election Policy - attraverso interviste con i dipendenti delle Big Techs e con i responsabili dei media digitali delle campagne per le elezioni primarie e generali degli Stati Uniti del 2016, integrate da

[77] In quel caso, entrambi gli sfidanti hanno fatto ampio ricorso alle tecnologie di Web Analytics per monitorare le reazioni degli elettori, adattando le proprie strategie, quasi in tempo reale, ai cambiamenti del sentiment dell'elettorato.

[78] Articolo cui si rinvia anche in ragione della ricchezza e varietà della bibliografia riportata.

osservazioni sul campo condotte durante la Convention Nazionale Democratica dello stesso anno - gli Autori hanno rilevato uno spiccato interesse di dette aziende per il marketing, le entrate pubblicitarie e la creazione di relazioni al servizio delle attività di lobbying. Inoltre, *Facebook*, *Twitter* e *Google* hanno dimostrato di plasmare attivamente la comunicazione delle campagne elettorali mediante la loro stretta collaborazione con gli staff politici, tanto da ipotizzare, per le stesse, il ruolo di "consulenti quasi digitali", poiché in grado di plasmare la strategia online, i contenuti e la corrispondente messa in atto. È interessante notare come, secondo Kreiss & McGregor, le piattaforme social meriterebbero di essere considerate quali agenti coinvolti a pieno titolo nei processi politici, in ogni caso, più di quanto sia stato apprezzato dalla letteratura precedente.

Sebbene l'ultimo decennio abbia assistito all'ascesa fulminea del c.d. Data-driven Campaigning come fulcro delle campagne politiche, le evidenze empiriche su tali pratiche risultano, al momento, ancora insufficienti per formulare robuste inferenze circa la pretesa "onnipotenza" dei dati stessi nel processo elettorale. Significativo, a tal proposito, risulta il lavoro di Baldwin-Philippi (2020), che sottolinea come la narrativa giornalistica circa l'uso della BDA si sia spesso basata su resoconti "gonfiati" in merito all'obiettività delle tecniche di analisi dei dataset, nella convinzione che un maggior numero di dati disponibili correli positivamente con una maggiore e migliore conoscenza dell'andamento del fenomeno elettorale. L'analisi critica del discorso sulla copertura collettiva della Data-driven Campaigning statunitense, negli anni 2008-2016 ha contribuito a depotenziare una simile narrativa – questione, peraltro, incidentalmente affrontata anche da una recentissima pubblicazione di Dommett, Barclay & Gibson (2023), nella quale gli Autori mettono in guardia da facili entusiasmi, sottolineando il crescente riconoscimento del fatto che le campagne elettorali "Data-driven" assumono forme diverse, in ragione di variabili come il contesto e le risorse economiche e finanziarie concretamente disponibili. Simili variazioni implicano differenti conseguenze sul piano democratico: mentre, infatti, taluni usi dei dataset sono da considerarsi legittimi (e, pertanto, consentiti), talaltri potrebbero sollevare preoccupazioni circa il rispetto dei principi di uno stato di diritto[79] (Bartlett, 2018; Bennett & Lyon, 2019; Kefford, Dommett, Baldwin-Philippi et al., 2023).

La questione appare controversa e assai dibattuta nella comunità accademica, come emerge dal lavoro di Simon (2019), il quale sostiene che, sebbene le campagne elettorali basate sui dati abbiano assunto importanza ormai da diversi anni, il know-how acquisito dalle aziende di settore appare tutt'ora modesto, segnatamente per ciò che riguarda la concezione dell'analisi dei dati, il targeting degli elettori e il loro ruolo nei processi elettorali. La principale criticità atterrebbe, in sostanza, al modo in cui tali questioni vengono concepite nella retorica di marketing dell'industria dell'analisi dei dati politici, evidenziando un fondamentale "scollamento" tra il discorso degli studiosi da un lato - spesso critico nei confronti delle affermazioni di dette aziende sull'efficacia dei loro metodi - e un immaginario dei dati altamente funzionale dall'altro, attivamente promosso dalla Political Data Analytics Industry e dai media compiacenti (Baldwin-Philippi, 2017). Secondo Dommett (2019), le ragioni di narrazioni tanto divergenti tra retorica e pratica sarebbero principalmente riconducibili agli incentivi che spingerebbero alcuni attori a mistificare il successo delle pratiche di campagna basate sui dati (Baldwin-Philippi, 2019) non ultimo il profitto derivante dalla vendita di

[79] Più in generale, è necessario riflettere sulle implicazioni di queste tendenze per la democrazia e sulla forma che potrebbe dover assumere qualsiasi risposta normativa in merito (Dommett, 2019).

applicazioni software dedicate (Bennett, 2016).

Ciò non significa, ovviamente, disconoscere il ruolo che la comunicazione assolve nell'ambito del processo elettorale (e politico, più ampiamente inteso): *"(…) the practices of political campaigns change as the communication environment changes"* sentenzia Stromer-Galley (2019, xi) nella Prefazione di un suo saggio di recente pubblicazione, dal titolo (evocativo) "Presidential Campaigning in the Internet Age". Del resto, le attività "data-driven" sono ormai considerate il *leitmotiv* nella gestione delle attuali campagne politiche ed elettorali di tutto il mondo, soprattutto da quando partiti politici ed attivisti hanno appreso l'uso dei dati per fornire campagne altamente mirate, strategiche e di successo (Anstead, 2017; Kefford, 2021; Kruschinski & Haller, 2017). Tuttavia, la persistente carenza di chiari parametri di riferimento, rispetto ai quali monitorare la forma e la portata delle campagne "data-driven", - si pensi alle crescenti preoccupazioni, più o meno giustificate dai risultati sperimentali, alimentate dal c.d. micro-targeting politico (Borgesius, Möller, Kruikemeier et al., 2018; Kreiss, 2017; Kruschinski & Haller, 2017) - suggerisce la necessità di una nuova comprensione concettuale ed empirica di tali pratiche, sia tra gli accademici che tra le autorità di regolamentazione (Dommett, 2019).

Le considerazioni che precedono dimostrano come gli impieghi della Big Data Analytics siano innumerevoli, sostanzialmente tesi alla profilazione dell'utente[80]: detta pratica, se per un verso è in grado di garantire all'utente/consumatore beni e servizi personalizzati - l'utilizzo appropriato dei dati può consentire agli individui di migliorare le previsioni, le indagini e le scelte (Shabana & Sharma, 2019) - per altro verso, risponde a logiche di potere e profitto proprie di operatori economici e politici presenti all'interno dell'ecosistema digitale globalizzato (Zuboff, 2019; Zuboff, Möllers, Wood et al., 2019), con conseguenze potenzialmente deleterie sul piano dei diritti individuali e dell'integrità delle istituzioni democratiche (Kapadia, 2020; Zuboff, 2022).

Più recenti, e tutt'ora in via di esplorazione, sono invece gli utilizzi del machine-learning da parte delle Pubbliche Amministrazioni: si pensi, ad esempio, alle disposizioni della Direttiva UE 2016/681 sull'impiego dei dati del codice di prenotazione aerea (PNR - Passenger Name Records) a fini di prevenzione di atti di terrorismo o di altri reati gravi, la quale espressamente prevede il ricorso al trattamento automatizzato di tali dati per la valutazione dei passeggeri sospetti e l'identificazione di quelli da sottoporre ad ulteriore verifica (questa volta mediante interventi non automatizzati, bensì umani) – questione che sarà oggetto di approfondita trattazione nel Capitolo 4, in tema di Big Data e Sicurezza Nazionale.

Oppure si pensi, ancora, alla diffusione della consulenza automatizzata in campo finanziario – trattasi della figura del c.d. robo-advisor, una sorta di consulente finanziario intelligente, a supporto della personalizzazione dei servizi e della gestione della tecnologia finanziaria (FinTech) - che ha indotto lo IOSCO (International Organizations of Securities Commissions, 2017) a proporre l'elaborazione di apposite Linee Guida per la mappatura dei rischi connessi alla diffusione di tali strumenti (Di Porto, 2017). In ogni caso, la robo-advisory - consulenza automatizzata sugli investimenti basata sul web (Anshari, Almunawar, & Masri, 2022; Fisch, Laboure & Turner, 2019) - ha incontrato una tiepida accoglienza da parte dei consumatori. Studi recenti suggeriscono che ciò potrebbe essere riconducibile a una combinazione di scarsa fiducia nelle banche, elevate aspettative di

[80] La capacità di profilatura dell'algoritmo (c.d. granularità) è direttamente proporzionale al volume delle "tracce digitali" prodotte dall'utente e da quelle raccolte attraverso i numerosi devices dello IoT (Di Porto, 2017).

trasparenza e generale incapacità o riluttanza a confrontarsi con le questioni relative agli investimenti (Cheng, Guo, Chen et al., 2019; Jung, Dorner, Weinhardt et al., 2018; Morana, Gnewuch, Jung et al., 2020; Oehler, Horn & Wendt, 2022; Tokic, 2018; Zhang, Pentina, & Fan, 2021).

Negli Stati Uniti, la FED [81] ha ipotizzato l'impiego del machine-learning per l'individuazione, in tempo reale, di pratiche di insider trading (Islam, Ghafoor, & Eberle, 2018). L'identificazione di attività di abuso di mercato a partire dal tracciamento delle operazioni di trading degli investitori risulta, invero, molto complessa, principalmente a causa del volume dei dati prodotti durante le transazioni (Mazzarisi, Ravagnani, Deriu et al., 2022). Pratiche del genere, che comportano l'impiego di algoritmi ad alta efficienza come il Gradient Boosting Decision Tree (GBDT), sono state di recente adottate anche nei mercati finanziari asiatici (Deng, Wang, Fu et al., 2021; Deng, Wang C., Wang M. et al., 2019; Deng, Wang, Li et al., 2019). Secondo la letteratura più recente, l'individuazione affidabile dell'insider trading rappresenta, a tutt'oggi, un ostacolo significativo, sia per la ricerca che per la pratica normativa. L'approccio d'indagine fondato sull'uso dell'apprendimento automatico sembra aumentare la capacità dei sistemi di sorveglianza di individuare valori anomali nelle transazioni (outliers), altrimenti sfuggenti (Lundblad, Yang & Zhang, 2022).

Anche il processo decisionale amministrativo è stato parzialmente rimesso alla regolamentazione robotica (Calo, 2017; Cobbe, 2019; Coglianese & Lehr, 2017; 2016). Gli algoritmi di apprendimento automatico stanno automatizzando compiti importanti in medicina, nei trasporti e nelle imprese, tanto che anche i funzionari governativi, stante l'accuratezza e la velocità che tali sistemi offrono, si affidano ad essi per supportare le principali funzioni del settore pubblico - amministrazione fiscale e supervisione normativa in primis (Coglianese & Lehr, 2019; Engstrom, Ho, Sharkey et al., 2020).

Orbene, posto che l'apprendimento automatico si è gradualmente esteso a tutti gli ambiti della società, l'emersione di preoccupazioni per l'intrusione di macchine algoritmiche in settori precedentemente rimessi al giudizio umano appare del tutto comprensibile, segnatamente nell'ipotesi della sostituzione degli esseri umani nella sfera governativa, come già da tempo documentato da Coglianese & Lehr (2017). I due Autori, dopo aver esaminato le conseguenze di un eventuale uso di strumenti decisionali robotici da parte delle agenzie governative, alla luce delle dottrine fondamentali e consolidate del diritto amministrativo e costituzionale statunitense - la dottrina della non delega, il giusto processo, la parità di protezione o i principi di ragionevolezza e trasparenza, in modo particolare – hanno concluso per la bontà di un impiego del machine-learning oculato e rispettoso dei parametri legali e convenzionali, foriero di benefici in termini di efficienza ed equità delle decisioni.

La questione, tuttavia, non è pacifica in dottrina, dove prevalgono opinioni discordanti e visioni a tratti pessimistiche. Berman, che rivendica "A Government of Laws and not of Machines" (questo il titolo del saggio del 2018), sottolinea criticità e limiti dell'impiego dell'Artificial Intelligence nell'attività di governo - primi fra tutti, opacità e arbitrarietà delle decisioni algoritmiche - invocando il coinvolgimento

[81] La Federal Reserve Bank (o FED) è la banca centrale responsabile della stabilità monetaria e finanziaria negli Stati Uniti. Fa parte di un sistema più ampio, noto come Federal Reserve System, composto da 12 banche centrali regionali, situate nelle principali città degli Stati Uniti (https://www.ig.com/it/glossario-trading/definizione-di-federal-reserve#:~:text=La%20Federal%20Reserve%20Bank%20(o,principali%20citt%C3%A0%20degli%20Stati%20Uniti).

del c.d. fattore umano nel processo decisionale. Altri Autori giungono a conclusioni del tutto simili, evidenziando come risultati empirici abbiano dimostrato che *"(…) Human-AI combination can be superior to AI acting alone (…)"* (Fagan e Levmore, 2019, 1). Non a caso, la definizione di nuovi tipi di interazione tra l'uomo e gli algoritmi di apprendimento automatico ("human-in-the-loop machine learning") integra la principale traiettoria di studio e di ricerca degli ultimi anni, a livello internazionale (Mosqueira-Rey, Hernández-Pereira, Alonso-Ríos et al., 2023; Munro & Monarch, 2021; Seidel, Berente, Lindberg et al., 2018; Wu, Xiao, Sun et al., 2022; Zanzotto, 2019).

Permangono, in ogni caso, timori generalizzati circa un "automated decision system" non sufficientemente regolamentato, soprattutto in ragione delle implicazioni sociali derivanti dall'impiego indiscriminato degli algoritmi di machine-learning – diritti e libertà individuali, opportunità sociali e public safety, in primis (Richardson, 2021) - da cui discende la pressante richiesta di un controllo giurisdizionale sui poteri decisionali automatizzati, segnatamente nel settore pubblico (Bayamlıoğlu & Leenes, 2018; Cobbe, 2019; Oswald, 2018).

La ricchezza e varietà della produzione scientifica sugli ambiti di applicazione dei Big Data è riassunta nei risultati riportati dalla poderosa analisi compiuta da Sunagar e colleghi nel 2020, che elenca i principali domini interessati dal fenomeno, ossia assistenza sanitaria, media e intrattenimento, IoT, settore manifatturiero e attività governativa. Grazie al ruolo assunto dai sistemi Big Data Systems, la medicina . personalizzata e l'analisi prescrittiva [82] stanno migliorando in modo significativo il settore dell'assistenza sanitaria. I ricercatori analizzano i dati per determinare il trattamento migliore per una particolare malattia, gli effetti collaterali dei farmaci, la previsione dei rischi per la salute, e molto altro. Le applicazioni mobili sui dispositivi sanitari e indossabili (Wearable Devices) favoriscono l'implementazione dei dati disponibili a un ritmo esponenziale. È possibile prevedere l'insorgere di una malattia attraverso la mappatura dei dati sanitari e dei dati geografici. Una volta prevista, è possibile gestire il contenimento dell'epidemia e pianificare l'eradicazione della malattia.

Le industrie dei media e dell'intrattenimento creano, pubblicizzano e distribuiscono i loro contenuti attraverso l'uso di nuovi modelli di business. L'esigenza dei clienti di visualizzare i contenuti digitali da qualsiasi luogo e in qualsiasi momento ha condotto all'introduzione di programmi TV online, canali on-demand ecc., pratiche grazie alle quali le case di produzione possono rivolgersi al pubblico con offerte personalizzate, così da incrementare i loro profitti (Bonner, Kureshi, Brennan et al., 2017; Kumar & Shindgikar, 2018). I siti di social media o i motori di ricerca analizzano intensamente i dati di cui possono disporre. Piattaforme come *Twitter* analizzano i tweet generati dai suoi utenti per identificare e confrontare i gruppi, analizzarne le abitudini o eseguire analisi del sentiment sul testo dei tweet. Allo stesso modo, *Facebook* è interessato al numero di "like" che una pagina ottiene nel corso del tempo e mantiene un contatore per gli URL raccomandati, per assicurarsi che passino meno di 30 secondi da un "clic" all'aggiornamento del relativo contatore. *Google* esegue costantemente il clustering del testo in *Google News*, così da accorpare notizie simili mostrandole affiancate nel

[82] L'analisi prescrittiva, che rappresenta una delle metodologie della BD Analytics, si avvale di simulazioni per elaborare suggerimenti con l'obiettivo di migliorare il processo decisionale dell'operatore. Essa tende a quantificare l'impatto delle decisioni future, consentendo alle organizzazioni di compiere scelte più efficienti ed evitare conseguenze negative. La stessa può essere utilizzata per il processo decisionale clinico in un ambiente sanitario basato sull'automazione dei dati e del flusso di lavoro in tempo reale (Shabana & Sharma, 2019; Tallman, Richardson, Rogow et al., 2023)

layout della pagina - solo per citare alcuni esempi di dette pratiche (Vossen, 2014).

I dispositivi IoT generano costantemente dati inviandoli, su base giornaliera, ad un server, che compie l'attività estrattiva di informazioni, così da implementare l'interconnettività dei dispositivi. La mappatura che ne deriva può essere impiegata dalle agenzie governative e da aziende del settore privato per potenziare le rispettive competenze. L'Internet of Things trova applicazione nei settori produttivi più disparati: dai sistemi di irrigazione intelligenti, alla gestione del traffico e della folla, nonché di altri eventi potenzialmente critici (Eikermann, Look, Roth et al., 2017).

Nel settore manifatturiero, i meccanismi di produzione predittiva possono contribuire ad aumentare l'efficienza, producendo più beni e minimizzando i tempi di inattività delle macchine – ciò comporta la disponibilità di un'enorme quantità di dati per tali industrie. Sofisticati strumenti di previsione seguono un processo organizzato per esplorare informazioni preziose su questi dataset. Tra i principali vantaggi dell'impiego delle applicazioni dei Big Data nelle industrie manifatturiere si annoverano l'alta qualità dei prodotti, il tracciamento di avarie del sistema, la pianificazione delle forniture, la previsione della produzione, l'aumento dell'efficienza energetica, i test e la simulazione di nuovi processi produttivi, nonché la personalizzazione della produzione su larga scala (Auschitzky, Hammer & Rajagopaul, 2014; Cui, Kara & Chan, 2020; O'Donovan, Leahy, Bruton et al., 2015).

Adottando i Big Data System, le agenzie governative possono conseguire l'efficienza in termini di costi, risultati e innovazioni. Poiché gli stessi dataset sono suscettibili di impiego in molteplici applicazioni, numerosi dipartimenti possono operare congiuntamente, massimizzando i benefici della collaborazione. Il governo centrale svolge un ruolo fondamentale nell'innovazione, agendo contestualmente in tutti i settori di interesse, tra i quali spiccano agricoltura, aviazione, sicurezza informatica e intelligence, prevenzione del crimine, e-commerce, rilevamento delle fake news, ricerca scientifica e farmacologica, previsioni meteorologiche e adempimenti fiscali (Al-Sai & Abualigah, 2017; Bertot, & Choi, 2013; Chen & Hsieh, 2014; Hardy & Maurushat, 2017; Kim, Trimi, & Chung, 2014; Löfgren & Webster, 2020; Pencheva, Esteve & Mikhaylov, 2020).

La sintetica rassegna della dimensione tecnologica dei Big Data si arricchisce di altri campi applicativi, rilevanti in tale contesto. Il primo è rappresentato dall'area dell'informatica statistica o statistica computazionale – disciplina al confine tra statistica e informatica, si occupa dello sviluppo di linguaggi di programmazione statistica e della progettazione e implementazione di algoritmi statistici, come quelli disponibili in pacchetti del tipo SPSS, SAS o STATA (Bartz–Beielstein, Parsopoulos & Vrahatis, 2004; Giudici, Givens & Mallick, 2013; Hoerl, Snee & De Veaux, 2014; Hunter, Krivitsky & Schweinberger, 2012; Piegorsch, Levine, Zhang et al., 2022).

La seconda area di interesse è il data mining, ovvero il processo di elaborazione di modelli (come regole di associazione o cluster) in grandi insiemi di dati – a partire dagli anni '90 del secolo scorso, il data mining ha acquisito popolarità come insieme di tecniche per estrarre conoscenza da dati grezzi (Han, Pei & Tong, 2022). La terza area, infine, corrisponde a quella della visualizzazione, relativa alla costruzione di rappresentazioni visive di dati numerici, testuali o geografici per rafforzare la cognizione umana e facilitare l'interpretazione dei dati o dei risultati di calcolo da essi derivanti (Agrawal, Kadadi, Dai et al., 2015; Ali, Gupta, Nayak et al., 2016; Bikakis, 2018; Bikakis, Papastefanatos & Papaemmanouil, 2019; Keim, Qu, & Ma, 2013).

Altri domini coinvolti dall'implementazione degli algoritmi per l'analisi dei Big Data sono quelli delle applicazioni aziendali e

ingegneristiche ad esse correlate; dell'assistenza sanitaria – segnatamente, quella associata alla pandemia COVID-19 a partire dai primi anni del 2020 – delle scienze comportamentali e delle Information and Communication Technologies (ITC), settori rispetto ai quali la comunità accademica internazionale sta manifestando un vivace interesse (Acharjya & Ahmed, 2016; Ahmed, 2017; Bhatt, Sengar & Pandey, 2022; Bendre & Thool, 2016; Li, Kong, Zheng et al., 2022; Pyne, Rao & Rao, 2016; Watson, 2019). Vero è che quelle sopracitate rappresentano aree scientifiche e tecnologiche in forte ascesa, destinate a rappresentare la prossima generazione di sistemi informatici basati sulla Artificial Intelligence. Trattasi di domini emergenti, in cui i processi trasformativi transdisciplinari in atto ne consentono la costante crescita, sia in termini di fondamenti teorici che di applicazioni empiriche (Martis, Gurupur, Lin et al., 2018; Nasser & Tariq, 2015).

Per quanto concerne le applicazioni aziendali e ingegneristiche correlate, basti pensare alla previsione, riportata da Nasser & Tariq quasi una decade fa, secondo la quale il 90% dei dirigenti riteneva che i dati sarebbero divenuti il quarto fattore di produzione per le aziende, essenziale al pari dei tradizionali fattori terra, lavoro e capitale (Power, 2014). L'analisi dei mercati bancari e finanziari asiatici compiuta da Shi (2022), unitamente alle evidenze derivanti dalle banche commerciali cinesi quotate in borsa, ha ampiamente confermato dette previsioni, soprattutto sul versante dell'impiego della BDA nella diversificazione del portfolio creditizio, come peraltro preconizzato da studiosi di settore (Chen, Shi, Wei et al., 2014).

Significative anche le applicazioni nell'ambito dell'ingegneria civile, ambientale e dei trasporti, grazie all'installazione, da parte delle compagnie ferroviarie, di sensori all'interno dei vagoni ferroviari per il tracciamento del grado di usura delle varie componenti tecniche, nonché di tutti gli eventi critici sopportati, così da ridurre la possibilità di sinistri ferroviari [83] (Pettit, Zarpelon Leao, Lock et al., 2022). L'Intelligenza Artificiale è stata applicata anche alla rete ferrotranviaria - con l'installazione di sensori in situ per il monitoraggio online degli eventi esterni (Tranos & Mack, 2019; Welch & Widita, 2019) - al pari della rete stradale, dotata di sistemi di rilevamento di eventi atmosferici e meteorologici avversi. L'uso generalizzato di tecnologie wireless ha condotto alla creazione dei c.d. sistemi di trasporto intelligenti (ITS): le attività di raccolta, archiviazione e analisi dei dati ha, infatti, consentito il potenziamento della pianificazione e della sicurezza nella gestione della mobilità (Mohammed, Arabnia, Qu et al., 2020; Stylianou, Dimitriou, & Abdel-Aty, 2019; Torre-Bastida, Del Ser, Laña et al., 2018).

Circa l'impiego della Big Data Analytics nel settore diagnostico, trattamentale e di assistenza sanitaria concomitante alla crisi pandemica da COVID-19, la letteratura è molto ampia e variegata, soprattutto per diversità di approcci, che spaziano da quello puramente clinico a quello manageriale. La complessità dell'argomento, unita alla mole di letteratura prodotta nell'ultimo triennio, richiederebbe una trattazione specifica, estranea alle finalità del presente lavoro. Sia consentito, pertanto, circoscrivere l'analisi ad una serie di considerazioni di carattere generale, emerse da un sommario esame delle principali pubblicazioni sul tema [84], indicizzate dal motore di ricerca *Google Scholar*.

[83] Le letture vengono memorizzate e analizzate, così da essere utilizzate in seguito, per identificare, ad esempio, i componenti suscettibili di riparazione/sostituzione prima di causare ulteriori danni.

[84] Sul piano metodologico, è stata operata una selezione di articoli primari sul tema, i cui contenuti sono stati successivamente ampliati mediante una serie di riferimenti secondari.

La letteratura esaminata concorda nel riconoscere la centralità delle tecnologie digitali nella prevenzione, nel controllo della diffusione della malattia e nella gestione sociale del fenomeno pandemico da COVID-19. La nuova generazione di tecnologie informatiche è stata, invero, ampiamente utilizzata nelle varie fasi della pandemia come importante supporto di base. Posto che la prevenzione e il controllo delle malattie infettive ad esordio improvviso dipendono da fattori come il controllo delle fonti di infezione, l'interruzione dei canali di trasmissione e la vaccinoprofilassi, i Big Data e l'Artificial Intelligence (AI) si sono dimostrati strumenti efficaci per identificare la fonte di infezione, assumendo un ruolo insostituibile nel tracciamento dei contatti e dei gruppi di popolazione sospetti [85] (Alsunaidi, Almuhaideb, Ibrahim, et al., 2021; Bragazzi, Dai, Damiani et al., 2020; Dong, Wu, Zhou et al., 2021; Haafza, Awan, et al., 2021; Hassanien, Dey, Elghamrawy et al., 2020; Mehta & Shukla, 2022; Park, 2022; Shuja, Alanazi, Alasmary et al., 2021).

L'analisi computazionale avanzata ha mostrato le sue potenzialità nella riduzione dei tempi di ricerca e sviluppo dei presidi vaccinali, oltre che nel miglioramento della qualità degli stessi. L'AI ha fornito un supporto significativo nell'elaborazione automatica di dati rilevanti provenienti da documentazione clinica, test e risultati di esami di strumentali e di laboratorio, tanto nella fase predittiva della progressione e della prognosi della malattia quanto nella raccomandazione di piani e strategie trattamentali (Awotunde, Folorunso, Jimoh et al., 2021; Pham, Nguyen, Huynh-The et al., 2020; Wu, Wang, Nicholas et al., 2020).

Oltre che rispetto al contenimento della crisi pandemica da COVID-19, le tecnologie digitali hanno dato prova del loro potenziale anche nella gestione di "major public health incidents" conformemente ai risultati dei lavori di Al-Sai, Husin, Syed-Mohamad et al., 2022; Corsi, de Souza, Pagani et al., 2021; Jia, Guo, Wang et al., 2020, concordi nel concludere per un valido contributo dei Big Data e della tecnologia AI alla prevenzione, diagnosi, trattamento e processo decisionale di gestione di futuri eventi critici di salute pubblica ad esordio improvviso (Dong, Wu, Zhou et al., 2021).

I vantaggi della BD Analytics nella ricerca scientifica, e nel settore medico in particolare, erano già stati segnalati in ambito diagnostico, con specifico riferimento alle patologie oncologiche. L'algoritmo LYNA (LYmph Node Assistant) - un applicativo in grado di analizzare immagini tissutali di pazienti oncologici, segnalando la presenza di metastasi cancerose anche di dimensioni ridotte - è stato oggetto di uno studio condotto da un pool di ricercatori della California nel 2019, fornendo risultati assolutamente incoraggianti: *"(…) LYNA achieved a slide-level area under the receiver operating characteristic (AUC) of 99% and a tumor-level sensitivity of 91% at 1 false positive per patient on the Camelyon16 evaluation dataset. We also identified 2 "normal" slides that contained micrometastases. When applied to our second dataset, LYNA achieved an AUC of 99.6% (…)*". Altrimenti detto, comparando la performance dell'equipe medica nella diagnosi di malattia, rispettivamente con e senza l'ausilio dell'algoritmo, la sensitività passava dall'83% (senza LYNA) al 91%, con il dimezzamento del tempo necessario per porre diagnosi (Liu, Kohlberger, Norouzi et al., 2019; Ngiam & Khor, 2019). Il titolo di un contributo di Nqaba Matshazi (2018) restituisce perfettamente la potenzialità dello strumento algoritmico: *"Google's AI's LYNA better than humans in detecting advanced breast cancer"*.

La rassegna in merito agli attuali impieghi delle tecnologie informatiche si conclude con

[85] Com'è stato opportunamente rilevato da Dong e colleghi (2021), ii risultati relativi all'applicazione dei Big Data e delle tecnologie di AI ad eventi improvvisi di salute pubblica mancano di una convalida della ripetibilità e dell'universalità,

il richiamo al versante del Law Enforcement, ambito nel quale l'uso del Big Data Analytics sembra destare la maggiore attenzione della comunità scientifica e accademica a livello internazionale. Di Porto (2017) riferisce dell'uso di detti strumenti tecnologici per finalità di polizia amministrativa come accade negli Stati Uniti, dove la programmazione di controlli e ispezioni ha acquisito maggiore efficacia da quando si avvale delle tecniche di machine-learning, con un notevole risparmio di risorse. Già da tempo, la città di Chicago utilizza simili tecnologie per definire il calendario delle ispezioni igieniche nei ristoranti – ciò che ha consentito l'invio selettivo degli ispettori sanitari, con priorità per gli stabilimenti a più alto rischio, permettendo l'accertamento di violazioni con una riduzione della tempistica del 25% rispetto al passato, secondo le stime del Commissario del *Chicago Department of Innovation and Technology* (Rabasca Roepe, 2019).

A livello federale, Coglianese & Lehr (2017) riportano la pratica dell'*Internal Revenue Service* (IRS) statunitense di avvalersi delle reti neurali per identificare possibili nuove aree di non compliance fiscale. Neuman & Sheu (2022), dal canto loro, chiariscono che, sebbene l'analisi dei Big Data possa essere considerata una "panacea" per l'IRS – la BDA, infatti, consente la profilazione dei contribuenti per accertarne più efficacemente la non conformità fiscale utilizzando l'AI e il machine-learning e riducendo, al tempo stesso, i costi dell'attività di accertamento - detta strategia non è scevra da profili di criticità, primi fra tutti la violazione della privacy, le pratiche di informazione scorretta e i pregiudizi ad essa sottesi, con effetti significativi sull'equità percepita (Black, Elzayn, Chouldechova et al., 2022; Brezina, Eberhartinger, & Zieser, 2021; Houser & Sanders, 2018). Sebbene le agenzie governative di tutto il mondo utilizzino algoritmi basati sui dati per allocare le risorse atte a garantire l'applicazione della legge - quand'anche tali algoritmi appaiano formalmente neutri rispetto a caratteristiche protette come la razza - è ampiamente diffusa la preoccupazione che essi possano gravare in modo sproporzionato sui gruppi vulnerabili. Gli outcomes di studi appena pubblicati evidenziano, purtroppo, come le scelte, apparentemente tecnocratiche, di progettazione algoritmica in materia di tassazione possano incorporare valori politici e compromessi (Elzayn, Smith, Hertz et al., 2023).

A New York, la prevenzione degli incendi negli edifici è gestita attraverso il Mayor's Office of Data Analytics (MODA), che fa uso del machine-learning per l'invio degli ispettori incaricati dei controlli preventivi. Il MODA è il centro di intelligence civica di New York City, che consente l'aggregazione e la successiva analisi dei dati provenienti da tutte le agenzie cittadine, per affrontare in modo più efficace i problemi di criminalità, sicurezza pubblica e qualità della vita. L'ufficio utilizza strumenti di analisi per assegnare priorità ai rischi in modo strategico, così da garantire una fornitura di servizi più efficiente, un più efficace rispetto delle leggi e un incremento della trasparenza dell'attività amministrativa (www.nyc.gov, 2023).

Inserita nella Carta della città di New York nel 2018, l'attività dell'Ufficio si articola in tre fasi principali, riassumibili nell'analisi dei dati per il miglioramento delle operazioni delle agenzie cittadine; nella fruibilità degli open data da parte dei cittadini e nella condivisione delle infrastrutture digitali preposte. Come riportato dal sito istituzionale, il lavoro di analisi svolto da MODA spazia da rapide interrogazioni di dati a iniziative strategiche a lungo termine, progettate per implementare il processo decisionale basato sui dati nell'amministrazione della città di New York. Nel corso di un progetto di analisi, MODA lavora a stretto contatto con l'Agenzia competente per comprendere il problema operativo e i dati associati, progettare un modello analitico applicabile e, infine, fornire informazioni utili per migliorare le

operazioni (modaprojects.cityofnewyork.us, 2023).

Ispirato ai principi dell'efficienza e dell'equità nella gestione delle maxi emergenze è stato l'intervento durante l'uragano "Sandy"[86] del 2012 – una tempesta tropicale che ha colpito i Caraibi e la costa orientale degli Stati Uniti nell'autunno del 2012, rendendo necessaria l'evacuazione di un ottavo della popolazione di New York City. In quella circostanza, la risposta è stata gestita attraverso l'impiego della Big Data Analytics, principalmente mediante la classificazione dei testi pubblicati dagli utenti sui social media (es. Twitter) per acquisire consapevolezza della situazione sul territorio in tempo reale (Yu, Huang, Qin et al., 2020; Neppalli, Caragea, Squicciarini et al., 2017; Shelton, Poorthuis, Graham et al., 2014; Wang & Zhuang, 2017). L'apporto di MODA si è rivelato di fondamentale importanza, dal momento che l'Ufficio cittadino ha integrato le risposte dei residenti ai sondaggi per l'allocamento delle risorse, destinandole proporzionalmente alla vulnerabilità delle vittime del disastro [87] (Kontokosta & Malik, 2018; Mudigonda, Ozbay, & Bartin, 2019).

L'impiego dei sistemi di machine-learning si è dimostrato cruciale anche per la valutazione della resilienza delle infrastrutture critiche in caso di eventi climatici avversi o calamità naturali di ampia portata, come quello analizzato dal case-study sull'uragano Florence a Wilmington, North Carolina, USA. Una complessa ricerca compiuta da Yuan e colleghi ha introdotto un framework abilitato all'Internet of People (IoP) per valutare la perdita di prestazioni della rete stradale durante i disastri, proponendo il corrispondente tasso di riduzione come parametro per valutare la resilienza della rete stradale [88]. Ebbene, i risultati hanno dimostrato come il framework abilitato all'IoP sia in grado di supportare efficacemente la protezione delle infrastrutture critiche per la costruzione di ambienti cittadini intelligenti e resilienti (Yuan, F., Liu, R., Mao et al., 2021).

Nello stesso filone di ricerca si collocano lavori simili per approccio, che evidenziano come l'impiego del patrimonio informativo derivante dall'analisi dei contenuti delle piattaforme digitali consenta una valutazione tempestiva e affidabile dell'impatto dei disastri sulle arterie stradali, necessaria per effettuare le evacuazioni, fornire servizi di emergenza e pianificare le attività di soccorso e recupero delle vittime. Nella ricerca di Chen, Wang & Ji (2020), due sono stati gli indicatori basati sui social media - ovvero il numero di tweet relativi all'impatto (NIT) e la geolocalizzazione dei tweet relativi all'impatto (GIT) - progettati per indicare, rispettivamente, la gravità e la localizzazione

[86] Da notare come gli uragani "Sandy", "Harvey" "Irma" e "Irene" siano stati ampiamente impiegati dalla comunità scientifica quali casi di studio per la comprensione del fenomeno in chiave migliorativa (Huang, Dong, Yesha et al., 2014; Yu, Huang, Qin et al., 2020; Pourebrahim, Sultana, Edwards et al., 2019; Preis, Moat, Bishop et al., 2013).

[87] Ancora in risposta all'uragano "Sandy", la società statunitense *Direct Relief* ha applicato la tecnologia dei Big Data per coordinare le attività di soccorso. Avvalendosi di software di analisi e mappatura forniti da partner tecnologici leader nel settore, Direct Relief ha distribuito le risorse appropriate nelle aree critiche, a partire dall'analisi della vulnerabilità sociale e del rischio sanitario e proseguendo con le indagini meteorologiche e il monitoraggio costante dello stato delle infrastrutture critiche (https://www.directrelief.org/emergency/hurricane-sandy/).

[88] In sostanza, attraverso l'analisi semantica dei dati estratti dai social media relativi alla rete stradale, lo studio ha identificato, innanzitutto, i percorsi stradali colpiti durante l'uragano; l'analisi della rete stradale ha, successivamente, consentito di generare una zona di servizio per ciascun ospedale presente nell'area di interesse prima e dopo l'evento climatico avverso. I tempi di percorrenza ponderati per la popolazione in corrispondenza di ciascuna zona di servizio sono stati confrontati prima e dopo l'uragano per indicare il grado di perdita di prestazioni.

dell'impatto dei disastri sulle rete autostradale, dimostrando un apprezzabile grado di affidabilità. La letteratura sul tema si sta sviluppando, nel corso degli ultimi anni, soprattutto negli Stati Uniti, dove eventi climatici dagli esiti potenzialmente catastrofici impongono una gestione delle emergenze rapida nella risposta ed efficiente nelle modalità, segnatamente rispetto ai danni alle c.d. infrastrutture critiche (Chen, & Ji, 2021; Fan, Yao, Mostafavi et al., 2019; Kocatepe, Ulak, Sriram et al., 2018; Roy, Hasan & Mozumder, 2020; Zhang, Yao, Yang et al., 2020; Zhu, Ozbay, Xie; 2016).

L'impiego di soluzioni basate sul machine-learning nella regolazione e nell'attività di Law Enforcement ampiamente intesa trova compiuta espressione nel programma EMBERS (Early Model Based Event Recognition Using Surrogates), un progetto sviluppato dell'Agenzia di Intelligence statunitense "Advanced Research Projects". Lo scopo del programma è quello di sviluppare previsioni di eventi critici – come disordini civili, malattie, proteste, focolai ed esiti elettorali - fornendo "intelligence preventiva" su detti eventi, mediante la scansione di indicatori open source e la costante ottimizzazione delle potenzialità dell'applicativo per rilevare nuove tipologie di indicatori (Chi, 2017; Discovery Analytics Center, 2016, 2014; Roff, 2020).

Trattasi di un sistema di analisi di Big Data su larga scala per la previsione di eventi sociali significativi, sulla base di grandi volumi di dati, prodotti in maniera continua e automatizzata, pubblicamente disponibili. EMBERS ha operato come *proof of concept* dall'agosto 2012 al luglio 2016[89](Doyle, Katz, Summers et al., 2016), raccogliendo quasi una dozzina di fonti di dati di dimensioni variabili - dai rapporti governativi settimanali a *Twitter* [90], da notizie, blog, eventi codificati dalla macchina, fino a tassi di cambio e prezzi dei prodotti alimentari - con un feed di informazioni completo che generava quotidianamente circa 19,2 gigabyte di dati provenienti da fonti spagnole, portoghesi e inglesi, con focus geografico sul Sudamerica. Il sistema era progettato per individuare segnali precursori nei flussi di social media (Kallus, 2014), utilizzando tali indicatori per guidare algoritmi statistici e di apprendimento automatico, così da generare le previsioni (Doyle, Katz, Summers et al., 2014).

I feed grezzi erano stati successivamente arricchiti utilizzando indicatori per la ricerca di persone, luoghi, organizzazioni e altre caratteristiche, come numeri, date e hashtag nel testo, geocodifica e analisi finale del sentiment collettivo – ciò che ha comportato un ampliamento del volume dei dati elaborati fino a 40 gigabyte al giorno. EMBERS è unanimemente considerato un esempio di "Anticipatory Intelligence", all'epoca ultima frontiera della BD Analytics (Doyle, Katz, Summers et al., 2014; Ramakrishnan, Butler, Muthiah et al., 2014).

Le classi di eventi che EMBERS era progettato per prevedere includevano conteggi di casi di malattie simil-influenzali, epidemie di malattie rare, elezioni, crisi politiche interne e disordini civili. Per quest'ultima categoria in particolare, EMBERS produceva previsioni dettagliate su eventi futuri, tra cui la data, il luogo, il tipo di evento (ad esempio, una protesta per i salari o per la sicurezza) e il sottogruppo di popolazione coinvolta nell'evento (ad esempio, educatori, operai, medici), fornendo anche indicazioni utili in merito agli elementi di incertezza implicati nella

[89] Nel periodo di operatività il sistema forniva circa 50 previsioni al giorno per i paesi dell'America Latina (Chi, 2017).

[90] Il sistema era in grado di elaborare una serie di dati rumorosi ad alto volume e ad alta velocità, come Twitter insieme a fonti di volume inferiore e di qualità superiore, come gli indicatori economici (Doyle, Katz, Summers et al., 2014).

previsione (Doyle, Katz, Summers et al., 2014).

Tra le principali considerazioni ad avere motivato la progettazione di EMBERS spiccava il necessario coinvolgimento di un team di lavoro multidisciplinare: otto istituti universitari di ricerca e due partner industriali che contribuivano con differenti competenze in discipline come informatica, apprendimento automatico, modellazione di malattie, scienze sociali, elaborazione linguistica e integrazione di sistemi. All'originaria attenzione rivolta ai paesi dell'America Latina aveva fatto seguito l'espansione alla regione del Medio Oriente e al Nord Africa (MENA) (Cadena, Korkmaz, Kuhlman et al., 2015). Al terzo anno di attività, il sistema era stato in grado di selezionare indicatori di eventi critici nei contenuti dei social media, producendo previsioni corrispondenti sia alla tempistica degli eventi che alla loro traiettoria in termini di dimensioni e intensità. Emblematica la previsione della serie di proteste in Brasile, nel giugno 2013, e delle rivolte guidate dagli studenti in Venezuela, nel febbraio 2014 (Doyle, Katz, Summers et al., 2014; Muthiah, Butler, Khandpur et al., 2016).

Sebbene sia sorta come una piattaforma di ricerca, EMBERS può avere molteplici usi, sia in ambito industriale che governativo: dagli avvisi su hotspot e potenziali interruzioni della catena di approvvigionamento di beni e servizi fino all'individuazione delle lamentele dei cittadini cui dare priorità, perché potenzialmente indicative di proteste imminenti. Del resto, lo studio dei disordini civili rappresenta un argomento-chiave per gli scienziati politici, poiché aiuta a cogliere un importante meccanismo di espressione dei cittadini (Muthiah, Huang, Arredondo et al., 2015). La gestione proattiva dei disordini civili richiede, tuttavia, una costante valutazione del rischio (Qiao, Li, Zhang, et al., 2017), che rende ragione dell'analisi del c.d. microblogging (es. tweet e messaggistica istantanea) – pratica riferita già da tempo dalla letteratura specialistica come fonte di informazioni preziose per la previsione di eventi critici (Alsaedi, Burnap & Rana, 2017; Hua, Chen, Zhao et al., 2013; Kireyev, Palen & Anderson, 2009; Morstatter, Kumar, Liu, et al., 2013).

Dopo 4 anni di operatività del sistema, EMBERS si è dimostrato un affidabile strumento di previsione di eventi di disordine civile in 10 paesi diversi e in tre lingue principali, ma ha evidenziato, altresì, profili di criticità, soprattutto sul piano etico. EMBERS, in quanto sistema di intelligence anticipatoria, vanta molteplici usi legittimi ma è anche una potenziale fonte di abusi, posto che il diritto di manifestare il proprio dissenso rappresenta un valore meritevole di tutela in qualsiasi paese ispirato a valori democratici. Per converso, nell'odierna società interconnessa, le proteste causano gravi disagi, come interruzioni nella catena di approvvigionamento, nella mobilità e in altri settori cruciali; pertanto, anticipare dette interruzioni si rivela di fondamentale importanza per garantire la sicurezza e l'affidabilità della vita civile (Muthiah, Butler, Khandpur et al., 2016)

Sulla scorta di tali considerazioni, la comunità scientifica continua ad interrogarsi in merito alle potenzialità e ai limiti che qualsiasi forma di "Intelligence anticipatoria" pone alla collettività e ai decisori politici, quando sia necessario condurre valutazioni strategiche circa minacce basate su notizie ed eventi globali (Halterman, Schrodt, Beger, et al., 2023; Hosseini, Chen, Yang et al., 2022; Sufi, 2022; Treverton & Gabbard, 2008; Triebe, Hewamalage, Pilyugina et al., 2021; Zhao, 2021).

4 Big Data, Sicurezza Nazionale e Minacce Asimmetriche

Ancor più complesso e articolato lo scenario che si delinea a seguito dell'impiego

del Big Data Analytics nel contesto della sicurezza nazionale, con specifico riferimento alla portata delle conseguenze che una gestione inappropriata di tale tecnologia emergente può provocare. È, invero, opinione diffusa che la comunità della sicurezza nazionale sia chiamata a gestire un approccio "all-hazards" anche rispetto alla governance dei Big Data, posto che la rapidissima generazione di enormi volumi di dati è tale da travolgere anche i tradizionali sistemi di gestione di database e software analitici delle Agenzie governative. Non a caso, le tendenze attualmente osservabili nel settore di BD chiamano in causa la combinazione di dati e sofisticate tecnologie di machine-learning, prima fra tutte l'Artificial Intelligence (Bao En, n.d.; Chi, 2017).

Un dettagliato studio in materia è stato rilasciato dal Centro di Ricerca "Australian Strategic Policy Institute" nel 2017, distinguendosi tanto per chiarezza d'impostazione quanto per ricchezza di risorse bibliografiche e richiami alla letteratura specialistica [91] È interessante notare, innanzitutto, l'approccio critico alla tematica, incentrato sull'esplorazione delle principali sfide correlate alla BDA: sebbene la comunità australiana della sicurezza nazionale nutra particolari aspettative derivanti dall'applicazione dei Big Data, pacifica è la convinzione che ciò comporti limitazioni specifiche e rischi aggiuntivi rispetto all'impiego di detta tecnologia emergente nel settore privato . Orbene, due sarebbero "i principali imperativi" a generare preoccupazione tra gli esperti: rispettivamente, la condivisione delle informazioni ("information sharing") e la gestione del sovraccarico delle medesime, in quanto provenienti da molteplici fonti ("master the all-source information overload") (Australian Strategic Policy Institute, 2017, 15; O'Hara & Bergin, 2009).

Sul piano della "information sharing", merita sottolineare come la principale censura mossa alla comunità di sicurezza nazionale degli Stati Uniti all'indomani degli attacchi alle Twin Towers riguardasse proprio lo iato tra la raccolta di informazioni e la (mancata) attività di investigazione preventiva da parte delle Forze di Polizia (Committee on Homeland Security and Governmental Affairs, 2011). Successivamente, gli investigatori avrebbero sostenuto che la comunità di Intelligence e quella di Law Enforcement disponevano di molteplici indicatori ("red flags") che, se opportunamente condivisi tra le Agenzie e analizzati in forma aggregata, avrebbero consentito di emanare un allarme tempestivo degli attacchi dell'11 settembre 2001 (National Commission on Terrorist Attacks upon the United States, 2004; Krebs, 2002; Schwartz, 2015). In realtà, quella della mancata condivisione delle informazioni e dell'assenza di coordinamento sembra costituire un fattore critico comune alla maggior parte dei c.d. "Intelligence failures" (Criminal Intelligence Service Canada, 2007). In ogni caso, la produzione scientifica sul tema si è, comprensibilmente, intensificata nello scenario globale "post 9/11 attacks", documentando un progressivo interesse per le applicazioni dei Big Data nel campo dell'antiterrorismo (Gundabathula & Vaidhehi, 2018; Huamaní, Alva, & Roman-Gonzalez, 2020; Verma, Malhotra, Verma et al., 2018; Nie & Sun, 2016; Singh, Chaudhary & Kaur, 2019; Talreja, Nagaraj, Varsha et al., 2017; Vajjhala, Strang, & Sun, 2015).

Strang & Sun (2017), in particolare, utilizzando il software di BD *Hadoop*, hanno evidenziato una relazione significativa tra ideologia terroristica e tipologia di attacco – scoperta, quest'ultima, che potrebbe rivelarsi cruciale rispetto alla pianificazione della sicurezza nazionale – al pari di altri Autori, che hanno enfatizzato l'importanza di

[91] Per una visione diacronica della tematica, si rinvia al Report della Australian Strategic Policy Institute (2010), al Report periodico del Data to Decisions Cooperative Research Centre (2017), al contributo di Street, Brady & Moroney (2008) e all'analisi "futuristica" di Behm (2007).

elaborare modelli predittivi e sistemi di allerta precoce antiterrorismo, così da implementare la gestione del rischio minimizzando il relativo margine di incertezza (Agarwal, Sharma & Chandra, 2019; Gohar, Butt & Qamar, 2014; Khorshid, Abou-El-Enien & Soliman, 2015; Kumar, Mazzara, Messina et al., 2020; Meng, Nie & Song, 2019; Soliman & Abou-El-Enien, 2019; Tolan & Soliman, 2015).

L'impiego di sistemi di apprendimento automatico (machine-learning) per la generazione di modelli predittivi di attacchi terroristici è stato oggetto di accurata analisi da parte di Verma e colleghi (2018), i quali hanno concluso per l'elevato grado di precisione dei modelli elaborati, sulla base dei quali è stata riscontrata una correlazione positiva tra la tipologia di attacco e il tipo di arma utilizzata. È interessante notare come, rispetto alle categorie di armi biologiche, chimiche e radiologiche, il modello sia in grado di prevedere, oltre alla tipologia di attacco, anche il profilo del potenziale autore. Con specifico riferimento all'attacco a infrastrutture critiche, la BDA ha evidenziato il frequente impiego di armi sconosciute o non convenzionali.

L'alta frequenza degli attacchi terroristici e l'intrinseca difficoltà di individuazione degli autori hanno motivato Pan (2021) ad elaborare un framework di classificazione basato sul machine-learning per identificare le organizzazioni terroristiche coinvolte. I risultati sperimentali dell'analisi statistica e quantitativa delle attività delle organizzazioni presenti nel Global Terrorism Database (GTD), dal 1970 al 2017, hanno evidenziato le eccellenti prestazioni dei modelli elaborati, con punte di valori fino al 97,16% per la previsione di ben 32 organizzazioni terroristiche con la più alta frequenza di attacchi perpetrati. Il framework di classificazione proposto si è dimostrato utile per la previsione accurata ed efficiente delle organizzazioni responsabili di attacchi, potendo essere esteso all'identificazione delle organizzazioni terroristiche tout court (Sachan & Roy, 2012), così da anticipare la soglia di rischio di eventi avversi. L'ottima performance derivante dall'applicazione congiunta di tecniche di data mining per l'analisi degli attacchi terroristici globali è stata, peraltro, avvalorata dai recenti risultati riportati dalla letteratura internazionale (Agarwal, Sharma & Chandra, 2019; ALfatih, Li, & Saadalla, 2019; Kumar, Mazzara, Messina et al., 2020; Mohammed & Karabatak, 2018; Singh, Chaudhary & Kaur, 2019; Talreja, Nagaraj, Varsha et al., 2017; Tolan & Soliman, 2015).

Studi più datati hanno riportato analisi comparative per la verifica dell'accuratezza predittiva di differenti algoritmi di apprendimento automatico - l'intento dichiarato dei ricercatori era quello di formulare previsioni attendibili in merito ai gruppi terroristici responsabili di attacchi in Medio Oriente e Nord Africa, nel quinquennio 2004-2008 - anticipando i risultati delle ricerche successive, ossia gli elevati livelli di efficienza garantiti dal machine-learning in ambito predittivo (Gundabathula & Vaidhehi, 2018; Khorshid, Abou-El-Enien & Soliman, 2015; Tolan & Soliman, 2015).

Recentissimi lavori sull'argomento, che si sono avvalsi della raccolta di dati relativi ad incidenti terroristici in Nigeria per formulare analisi predittive sul territorio nazionale applicando modelli di apprendimento automatico (MLM), hanno confermato la bontà della performance dell'algoritmo per il conseguimento dell'obiettivo dichiarato. Conformemente a tale modello, il successo di un attacco terroristico dipenderebbe da fattori - listati in ordine di rilevanza - quali il numero di autori, il tipo di attacco, il tipo di arma, il target (ossia, la tipologia di vittima prescelta), il numero di vittime e le caratteristiche tecniche dell'evento critico (Odeniyi, Adeosun & Ogundunmade, 2022).

Huamaní, Alva & Roman-Gonzalez (2020) a partire dal GTD Dataset - il già citato database contenente una raccolta sistematica degli attacchi terroristici globali,

a partire dal 1970 fino all'ultimo anno di attività registrata, ovvero il 2018 – si sono avvalsi delle tecniche di AI per prevedere possibili attacchi terroristici, grazie all'impiego di specifici modelli algoritmici di classificazione. Posta la necessità uno studio sistematico delle applicazioni dei Big Data al settore dell'antiterrorismo (Nie & Sun, 2016), l'analisi predittiva ha dimostrato la necessità di una conoscenza approfondita di elementi peculiari, quali il numero di incidenti terroristici nel mondo, le tipologie di attacchi più frequenti oltre al numero di sequestri perpetrati per area territoriale – ciò che potrebbe favorire una risposta mirata ed efficace, da parte della comunità internazionale, rispetto a potenziali, futuri eventi terroristici a livello globale. Vero è che l'elaborazione di modelli predittivi di minacce asimmetriche alla sicurezza nazionale rappresenta una sfida impegnativa e, al tempo stesso, un'area di ricerca promettente, come documentato dal crescente interesse accademico per la complessa tematica maturato nell'ultima decade (Akhgar, Saathoff, Arabnia et al., 2015; Krishnamurthy & Desouza, 2014; Landon-Murray, 2016; Lyon & Wood, 2020; Van Puyvelde, Coulthart & Hossain, 2017).

L'elaborazione di efficaci indicatori correlati ai c.d. warning systems richiede, tuttavia, il coordinamento di disparate fonti di informazione – quello che è stato descritto come "effetto puzzle" o "Humpty Dumpty", riferendosi alla necessità di raccordare frammenti informativi disseminati tra i documenti prodotti da molteplici organizzazioni di intelligence e agenzie governative. Si tratta, in sostanza, di delineare un quadro sufficientemente accurato così da fornire un avvertimento affidabile ai potenziali stakeholder, utilizzando un approccio "all-sources" e "all-hazards" (Chi, 2017; Criminal Intelligence Service Canada, 2007; Lindsay, 2016; Podesta, Pritzker, Moniz et al., 2014; Schneier, 2013; Symon & Tarapore, 2015). Di certo, il cambio di paradigma successivo agli eventi dell'11 settembre 2001 verso la condivisione delle informazioni ha condotto ad una più stretta collaborazione tra agenzie di intelligence, forze dell'ordine e autorità di regolamentazione, ora considerate come parte di una più ampia *"comunità di sicurezza nazionale"* (Australian Strategic Policy Institute, 2017, 16; National Commission on Terrorist Attacks upon the United States, 2004).

Correlata alla tematica sopra enunciata, appare l'esigenza di padroneggiare il sovraccarico di informazioni derivante dall'ampiezza e dalla velocità dei dati che la comunità di sicurezza nazionale ha il compito di gestire. Non solo l'enorme flusso di dati minaccia di sopraffare gli analisti, ma la stessa comunità di sicurezza nazionale affronta un'ulteriore sfida, quella rappresentata dalla varietà del patrimonio informativo. Posto che le Agenzie di sicurezza nazionale sono investite di una serie di incarichi a livello strategico, tattico e operativo, l'attività analitica coinvolgerà tanto dati strutturati – si pensi ai dati sull'immigrazione, visti, voli, arrivi marittimi, flussi commerciali, feed di social media, telecomunicazioni, metadati e-mail, carte di credito, conti bancari, acquisti al dettaglio, conti telefonici e internet, metadati dei fornitori di servizi – quanto dati non strutturati, come riprese ISR ["Intelligence, Surveillance and Reconnaissance", (Porche III, Wilson, Johnson et al., 2014] e dei droni, sonar passivi, sistemi di ecolocalizzazione dei colpi d'arma da fuoco, sistemi di traffico, filmati di sorveglianza, sistemi di analisi applicati ai testi dei social media e sistemi di analisi del traffico - un elenco meramente indicativo e destinato necessariamente ad espandersi. Ebbene, queste fonti di dati dovranno essere opportunamente gestite dalla comunità di sicurezza nazionale per mantenere il proprio vantaggio in termini di intelligence (Australian Strategic Policy Institute, 2017; Chi, 2017; Landon-Murray, 2016; Reilly, 2015).

Con l'aumento esponenziale dei dati open source provenienti dai social media e il crescente monitoraggio governativo, la scienza dei Big Data appare strettamente

allineata con le politiche di sicurezza nazionale e con la comunità di Intelligence (Jani & Soni, 2018). Negli ultimi anni, il governo degli Stati Uniti ha attribuito grande importanza allo sviluppo di una strategia nazionale per incorporare i Big Data nel processo decisionale, come documentano i due rapporti pubblicati dalla Casa Bianca (White House, 2014; 2016). D'altronde, per la Comunità di Intelligence di una nazione, i dati e le loro analisi sono centrali per adempiere alla mission istituzionale di anticipazione degli eventi al servizio delle priorità di sicurezza nazionale soprattutto per assolvere ai compiti imposti dalla Anticipatory Intelligence, il cui l'obiettivo principale è quello di ottenere previsioni più accurate e tempestive su eventi di natura (geo)politica, sociale, criminosa, ecc. creando schemi anticipatori affidabili a partire dall'analisi ragionata del dato storico (Richelson, 2002). Progetti specifici sulla modellazione delle minacce, sull'analisi delle reti sociali e sulla trasformazione della tecnica analitica e predittiva sono allo studio già da tempo (Aradau & Blanke, 2016; Doyle, Katz, Summers et al. 2014; Richey, 2015).

Sebbene l'applicazione delle scienze dei Big Data alla Intelligence Community (IC) appaia evidente, soprattutto in un'ottica di salvaguardia della sicurezza nazionale, i BD costituiscono una di quelle (rare) tecnologie emergenti il cui sviluppo nell'industria e nel mondo accademico si è verificato ad un ritmo ben più elevato rispetto alla ricerca nel campo della difesa (Stato Maggiore della Difesa, 2022; Symon & Tarapore, 2015) – ciò che ha imposto alle agenzie governative preposte un tempestivo e costante upgrade. Nonostante l'adozione di politiche a sostegno dell'impiego dei Big Data nei processi decisionali governativi, tale pratica non è risultata scevra da limitazioni e cautele, ben delineate da Jani & Soni (2018): affidarsi ad esperti per un uso efficiente della BD Technology; colmare le lacune conoscitive nel settore BDA; utilizzare le infrastrutture digitali consci della natura dinamica e in costante evoluzione dei Big Data – ciò implica l'adozione di decisioni dalla durata temporale limitata – e consapevolezza circa le forti implicazioni etiche che la BD Analytics comporta a carico dei cittadini, segnatamente in termini di sorveglianza (Matzner, 2016).

Il tema della sorveglianza associato alla Big Data Technology ricorre frequentemente nella produzione scientifica dell'ultima decade: Lyon definisce la sorveglianza come *"(…) l'attenzione sistematica, routinaria ai dettagli personali per un determinato scopo (come la gestione, l'influenza o l'acquisizione di diritti (…)"* (Lyon 2014: 2), mentre Degli Esposti estende questa definizione alla c.d. dataveillance, neologismo coniato per indicare *"(…) il monitoraggio sistematico di persone o gruppi, attraverso sistemi di dati personali, al fine di regolare o governare il loro comportamento"* (Degli Esposti 2014: 210).

Matzner (2016) introduce il concetto della "sorveglianza prospettica"[92], evocando l'idea di una raccolta indiscriminata di dati con futuro potenziale di vigilanza governativa sui cittadini: si pensi alle misure dei controlli di frontiera, che accederebbero ad un'ampia gamma di dati e modelli, non ultimi *"alerts for firearms and criminal records, missing and wanted persons, stolen vehicles and witnesses"* (Adey 2012: 196-7). In casi simili, è di tutta evidenza il conflitto tra istanze entrambe meritevoli di tutela, rispettivamente, il diritto alla riservatezza del cittadino e la salvaguardia della sicurezza nazionale (Antares Fumagalli, 2018).

[92] *"Concerning surveillance, this leads to a phenomenon which could be called "prospective surveillance": huge databases that are just stored for the time being—but with the possibility" to be used for purposes of surveillance at any time in the future"* (Matzner, 2016: 199).

Per comprendere i timori che hanno turbato i ricercatori di Surveillance Studies fin dagli albori del nuovo millennio, basti pensare all'immagine del "Big Brother", figura potente ma ritenuta insufficiente a definire potenzialità e rischi della *"data-based surveillance"* (Matzner, 2016: 198): si tratterebbe, pur sempre, di una figura centralizzata e, pertanto, non paragonabile, negli effetti, alla pervasività del controllo "strisciante" operato dalla BD Technology [*"(…) the pervasiveness of computing machines (…)"*, Matzner, 2016: 205]. Haggerty ed Ericson, in quello che sarebbe diventato uno dei lavori fondamentali per la ricerca sulla sorveglianza basata sui dati, scrivono, infatti: *"In the intervening decades, however, the abilities of surveillance technologies have surpassed even his [Orwell's] dystopic vision. Writing at the cusp of the development of computing machines, he could not have envisioned the remarkable marriage of computers and optics which we see today"* (Haggerty & Ericson 2000: 606). Come a dire, uno sviluppo futuro che andrebbe ben oltre le più nefaste previsioni in termini di sorveglianza e controllo sociale. Vero è che i Big Data continuano a rappresentare nuove e complesse sfide per le teorie della sorveglianza: maturare una conoscenza concreta circa le modalità di impiego dei dati da parte delle Agenzie governative richiede, invero, competenze elevate per districare accordi tecnologici e istituzionali sempre più opachi e intricati (Kazansky, 2021; Kazansky & Milan, 2021; Saetnan, Schneider & Green, 2018).

La comunità di sicurezza nazionale australiana, particolarmente sensibile rispetto alla tematica in oggetto, include, tra le potenziali applicazioni dei Big Data al settore della National Security, una serie di profili ampi e articolati, oltre alla citata integrazione delle informazioni condivise: riconoscimento e tracciamento delle entità; analisi predittiva; generazione di nuove ipotesi di conoscenza; infine, sicurezza e governance nazionale preventiva e predittiva (Australian, Strategic Policy Institute, 2017). Rispetto all'item delle informazioni condivise, la tendenza consolidata ormai da tempo è quella dell'integrazione dei dati (Data Fusion), consistente nell'impiego di sistemi automatizzati che, generando l'unione di feed di sensori e contenuti informativi provenienti da fonti eterogenee, elabora l'immagine di un'entità, target o altro oggetto di interesse. Trattasi di una metodologia che riproduce, avvalendosi di algoritmi di AI, l'attività attesa da un analista-tipo, soprattutto nel contesto dell'Intelligence post 9/11 (Best Jr., 2007; Chermak, Carter, J., Carter, D. et al., 2013; Johnson, 2007; Miller, 2005; Monahan & Palmer, 2009).

Uno degli ambiti di sviluppo più recente è quello della sorveglianza urbana, in cui algoritmi di ultima generazione, sfruttando l'Intelligenza Artificiale e la fusione di dati/informazioni, sono in grado di fornire consapevolezza situazionale e risposta tempestiva alle emergenze territoriali. Il processo decisionale automatizzato permette una valutazione delle minacce in tempo reale, con un enorme impatto su molteplici fronti di rilevanza nazionale, che spaziano dalla salvaguardia della National Security al monitoraggio delle infrastrutture critiche (Blasch, Pham, Chong et al., 2021; Chen, N., Chen, Y., You et al., 2016; Khan, Nazir, García-Magariño et al., 2021: Munir, Kwon, Lee et al., 2021; Zhang, Song, Du et al., 2018).

Centri di Data Fusion, ormai onnipresenti[93], operano tramite ufficiali di collegamento che

[93] A livello nazionale, gli attacchi terroristici alle Twin Towers hanno giustificato la creazione del Dipartimento di Homeland Security (DHS) nel 2002, che ha incorporato 22 agenzie governative, impiegando oltre 230.000 persone. A livello statale e locale, il DHS ha costruito una solida rete di "Centri di fusione" per diffondere e analizzare i dati su individui o attività sospette, assistere nelle indagini e identificare potenziali minacce. Poiché dette strutture devono affrontare il difficile compito di armonizzare gli imperativi di sicurezza nazionale con le esigenze della polizia locale, esse sono attendibili rivelatori di criticità

hanno accesso alle informazioni classificate delle reti informative delle loro organizzazioni e condividono dette informazioni con ciascuna delle altre reti informative. Ciò permette agli investigatori di raccordare record di dati di natura diversa e provenienti da fonti eterogenee: assegni di assistenza e disoccupazione, licenze di armi da fuoco, informazioni di noleggio auto, rapporti creditizi, indirizzi e numeri di telefono, informazioni sulla clientela del Banco dei pegni, indagini postali, dati di salute pubblica, dati di indagini di polizia, rapporti di furto di identità e di attività sospette, dati relativi a libertà vigilata e condizionale, informazioni provenienti da dipartimenti di polizia e penitenziari e molto altro (Hollin, 2015; Monahan & Regan, 2012).

I potenziali vantaggi della Data Fusion sono innegabili, posto che l'automatizzazione del processo di fusione consente il collegamento di feed di dati complementari, ridondanti e cooperativi di interesse comune - come quelli che tracciano la stessa entità o sono fisicamente localizzati nello stesso sito - fornendo una tassonomia relativamente completa di fonti e fatti, che possono essere combinati in un'immagine di intelligence armonica ed integrata (Durrant-Whyte, 1988). Non sono mancate, tuttavia, voci dissonanti in ambito accademico: alcuni Autori statunitensi hanno, infatti, denunciato il progressivo fallimento del concetto di "National Criminal Intelligence Sharing Plan (NCISP)", unitamente all'insorgere di una politica di sorveglianza diffusa e indiscriminata - quando non addirittura lesiva delle libertà fondamentali dei cittadini - associata ad aree di opacità algoritmica difficilmente penetrabili (Monahan, 2009; Monahan & Regan, 2013; Peterson, 2022; Taylor & Russel ,2012).

Un interessante caso di studio nel settore della *information integration* è senz'altro rappresentato da "Palantir Technologies", un servizio di ricerca di informazioni che, mediante un software avanzato ("Palantir Gotham"), setaccia le banche dati disponibili e combina le informazioni estratte dai molteplici dataset elaborandone una rappresentazione visiva (sociogramma) o grafica (network) (Mazzei & Noble, 2017; Payne, Solomon, Sankar et al., 2008; Risen, & Lichtblau, 2013; Vance & Stone, 2011; Wright, Payne, Steckman et al., 2009). "Palantir Gotham" rientra a pieno titolo nel circuito della *"information-sharing architecture"*, atta a garantire l'efficace funzionamento della *"Multilateral Intelligence Collaboration"* agita da Australia, Canada, Regno Unito, Stati Uniti e Nuova Zelanda, dettagliatamente descritta da McGruddy (2013: 219) e successivamente ripresa dalla letteratura internazionale (Chan, Sanders, Bennett Moses et al., 2022; Lander, 2004).

I casi d'uso di "Palantir Gotham" sono ampi e variegati, sia nel settore dell'Intelligence – si pensi all'eventualità di visualizzare la daily routine di un target di interesse in base ai suoi acquisti, comunicazioni, transazioni finanziarie, alloggio, uso del veicolo, prenotazioni di mezzi di trasporto, reti di contatti unitamente ad altri dati e relazioni (Australian, Strategic Policy Institute, 2017; Erwin, 2013; Kulshrestha, 2016; Munn, 2017; Winston, 2018) – sia nel settore Law Enforcement, in cui le Forze dell'ordine potrebbero utilizzare il sistema come piattaforma unificata per la gestione dei casi d'indagine in maniera olistica e integrata (Khurana, Basney, Bakht et al., 2009; Munn, 2017; Rodriguez & Nayor, 2022). Le Agenzie di informazione finanziaria, dal canto loro, potrebbero applicarla alla frode finanziaria, illecito che comporta un frequente scambio interattivo tra le persone coinvolte nelle reti criminali (Obolentsev & Yushchenko, 2019)

Sul piano tecnico, "Palantir" presenta un ulteriore vantaggio, poiché offre la capacità di integrazione di molteplici database con

con l'apparato emergente di sorveglianza statale (Monahan & Regan 2013).

un'interfaccia-utente che richiede solo query in linguaggio comune (in luogo del linguaggio di programmazione) e con risposte quasi in tempo reale, potendo elaborare i dati ad una velocità di oltre 50.000 variabili alla volta. Le tecniche di aggregazione e visualizzazione dei dati consentono, inoltre, la visione dei dati aggregati e possono essere utilizzate per analisi più complesse, ossia per enfatizzare intuizioni nascoste, generando così nuove forme di conoscenza (Burns, 2015; Dimitrakopoulos, 2017; Lev-Ram, 2016).

Benché l'importanza strategica dell'Intelligenza Artificiale e delle sue applicazioni per la sicurezza nazionale sia universalmente riconosciuta (Mikhailov, 2023), le critiche mosse alla piattaforma "Palantir Gotham" non possono essere sottaciute. La dottrina recente ha, infatti, sollevato censure sul piano dei rischi per la libertà individuale - l'azienda fornisce soluzioni informatiche per l'integrazione e il monitoraggio dei dati a Forze di polizia e Agenzie governative, Organizzazioni umanitarie e Aziende private, attuando pratiche di sorveglianza opache (Iliadis & Alker, 2022) – e per i diritti umani in senso ampio, con specifico riferimento al controllo delle politiche migratorie (Molnar, 2022; 2021(a); 2021(b); 2019(a); 2019(b); Naranjo & Molnar, 2019). Per giunta, il ricorso ad una piattaforma aziendale privata configurerebbe una sorta di pericolosa esternalizzazione (outsourcing) delle attività di polizia: altrimenti detto, l'affidamento al settore privato, finalizzato all'applicazione della legge, costituirebbe una seria minaccia per la legislazione sulla privacy a livello globale. Orbene, posto che le Agenzie governative rappresentano uno dei maggiori fruitori dei dati raccolti dai fornitori (privati) di tecnologia, i regimi di riservatezza dei dati che consentono il flusso di informazioni personali alle Agenzie governative non proteggerebbero efficacemente la privacy individuale (Colbary, 2021; Fairclough, 2016; Privacy International, 2020).

Sul piano del riconoscimento e monitoraggio di entità denominate, gli algoritmi di apprendimento automatico non supervisionati possono fornire riepiloghi di dati non strutturati per raggruppamento in cluster di informazioni con significato semantico simile. Detta funzionalità trova frequente applicazione nel dominio di sicurezza nazionale: un buon programma di analisi di testo, dotato di un efficace algoritmo di ML, per esempio, può leggere velocemente serie di trascrizioni e documenti, identificando i bit di testo rilevanti per un'indagine in corso (Asharef, Omar, Albared et al., 2012; Asmai, Salleh, Basiron et al., 2018; Australian Strategic Policy Institute, 2017; Shabat & Omar, 2015; Shabat, Omar & Rahem, 2014). In realtà, il riconoscimento mirato di entità denominate mediante l'utilizzo dell'apprendimento automatico è una funzionalità sfruttata dagli esperti da almeno un ventennio (Mollá, Van Zaanen & Cassidy, 2007; Zhang, Pan & Zhang, 2004), segnatamente nel settore biomedico e tossicologico-forense per l'identificazione di sostanze e/o composti farmacologici di interesse a partire dall'analisi di testi (Ali, Masood, Riaz et al., 2022; Armengol-Estapé, Soares, Marimon et al., 2019; Govindarajan, Mustafa, Kiyosov et al., 2023; Liu, Gao, Guo et al., 2021; Sun & Yang, 2019; Sun, Yang, Wang et al., 2021).

Le tecniche di riconoscimento di entità nominate [Named Entity Recognition (NER)] sono validamente impiegate, da una decade circa, anche per la visualizzazione di strutture, reti sociali o movimenti di massa nonché per il monitoraggio di entità-target all'interno delle reti stesse: ne è un esempio l'analisi dei social network condotta per monitorare le reti terroristiche attive nella Striscia di Gaza (Lotan, 2014) o per il rilevamento e l'analisi dei trends terroristici tout-court (Kade & Dhande, 2017; Kim, Pottenger & Behe, 2018; Oladejo & Onyemenam, 2019; Xia & Gu, 2019; Zamin, 2009). L'utilità per gli analisti è evidente, permettendo loro la rapida esplorazione dei lead fino agli obiettivi di interesse, come i nodi importanti all'interno delle comunità o altro (Asgari-Chenaghlu, Feizi-Derakhshi, Farzinvash et al., 2022; Hung & Chang,

2021; Li, Sun, Han et al., 2020; Lowe & Matthee, 2020; Simran, Sriram, Vinayakumar et al., 2020; Stalph & Heravi, 2021).

Una concreta applicazione di NER in Australia è offerta dal *Data to Decisions Cooperative Research Centre,* che fornisce un sistema di analisi di immagini satellitari. Il progetto *Immersive Intelligence Pod* osserva i set di dati geospaziali, visualizza le entità e il modo in cui esse convergono, co-localizzano (ossia, si incontrano) e divergono (si allontanano) in luoghi diversi e nel corso di differenti periodi di tempo. Il progetto mira a identificare modelli comportamentali daily routine delle entità di interesse, nonché relazioni e reti complesse di cui le stesse fanno parte. La tecnologia è stata concessa in licenza alla società "GIS Esri Australia", che dal 2017 la sta sviluppando per il Dipartimento della Difesa (Australian Strategic Policy Institute, 2017; Dimitrakopoulos, 2017; Jeffries, 2017; Kearns, 2015; Russel 2017).

Nel dominio della National Security è ormai consolidato l'utilizzo di indicatori automatizzati in grado di elaborare profili comportamentali, finanziari e di altra natura, o di attori rivelatisi precedentemente pregiudizievoli, per identificare potenziali minacce imminenti: si tratta dell'analisi predittiva[94], attività compiuta dalle Agenzie di sicurezza nazionale per l'elaborazione di modelli di rischio (Eckerson, 2007; Kumar & Garg, 2018; Roff, 2020). Si tratta, in sostanza, di un'attività compiuta mediante una varietà di tecniche predittive e modelli statistici che, opportunamente combinati, producono intuizioni per il futuro [*"insights for the future"* (Koukaras & Tjortjis, 2019)], una funzione cruciale per gli analisti impegnati nella mappatura delle minacce asimmetriche alla sicurezza nazionale (Australian Strategic Policy Institute, 2017). Vero è che una componente essenziale per l'analisi condotta dall'Intelligence è dedurre una spiegazione a partire da dati incerti, contraddittori e incompleti (McDermott, Veinott, Eusebi et al., 2021; Schwartz, 2015).

La correlazione della Predictive Analytics con il profilo "information-sharing" è facilmente intuibile, posto che i flussi di dati condivisi, raccolti, ordinati e analizzati mediante l'impiego delle BD Technologies costituiranno il "mattone tecnologico" dell'attività predittiva vera e propria (Chi, 2017; Koukaras & Tjortjis, 2019; Schneier, 2013). Inoltre, l'analisi dei dataset relativi ad eventi precedenti mediante l'uso di tecniche di data-mining – metodologie aventi ad oggetto l'estrazione di informazioni utili da grandi quantità di dati (Larose, 2015; McCue, 2006) - può condurre alla scoperta di relazioni non ovvie, che possono essere validamente impiegate nell'elaborazione di modelli predittivi attendibili (Fingar, 2009; Myne, Leahy & Soklaski, 2022; Ongsulee, Chotchaung, Bamrungsi et al., 2018; Scoblic, 2018), oggi più che mai sul versante della mitigazione delle cyber-minacce (Chen, Yan, Wu et al., 2021; Jemal, Cheikhrouhou, Hamam et al., 2020; Joshi & Geetha, 2014; Kamtuo & Soomlek, 2016; Uwagbole, Buchanan & Fan, 2017).

Una branca dell'analisi predittiva in forte ascesa nell'ultimo quindicennio è quella rappresentata dalla *Technosocial Predictive Analytics (TPA)* - metodologia di supporto al ragionamento anticipatorio che, avvalendosi di un approccio multidisciplinare all'analisi strategica e alla risposta ad eventi critici, rende possibile uno sforzo decisionale concertato da parte di analisti e responsabili politici (Sanfilippo, Cowell, Malone et al., 2009). Trattasi di un approccio predittivo che considera l'impatto sulla sicurezza nazionale di eventi i quali, comportando l'interazione di processi complessi - cambiamenti climatici, malattie infettive emergenti, affidabilità energetica, terrorismo, proliferazione nucleare, disastri naturali e provocati dall'uomo, vulnerabilità geopolitiche, sociali

[94] Previsione, in questo contesto, significa utilizzo delle informazioni disponibili per prevedere, o inferire, in maniera probabilistica, informazioni non disponibili al momento attuale (Chi, 2017).

ed economiche, ecc. - minacciano la salute, la sicurezza e la crescita sostenibile della società globale (Boulos, Sanfilippo, Corley et al., 2010; Madison, Cowell, Butner et al., 2012; Malone, Izaurralde, Thomson et al., 2009; Sanfilippo, Butner, Cowell et al., 2011; Wong, Leung, Lu et al., 2009). Di fronte a simili sfide, quello che la dottrina definisce "ragionamento anticipatorio integrato" [*"integrated anticipatory reasoning"* (Boulos, Sanfilippo, Corley et al., 2010: 19] deve necessariamente tradursi in un'attitudine costante e condivisa da parte della comunità di sicurezza nazionale (Karuna, Rana & Purohit, 2017; Mageto, 2021; Sanfilippo, Gilbert & Greaves, 2012).

Tra gli esperti in materia, analisti e decisori politici è maturata una crescente consapevolezza in merito all'importanza dell'interazione combinata di fattori naturali e antropici per affrontare il processo decisionale strategico in modo proattivo. L'Analisi Predittiva Tecnosociale è da tempo riconosciuta come un campo autonomo di indagine e sviluppo dell'interazione uomo-macchina, in cui data-mining e machine-learning svolgono un ruolo predominante nelle sue attuali applicazioni (Sanfilippo, Gilbert & Greaves, 2012). Studi multilivello, che assumono una vasta gamma di fattori biologici, familiari, comunitari, socio-culturali, ambientali, politici e macroeconomici, sono divenuti imprescindibili per prevenire e mitigare l'emergere di minacce alla salute collettiva come l'obesità infantile (National Institutes of Health, 2008) e la tossicodipendenza (National Institute on Drug Abuse, 2007) Allo stesso modo, la comprensione integrata del contesto infrastrutturale, sociale e culturale in cui operano movimenti sociali devianti è essenziale per l'inquadramento di condotte aggressive e violente (Wiktorowicz, 2004). Simili evidenze impongono l'adozione di processi decisionali di carattere anticipatorio in cui il giudizio umano si embrica con le inferenze automatizzate del machine-learning, generando quella sorta di "modellazione predittiva integrata" ampiamente descritta nella letteratura specialistica (Barral, Pinet, Tacnet, et al., 2023; McDermott, Veinott, Eusebi et al., 2021; Fuchs, Passarella & Conti, 2022; Riensche & Whitney, 2012; Sanfilippo, Riensche, Unwin et al., 2010).

La generazione di nuove ipotesi di conoscenza sul versante della gestione delle minacce alla sicurezza nazionale integra un'ulteriore potenzialità derivante dall'impiego del Big Data Analytics: il "data-mining [95] approach" permette, infatti, l'estrazione di dati significativi per la scoperta di correlazioni, modelli e trends in precedenza non considerati, o addirittura, non rilevabili dalla mente umana (DeRosa, 2004), favorendo l'identificazione di nuovi indicatori di eventi rilevanti nel "Homeland Security Domain" (Chi, 2017; Seifert, 2004). I Big Data, nello specifico, prometterebbero un duplice vantaggio per l'analisi predittiva: gli algoritmi di apprendimento supervisionati consentirebbero la strutturazione degli indicatori in allarmi automatici mentre gli algoritmi senza supervisione potrebbero individuare nuovi indicatori e avvertimenti nel "rumore" dei Big Data, consentendo la creazione di modelli predittivi inediti (Australian Strategic Policy Institute, 2017; DeRosa, 2004).

La combinazione di pratiche di fusione dei dati, analisi dei social network e indicatori predittivi ha generato grandi speranze negli operatori della sicurezza nazionale: in un'ottica anticipatoria di eventi probabili e di monitoraggio di potenziali minacce, l'idea è

[95] Il data-mining è una tecnica che utilizza le tecnologie informatiche per identificare schemi e connessioni, precedentemente non rivelati, tra diversi punti dei dati esistenti, con l'obiettivo di predire il comportamento futuro (Ramasastry, 2005). Thuraisingham (2004:191) lo definisce *"(…) the process of posing queries and extracting useful patterns or trends often previously unknown from large amounts of data using various techniques such as those from pattern recognition and machine learning"*.

quella di consentire, se necessario, l'adozione di misure preventive, contenitive o di minimizzazione del rischio. Ciò vale, principalmente, per minacce asimmetriche come attività terroristiche, il cui contrasto, in chiave preventiva, può essere perseguito mediante l'impiego del data-mining per l'identificazione di pattern comportamentali insoliti, a partire dalla raccolta di dati sulle persone, esame dei messaggi di posta elettronica, conversazioni telefoniche e altre attività di sorveglianza indiretta (DeRosa, 2004; Lakshmi & Raghunandhan, 2011; Maxwell, 2005; Ramasastry, 2005; Seifert, 2004; Thuraisingham, 2009, 2002; Thuraisingham, Khan, Masud et al., 2008) In un siffatto contesto, assume rilievo non solo la minaccia terroristica armata, bensì anche le minacce informatiche a reti, sistemi e infrastrutture critiche (cyber-terrorism) nonché quelle relative all'impiego di agenti biologici, chimici, radiologici e nucleari (CBRNe-threats) (Thuraisingham, 2004).

Trattasi di rischi di portata globale, rispetto ai quali le tecniche di estrazione e analisi automatizzata dei dati sono state unanimemente riconosciute, dalla comunità scientifica internazionale post 9/11, come potenti strumenti di rilevamento e contrasto (Bridgelall, 2022; Chen, Reid, Sinai et al., 2008; DeRosa, 2004; Ferooz, Hassan, Awan et al., 2021; Ganor, 2021; Yang, 2022; Kumar, Mazzara, Messina et al., 2020; Labib, Rizka & Shokry, 2020; Li, Jia, Liu et al., 2022; Olabanjo, Aribisala, Mazzara et al., 2021; Uche, Tsopze & Ebem, 2020; Shaikh, Wang, Liu et al., 2007; Skillicorn & Vats, 2007; Will, Memon & Gniadek, 2011). Il recentissimo contributo di Saini & Bansal (2023), che presenta una revisione sistematica dell'analisi delle reti terroristiche [Terrorist Network Analysis (TNA)] – trattasi di un campo di indagine che include diversi sottodomini, come la scansione dei dati su gruppi terroristici e caratteristiche di attacco nonché l'analisi comportamentale e predittiva per l'elaborazione di idonee contromisure – ha confermato il ruolo prioritario delle tecniche di data-mining e machine-learning, applicate al setacciamento sistematico dei social network, al fine di contrastare il fenomeno terroristico.

Nello stesso filone di ricerca si colloca anche il lavoro di Ali e colleghi (2023) che - posto il massiccio aumento delle attività illecite sulle piattaforme del Dark Web (He, He & Li, 2019; Saini & Bansal, 2019), implicante anche la diffusione dell'estremismo su larga scala (Alrasheed & Rigato, 2019; Liggett, Lee, Roddy et al., 2020) - propongono un promettente approccio per generare modelli testuali dalle discussioni sui forum terroristici del Dark Web [96] mediante l'utilizzo di tecniche di estrazione di dati. Ebbene, i modelli generati si sono dimostrati efficaci sia per l'identificazione di membri influenti all'interno dei gruppi terroristici che per l'estrazione di preziose informazioni su forum di discussione, post e argomenti critici di interesse per gli analisti di settore. A risultati sostanzialmente sovrapponibili sono pervenuti anche altri recentissimi contributi (Alghamdi & Selamat, 2022; Chaudhary & Bansal, 2022; Goyal, Saini, & Bansal, 2019; Saini, 2023). Va, tuttavia, osservato che, per sviluppare efficaci strumenti di contrasto, gli specialisti del data-mining devono necessariamente collaborare con esperti di antiterrorismo: altrimenti detto, un corretto impiego delle tecniche di data-mining non può prescindere da una buona comprensione delle minacce che si intendono mitigare (Davies, 2016; Thuraisingham, 2004).

Le tecniche di estrazione dei dati manifestano le loro potenzialità sia nella rilevazione di minacce asimmetriche di natura complessa e multidominio – emblematico è l'esempio della minaccia terroristica nelle sue molteplici dimensioni –

[96] Numerosi website che diffondono contenuti di terrorismo ed estremismo usano lingue diverse dall'inglese, rendendo così laborioso per gli investigatori capire il contenuto del sito web di potenziale interesse (Alkhatib & Basheer, 2019; Alrasheed & Rigato, 2019).

sia rispetto a minacce riconducibili a cause antropiche non intenzionali (es. errori umani) o disastri naturali (es. uragani, terremoti, incendi con conseguenti interruzioni di energia elettrica o della supply chain), come da tempo evidenziato da Thuraisingham (2004). L'impiego del data-mining può, infatti, consentire l'elaborazione di modelli predittivi di catastrofi naturali a partire dall'analisi di dati geologici: si pensi all'enorme utilità della previsione dell'imminente verificarsi di un terremoto, con conseguente evacuazione anticipata delle zone a rischio. Lo stesso dicasi per l'analisi dei dati meteorologici, che può rivelarsi cruciale nella mitigazione dei rischi per la collettività, implementando la rapidità delle risposte di emergenza in caso di eventi naturali estremi (Au, Curcin, Ghanem et al., 2004; Deparday, Gevaert, Molinario, et al., 2019; Goswami, Chakraborty, Ghosh et al., 2018; Li, Xie, Zeng et al., 2017; Linardos, Drakaki, Tzionas et al., 2022; Zagorecki, Johnson & Ristvej, 2013).

In simili circostanze, il data mining può essere validamente utilizzato anche per analizzare casi precedenti, creando scenari ipotetici, simulazioni realistiche e modelli a partire dai quali addestrare operatori/first responder in caso di futuri eventi avversi (Refonaa, Lakshmi, & Vivek, 2015; Zheng, Wang, & Zheng, 2017). La letteratura sul tema è molto ampia, soprattutto in ragione dei disastri naturali correlati alle recenti crisi climatiche: migliaia di vite umane vengono perse ogni anno, in tutto il mondo, oltre a danni significativi a flora, fauna e proprietà, a causa di eventi avversi come terremoti, inondazioni, tsunami, uragani e altre tempeste tropicali, frane, nubifragi, ondate di calore e incendi boschivi (Goswami, Chakrabort, Ghosh et al., 2018; Ramadhan, 2017; Supriyadi, Windarto, & Soemartono, 2018).

Qualunque sia il dominio di applicazione degli strumenti di DM e ML, è la fase finale del processo di data-mining a generare le maggiori criticità, posto che la stessa richiede l'interpretazione dei risultati dell'attività analitica e l'adozione di decisioni in merito all'impiego di dette risultanze (DeRosa, 2004). Problematiche assai delicate possono emergere nella fase conclusiva del data-processing, prima fra tutte, la misura in cui l'adozione della scelta dipende dai risultati dell'analisi automatizzata dei dati – altrimenti detto, il coinvolgimento della componente umana all'interno del processo decisionale, ciò che viene tecnicamente definito "human-in-the-loop machine-learning". La questione è molto dibattuta in ambito accademico, rappresentando una delle ultime frontiere dell'interazione uomo-macchina, in forte ascesa per quanto riguarda la ricerca e lo sviluppo di applicazioni nel mondo reale (Chai & Li, 2020; Cui, Koppol, Admoni et al., 2021; Monarch & Munro, 2021; Xin, Ma, Liu et al., 2018).

Sul piano puramente tecnico, l'inclusione della conoscenza dell'utente nel sistema di machine-learning, mediante l'immissione di dati utili all'addestramento dell'algoritmo, si è rivelata determinante in un'ottica di potenziamento dell'automazione e miglioramento delle prestazioni (Mosqueira-Rey, Hernández-Pereira, Alonso-Ríos et al., 2023; Waytowich, Goecks & Lawhern, 2018; Wu, Xiao, Sun et al., 2022). Ciononostante, le preoccupazioni di ricercatori e utenti in merito all'usabilità di tali sistemi, unitamente alle crescenti interazioni uomo-AI, stanno alimentando il settore emergente dell'apprendimento automatico centrato sull'uomo, noto come modello "Human-Centered Machine Learning" (HCML), un'area di ricerca che promette di coniugare le possibilità tecnologiche con i bisogni e i valori umani. In sostanza, il modello HCML integra le innovazioni tecniche del ML con valori sociali come equità, uguaglianza e giustizia. L'obiettivo è ambizioso, includendo la progettazione di algoritmi equi e trasparenti, il processo decisionale "Human-in-the-Loop", la progettazione per le collaborazioni uomo-AI e l'esplorazione degli impatti sociali del machine-learning nel contesto globale (Fiebrink & Gillies, 2018; Kaluarachchi, Reis & Nanayakkara, 2021;

Ramos, Suh, Ghorashi et al., 2019; Riedl, 2019; Sperrle, El-Assady, Guo et al., 2021). Si tratta di quello che Chancellor (2023) definisce come un insieme di buone pratiche per la creazione, valutazione, distribuzione e critica di sistemi di machine-learning, capace di bilanciare l'innovazione tecnica con un'attenzione equivalente alle preoccupazioni umane e sociali. Tuttavia, nonostante gli sforzi in atto, non sono al momento disponibili linee guida unificanti sul significato concreto di "human-centered", né un insieme di valori condivisi dalla comunità internazionale (Chancellor, Baumer & De Choudhury, 2019; Ramos, Suh, Ghorashi et al., 2019; Shneiderman, 2020; Pelillo & Scantamburlo, 2021).

L'applicazione delle tecniche di data-mining da parte di attori governativi, segnatamente in un contesto di contrasto alla minaccia terroristica, pone in risalto l'annoso dilemma "Privacy vs National Security": la questione è senz'altro spinosa, poiché impone il contemperamento di due istanze di pari rango costituzionale e, pertanto, entrambe meritevoli di tutela (Bignami, 2007; Committee on Technical and Privacy Dimensions of Information for Terrorism, Prevention and Other National Goals, National Research Council, 2008; Duhigg, 2012; Lindsay, 2016; Rosenzweig, 2006; Schwartz, 2011). I sostenitori dei diritti civili hanno denunciato l'illiceità dell'uso governativo dei dati provenienti dal settore privato, posto che le Agenzie governative hanno accesso ad ampi "dossier digitali" gestiti da intermediari di dati commerciali – data brokers, nello specifico, dai quali dette Agenzie acquistano sistematicamente dataset dei consumatori [97] per molteplici scopi, non ultimo quello di individuare i recapiti di utenti colpiti da misure personali restrittive pendenti (Lamdan, 2019). Orbene, l'applicazione di tecniche di DM consente l'estrazione di una pluralità di informazioni su schemi comportamentali e daily routine degli utenti (Rothstein, 2013; Weiss, 2004) che, se opportunamente combinate con informazioni estratte da altre fonti, potrebbero consentire il "setacciamento" della sfera privata dei cittadini aggirando le procedure di rilascio di mandati di perquisizione e vanificando, di fatto, la funzione garantista ad esse sottesa – quello che, in dottrina, è stato definito *"the adverse inference issue"* (Birrer, 2005; Cate, 2008; DeRosa, 2004; Ramasastry, 2005; Renke, 2005; Rosenzweig, 2009; Solove, 2008; Solow-Niederman, 2022; Taipale, 2003).

Il dibattito tra sicurezza (nazionale) e libertà (individuale) si è, comprensibilmente, intensificato tra i cultori della materia all'indomani dei tragici eventi dell'11 settembre 2001 - data a partire dalla quale le Agenzie governative occidentali hanno iniziato a manifestare uno spiccato interesse per il data-mining, questo nuovo strumento tecnologico per l'individuazione preventiva di attori potenzialmente pericolosi, mediante la creazione di "profili" a partire dalla raccolta di dati personali associati a modelli comportamentali ritenuti sospetti. La complessità della questione è sapientemente riassunta da Solove (2008: 345), il quale afferma: *«We live in an "age of balancing," and the prevailing view is that most rights and civil liberties are not absolute (…). Thus, liberty must be balanced against security. But there are systematic problems with how the balancing occurs that inflate the importance of the security interests and diminish the value of the liberty interests».* Secondo la dottrina più accreditata, il bilanciamento tra sicurezza e libertà non può prescindere dalla valutazione di due componenti, rispettivamente, la gravità della minaccia alla sicurezza e l'efficacia delle misure per mitigare la minaccia stessa. Detta valutazione, tuttavia, appare tutt'altro che lineare: *"Assessing the risk of harm from terrorism is very difficult because terrorism is*

[97] Senza contare che, secondo le stime riportate da Tien (2004), le Agenzie e i Dipartimenti federali statunitensi detengono enormi quantità di informazioni sulle transazioni e le attività delle persone" - quasi 2000 banche dati agli inizi del nuovo millennio.

such an irregular occurrence and is constantly evolving" (Solove, 2008: 351).

La materia è stata oggetto di vivace dibattito nella comunità accademica internazionale, che ha sottolineato la necessità della costruzione di un'adeguata cornice giuridica, capace cioè di assicurare un impiego equilibrato dei dati personali all'interno di efficaci e condivisi programmi di sicurezza nazionale. Il dinamismo degli esperti di settore è ampiamente documentato dalla fiorente produzione scientifica della prima decade degli anni Duemila, di cui vengono di seguito riportati solo i contributi principali (Bignami, 2007; Cate, 2008; Chesney, 2002; Donohue, 2005; Fulda, 2000; Kris, 2006; McCarthy, 2002; Rackow, 2001; Rubinstein, Lee, & Schwartz, 2008; Slobogin, 2008; Solove, 2007, 2001; Taipale, 2003).

È interessante notare come, nell'ambito della produzione scientifica a cavallo tra la prima e la seconda decade degli anni Duemila, si assista ad una presa di posizione tendente a considerare i due aspetti della Privacy e della Sicurezza Nazionale non necessariamente contrapposti bensì concorrenti e, in ogni caso, ben distinti l'uno dall'altro. Il principale sostenitore della tesi in oggetto è Dereck Bambauer che, nel saggio del 2013 dal titolo *"Privacy versus Security"*, denuncia la tendenza dei giuristi a confondere privacy e sicurezza, laddove i due termini della questione possano (e, soprattutto, debbano) essere trattati separatamente. L'approccio suddetto produrrebbe significative ripercussioni sul piano pratico, non ultima quella di sanzionare con differente grado di severità i *"security failings"* (Bambauer, 2013: 667) rispetto alle inosservanze a carico della privacy – ciò che chiama necessariamente in causa valutazioni di politica criminale.

Nel corso dell'ultima decade si è assistito ad un progressivo intensificarsi dei timori sul versante garantista, soprattutto in ragione del rischio di errore derivante dall'impiego delle tecnologie di analisi predittiva della condotta criminosa (Ferguson, 2015; Goel, Shroff, Skeem, et al., 2021). Taluna dottrina ha ravvisato una profonda ambivalenza della collettività rispetto all'impiego degli algoritmi predittivi, ipotizzando una sorta di "avversione algoritmica" rispetto all'area della giustizia penale, per ridurre la quale potrebbero rivelarsi utili correttivi quali trasparenza e maggiore affidabilità dei processi di machine-learning; epurazione da pratiche discriminatorie; accuratezza nella valutazione della condotta individuale della persona di interesse e, non ultimo, mantenimento della componente umana quale decisore finale nel processo decisionale ("Human-in-the-Loop"). Ciò consentirebbe di valorizzare le potenzialità dell'integrazione degli algoritmi predittivi nel circuito del Law Enforcement e della giustizia penale: dal miglioramento della qualità dei giudizi espressi dagli agenti di polizia e dagli organi giurisdizionali, al risparmio delle risorse delle Forze dell'ordine fino, auspicabilmente, alla riduzione dell'irrazionalità e dei pregiudizi che affliggono il sistema giudiziario ampiamente inteso (Simmons, 2016).

Gli orientamenti della dottrina attuale sembrano oscillare tra i timori legati all'abuso della discrezionalità da parte delle Agenzie statali/governative e i margini di errore degli algoritmi predittivi (e, più in generale, delle applicazioni di DM e ML), con una marcata diffidenza verso pratiche di *"individualized suspicion"* (Berman, 2019: 463) avulse dal rispetto dei valori della dignità umana e dell'autonomia decisionale degli organi giurisdizionali, principi fondamentali dello stato di diritto (Ball, 2018; Barrett, 2017; Deeks, 2018; Deskus, 2018; Henderson, 2017; Lyn, 2020; Miller, 2014; Rich, 2016; Simmons, 2018; Talapina, 2022).

L'analisi predittiva compiuta mediante Big Data Technologies ha rinvigorito una delle questioni più controverse degli studi attuali in materia di sicurezza nazionale, quella della "predictive policing", intesa come l'attività di polizia che si avvale di Big Data e sistemi algoritmici per generare previsioni,

nel futuro prossimo, su persone e luoghi che si ipotizza possano essere coinvolti (o subire) un evento criminoso. I vantaggi dichiarati della polizia predittiva si concentrano sulla capacità della tecnologia di implementare attività di policing preventiva mediante automazione dell'intero processo decisionale: detta applicazione può consentire, cioè, alle Forze dell'ordine, di garantire la sicurezza pubblica con maggiore efficienza, ossia con attività mirate, minor dispendio di risorse (umane ed economiche) e impiego di modalità meno restrittive a carico dei potenziali autori di reato (Berk, 2021).

L'uso dell'Intelligenza Artificiale per generare ipotesi criminodinamiche localizzate nel tempo e nello spazio è, in larga parte, il risultato di un esercizio di statistica spaziale, che si colloca in seno al c.d. profiling geografico [98] , tecnica di profilazione ampiamente utilizzata dalla scienza criminologica anglosassone per circoscrivere un'area territoriale all'interno della quale è altamente probabile il verificarsi di una specifica attività criminosa (Picozzi & Zappalà, 2002). Uno degli sviluppi più diffusi e promettenti dell'ultimo ventennio è rappresentato dal "crime mapping", favorito dall'utilizzo del GIS *(Geographic Information System)* e delle moderne tecnologie algoritmiche, che consentono la memorizzazione su database, con successiva analisi e visualizzazione, dei dati relativi a determinate coordinate spaziali. Risale al 1996 l'istituzione del "Crime Mapping Research Center" (CMRC) presso il "National Institute of Justice" statunitense, nel cui ambito sono stati sviluppati software dedicati per il tracciamento degli "hot spot of crime", in grado di fornire informazioni dettagliate circa la concentrazione di attività criminali in aree territoriali circoscritte quali, ad esempio, il contesto urbano (Canter & Larkin, 1993; Rossmo, 2012; Rossmo & Rombouts, 2016, Rossmo & Baeza, 1998). Nonostante i recenti sviluppi in materia, la profilazione geografica presenta innegabili criticità, soprattutto in termini di (mancata) correlazione tra complessità della strategia e accuratezza del modello predittivo, imponendo una serie di riflessioni sulle implicazioni per le politiche e le procedure della polizia, nonché per la comprensione del processo decisionale umano da parte degli analisti (Snook, Zito, Bennell et al., 2005).

L'attività di polizia costituisce uno dei principali contesti organizzativi in cui l'uso della sorveglianza dei Big Data risulta in costante crescita. La produzione scientifica annovera, tra i motivi della proliferazione del "data-driven policing", l'incremento di efficienza e affidabilità del sistema; il miglioramento della previsione e prevenzione dei comportamenti criminosi, con conseguente riduzione dei tassi di criminalità; il potenziamento del meccanismo di responsabilizzazione delle forze dell'ordine, anche attraverso la mitigazione del rischio di pratiche discriminatorie (Berk, 2021). Nell'ultima decade, l'analisi predittiva è stata utilizzata per una vasta gamma di attività istituzionali delle Forze dell'ordine; oltre ai modelli predittivi di futuri eventi criminosi (Perry, 2013), infatti, si riscontrano modelli previsionali attinenti a soggetti con maggiori probabilità di coinvolgimento in episodi di violenza armata (Papachristos, Hureau, & Braga 2013) e modelli identificativi di operatori delle Forze dell'ordine con maggiori probabilità di impegnarsi in condotte a rischio (The White House -Executive Office of the President, 2016).

L'uso dell'Intelligenza Artificiale nell'ambito delle attività di polizia è oggetto di

[98] Particolare attenzione all'ambiente fisico (e sociale) nell'analisi del crimine fu posta da Shaw (1929), nei suoi studi sistematici sugli ambienti urbani degradati ad alto tasso di criminalità. Tali ricerche vennero proseguite da un gruppo di studiosi, a partire dagli anni Trenta del secolo scorso, che assunse la denominazione di "Scuola di Chicago": i sociologi appartenenti a questa corrente criminologica indicarono con l'espressione "aree criminali" le zone delle città nelle quali si sviluppava la maggior parte della criminalità urbana. Successivamente, tale approccio si estese anche allo studio della delinquenza giovanile (Shaw & McKai, 1942).

preoccupazioni diffuse, soprattutto da quando il processo decisionale basato sui Big Data ("big data-driven decision-making") è stato sistematicamente incorporato nelle pratiche di applicazione della legge (Joh, 2017; Okidegbe, 2019; Selbst, 2017; Shapiro, 2019; Southerland, 2020; Završnik, 2020). La sorveglianza è sempre più mediata tecnologicamente e le tecnologie emergenti la rendono possibile su una scala senza precedenti (Ericson & Haggerty, 1997; Lyon, 1994). Con lo sviluppo dell'informatica, inoltre, la sorveglianza di massa è emersa accanto alla comunicazione di massa (Rule 1974).

I timori manifestati dalla dottrina più sensibile a tematiche garantiste investono principalmente l'accuratezza, l'equità e la trasparenza dei sistemi algoritmici applicati ai modelli di polizia predittiva. Trattasi, invero, di profili di criticità per i quali nessuna soluzione tecnica è ipotizzabile, dovendosi ricercare un contemperamento di interessi attraverso processi politici e legislativi che raggiungano un equilibrio accettabile tra le priorità in competizione (Berk, 2021).

Tali preoccupazioni sembrano essere fondate, posto che, sebbene parte del fascino dei Big Data risieda da sempre nella loro promessa di un processo decisionale meno discrezionale e più obiettivo (Espeland & Vannebo 2007; Hacking 1990; Porter, 1995), le nuove piattaforme analitiche e le tecniche algoritmiche sono impiegate in contesti organizzativi preesistenti (Barley 1986, 1996; Kling 1991) e incarnano gli scopi dei loro creatori (Boyd & Crawford 2012; Kitchin, 2014). Rimane, pertanto, da chiedersi fino a che punto l'adozione di analisi avanzate ridurrà inefficienze e disuguaglianze organizzative, o al contrario, favorirà il radicamento delle tradizionali dinamiche di sorveglianza e potere all'interno delle preesistenti Organizzazioni (Southerland, 2020). A tal proposito, l'approccio proposto da Gitelman (2013) sembra offrire interessanti spunti di riflessione: l'Autrice, infatti, nel suo saggio, dal titolo evocativo *"Raw Data is an Oxymoron"*, sottolinea come molteplici episodi nell'evoluzione del concetto di "data" - dai primi problemi matematici moderni all'odierna, ineludibile "Data Surveillance" – testimoniano la dipendenza dei dati dal contesto a cui essi afferiscono: i dati, insomma, sarebbero tutt'altro che "grezzi" e, come tali, non dovrebbero essere percepiti come una risorsa naturale, bensì come il risultato di un processo culturale e, in quanto tali, interpretati alla luce del framework all'interno del quale sono stati generati.

L'annosa questione del "data-driven policing" è oggetto di specifica trattazione da parte di Brayne (2017) che, nel suo lavoro *"Big Data Surveillance: the case of Policing*", indaga approfonditamente la correlazione tra BD Analytics e il concetto di "policing", nella sua triplice dimensione di controllo, vigilanza e mantenimento dell'ordine pubblico. Secondo l'Autrice, l'impiego di sistemi algoritmici comporterebbe un cambio di paradigma, precisamente un viraggio dalla sorveglianza tradizionale all'attività di Intelligence – attività fondamentalmente predittiva, che comporta la raccolta di dati, l'identificazione di modelli sospetti, luoghi, attività e individui, intervenendo preventivamente sulla base delle informazioni acquisite. Fisiologica conseguenza di ciò, sarebbe il passaggio da una modalità reattiva ad un approccio predittivo rispetto alla criminalità: la polizia predittiva sarebbe, in sostanza, un'estensione della "Hot Spot Policing" [99]

[99] Nei primi anni 1980, negli Stati Uniti, di fronte all'inefficacia delle strategie reattive nella riduzione della criminalità, si è assistito ad un cambio di paradigma verso strategie di polizia più proattive e orientate alla risoluzione dei problemi, tra cui lo Hot Spots Policing (HSP). Gli "hot spots" sono aree territoriali con livelli costantemente elevati di criminalità e disordini: la HSP, pertanto, è quella strategia di riduzione della criminalità basata sull'idea che le condotte criminose non siano distribuite uniformemente sul territorio ma siano localizzate all'interno di

resa attualmente possibile dalla densità temporale dei Big Data, ossia dalle osservazioni ad alta frequenza [100] che la tecnologia algoritmica consente. Ne è un chiaro esempio il software progettato da "PredPol" (una società di polizia predittiva) e impiegato dal LAPD (Los Angeles Police Department) a partire dal 2012. "PredPol" utilizza un algoritmo basato sul modello "near-repeat" [101], il quale classifica la zona immediatamente adiacente al *locus commissi delicti* come area ad aumentato rischio per la commissione di crimini futuri: il software, nello specifico, impiega una triade di input - tipo di reato passato, luogo e ora del crimine - per identificare le aree in cui è più probabile che future attività criminose possano essere perpetrate.

Il cambio di paradigma dell'attività di policing si estenderebbe anche al viraggio da sistemi "query-based" [102] a sistemi "alert-based" [103], con evidenti implicazioni per la struttura della sorveglianza. Vale la pena notare, infatti, come i sistemi basati sugli avvisi integrano, invece di sostituire, i sistemi basati su query. Posto che le ricerche rappresentano tutt'ora modalità tipiche dei meccanismi informativi delle Forze dell'ordine, una delle trasformazioni apportate dalle Big Data Techs attiene alla circostanza che le query stesse stanno assurgendo al rango di "dati", con intuibile ampliamento delle reti di controllo degli utenti da parte del sistema di giustizia penale.

A ciò consegue un abbassamento delle soglie di inclusione all'interno dei database delle Forze dell'ordine: se, infatti, per lungo tempo, in seno ad essi sono confluite esclusivamente informazioni su soggetti sottoposti a misure restrittive della libertà personale o a condanne penali, detti sistemi hanno assistito ad un ampliamento dei set di dati oggetto di inclusione, fino a ricomprendervi anche destinatari di misure personali temporanee, come dimostra la proliferazione di "stop-and-frisk[104] database" nel panorama della giustizia penale statunitense. Ora, poiché le attuali piattaforme analitiche costituiscono parte integrante delle attività di polizia – basti richiamare il già citato software di "Palantir" che, integrando molteplici fonti di dati, permette una ricerca incrociata tra database privati e/o istituzionali (Burns, 2015) - anche informazioni relative a cittadini fino ad ora estranei al circuito penale possono essere agevolmente codificate dagli operatori attraverso sofisticate analisi di rete.

I rischi di un'estensione latente della governance e delle capacità di controllo sociale del sistema di giustizia penale sono enfatizzati da Couchman (2019) che, nel suo lavoro *"Policing by Machine:*: *Predictive Policing and the Threat to our Rights"* analizza criticamente l'applicazione degli algoritmi di Artificial Intelligence al settore di Law Enforcement. L'assunto fondamentale è che le nuove pratiche di sorveglianza digitalizzate amplino il raggio d'azione delle Forze dell'ordine, generando

aree circoscritte. Ebbene, concentrare le risorse di Law Enforcement su dette aree può consentire di fronteggiare la criminalità in modo meno dispendioso e più efficiente (Brayne, 2017).

[100] Il riferimento implicito è al paradigma delle "3 V" dei Big Data, ossia Velocità, Volume e Varietà, descritto nella prima sezione del presente lavoro.

[101] I "near-repeat events" (letteralmente, "eventi quasi-ripetuti") sono crimini che si verificano in prossimità, spaziale e temporale, di un evento criminale originario: trattasi di reati della stessa natura del primo evento che, tuttavia, non coinvolgono la medesima vittima (Weisel, 2005).

[102] Per "query-based" si intendono quei database a cui gli utenti inviano richieste di informazioni nella forma di una ricerca (query): tipico esempio è l'inserimento dei dati di un veicolo durante un controllo di polizia stradale.

[103] . Nei sistemi basati su avvisi ("alert-based") gli utenti ricevono notifiche in tempo reale quando determinate variabili, o configurazioni di variabili, sono presenti nel pool di dati del sistema stesso.

[104] Letteralmente, "fermo e perquisizione": si tratta di misure temporanee di limitazione della libertà personale (sostanzialmente equivalenti al fermo di polizia) conseguenti all'accertamento di infrazioni durante ordinarie attività di controllo del territorio compiute dalle Forze dell'ordine.

una forma tecnologica di "allargamento della rete" (Cohen, 1985) - in questo caso, delle maglie della giustizia penale - soprattutto mediante l'impiego di strumenti investigativi digitali di tipo "dragnet"[105]. Ciò faciliterebbe la creazione di reti di sorveglianza secondaria a carico di ampie fasce di popolazione suscettibili di divenire, loro malgrado, obiettivi di future attività di intelligence non legittimamente giustificate (Brayne, 2017). Questi ed altri sono i timori che assillano attualmente la comunità scientifica internazionale (Amoore & Raley, 2017; Andrejevic, 2017; Brayne, & Christin, 2021; Dencik, Hintz & Carey, 2018; Lavorgna, & Ugwudike, 2021; Shapiro, 2019; Van Brakel, 2016).

L'ultima frontiera della critica mossa al "data-driven policing" è rappresentata dall'indagine dei danni tecno-sociali provocati dagli algoritmi di polizia predittiva attraverso le attività convenzionali di applicazione della legge. Secondo questa impostazione, i falsi ideali di applicazione oggettiva, neutrale e non discrezionale dell'AI, tesi a promuovere la coerenza del sistema, sarebbero colpevoli dell'erosione della responsabilità per le decisioni adottate mediante processi automatizzati, con gravi ripercussioni sulle relazioni sociali ad ampio spettro (Wood & Warren, 2022). Nello stesso filone di ricerca si colloca l'originale lavoro di Egbert & Mann (2021) che, basandosi su dati empirici relativi all'utilizzo di software di previsione della criminalità in Germania e Svizzera, definiscono la "predictive policing" come una sorta di "assemblaggio socio-tecnico", comprensivo non solo dei modelli algoritmici (componente tecnica del sistema), bensì anche dell'attuazione delle conseguenti previsioni da parte degli organi di polizia (componente sociale, ovvero umana). Consci del potenziale discriminatorio delle pratiche tecno-sociali – tematica più volte emersa nel discorso accademico recente (Hopster, 2021; Skoff & Rollo, 2022) - gli Autori sollecitano una maggiore attenzione ai contesti storici e socio-politici da cui emergono le tecnologie predittive, proponendo un apprezzabile approccio olistico alla complessa tematica.

Considerazioni dello stesso tenore sembrano essere condivise dagli operatori di polizia, come emerge dalla ricerca di Sandhu & Fussey (2021), secondo cui la maggior parte di essi avrebbe maturato una consapevolezza dettagliata circa i limiti delle tecnologie predittive, in particolare di quelli causati da errori e distorsioni nei dati di input. Tale consapevolezza avrebbe indotto molti funzionari a sviluppare un atteggiamento scettico nei confronti delle tecnologie predittive tradottasi, in taluni casi, in un'esplicita riluttanza al loro impiego. Da questo lavoro emerge, inoltre, come le affermazioni sulla capacità del software predittivo di neutralizzare la soggettività dell'attività di polizia trascurino i continui sforzi degli operatori per garantire un'applicazione equa e mediata delle previsioni algoritmiche.

In ogni caso, al netto delle criticità evidenziate, i professionisti della sicurezza hanno progressivamente adottato linguaggio e metodi dell'informatica a scopo di previsione: i dispositivi digitali e i Big Data, in particolare, sembrano offrire risposte rispetto ad un'ampia gamma di problemi di (in)sicurezza, promettendo valide intuizioni su futuri sconosciuti (Aradau & Blanke, 2017; De Chant, 2014). Vero è che il "policing environment" successivo agli attentati terroristici dell'11 settembre 2001 si è dimostrato particolarmente sfidante, poiché caratterizzato da criminalità organizzata transnazionale, terrorismo globale ed estremismo interno. La necessità di convergenza tra intelligence criminale e sicurezza nazionale è divenuta, oggi più che mai, un imperativo non differibile (Gkougkoudis, Pissanidis & Demertzis, 2022) mentre, com'è stato efficacemente

[105] Letteralmente, "rete a strascico": trattasi di software che raccolgono informazioni in maniera indiscriminata, ossia su un numero indefinito di soggetti, piuttosto che focalizzarsi su sospettati di coinvolgimento in attività criminose.

sottolineato, *"Law enforcement can no longer afford to respond to contemporary and future problems with the solutions of yesterday"* (Peterson, 2005: vii).

Nel frattempo, il processo di legittimazione della raccolta e del trattamento delle informazioni da parte degli organismi ufficiali delle Forze dell'ordine, avvalendosi di Tecnologie Emergenti come Big Data e Intelligenza Artificiale, è stato implementato dalle Law Enforcement Agencies (LEA) a livello internazionale, prima fra tutte la LEA britannica, seguita dalla controparte statunitense, con l'obiettivo dichiarato di contrastare la minaccia terroristica (Gkougkoudis, Pissanidis & Demertzis, 2022; LeCates, 2018). Molti sono stati i Paesi europei ad avere aderito a siffatti modelli di gestione preventiva delle informazioni individuali, ossia ancor prima del formale coinvolgimento dei soggetti nel circuito di giustizia penale, anche solo sotto la vaga condizione di un probabile ma elevato rischio criminale (Carter, 2009).

Questa tendenza verso una forma di polizia c.d. "information-defined" (Gkougkoudis, Pissanidis & Demertzis, 2022: 146) ha trovato rapidamente credito presso progettisti e professionisti della sicurezza, ma anche presso teorici e accademici, che l'hanno presto identificata come una delle innovazioni più importanti del XXI° secolo nel settore del Law Enforcement, denominandola "Intelligence-Led Policing" (ILP) (Innes & Graef, 2012; Manning, 1992).

Un esame approfondito della letteratura specialistica (per la quale si rinvia all'ampia bibliografia curata da Gkougkoudis, Pissanidis & Demertzis, 2022), al fine di individuare le principali tecnologie ILP attualmente in uso, ha consentito di elencarne almeno una decina: al vertice della lista si colloca l'Intelligenza Artificiale la quale, utilizzando algoritmi di "deep learning" che addestrano le macchine all'analisi dei Big Data per l'elaborazione di modelli di "Risk Assessment", consente alle LEA una distribuzione razionale del personale di polizia in base alle aree di criminalità preventivamente identificate. Anche i software di riconoscimento facciale possano rivelarsi uno strumento prezioso per la prevenzione del crimine, grazie alla capacità di identificare potenziali terroristi e rintracciare criminali e/o persone scomparse [106] . La biometrica consente, l'analisi di caratteristiche biometriche e caratteristiche comportamentali del soggetto di interesse, tra cui il rilevamento delle emozioni, il riconoscimento vocale, l'analisi dell'andatura, il riconoscimento dell'iride e delle vene del polso, le impronte palmari e finanche il battito cardiaco.

La robotica, con lo sviluppo di auto a guida autonoma e di telecamere di nuova generazione, può garantire l'accesso visivo e audio a scene del crimine considerate troppo pericolose o difficili da raggiungere per gli operatori umani - la Cina ha introdotto, nel 2016, un robot ancora in fase di sviluppo, che sarebbe già stato utilizzato per pattugliare banche, aeroporti e scuole – mentre sistemi di eco-localizzazione di spari d'arma da fuoco come "ShotSpotter" possono coadiuvare gli analisti della polizia nell'identificazione del luogo dell'evento, così da indirizzare tempestivamente gli operatori sulla scena.

La termografia è uno strumento ampiamente utilizzato nella sorveglianza delle frontiere, poiché particolarmente efficace al buio. Le termocamere sfruttano la tecnologia di imaging a infrarossi per rilevare il calore emesso da oggetti animati, come esseri umani e animali, entro un raggio di 12 Km - ciò che consente il mantenimento della sicurezza lungo i confini nazionali, anticipando i movimenti di massa in caso di imminenti crisi migratorie.

La tecnologia "Automatic License (or Number) Plate Recognition" (ALPR o ANPR), ampiamente utilizzata dagli esattori di pedaggi per identificare automaticamente

[106] :Tali tecnologie sono considerate tra le più controverse del XXI° secolo, per le criticità che presentano in termini di rispetto delle libertà fondamentali dell'individuo (Couchman, 2019).

numeri e lettere sulla targa di un'auto, è ora sfruttata dalle forze dell'ordine per rintracciare veicoli rubati o conducenti colpiti da mandati di cattura, nonché per localizzare persone dichiarate scomparse (Amber Alerts), così come per il tracciamento di veicoli impiegati in attività illegali. Anche i sistemi TVCC sono diventati sempre più comuni tra le autorità di polizia, in quanto forniscono importanti prospettive di sorveglianza fungendo, al contempo, da strumento per la prevenzione (con funzione deterrente) e l'attività investigativa post-delictum. Progettata per non captare le conversazioni né catturare immagini nitide delle interazioni pubbliche, la tecnologia video è cruciale per il monitoraggio del comportamento delle persone, coprendo aree pubbliche e garantendo la raccolta di prove utili in caso di eventi criminosi.

Un'applicazione specifica dei sistemi di videosorveglianza è costituita dalle "Enhanced Body-Worn Cameras", ossia le telecamere indossabili di nuova generazione: in grado di registrare interazioni e contatti dell'operatore di polizia in servizio, sono in dotazione alle LEA per assicurare trasparenza e responsabilità nell'espletamento delle funzioni istituzionali. Posto che le tracce digitali sono suscettibili di supervisione, esse rappresenterebbero un'applicazione concreta della politica di sorveglianza dell'operato delle forze di polizia e di valutazione della correttezza delle procedure, con funzione garantista per i cittadini; ciononostante, l'uso di telecamere indossabili da parte della polizia ha avviato un recente dibattito circa l'impatto di questa tecnologia rispetto alla privacy e la protezione dei dati dei cittadini protagonisti delle interazioni con i predetti operatori.

Un ultimo impiego degno di nota è quello dei cc.dd. droni i quali, dal monitoraggio del traffico alla sorveglianza delle frontiere, si sono rivelati un ottimo strumento nel coadiuvare gli sforzi della polizia nel contrasto alla criminalità. L'uso della forza non letale da parte di aeromobili a pilotaggio remoto è già consentito dalle legislazioni di alcuni Paesi degli Stati Uniti, tanto che le pattuglie frontaliere sarebbero in grado di armare le squadriglie di droni in dotazione per immobilizzare potenziali sospetti o inviarli contro obiettivi di interesse. Nel panorama europeo, particolare interesse ha suscitato lo "Unmanned Aircraft Service" della polizia ellenica: fondato nel 2017 e incaricato di monitorare il territorio nazionale, trasmettendo informazioni alle Forze di polizia terrestri per la prevenzione e repressione della criminalità, esercita il controllo dell'immigrazione clandestina nelle aree di confine; gestisce l'ordine pubblico e il traffico stradale; fornisce supporto al corpo dei Vigili del Fuoco nella gestione di incendi, calamità naturali, alluvioni, terremoti o incidenti gravi. Il Servizio vanta una dotazione di nove UAV, recentemente utilizzati per monitorare il traffico nei periodi di quarantena imposti durante la pandemia da COVID-19: alcuni esemplari, infatti, sono dotati di strumentazioni avanzate, come telecamere con termografia e mappatura tridimensionale per lo svolgimento di compiti specifici o in condizioni ambientali avverse.

Allo stato attuale, il connubio tra LEA e Tecnologie Emergenti è divenuto indissolubile, anche grazie all'introduzione di nuovi strumenti per mitigare le minacce asimmetriche (Custers, 2012; Manning, 1992; Nunn, 2001). Sebbene la maggior parte delle procedure di sorveglianza post 9/11 - in particolare la raccolta, distribuzione ed elaborazione delle informazioni con il massiccio impiego di Big Data Technologies – abbiano costituito l'approccio più dinamico ai nuovi scenari di minaccia, è innegabile la loro natura "opaca" e controversa (Dignum, 2018; Islam & Zahidul, 2015; LeCates, 2018; Seele, 2017; Spiegel, 2018; Wessel, 2020).

Conclusioni

Il presente lavoro di ricerca ha consentito di enfatizzare alcuni elementi degni di nota nel complesso panorama della

Governance dei Big Data, delle loro potenzialità e criticità di impiego e, in senso più ampio, dell'impatto delle Tecnologie Emergenti suddette sul versante della sicurezza nazionale. Di certo, i Big Data hanno molteplici ripercussioni sulla società odierna, permettendo attività del tutto inedite: dall'accesso alla conoscenza e alla comunicazione globale alla fornitura di servizi e infrastrutture. Tuttavia, al pari delle altre Tecnologie Emergenti Dual Use, i BD sono in grado di esacerbare le minacce alla sicurezza nazionale attualmente esistenti, generandone, al tempo stesso, di nuove e imprevedibili (Australian Strategic Policy Institute, 2017). I Big Data, invero, possono essere utilizzati come sistemi d'arma in teatri di conflitto, fornendo informazioni strategiche ai competitors e capacità di puntamento cinetico (Stato Maggiore della Difesa, 2022), ma hanno, altresì, l'attitudine ad implementare meccanismi di oppressione agiti dagli apparati di Law Enforcement a danno dei cittadini, perpetuando cicli discriminatori a carico di fasce deboli della popolazione (Brayne, 2017; Couchman, 2019). È, peraltro, opinione consolidata il loro porsi a fondamento della guerra dell'informazione e dell'interferenza politica e sociale (Hammond-Errey, 2022).

I progressi senza precedenti della tecnologia digitale hanno prodotto una rivoluzione che sta concretamente trasformando la scienza e la società. I Big Data sono stati generati rapidamente in numerose discipline, come l'economia, le scienze, l'ingegneria, la medicina, la biologia e le scienze umane. Spesso accompagnati da un gran numero di caratteristiche e da un grande volume di osservazioni - ricercatori e operatori del settore hanno distinto, caratterizzato ed esplorato i BD in termini di volume, velocità, varietà, valore, viralità, volatilità, visualizzazione, viscosità e validità, identificando, complessivamente, ben 17 items ad essi afferenti (Panimalar, Shree, & Kathrine, 2017) - il loro valore risiede in un'analisi efficace che impiega metodi di inferenza statistica e di apprendimento automatico scalabili ed efficienti dal punto di vista computazionale. Lo sviluppo di nuovi metodi e strumenti statistici di gestione dei Big Data si è reso necessario quando le tecnologie tradizionali hanno rivelato la loro inettitudine alla gerenza di dataset di ampiezza e complessità inedite, soprattutto in imprese come banche, internet business, assicurazioni, industria manifatturiera e molto altro (Li, Kong, Zheng et al., 2022).

L'enorme mole di dati raccolti viene elaborata algoritmicamente e tradotta in "comportamenti" mentre la sorveglianza si trasforma, impercettibilmente, in controllo: questa l'analisi compiuta dagli esperti di settore maggiormente garantisti e attenti alle esigenze di tutela dei diritti fondamentali dei cittadini (Kapadia, 2020; Zuboff, 2022; Zuboff, Möllers, Wood et al., 2019). La tendenza tecnologica, scientifica e sociale alla raccolta e all'analisi di enormi volumi di dati ha generato quantità destabilizzanti di informazioni, in grado di sfidare le norme etiche comunemente accettate. I Big Data, di fatto, rimangono "un'idea confusa" [*"a fuzzy idea"* (Mittelstadt & Floridi, 2016: 445)], emergente in contesti sociali, scientifici ed economici talvolta apparentemente correlati solo dalle dimensioni abnormi dei dataset considerati, rispetto ai quali la comprensione delle implicazioni etiche, da parte degli stakeholders, appare in colpevole ritardo. Una meta-analisi della letteratura compiuta da Mittelstadt & Floridi (2016) ha individuato una serie di aree-chiave critiche, rilevanti sul piano etico, tra cui gli effetti dei Big Data sulle relazioni fiduciarie tra titolari e custodi dei dati (ad es. ricercatori, organizzazioni commerciali, archivi) [107]; la

[107] Quando i Big Data sono intesi come una forma di business basata sulla vendita e l'elaborazione di dati per un vantaggio commerciale, è forse inappropriato aspettarsi un rapporto basato sulla fiducia o sulla professionalità tra soggetti e custodi (Terry, 2012). Diversamente nel contesto medico, dove il paziente è "disvelato" nella sua

necessità di distinguere tra pratiche "accademiche" e "commerciali" in termini di danno potenziale per gli interessati; l'annosa questione della proprietà intellettuale derivante dall'analisi di set di dati aggregati nonché la difficoltà di fornire concreti strumenti di accesso ai dati ai titolari che non dispongono delle risorse necessarie.

La letteratura esaminata ai fini della presente ricerca riporta numerose preoccupazioni concernenti la riservatezza degli individui, alcune delle quali relative a concetti specifici come l'autonomia o la libertà di informazione del paziente/utente/consumatore. Taluni Autori hanno discusso il tema della privacy in termini di "invasività", collegata principalmente all'analisi di insiemi di dati aggregati, con specifico riferimento a quelli derivanti da fonti di geolocalizzazione o da query su Internet, anche quando tali dati siano stati anonimizzati (Markowetz, Błaszkiewicz, Montag et al., 2014; Moore, Xhafa, Barolli et al., 2013; Shilton, 2012). Altri timori emergono dall'analisi della letteratura di settore - come l'obsolescenza del software, la presenza di malware e la vulnerabilità dei supporti fisici – quali potenziali minacce alla conservazione dei dati e, in sostanza, alla tutela della sfera personale dell'utente. Gli stessi processi di anonimizzazione sono sotto accusa, stante la possibilità di una nuova identificazione del titolare mediante riferimenti incrociati con i dati riguardanti background etnico, dati di geolocalizzazione, cartelle cliniche o ulteriori metadati (Hayden, 2012).

La complessità dei Big Data genera ulteriori criticità correlate sul piano etico, prima tra tutte la tendenza – rilevabile principalmente nei mass media e nel settore commerciale - ad enfatizzarne l'obiettività, intesa come la capacità di rivelare verità oggettive senza la necessità di interpretazione umana [108] (Crawford, Gray & Miltner, 2014). Un simile approccio crea crescenti preoccupazioni, soprattutto in merito alla giustificazione delle pratiche di aggregazione e manipolazione secondaria dei dati, quando tali pratiche siano considerate funzionali a garantire rilevanti scoperte scientifiche. Le considerazioni formulate da Mittelstadt & Floridi (2016) restituiscono perfettamente lo scenario prospettato da aspettative irrealistiche in merito alla (pretesa) oggettività dei Big Data: *"At each step the data undergoes a transformation by passing through an interpretive framework, yet custodians act as though it remains an objective analogue of reality. What is or may be relevant depends on the questions being asked, which in turn depend on the purposes for which the investigation is being developed. Only a clear understanding of the purposes can ground a rational determination of the levels of abstraction at which the data are queried. The need for human intelligence is actually increasing the more data become available, in order to know which sensible questions to ask and what answers actually make sense"*. Altrimenti detto, in nessun caso i Big Data possono ritenersi una rappresentazione neutrale della realtà che descrivono, necessitando piuttosto di un'interpretazione contestualizzata (Busch, 2014), anche alla luce di obiettivi e finalità sottesi all'impiego dei dataset considerati (Hoffman & Podgurski, 2013).

Forti critiche sorgono in merito ai rischi derivanti dal fenomeno definito da taluni Autori "prosumption" (letteralmente, prosciugamento), termine usato per indicare la produzione e il consumo simultanei di

intimità ai custodi attraverso i propri dati. In medicina, infatti, è stato dimostrato come una maggiore dipendenza dalle rappresentazioni dei dati dei pazienti causate dall'adozione di pratiche di Big Data possa creare lacune nelle cure o nei rapporti medico-paziente (Beauchamp & Childress, 2009; MacIntyre, 2007; Pellegrino & Thomasma, 1993).

[108] Secondo Puschmann & Burgess (2014: 1699), il significato dei dati *"(…) is already there, just waiting to be uncovered"*.

contenuti digitali [109] (Beer, 2009, Beer & Burrows, 2010; Ritzer, Dean & Jurgenson, 2012). Divenuti importanti fonti di informazioni commerciali per i "capitalisti dei nuovi media" (Gehl, 2011: 1230), i dati generati dalla "prosumption" sono doppiamente apprezzati: perché raccolti in tempo reale su piattaforme social e come sottoprodotto del comportamento digitale degli utenti, piuttosto che direttamente, tramite sondaggi o interviste mirate (Lupton, 2014). Tali dati, in special modo impiegati per la profilazione degli utenti a scopo commerciale, costituiscono, al tempo stesso, un prezioso bacino cui attingono le Agenzie governative per tracciare il comportamento dei gruppi sociali o di particolari popolazioni all'interno della comunità (Adkins & Lury, 2011, Beer, 2009, Beer & Burrows, 2013, Boyd & Crawford, 2012). Lupton (2014) descrive questo fenomeno in termini di metriche analitiche in grado di rendere visibili caratteristiche collettive non altrimenti percepibili, al fine di effettuare valutazioni sulle prestazioni di persone e gruppi in contesti scolastici, lavorativi, sanitari o terapeutici (Ruppert, 2012; Ruppert, Law & Savage, 2013). Com'è stato acutamente sostenuto, il profiling può assumere rapidamente implicazioni di sorveglianza (Bonilla, 2014).

Problematiche relative ad un "eccesso di sorveglianza" sono state evidenziate a seguito dell'impiego di Big Data Analytics nel settore del Law Enforcement dove, a fronte di una maggiore efficienza nella pianificazione delle operazioni, è emerso l'esacerbarsi di bias nella selezione di gruppi sociali da sottoporre ad attività istituzionali di sorveglianza - fenomeno noto come "over-policing" (Dobbe, Dean, Gilbert et al., 2018; Ensign, Friedler, Neville et al., 2018; Ugwudike, 2022). Inoltre, l'integrazione di banche dati inter-istituzionali [110] può tradursi in un ampliamento indiscriminato dei soggetti sottoposti ad attività di controllo, con inclusione di individui precedentemente estranei al circuito penale (Brayne, 2017; Chi, 2017; Di Porto, 2017).

Inoltre, i dati storici sulla criminalità, che alimentano i sistemi algoritmici, possono risultare viziati da incompletezza o incorporare discriminazioni, anche non intenzionali[111], perpetuando una "conferma" dei trends precedenti – quelli che, in gergo, vengono definiti "runaway feedback loops", o cicli di retroazione[112] (Alikhademi, Drobina, Prioleau et al., 2022; Ensign, Friedler, Neville et al., 2017; Martin, Prabhakaran, Kuhlberg et al., 2020). Verosimilmente, ciò si traduce in una diversione di risorse da aree ed individui che si sottrarrebbero alla sorveglianza [effetto c.d. "under-policing" (Mashiat, Gitiaux, Rangwala et al., 2023)] in quanto privi di una "storia" pregressa di contatti con le istituzioni (Couchman, 2019; Home Affairs Select Committee, 2009; Kinsman & Wong, 2023).

---

[109] Rientrano nel fenomeno predetto sia i contenuti generati involontariamente dagli utenti (c.d. traccia digitale) sia quelli intenzionalmente caricati su piattaforme di social media dagli utenti stessi (Lupton, 2014).

[110] Cioè, afferenti a diverse Amministrazioni pubbliche e persino a soggetti privati.

[111] In sostanza, gli individui a basso reddito tendono ad avere uno "score" più elevato di quelli a reddito più alto e, dunque, ad essere assoggettati a maggiori controlli (Di Porto, 2017).

[112] I cicli di feedback (o retroazione) limitano e falsificano l'azione predittiva. Il meccanismo è piuttosto articolato. Se un soggetto è sospettato, l'aumento di attenzione nei suoi confronti si tradurrà in una "sovrarappresentazione" di infrazioni minori (che, altrimenti, passerebbero inosservate) tanto da assurgere ad items negativi nei suoi record, innescando una sorveglianza o un monitoraggio più aggressivi con conseguente inasprimento di sanzioni e ulteriore peggioramento del record individuale. Questi "errori di condotta", siano essi derivanti da una sovrarappresentazione o da un'eccessiva sorveglianza, possono condurre ad un circolo vizioso di feedback negativo, con il rischio di "over-policing" e radicalizzazione. Tale processo è dovuto principalmente al sistema automatizzato, aggregato e accelerato garantito dalle tecnologie dei Big Data (Bishop, 2006).

Ulteriori effetti di feedback – questa volta agiti su gruppi sociali e comunità online - sono riconducibili all'analisi predittiva compiuta dalla tecnologia algoritmica. La proliferazione di "bolle di filtro" ("filter bubbles") a tipo "determinismo informativo" ne rappresenta un valido esempio: trattasi di ipotesi in cui la costante personalizzazione del web inserisce gli utenti in specifiche comunità virtuali, indirizzandole verso determinati prodotti i quali, a loro volta, costituiranno il fondamento per ulteriori profilazioni, in una sorta di ciclo senza fine. Detto fenomeno concorre alla creazione delle cc.dd. camere dell'eco ("echo chambers") – ambienti virtuali al cui interno gli utenti rafforzano le proprie prospettive e pregiudizi, in seno a comunità chiuse - generando un effetto di intensificazione delle credenze individuali e di gruppo, con un meccanismo simile a quello della radicalizzazione. Una siffatta segmentazione di gruppi sociali online è stata identificata come una minaccia al discorso pubblico e all'apertura del dibattito informato, oltre ad enfatizzare il divario nelle fonti di informazione, comunicazione e collegamenti di rete che l'eccessiva personalizzazione può creare, unitamente ai crescenti rischi di estremismo derivanti dalla frammentazione delle comunità pubbliche (Barberá, 2020; Bright, 2016; Bruns, 2017; Cinelli, Morales, Galeazzi et al., 2020; Choi, Chun, Oh et al., 2020; Garimella, De Francisci Morales, Gionis et al., 2018; Ross Arguedas, Robertson, Fletcher et al., 2022; Terren & Borge-Bravo, 2021; Zeitzoff, Kelly & Lotan, 2015).

Comprendere il concetto di "Big Data" e il suo utilizzo efficiente per aumentare la produttività è diventata una questione di interesse primario per ogni organizzazione su piccola e grande scala, posto che anche una ridotta quantità di dati può rivelarsi una preziosa risorsa in termini di asset strategico aziendale (Banu & Yakub, 2020). Tanto premesso, questioni di natura giuridica in materia di Big Data Governance investono sia la disciplina del consenso individuale alla prima raccolta dei dati sia gli impieghi successivi dei dati predetti da parte delle Amministrazioni - nella misura in cui le stesse potranno accedervi anche in assenza di apposita autorizzazione del titolare – in ossequio a quella che i giuristi statunitensi definiscono "Third-Party Doctrine" (Gee, 2020; Haxhiu, 2020; Jacobi & Stonecipher, 2021; Richards, 2016; Ormerod & Trautman, 2017; Posadas Jr., 2017). Del resto, le attività di raccolta massiva, elaborazione e analisi dei dati digitali rivestono, allo stato attuale, un ruolo cruciale e del tutto inedito: mai prima d'ora, infatti, la scienza dei Big Data aveva dominato settori-chiave come quello sanitario, finanziario, industriale, accademico e della difesa a livello globale in maniera così diffusa e pervasiva (Jani & Soni, 2018). Ciononostante, i progressi nell'applicazione e negli usi commerciali dei Big Data sembrano sovrastare in modo significativo la regolamentazione giuridica del fenomeno, che appare perciò costantemente inadeguata e, per taluni versi, anacronistica rispetto ai rapidissimi processi evolutivi digitali (Hammond-Errey, 2022).

Un numero crescente di aziende sfrutta gli enormi effetti commerciali che i Big Data possono garantire, unitamente all'impatto sulla pubblicità, sul commercio e sulla Business Intelligence. L'abbondanza di dati (e, talvolta, la loro assenza) consente, infatti, alle aziende digitali che ne entrano in possesso di formulare inferenze su credenze, valori, preferenze, stato psicologico e dettagli intimi di chi li produce - non ultimi, il sentiment e le vulnerabilità delle persone - a partire dall'aggregazione di dati raccolti da attività umane apparentemente insignificanti. In breve, i Big Data hanno "slatentizzato" il valore economico delle informazioni personali e identificabili (Harford, 2014), generando il fenomeno giuridico della c.d. patrimonializzazione dei dati personali (Casalini, 2021; De Franceschi. 2017; Parenzo, 2021; Ricciuto, 2020, 2018; Senigaglia, 2020; Thobani, 2018). Com'è stato efficacemente sostenuto da Tosi (2019) *«proprio le straordinarie potenzialità delle nuove tecnologie esigono (…) uno statuto di regole capace di restituire alla persona quella centralità altrimenti negata*

dall'economia fondata sullo sfruttamento dei dati: materia prima di un nuovo capitalismo estrattivo alimentato da frammenti, spesso delicatissimi, della nostra vita. Uno statuto di regole funzionale anche alla stessa economia digitale, il cui sviluppo (…) dipende dalla fiducia riposta dai cittadini nelle sue dinamiche e quindi, in primo luogo, nel modo in cui i dati di ciascuno siano gestiti, essendo essi stati resi "un'arma con efficienza militare"». Tuttavia, il fatto che le capacità tecniche e analitiche, essenziali per il funzionamento dell'odierna società digitalizzata, siano concentrate nelle mani di un esiguo numero di entità commerciali (Big Tech) sembra minare in radice la centralità del cittadino-utente [113] (Zuboff, 2019) – ciò che ha imposto un progressivo rafforzamento della disciplina in materia di tutela del consumatore, soprattutto da quando studi empirici hanno evidenziato scarsa diligenza, da parte degli utenti, nei confronti di complesse e articolate informative privacy, con conseguenti transazioni commerciali soggette a forti asimmetrie informative (Piretti, 2020).

Minacce alla democraticità dei processi decisionali conseguenti all'impiego delle Big Data Technologies sono state ampiamente denunciate degli esperti di settore. I dati rappresentano la "nuova moneta" della società digitale: questa affermazione, rilasciata da numerosi CEO e dirigenti di società di marketing, ha permeato il mondo elettorale e della politica democratica dell'ultima decade (Bartlett, 2018). La mappatura dell'ascesa del "Technology-Intensive Campaigning" (Kreiss, 2016) e dell'uso dei Big Data da parte dei partiti (Nickerson & Rogers, 2014) è divenuta una preoccupazione diffusa a livello accademico, con le attività "Data-Driven Campaigning" (DDC) attualmente considerate dominanti nelle elezioni moderne (Dommett, 2019; Kefford, 2021; Kruschinski & Haller, 2017), grazie all'impiego di pratiche specifiche, come il micro-targeting e la profilazione degli elettori.

Lo scenario futuro sarà, verosimilmente, caratterizzato da un incremento dell'uso di sistemi decisionali automatizzati, soprattutto nell'ambito del settore pubblico, imponendo il contemperamento di istanze di pari rango giuridico – risultato, quest'ultimo, che richiede di coniugare i principi del Diritto Amministrativo e del c.d. Data Protection con la comprensione tecnica dei sistemi di machine-learning, anche al fine di identificare le aree in cui potrebbero rendersi necessarie ulteriori ricerche e approfondimenti (Cobbe, 2019; Oswald, 2018). Esempi di supporto decisionale automatizzato, impiegato da attori statali, sono rilevabili in Danimarca, dove la Pubblica Amministrazione si fonda ampiamente sul trattamento di grandi quantità di dati relativi al singolo utente, ricorrendo sempre più spesso ad analisi predittive per l'identificazione di aree specifiche di intervento, quali la frode e la vulnerabilità sociale, come parte integrante del processo di decision-making. A tal proposito, gli accademici raccomandano l'adozione di un approccio alla "governance intelligente" più critico e incentrato sull'uomo, pena la promozione di una "tecnocrazia digitale" che considera i cittadini alla stregua di "punti dati" adatti per il calcolo e la revisione, piuttosto che come individui titolari di diritti inalienabili (Jørgensen, 2023).

L'intrusione di macchine algoritmiche in ambiti esistenziali precedentemente regolamentati dal giudizio umano – la visione di un futuro in cui le agenzie

[113] In sostanza, i dati forniti dall'utente/consumatore non sono strettamente necessari all'esecuzione del contratto sottoscritto con il gestore della piattaforma digitale, ad esempio: piuttosto, servono a quest'ultima per immetterli in un processo di profilazione necessario per raggiungere determinati fini commerciali, e cioè incrementare la vendita di determinati beni e/o servizi, attraverso una pubblicità più mirata, confacente alle caratteristiche proprie dell'utente (Cerrone, 2022).

governative potrebbero legiferare con l'impiego di robot evoca immagini distopiche di individui che cedono la loro libertà al controllo di "signori computerizzati" (Coglianese & Lehr, 2019) – dovrà necessariamente passare "sotto la lente" delle dottrine fondamentali e consolidate del Diritto Amministrativo e Costituzionale, attagliandosi ai parametri legali convenzionali. Dovrà, in ogni caso, essere bandita qualsiasi applicazione "disinvolta" della governance algoritmica, a favore di un uso responsabile di essa da parte delle autorità governative, così da mitigare l'opacità dei sistemi di machine-learning con opportuni accorgimenti tecnici e mantenere la componente umana nel processo decisionale. Il rispetto di tali condizioni consentirebbe, così, alla governance algoritmica di soddisfare sia le esigenze legali che quelle di trasparenza dello Stato di diritto (Coglianese & Lehr, 2017).

Nell'ultimo quinquennio, l'Unione Europea ha concentrato le sue risorse sul miglioramento del controllo delle frontiere e sulla mitigazione dei rischi per la sicurezza connessi al terrorismo transfrontaliero e alla criminalità transnazionale – settori in cui gli Stati membri hanno dimostrato un'apprezzabile flessibilità nell'adozione di Tecnologie Emergenti in grado di fornire un'accurata identificazione degli individui, al fine di monitorarne la mobilità e ridurre i fenomeni criminosi (Gkougkoudis, Pissanidis, Demertzis, 2022). Allo stato attuale, la tecnologia digitale è incorporata nelle pratiche penali in Scozia e, sebbene il ruolo della predetta nel sistema di giustizia penale e la c.d. Datafication[114] delle pratiche penali stiano diventando aree di crescente interesse all'interno della ricerca criminologica, ben poco si conosce in merito a banche dati, valutazioni e dispositivi, nonché al concreto impatto della governance algoritmica sulle esperienze punitive a lungo termine (Casey, 2021).

Il discorso accademico sulle "Tecnologie Emergenti e Dirompenti" è stato fortemente influenzato dalla c.d. Teoria dell'Innovazione Dirompente, originariamente elaborata per analizzare le conseguenze delle interruzioni tecnologiche sui mercati e sulle imprese (Bresciani, 2016; Downes & Nunes, 2017). Poco o nulla, tuttavia, è stato indagato in merito a quello che è stato definito il "grado di disgregazione tecnosociale", posto che le tecnologie algoritmiche possono interrompere le relazioni sociali, le istituzioni, i paradigmi fondamentali della società civile, i valori e persino la natura della cognizione e dell'esperienza umana - domini colpevolmente trascurati nella produzione scientifica in materia Le dinamiche sociali, morali ed esistenziali della disgregazione tecnosociale risultano, a tutt'oggi, ampiamente sconosciute (Hopster, 2021). Simili considerazioni gettano luce sull'annosa questione dell'interazione tra Tecnologia e Società e, in ultima analisi, tra Uomo e Macchina nell'ambiente di lavoro digitalizzato, comprendente due sottosistemi distinti ma interconnessi: uno che rappresenta la dimensione tecnologica del lavoro e l'altro, che ne incarna la dimensione sociale. L'integrazione dei due sottosistemi genera un complesso insieme socio-tecnico capace di produrre risultati economici, come profitto ed efficienza, oltre a risultati umani, quali impegno e benessere. Orbene, a fronte del dominio delle componenti tecnologiche dell'Intelligenza Artificiale sul versante sociale del lavoro, i relativi effetti sul benessere e l'impegno del prestatore d'opera risultano, a tutt'oggi, poco compresi (Fischer, Wunderlich & Baskerville, 2023).

Alle numerose (e innegabili) opportunità offerte dalla diffusione di Tecnologie Emergenti come i Big Data si oppongono notevoli rischi, spesso ignorati o sottostimati, destinati ad investire assiomi a fondamento degli ordinamenti giuridici ed economici nazionali, oltre che ai rapporti di forza tra

[114] La c.d. dataficazione ("Datafication") è la tendenza tecnologica a trasformare gli aspetti della vita degli individui in dati, da cui estrapolare informazioni dotate di valore intrinseco (Biltgen & Ryan, 2016).

attori geopolitici (Antares Fumagalli, 2018). Tra le principali matrici di criticità, sul versante privatistico, spiccano l'asimmetria informativa; la dispersione del controllo sui dati personali e la perdita di qualità del consenso; l'eccesso di esposizione dell'utente e la conseguente deduzione invasiva di profili; infine, la pluralità dei titolari del trattamento dati, la molteplicità delle finalità [115] e il difetto di trasparenza del trattamento medesimo. Trattasi di forme emergenti di vulnerabilità connesse alla gestione massiva di dati personali in contesti di forte inferenza informativa (Prestipino, 2017).

Secondo Campo, Martella & Ciccarese (2018), la costruzione algoritmica delle timeline incide in maniera impercettibile sulla fruizione dei contenuti e sulla gestione della propria rete sociale di contatti (Rader, 2017) nonostante, nell'esperienza d'uso, i social network sembrino fondarsi sul mito dello "us now". Un mito particolarmente seduttivo in verità, secondo cui dette comunità formerebbero una sorta di "collettività naturale": *«(…) on this story, media institutions, at least in their normal form, drop out altogether from the picture of "what is happening". This myth offers a story focussed entirely on what "we" do when, as humans like to, we keep in touch with each other»* (Couldry, 2015: 620). Questa narrativa supporta l'illusione dell'immediatezza delle relazioni sui social, mentre le predette risultano mediate, in primo luogo, dagli algoritmi della piattaforma stessa, la cui logica può indurre gli utenti ad una maggiore esposizione a contenuti coerenti con la propria visione del mondo, alimentata da quei meccanismi centrali nella creazione delle "echo-chambers" e delle "filter bubbles" (O'Hara & Stevens, 2015; Messingschlager & Holtz, 2020; Pariser, 2012).

Il tema della polarizzazione e della formazione delle cc.dd. echo-chambers è argomento di dibattito tra gli studiosi dei social media da almeno un decennio (Barberá, Jost, Nagler et al., 2015; Conover, Gonçalves, Flammini et al. 2012), con risultati talvolta contrastanti. In ogni caso, i processi enfatizzati dagli esperti di settore si manifestano come una complessa combinazione di azioni umane e algoritmiche capaci di rafforzare le opinioni soggettive all'interno di vere e proprie "bolle informative", create sulla base di preferenze, interessi e abitudini dell'utente, che i sistemi di machine-learning elaborano nella forma di pattern comportamentali. Tali processi evidenziano il rapporto tra azione algoritmica e riproduzione del contesto sociale in termini di riproposizione e amplificazione delle discriminazioni insite nella società e nei dataset che dovrebbero, in qualche modo, rappresentarla. Vero è che, dalla letteratura scientifica recente e dai saggi divulgativi - così come dal discorso pubblico generale (Antwan, 2016; Miller, 2015; Devlin, 2016; Turner Lee, 2018) – affiora il rischio di riprodurre, esasperandoli, processi e visioni del mondo egemoniche nel tessuto sociale complessivo, con specifico riferimento ai casi di discriminazione [116] diffusi negli ambiti più disparati: dalla concessione di un prestito alla selezione per un incarico professionale; dalla profilazione su base etnica o di genere per obiettivi pubblicitari fino alla scelta di escludere i quartieri cittadini più disagiati dai servizi di *Amazon Prime* (Antwan, 2016; Cheney-

[115] Questa vulnerabilità attiene all'incapacità di mantenere nel tempo la finalità originaria di un trattamento o di associarne una compatibile; trattasi di una problematica intrinseca a tutti gli ambiti informativi data intensive, distinti da scambi di dati voluminosi, pervasivi e continui nel tempo (Antares Fumagalli, 2018).

[116] L'individuo è soggetto ad un crescente fenomeno di c.d. cross-medialità, cioè di uso congiunto di mezzi d'informazione tradizionali e di varie modalità offerte dal web, soprattutto attraverso motori di ricerca, come *Google* e *Yahoo*, e social network come *Facebook* e *Twitter*. L'utilizzo di queste informazioni, all'interno di una strategia di marketing mirata, ad esempio, rende possibile l'attuazione di pratiche di discriminazione di prezzo, cioè di differenziazione dei prezzi in funzione della tipologia del consumatore (Nicita & Delmasto, 2019).

Lippold, 2017, 2011; O'Neil, 2020; Pedreschi, Giannotti, Guidotti et al., 2018; Umoja Noble, 2018).

Rispetto al dibattito presente in letteratura, il problema della riproduzione del contesto sociale e della retroazione generata dagli algoritmi sembra articolarsi in almeno tre punti principali, ovvero la costruzione del dato, le difficoltà nella rimozione (o riduzione) degli aspetti discriminatori esistenti nei dataset considerati nonché l'interpretazione dei risultati derivanti dal processo analitico dei dati medesimi (Cardon, 2016; Christl & Spiekermann, 2016; O'Neill, 2020). Inoltre, la "governance algoritmica", intesa come la possibilità che i processi informatici assumano decisioni a proposito di ambiti socialmente sensibili in forma del tutto automatizzata – dunque, in assenza dell'intervento umano e della possibilità di ricostruire i passaggi intermedi che hanno condotto ad una determinata scelta – evoca il concetto della *"governance by numbers"* (Katz, 2017), alimentata da un potere che sembra esercitarsi in maniera impersonale, astratta e oggettiva, guidata solo da criteri di efficienza tecnica (Visentin, 2019). Detta impostazione teorica chiama in causa la c.d. fairness degli algoritmi, ovvero la loro capacità di ridurre e rimuovere le discriminazioni rispetto a determinati attributi sensibili. Orbene, studi sul tema mostrano chiaramente come determinate idee e visioni del mondo strutturino necessariamente le procedure di formalizzazione: il modo di intendere la discriminazione incide, cioè, sulla codificazione matematica della fairness. Altrimenti detto, la dimensione discrezionale, e dunque politica, è una componente fisiologica dell'algoritmo (Campo, Martella & Cicarrese, 2018; Galeotti, 2018).

L'uso dei Big Data nel settore della National Security deve superare sfide di complessità, in termini di qualità e quantità dei dati: l'attività di BD Analytics, infatti, richiede l'impiego di una pipeline di tecnologie e l'impegno costante da parte di un team di analisti multidisciplinari oltre ad una robusta integrazione dei sistemi informatici. L'analisi predittiva, in particolare, impone la raccolta di enormi volumi di dati quantitativamente e qualitativamente conformi, posto che anche gli algoritmi di classificazione più accurati soffrono dell'incidenza, statisticamente significativa, di falsi positivi e falsi negativi (Chi, 2017). Da ciò deriva la formulazione di una serie di raccomandazioni da parte della comunità di sicurezza nazionale, prime fra tutte la massimizzazione dei benefici analitici dei Big Data e lo sfruttamento del loro valore in rete, trattando i dati raccolti come un asset strategico e investendo nelle risorse umane, nei processi e nelle tecnologie informatiche. Mantenere il controllo sui Big Data è di fondamentale importanza, soprattutto per garantirne l'integrità dinnanzi a possibili condotte di "adversarial evasion" ovvero di elusione dei sistemi di sicurezza di rete da parte di cyber-criminali [117] (Australian Strategic Policy Institute, 2017).

[117] Gli attori avversi possono tentare di sconfiggere o, peggio, trasformare gli algoritmi di apprendimento automatico a loro vantaggio "attaccando" l'algoritmo stesso. Ciò comporta quello che è stato definito "innalzamento del rumore di fondo": un avversario, cioè, bombarda un modello di apprendimento con falsi positivi, il che induce gli analisti ad innalzare la soglia di allerta, generando un punto "sotto la soglia", all'interno del rumore, nell'ambito del quale gli aggressori possono strutturare il loro comportamento per evitare di essere intercettati dai sistemi di sicurezza. Vengono utilizzate quelli che, in gergo, si definiscono "tecniche contraddittorie" per "avvelenare" i dati di addestramento di un algoritmo, così da minarne la capacità di riconoscere accuratamente un'immagine - ad esempio, inserendo alcuni pixel scoloriti in un'immagine, si può ingannare l'algoritmo, facendo in modo che questo classifichi erroneamente un panda come un gibbone con una fiducia estremamente elevata – con gravi implicazioni per l'affidabilità e la veridicità degli algoritmi e delle decisioni fondate su database. L'apprendimento automatico si basa, in larga parte, sul presupposto che i dati di formazione descrivano adeguatamente i fenomeni sottostanti

Sebbene preoccupazioni in merito ad attività manipolatorie perpetrate da attori malevoli non siano inedite per la comunità della sicurezza nazionale, un'altra fonte di rischio avversario è stata recentemente ravvisata nella crescente "democratizzazione" dei sistemi di machine-learning e delle tecnologie dei Big Data, sempre più spesso disponibili su base "open source" e gratuite. Ciò favorisce la diffusione di nuove forme di criminalità informatica su vasta scala, grazie anche all'emersione di data-scientists con approcci analitici DIY ("Do-It-Yourself"): si tratterebbe di attori criminali in grado di valutare un'ampia varietà di organizzazioni, così da identificare "obiettivi di alto valore" – anche sfruttando falle nella catena della sicurezza informatica - come parte di un modello di business illegale che fornisce l'acquisizione di obiettivi sensibili quale servizio dietro corrispettivo[118] (Chua, 2023; Holt, 2016; Porcedda & Wall, 2019).

La produzione scientifica esaminata ai fini della presente ricerca è unanime nel riconoscere il significativo cambiamento tecnologico apportato dai Big Data nel campo della produzione di sicurezza, con specifico riferimento ai differenti sottocampi (pubblici e privati) afferenti alle Forze dell'Ordine e alle Organizzazioni di Intelligence deputate al mantenimento dell'ordine pubblico, alla prevenzione della criminalità, al rispetto delle leggi, alla tutela degli individui e alla salvaguardia della proprietà (Valverde, 2014). Tuttavia, nonostante le promesse dei Big Data di migliorare l'efficienza e l'efficacia degli operatori della sicurezza nazionale, ancora troppo poco si conosce in merito all'utilizzo concreto e alla percezione di detta tecnologia maturata dagli "addetti ai lavori". Vero è che gli stakeholders vantano aspettative differenti circa le potenzialità dei Big Data quali tecnologie a servizio della sicurezza, oscillanti tra rilevamento e indagine (post-delictum) e prevenzione o mitigazione del rischio (ante-delictum), oltre a migliori capacità di monitoraggio dei fenomeni criminosi in real time (Brayne, 2017; Chan & Bennett Moses, 2016).

La revisione della letteratura sul tema ha evidenziato il supporto fornito dai recenti sviluppi tecnologici all'intenso impegno dei governi moderni nella sorveglianza degli individui, unanimemente riconosciuto come parte integrante del percorso verso la globalizzazione e strettamente connesso alle attuali condizioni geopolitiche. Benché gli eventi dell'11 settembre 2001 non possano considerarsi geneticamente correlati agli attuali trends di sorveglianza (in essere già

impiegati per addestrare il sistema: ipotesi, ovviamente violata se i test di addestramento vengono intenzionalmente alterati. Esistono aree in cui le tecniche contraddittorie possono essere utilizzate per eludere il rilevamento e i sistemi di monitoraggio: si pensi ai sistemi di sorveglianza, ad esempio, dove tali tecniche potrebbero causare l'eliminazione completa di parti vitali dei dati video, oppure, potrebbero alterare i sistemi di analisi del testo utilizzati per il tracciamento di entità in passaggi scritti o per la valutazione del sentiment degli utenti sui social media. Vulnerabilità di sistema sono state evidenziate nelle pratiche di avvelenamento, offuscamento o elusione di algoritmi di machine-learning che analizzano il clustering di malware comportamentale per la sicurezza informatica contro famiglie di malware polimorfiche. Ebbene, i ricercatori hanno evidenziato come anche una corruzione del 3% dei dati possa sovvertire completamente il processo di clustering, minando integrità e sicurezza dei database nazionali. Tutto ciò si traduce, peraltro, in una erosione dei vantaggi dell'automazione algoritmica, richiedendo controllo umano e revisione costante dei processi decisionali automatizzati (Biggio, Rieck, Ariu et al., 2014; Goodfellow, 2015; Laskov & Lippman, 2010).

[118] In questa forma di criminalità informatica vengono utilizzate massicce fughe di dati e violazioni per costruire modelli ad alto valore, ossia, "obiettivi da vendere" a criminali informatici pronti a commettere cyber-attacchi a danno del target identificato (Mell, 2012; Odabas, Holt & Breiger, 2017). A tal proposito, Chua, nel suo recentissimo saggio (2023: 140) parla di *"online stolen data marketplaces"*, attualmente relegati nel Dark Web, che garantisce l'anonimato dell'utente, complice anche la mancata indicizzazione dei siti sui motori di ricerca.

da tempo), essi sono, comunque, da considerarsi una delle principali cause dell'intensificazione del controllo dei cittadini su vasta scala. Procedure di sorveglianza come la raccolta, distribuzione ed elaborazione delle informazioni a partire da quella data, infatti, hanno integrato l'approccio più dinamico – e, al tempo stesso, il più opaco e controverso - alla gestione della minaccia terroristica. Degna di nota, a tal proposito, è la recente tendenza verso quella forma di policing guidata dall'informazione, nota come "Intelligence-Led Policing" (ILP)[119], considerata "the last frontier" del contrasto al terrorismo transnazionale. Non è un caso che tutte le recenti definizioni di ILP enfatizzino l'importanza della raccolta sistematica di informazioni (non solo criminologiche) e dell'analisi dei dati rilevanti al fine di concorrere all'elaborazione di un quadro decisionale ampio e articolato, con l'obiettivo finale della prevenzione del fenomeno criminoso sia attraverso una strategia gestionale dei fattori di rischio generici sia mediante un'efficace repressione individualizzata, ossia condotta in maniera mirata nei confronti dei criminali ritenuti ad elevata pericolosità sociale (Gkougkoudis, Pissanidis & Demertzis, 2022).

Numerosi Report rilasciati da Agenzie specializzate indicano che l'attenta adozione della ILP e delle Tecnologie Emergenti - come l'Intelligenza Artificiale, la biometria potenziata e gli strumenti di sorveglianza nel contesto della policing, del controllo delle frontiere e della sicurezza nazionale - potrebbe offrire numerosi vantaggi per il futuro della polizia moderna. Tuttavia, poche sono le prove circa la portata della loro applicazione per confermare incontestabilmente sia i benefici attesi che i rischi derivanti dall'uso dei Big Data nei Paesi dell'Unione Europea, a causa della mancata applicazione uniforme e su larga scala di detta tecnologia. La carenza di letteratura specifica sul tema impone, pertanto, ulteriori e approfonditi studi sull'argomento (Gkougkoudis, Pissanidis & Demertzis, 2022), che si collocherebbero nell'alveo dell'attività di ricerca sottesa agli sforzi statali per la gerenza della "rivoluzione tecnologica" in atto, testimoniata dalle recentissime iniziative normative europee in materia di Big Data Governance (European Commission, Directorate-General for Migration and Home Affairs, 2021; Renda, Arroyo, Fanni et al., 2021; Seele, 2017).

I principali trends di sviluppo in materia di Big Data, così come emersi dalla letteratura esaminata ai fini del presente lavoro, hanno evidenziato la necessità dell'implementazione di processi di Data Governance per l'intero ciclo di vita degli stessi (Data Life Cycle) mediante capacità di raccolta, memorizzazione e trattazione; disponibilità di sistemi di analisi e predizione, di supporto al processo decisionale (Data Analytics); applicazione di adeguate forme di protezione che ne impediscano l'uso indiscriminato; definizione di nuove regole etico-giuridiche che bilancino esigenze di privacy e disponibilità delle informazioni.

Le recenti Tecnologie Emergenti e Dirompenti (EDTs) - come i veicoli autonomi, i sistemi d'arma autonomi, la tecnologia blockchain, il ride-sharing, la genomica e l'Internet of Things (IoT), solo per citarne alcune [120] - hanno innescato profondi

[119] Secondo molti accademici, l'ILP è parte di una filosofia concettuale di policing o, come la definisce Sheptycki (2005), lo sforzo tecnologico per gestire le informazioni su minacce e rischi, al fine di orientare strategicamente la missione di polizia. Ratcliffe (2008), insieme all'Organizzazione per la Sicurezza e la Cooperazione in Europa (OSCE 2021), seguono una terminologia più orientata al business, affermando che l'ILP si sta evolvendo in un modello manageriale di decisioni di allocazione delle risorse basate su prove attraverso l'assegnazione delle priorità.

[120] Stampa 3D, Intelligenza Artificiale, robot automatizzati, veicoli autonomi, Big Data Analytics, blockchain, droni, veicoli elettrici e Internet of Things sono stati identificati come c.d. Tecnologie Emergenti dalla letteratura specialistica attuale (Dong, Akram, Andersson, et al., 2021).

cambiamenti in grado di modificare in maniera incisiva i sistemi socio-economici esistenti e, sebbene dette tecnologie generino conseguenze talvolta non prevedibili a priori potendo comportare inedite forme di rischio, negarne le opportunità di miglioramento dell'efficienza economica e della qualità della vita costituirebbe una grave forma di miopia intellettuale (Taeihagh, Ramesh, & Howlett, 2021). I seri problemi che l'aumento del ritmo dell'innovazione tecnologica pone ai governi, chiamati perciò ad affrontare le sfide imposte dalla velocità e dalla portata delle trasformazioni in molteplici settori, non può certamente tradursi in un rifiuto aprioristico dell'uso delle EDTs, le cui potenzialità costituiscono, ormai, un dato fattuale incontrovertibile (Stato Maggiore della Difesa, 2022).

La recente ondata di Tecnologie Dirompenti è stata influenzata dallo sviluppo di varie tecnologie decisionali come *Crowdsourcing* e *Big Data Analytics* (Athey 2017; Prpić, Taeihagh, & Melton, 2015; Taeihagh 2017) capaci di generare, al contempo, sia nuovi problemi che ampie opportunità per i governi. Basti pensare alle risposte politiche alla recentissima pandemia di Covid-19, che hanno spaziato dall'uso di tecnologie consolidate come il GPS per migliorare l'applicazione della quarantena all'introduzione di applicazioni mobili che utilizzano il Bluetooth per migliorare il contatto-tracciamento, fino all'impiego di tecnologie più sofisticate, che utilizzano una combinazione di dati provenienti dalla telefonia mobile e apprendimento automatico per sviluppare grafici sociali, così da misurare la propensione e la frequenza con cui le persone si incontrano o transitano attraverso una data posizione geografica (Taeihagh, Ramesh, & Howlett, 2021).

Tecnologie aventi carattere di "Innovazione Dirompente" (*Disruptive Innovation*) hanno il potenziale di modificare positivamente il modo in cui le vite umane interagiscono tra loro, al pari delle tendenze del mercato e di altri aspetti della daily routine, compresi i trasporti e le comunicazioni, come dimostra una recente rassegna bibliografica sistematica su tendenze e opportunità di ricerca nell'era delle Tecnologie Dirompenti, che riconosce l'impatto delle stesse sulla logistica e sui trasporti, principalmente a supporto delle decisioni nel settore gestionale (Ab Rahman, Hamid, & Chin, 2017). L'impatto delle EDTs in contesti critici come COVID-19 è stato oggetto di un recentissimo contributo di Kiani Mavi e colleghi (2022), i quali, a partire dalla individuazione di cinque temi emergenti - operazioni di trasporto, innovazione tecnologica, economia dei trasporti, politica dei trasporti, resilienza e gestione dei disastri – hanno evidenziato la positiva influenza dell'impiego di tecniche di ottimizzazione e simulazione e, più recentemente, degli approcci di Intelligenza Artificiale (AI) e Machine Learning (ML) nella risoluzione di problematiche attinenti al trasporto merci. In particolare, è emerso come le innovazioni dell'automazione abbiano coinvolto anche le catene di trasporto e di fornitura, al pari delle Information & Communication Technologies (ICT) - realtà risultate entrambe efficaci nella costruzione di catene di approvvigionamento altamente resilienti; (Efthymiou & Ponis, 2021; Jafari, Azarian & Yu, 2022; Lagorio, Zenezini, Mangano, et al., 2022; Sun, Yu, Solvang, et al., 2021; Wang & Sarkis, 2021; Woschank, Rauch, & Zsifkovits, 2020).

L'aumento della domanda alimentare nelle aree urbane, unito a pratiche agricole convenzionali insostenibili e alla riduzione delle terre coltivabili, ha condotto allo sviluppo dell'agricoltura urbana, con l'impiego di tecniche inedite (quali agricoltura verticale e indoor, idroponica, aeroponica e acquacoltura), la cui applicazione impone l'uso di innovazioni tecnologiche in grado di sfruttarne appieno potenziale e benefici - primi fra tutti, IOT, AI, robotica, blockchain, energie rinnovabili, modificazione genetica e nanotecnologie - con innumerevoli vantaggi in termini di riduzione della povertà, aumento della sicurezza alimentare e della ecosostenibilità (Butt, Yaqub, Hammad, et al., 2019;

Dhingra, S., Dhingra, A., & Gupta, 2022; Kalantari, Mohd Tahir, Mahmoudi Lahijani, et al., 2017; Krishnan, & Swarna, 2020; Ng & Mahkeswaran, 2021; Popkova, 2022; Sharma, Dhanda & Verma, 2023; Virk, Noor, & Fiaz, 2020). Appare, pertanto, ragionevole associare il concetto delle Tecnologie Emergenti a quello del progresso sostenibile, come emerge da recenti lavori in materia di *Smart Cities*, in cui la visione della c.d. città intelligente combina differenti EDTs per generare un ecosistema in grado di ridurre drasticamente il consumo energetico dei dispositivi di ultima generazione. Originariamente sorta come alternativa per la risoluzione di complessi problemi urbani come obsolescenza strutturale, congestione del traffico, carenze energetiche, inquinamento ambientale e criminalità, la *Smart City* è attualmente considerata l'emblema della Quarta Rivoluzione Industriale, da attuarsi mediante dati, rete e Intelligenza Artificiale (AI). Il recente incremento dei tentativi di risoluzione dei problemi urbani utilizzando la tecnologia delle reti, sia a livello nazionale che internazionale, ha reso il mercato delle cc.dd. città intelligenti un motore di crescita innovativo, incentrato su energia, trasporti e sicurezza, che ricorre ad un massiccio impiego della Tecnologia dell'Informazione e della Comunicazione (ICT) come AI, Big Data, e rete 5G (Ahmed, Zhang, Jeon et al., 2022; Alaeddini, Hajizadeh, & Reaidy, 2023; Jo, Sharma, Sicato et al., 2019; Kumar, Jain, Yie et al., 2023; Samuel, Javaid, Alghamdi et al., 2022; Sharma, Podoplelova, Shapovalov et al., 2021; Singh, Sajid, Gupta et al., 2022; Singh, Sharma, Yoon et al., 2020; Yuvaraj, Praghash, Logeshwaran et al., 2023).

Nell'ultima decade, la Medicina di Laboratorio ha assistito all'avvento di numerose tecnologie innovative e dirompenti, che hanno permesso l'esecuzione di nuovi test su larga scala, implementando il settore della diagnostica, favorendo previsioni più accurate della prognosi di malattia e una migliore gestione del paziente. Alcuni esempi di innovazione dirompente in campo biomedico includono l'analisi continua di flusso, la PCR e l'uso della spettrometria di massa di MALDI-TOF per l'identificazione dell'agente patogeno, ai quali si accompagnano progetti di ricerca e sviluppo che, già da tempo, considerano lo smartphone come hub della medicina del futuro, in grado di eseguire esami di laboratorio di routine (con aggiunte hardware adeguate) in tempo reale. L'obiettivo finale è quello della raccolta dei dati biosensoriali del paziente, in vista dell'elaborazione delle informazioni cliniche da parte di una sorta di "assistente medico" virtuale, operante con il supporto del Cloud Computing e di algoritmi convalidati (Khatab & Yousef, 2021; Nam, 2015; Rifai, Topol, Chan et al., 2015). Allo stato attuale, spicca la crescente disponibilità e l'uso biomedico dell'Intelligenza Artificiale, che impone, tuttavia, un ripensamento del ruolo dello specialista di laboratorio nell'ambito dell'assistenza sanitaria e il valore aggiunto che lo stesso può fornire alla cura del paziente, poste le richieste in rapida evoluzione che l'adozione di dette tecnologie impone (Cadamuro, 2023).

La pratica della digitalizzazione della patologia è un altro esempio di tecnologia dirompente, con un esteso ambito di applicazione e innumerevoli vantaggi, primo fra tutti quello della sostituzione di batterie di test di laboratorio, che richiedono strumenti e competenze costosi e specializzati, con un'unica tecnologia più economica. L'uso dell'AI rappresenta una promettente innovazione dirompente, capace di trasformare il futuro della Patologia e della Medicina di Laboratorio, cui si affianca l'uso dei social media nei settori della pratica clinica, della formazione e delle pubblicazioni scientifiche. Numerose sono le ragioni per incoraggiare l'impiego di dette innovazioni in siffatti ambiti, tra cui l'aumento dei costi dell'assistenza sanitaria, la necessità di una migliore accessibilità alle cure diagnostiche, oltre alla crescente domanda della diagnostica di precisione. Non mancano, tuttavia, una serie di sfide che necessitano di essere affrontate: dalla considerevole resistenza alle innovazioni dirompenti da

parte degli attuali fornitori di tecnologia e degli enti governativi alle perplessità dei fornitori di assistenza sanitaria e delle compagnie assicurative (Church & Naugler, 2022; Greaves, Bernardini, Ferrari et al., 2019; Khatab & Yousef, 2021; Wilson, Steele, & Adeli, 2022) e, ancor prima, l'assenza di una definizione univoca di "Tecnologia Emergente" (*Emerging Technology* – ET) applicata al settore della Medicina di Laboratorio (Greaves, Kricka, Gruson et al., 2023).

A fronte di teorici secondo cui tali innovazioni renderebbero le tecnologie esistenti e le relazioni sociali obsolete o radicalmente alterate, provocando perturbazioni sociali, economiche ed ecologiche dirette e indirette (Meissner, Gokhberg, & Saritas, 2019; Thomas, 2019), spingendosi ad avanzare criteri per valutare il "grado di disgregazione sociale" delle singole ET (Schuelke-Leech, 2018; Hopster, 2021), altri Autori riconoscono nelle tecnologie dirompenti un fenomeno che inaugura una nuova fase nello sviluppo delle forze produttive, ponendo l'accento sulla natura creativa delle stesse, poiché in grado di implementare cicli tecnologici inediti. L'aumento degli investimenti lordi e il miglioramento dell'aspettativa di vita sarebbero tra i principali fattori responsabili dell'incremento delle prestazioni economiche. Di certo, le tecnologie digitali sono responsabili di cambiamenti epocali nella velocità di funzionamento dell'economia globale: Internet e i dispositivi digitali sarebbero, pertanto, da considerarsi autentici motori di crescita economica (Afonasova, Panfilova, Galichkina et al., 2019; Jovanović, Dlačić & Okanović, 2018; Melnyk, Dehtyarova, Kubatko et al., 2019).

Analisi approfondite sull'impatto economico delle tecnologie digitali in Europa sottolineano come l'accessibilità alle strutture ICT costituisca solo una condizione preliminare per la creazione di una società digitalizzata, mentre il "livello" e la "qualità" nell'uso di tali tecnologie - così come le condizioni che facilitano o ostacolano l'empowerment digitale - sembrano svolgere un ruolo ben più significativo nel processo in atto. In particolare, le evidenze econometriche mostrano come l'uso delle ICT (l'empowerment digitale, in primis) eserciti i principali effetti economici sull'occupazione, favorendo l'inclusione di gruppi svantaggiati nel mercato del lavoro. Se ben orientato, dunque, il processo di digitalizzazione può guidare la crescita della produttività e dell'occupazione, le quali, associate a politiche inclusive, possono contribuire efficacemente a colmare il divario tra le fasce più favorite e quelle più svantaggiate della popolazione (Evangelista, Guerrieri & Meliciani, 2014; Katz, Koutroumpis, & Martin Callorda, 2014; Vasilescu, Serban, Dimian et al., 2020). Analisi empiriche recenti, che hanno valutato la correlazione tra il tasso di crescita degli investimenti materiali e immateriali e le diverse misure di crescita della produttività, mostrano una significativa correlazione tra gli investimenti immateriali e la produttività del lavoro nel periodo successivo alla crisi finanziaria del 2008. Analogamente, entrambe le misure di crescita della produttività appaiono correlate con una combinazione di investimenti materiali e immateriali che includono ICT, software e database[121] (Bertani, Raberto &

[121] A conferma di ciò, per converso, militano studi di settore che individuano i principali fattori responsabili dei ritardi dell'Europa nel recepire gli effetti di crescita delle tecnologie digitali nella natura immateriale delle stesse – diversa, cioè, dal carattere tangibile dell'industria manifatturiera tradizionale – che impone una politica correlata ad hoc. I fallimenti del mercato vengono, inoltre, ricondotti alle esternalità derivanti da analfabetizzazione informatica della popolazione e debolezze nei sistemi di sicurezza (Gruber, 2017). Altri Autori rilevano come, nell'ultimo quindicennio, la crescita della produttività nelle economie avanzate abbia subito un marcato rallentamento, originando il paradosso della produttività della Nuova Economia Digitale – ossia, un aumento della spesa aziendale per gli asset ICT e i servizi digitali senza un aumento corrispondente della

Teglio, 2020; Corejova & Chinoracky, 2021; Sidorenko & Khisamova, 2020).

Se l'economia globale ipercompetitiva del XXI secolo costituisce un centro di innovazione, tecnologia, talento, competenze, velocità, efficienza, produttività e soddisfazione, di certo il capitale umano ne rappresenta un elemento di crescita in chiave di sostenibilità. I risultati di una ricerca in corso – con lo scopo di individuare eventuali correlazioni tra il benessere della popolazione di 11 paesi dell'Europa Centrale e Orientale (PECO) membri UE e le componenti della tendenza alla digitalizzazione, compresa la nuova industria del cloud umano, l'ICT e la connettività IOT - hanno mostrato una correlazione positiva tra le variabili dipendenti e indipendenti, confermando che la digitalizzazione dell'economia e il capitale umano sviluppato condurranno all'aumento del benessere della popolazione (Chinoracky & Corejova, 2021; Grigorescu, Pelinescu, Ion et al., 2020; Wysokińska, 2021).

Nei 27 paesi dell'UE, l'importanza della sostenibilità della trasformazione digitale (SOSDIT) è amplificata dalla necessità del bilanciamento della crescita economica con la coesione sociale, in un'ottica di raggiungimento degli obiettivi di sviluppo sostenibile (SDG). Sebbene studi di settore concordino nel riconoscere il generico impatto positivo sulla promozione sociale di fattori come connettività, capitale umano, uso dei servizi Internet, integrazione della tecnologia digitale e servizi pubblici digitali, un'attenta analisi suggerisce ai policymakers l'adozione di soluzioni concretamente attagliate al contesto locale. Benché la trasformazione digitale della società sia in atto globalmente, infatti, i progressi ad essa relativi sono suscettibili di notevoli variazioni, in ragione del differente livello regionale di progettualità in materia di innovazione sociale digitale (Imran, Liu, Wang et al., 2022; Nagy & Somosi, 2022; Nosratabadi, Atobishi, & Hegedűs, 2023). A tal proposito, prove empiriche attestanti l'importanza delle tecnologie digitali nella promozione della crescita socio-economica si desumono attingendo alla letteratura che sfrutta set di dati a livello nazionale anche in contesti extraeuropei. Studi condotti in Australia sull'impatto economico a lungo termine delle tecnologie digitali - che utilizzano l'indice di pervasività della telefonia mobile e dell'uso di Internet quali indicatori generali - suggeriscono come, tra il 2004 e il 2014, la diffusione delle ICT abbia migliorato significativamente la produzione economica interna ed estera, contribuendo ad una crescita del PIL pro capite pari al 5,8% circa (Qu, Simes, & O'Mahony, 2017). Ciò risulta confermato anche dagli outcomes di una recentissima analisi econometrica finalizzata a quantificare l'impatto della trasformazione digitale sul settore socio-economico condotta da Tudose e colleghi (2023) su un campione di 46 paesi, selezionati in base all'entità del loro reddito nazionale lordo pro capite, oltre che dai trends di una serie di ricerche a livello internazionale (Gherghina, Paşa, & Onofrei, 2021; Yoo & Yi, 2022; Zolkover, Petrunenko, Iastremska et al., 2022).

La fiorente letteratura sul tema evidenzia l'esistenza di un ampio dibattito in seno agli studiosi di Information Technology (IT) e letteratura economica in materia di digitalizzazione e crescita economica globale. Mentre molti studi empirici hanno acclarato il potenziale positivo delle IT, altri hanno rivelato un impatto dannoso delle stesse sulla promozione sociale. Sulla scia di una poderosa analisi compiuta in 59 Paesi siti in 7 aree regionali, impiegando analisi di correlazione e regressione nel triennio 2018-2020, è possibile affermare l'esistenza di una correlazione positiva tra crescita economica e digitalizzazione, ad eccezione dell'Africa settentrionale, dell'Asia occidentale e delle regioni dell'Africa subsahariana (Niranga, Sedera & Sorwar, 2022). Vero è che enormi

produttività. A detti ritardi è imputato il lento emergere degli effetti della produttività derivanti dalla trasformazione digitale (Van Ark, De Vries & Erumban, 2021).

trasformazioni digitali stanno investendo l'economia globale, arrecando benefici in termini di sviluppo economico, produttività del lavoro e occupazione ma proponendo, al contempo, molteplici sfide da affrontare sul piano etico e giuridico.

La necessità di conciliare le istanze di protezione della sicurezza nazionale e internazionale con la tutela delle libertà civili (il diritto alla riservatezza dei cittadini, in primis) incide significativamente sul nucleo del moderno Stato di diritto, generando quello che, nel discorso accademico recente, è stato definito "il dilemma Privacy vs. National Security". Negli ultimi tempi, si è registrato un crescente interesse per l'utilizzo del Data Mining nelle attività di contrasto al terrorismo, in ragione della sua capacità di individuare pattern comportamentali insoliti, attività sovversive e condotte fraudolente. Orbene, se i benefici che dette applicazioni possono apportare, in termini di salvaguardia di vite umane, appare incontestabile, esse rappresentano, tuttavia, una severa minaccia per la privacy degli individui, traducendosi in condotte intrusive della sfera personale, atte ad estrarre informazioni utili alla elaborazione di modelli predittivi di future condotte criminose. Il quesito intorno al quale ruota il dibattito attuale – e, al tempo stesso, una delle principali sfide per tecnologi, sociologi, giuristi, attivisti dei diritti umani ma, altresì, per criminologi ed esperti di sicurezza internazionale – attiene alle modalità necessarie per garantire, al contempo, la riservatezza dei cittadini e la sicurezza delle nazioni, in un contesto fortemente globalizzato: quale interesse sacrificare e, soprattutto, in quale misura?

Le problematiche etico-giuridiche correlate all'impiego di queste potenti tecnologie generano timori del tutto condivisibili, tanto nella comunità accademica quanto nel dibattito pubblico. Tuttavia, si tratta di strumenti troppo preziosi per essere aprioristicamente respinti. Al contrario, abbracciarli in maniera acritica, in assenza di apposite linee guida o adeguati controlli in merito al loro impiego comporterebbe l'assunzione di un rischio enorme ed inaccettabile. I responsabili politici e i decision-maker dovranno progressivamente acquisire una maggiore comprensione degli strumenti di Data Mining e di analisi automatizzata dei dati, implementando politiche idonee ad incoraggiare un uso responsabile ed eticamente sostenibile delle tecnologie dei Big Data.

Affrontare le sfide poste dalle Tecnologie Emergenti e Dirompenti richiede primariamente l'applicazione di una serie di strumenti atti a definire obiettivi politici, stabilire relazioni di governance e norme all'interno del settore, come la letteratura scientifica da tempo sostiene (Chapman, 2003; Howlett, Mukherjee, & Woo, 2015; Taeihagh, Givoni, & Bañares-Alcántara, 2013), senza trascurare il forte impatto che l'ampia diffusione delle EDTs dimostra di esercitare sul processo mentale e decisionale umano (Jalilovich, 2021; Jirovsky & Jirovsky, 2020; Vardan, Sofya, Nane et al., 2023). Come è stato acutamente sostenuto in occasione della *International Scientific and Practical Conference "Environmental Risks and Safety in Mechanical Engineering" (ERSME-2023 Rostov-on-Don)*, le tecnologie di Intelligenza Artificiale – e, in senso ampio, le EDTs - sono in grado di risolvere problemi complessi, garantendo lo sviluppo tecnico ed economico della società. Allo stesso tempo, tuttavia, al pari di ogni altra innovazione significativa, richiedono un'analisi attenta ed equilibrata in merito alla loro integrazione nei differenti ambiti della società umana. È richiesto, pertanto, ai policymakers uno sforzo serio e coordinato non solo per sviluppare e diffondere le tecnologie stesse, ma altresì per formare la disponibilità della compagine sociale, delle imprese e degli attori istituzionali a soluzioni tecnologiche quantomai prorompenti e innovative (Yadrovskaia, Porksheyan, Petrova et al., 2023).

"L'analisi degli scenari futuri (2040+) indica un ruolo sempre più determinante e pervasivo delle tecnologie emergenti e dirompenti che modificheranno in maniera sostanziale la società, l'economia, la politica e la dimensione della sicurezza e della difesa nazionale ed internazionale. Lo sviluppo tecnologico, caratterizzato da un andamento esponenziale, procede così rapidamente da non dare l'opportunità di comprenderne il cambiamento, tantomeno le conseguenze correlate. Si rende, quindi, necessario un approccio proattivo e condiviso tra attori istituzionali, ambiente accademico, mondo industriale e della ricerca per colmare la contrapposizione fra il ciclo di vita delle tecnologie e le tempistiche di sviluppo e di approvvigionamento. La capacità di sviluppare ed implementare tali tecnologie pone poi l'attenzione sulle questioni di "sovranità tecnologica" come componente essenziale dell'indipendenza di uno Stato e fondamentale strumento a sostegno del proprio livello di ambizione strategica rispetto ai principali competitors"

L'impatto delle Emerging & Disruptive Technologies (EDTs) sulla Difesa, Stato Maggiore della Difesa, 2022

Bibliografia

Abraham, R., Schneider, J., & vom Brocke, J. (2023). A taxonomy of data governance decision domains in data marketplaces. *Electronic Markets*, *33*(1), 22.

Ab Rahman, A., Hamid, U. Z. A., & Chin, T. A. (2017). Emerging Technologies with disruptive effects: A review. *Perintis eJournal*, *7*(2), 111-128.

Acharjya, D. P., & Ahmed, K. (2016). A survey on big data analytics: challenges, open research issues and tools. *International Journal of Advanced Computer Science and Applications*, *7*(2), 511-518.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic Literature*, *54*(2), 442-492.

Adey, P. (2012). Borders, Identification and Surveillance. In D. Lyon, K. Ball, & K. Haggerty, *Routledge Handbook of Surveillance Studies* (pp. 193-201). Routledge.

Adkins, L., & Lury, C. (2011). Introduction: Special Measures. *Sociological Review*, *5*(2), 5-23.

Afonasova, M. A., Panfilova, E. E., Galichkina, M. A., & Ślusarczyk, B. (2019). Digitalization in economy and innovation: The effect on social and economic processes. *Polish Journal of Management Studies*, *19*(2), 22-32.

Agrawal, R., Kadadi, A., Dai, X., & Andres, F. (2015, October). Challenges and opportunities with big data visualization. In *Proceedings of the 7th International Conference on Management of Computational and Collective Intelligence in Digital Eco Systems* (pp. 169-173).

Agarwal, P., Sharma, M., & Chandra, S. (2019, August). Comparison of machine learning approaches in the prediction of terrorist attacks. In *2019 Twelfth International Conference on Contemporary Computing (IC3)* (pp. 1-7). IEEE

Ahi, A. A., Sinkovics, N., Shildibekov, Y., Sinkovics, R. R., & Mehandjiev, N. (2022). Advanced technologies and international business: A multidisciplinary analysis of the literature. *International Business Review*, *31*(4), 101967.

Ahmed, S. E. (Ed.). (2017). *Big And Complex Data Analysis: Methodologies And Applications*. Springer.

Ahmed, I., Zhang, Y., Jeon, G., Lin, W., Khosravi, M. R., & Qi, L. (2022). A blockchain-and artificial intelligence-enabled smart IoT framework for sustainable city. *International Journal of Intelligent Systems*, *37*(9), 6493-6507.

Ajah, I. A., & Nweke, H. F. (2019). Big data and business analytics: Trends, platforms, success factors and applications. *Big Data and Cognitive Computing*, *3*(2), 32.

Akhgar, B., Saathoff, G. B., Arabnia, H. R., Hill, R., Staniforth, A., & Bayerl, P. S. (2015). *Application Of Big Data For National Security: A Practitioner's Guide To Emerging Technologies*. Butterworth-Heinemann.

Alaeddini, M., Hajizadeh, M., & Reaidy, P. (2023). A Bibliometric Analysis of Research on the Convergence of Artificial Intelligence and Blockchain in Smart Cities. *Smart Cities*, *6*(2), 764-795.

Albergaria, M., & Jabbour, C. J. C. (2020). The role of big data analytics capabilities (BDAC) in understanding the challenges of service information and operations management in the sharing economy: Evidence of peer effects in libraries. *International Journal of Information Management*, *51*, 102023.

ALfatih, M., Li, C., & Saadalla, N. E. (2019, July). Prediction of groups responsible for terrorism attack using tree based models. In *Proceedings of the 2019 International Conference on Artificial Intelligence and Computer Science* (pp. 320-324).

Alghamdi, H., & Selamat, A. (2022). Techniques to detect terrorists/extremists on the dark web: a review. *Data Technologies and Applications*, *56*(4), 461-482.

Ali, S. M., Gupta, N., Nayak, G. K., & Lenka, R. K. (2016, December). Big data visualization: Tools and challenges. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 656-660). IEEE.

Ali, F., Basheer, R., Kawas, M., & Alkhatib, B. (2023). Towards Detecting Influential Members and Critical Topics from Dark Web Forums: A Data Mining Approach. *Journal of Information and Organizational Sciences*, *47*(1), 1-20.

Ali, S., Masood, K., Riaz, A., & Saud, A. (2022, February). Named entity recognition using deep learning: A review. In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)* (pp. 1-7). IEEE.

Alikhademi, K., Drobina, E., Prioleau, D., Richardson, B., Purves, D., & Gilbert, J. E. (2022). A review of predictive policing from the perspective of fairness. *Artificial Intelligence and Law*, *30*, 1-17.

Al-Jarrah, O. Y., Yoo, P. D., Muhaidat, S., Karagiannidis, G. K., & Taha, K. (2015). Efficient machine learning for big data: A review. *Big Data Research*, *2*(3), 87-93.

Alkhatib, B. & Basheer, R.S. (2019). Mining the Dark Web: A Novel Approach for Placing a Dark Website under Investigation. *International Journal of Modern Education and Computer Science, 11*(10), 1- 13.

Alsaedi, N., Burnap, P., & Rana, O. (2017). Can we predict a riot? Disruptive event detection using Twitter. *ACM Transactions on Internet Technology*, *17*(2), 1-26.

Al-Sai, Z. A., & Abualigah, L. M. (2017, May). Big data and E-government: A review. In *2017 8th International Conference on Information Technology (ICIT)* (pp. 580-587). IEEE.

Al-Sai, Z. A., Husin, M. H., Syed-Mohamad, S. M., Abdin, R. M. D. S., Damer, N., Abualigah, L., & Gandomi, A. H. (2022). Explore big data analytics applications and opportunities: A review. *Big Data and Cognitive Computing*, *6*(4), 157.

Alsunaidi, S. J., Almuhaideb, A. M., Ibrahim, N. M., Shaikh, F. S., Alqudaihi, K. S., Alhaidari, F. A., ... & Alshahrani, M. S. (2021). Applications of big data analytics to control COVID-19 pandemic. *Sensors*, *21*(7), 2282.

Amato Mangiameli, A. (2022). Intelligenza artificiale, big data e nuovi diritti. *Rivista Italiana di Informatica e Diritto*, *4*(1), 93-101.

Amoore, L. (2019). Introduction: Thinking with algorithms: Cognition and computation in the work of N. Katherine Hayles. *Theory, Culture & Society*, *36*(2), 3-16.

Amoore, L., & Raley, R. (2017). Securing with Algorithms: Knowledge, Decision, Sovereignty. *Security Dialogue*, *48*(1), 3-10.

Anderson, J. (2020). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power by Shoshana Zuboff. *Archivaria*, *90*(1), 192-195.

André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., ... & Yang, H. (2018). Consumer choice and autonomy in the age of artificial intelligence and big data. *Customer Needs and Solutions*, *5*, 28-37.

Andrejevic, M. (2017). To Preempt a Thief. *International Journal of Communication*, *11*, 879-896.

Andrejevic, M., & Gates, K. (2014). Big Data Surveillance: Introduction. *Surveillance & Society*, *12*(2), 185-196.

Aneesh, A. (2009). Global labor: Algocratic modes of organization. *Sociological Theory*, *27*(4), 347-370.

Angwin, J. (2014). *Dragnet Nation: A Quest For Privacy, Security, And Freedom In A World Of Relentless Surveillance*. MacMillan.

Anshari, M., Almunawar, M. N., & Masri, M. (2022). Digital twin: Financial technology's next frontier of robo-advisor. *Journal of Risk and Financial Management*, *15*(4), 163.

Asgari-Chenaghlu, M., Feizi-Derakhshi, M. R., Farzinvash, L., Balafar, M. A., & Motamed, C. (2022). CWI: A multimodal deep learning approach for named entity recognition from social media using character, word and image features. *Neural Computing and Applications*, 1-18.

Antares Fumagalli, D. (2018). Privacy e sicurezza nazionale. L'evoluzione della Data Protection da diritto dell'individuo a interesse pubblico. *Gnosis, 2*, 131-145.

Anstead, N. (2017). Data-driven campaigning in the 2015 United Kingdom general election. *The International Journal of Press/Politics,* ,22(3), 294-313.

Appel, R. E., & Matz, S. C. (2021). Psychological targeting in the age of Big Data. In D. Wood, S.J. Read, P.D. Harms, and A. Slaughter (Eds.) *Measuring and Modeling Persons and Situation* (pp. 193-222). Academic Press.

Aradau, C., & Blanke, T. (2017). Politics of prediction: Security and the time/space of governmentality in the age of big data. *European Journal of Social Theory*, *20*(3), 373-391

Aragona, B., & Felaco, C. (2019). Big data from below. Researching data assemblages. *Tecnoscienza: Italian Journal of Science & Technology Studies*, *10*(1), 51-70.

Ardito, L., Scuotto, V., Del Giudice, M., & Petruzzelli, A. M. (2019). A bibliometric analysis of research on Big Data analytics for business and management. *Management Decision*, *57*(8), 1993-2009

Armengol-Estapé, J., Soares, F., Marimon, M., & Krallinger, M. (2019). PharmacoNER Tagger: a deep learning-based tool for automatically finding chemicals and drugs in Spanish medical texts. *Genomics & Informatics*, *17*(2), e15.

Asghar, Z., Ali, T., Ahmad, I., Tharanidharan, S., Nazar, S. K. A., & Kamal, S. (2019). Sentiment analysis on automobile brands using Twitter data. In *Intelligent Technologies and Applications: First International Conference, INTAP 2018, Bahawalpur, Pakistan, October 23-25, 2018, Revised Selected Papers 1* (pp. 76-85). Springer.

Asharef, M., Omar, N., Albared, M., Minhui, Z., Weiming, W., & Jingjing, Z. (2012). Arabic named entity recognition in crime documents. *Journal of Theoretical and Applied Information Technology*, *44*(1), 1-6.

Asmai, S. A., Salleh, M. S., Basiron, H., & Ahmad, S. (2018). An enhanced Malay named entity recognition using combination approach for crime textual data analysis. *International Journal of Advanced Computer Science and Applications*, *9*(9), 474-483.

Athey, S. (2017). Beyond Prediction: Using Big Data for Policy Problems. *Science*, *355*(6324), 483-485.

Athmaja, S., Hanumanthappa, M., & Kavitha, V. (2017, March). A survey of machine learning algorithms for big data analytics. In *2017 International Conference On Innovations In Information, Embedded And Communication Systems (ICIIECS)* (pp. 1-4). IEEE.

Au, A. K. T. P., Curcin, V., Ghanem, M., Giannadakis, N., Guo, Y., Jafri, M., ... & Zhang, Y. (2004, September). Why grid-based data mining matters? fighting natural disasters on the grid: from SARS to landslides. In *UK E-science All-hands Meeting (AHM 2004), Nottingham, UK* (pp. 121-126).

Awotunde, J. B., Adeniyi, E. A., Ogundokun, R. O., & Ayo, F. E. (2021). Application of big data with Fintech in financial services. In P. Moon Sub Choi, and S.H. Huang (Eds.) *Fintech with Artificial Intelligence, Big Data, and Blockchain* (pp. 107-132). Springer.

Awotunde, J. B., Folorunso, S. O., Jimoh, R. G., Adeniyi, E. A., Abiodun, K. M., & Ajamu, G. J. (2021). Application of artificial intelligence for COVID-19 epidemic: an exploratory study, opportunities, challenges, and future prospects. *Artificial Intelligence for COVID-19*, 47-61.

Balakrishnan, S., & Rahul, R. (2018, November). Big Data in Business Intelligence. *CSI Communications, 42*(8), 21-23.

Baldwin-Philippi, J. (2020). Data ops, objectivity, and outsiders: Journalistic coverage of data campaigning. *Political Communication*, *37*(4), 468-487.

Ball, K. (2019). Review of Zuboff's the Age of Surveillance Capitalism. *Surveillance & Society*, *17*(1/2), 252-256.

Ball, W. D. (2018). The Plausible and the Possible: A Bayesian Approach to the Analysis of Reasonable Suspicion. *Am. Crim. L. Rev.*, *55*, 511.

Ball, K., & Webster, F. (2003). *The Intensification Of Surveillance: Crime, Terrorism And Warfare In The Information Era*. Pluto Press.

Ballatore, A. & Natale, S. (2018). Fallimenti, controversie e il mito tecnologico dell'Intelligenza Artificiale. In P. Magaudda & G. Balbi. *Fallimenti digitali: Un'archeologia dei «nuovi» media* (pp. 137-149). Unicopli.

Bambauer, D. E. (2013). Privacy versus Security. *J. Crim. L. & Criminology*, *103*, 667.

Banu, A., & Yakub, M. (2020). Evolution of big data and tools for big data analytics. *Journal of Interdisciplinary Cycle Research*, *12*(10), 309-316.

Barberá, P. (2020). Social Media, Echo Chambers, and Political Polarization. In N. Persily, and Tucker, J. (2006). *Social Media and Democracy: The State of the Field and Prospects for Reform* (pp. 34-55). Cambridge University Press.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting from left to right: Is online political communication more than an echo chamber?. *Psychological Science*, *26*(10), 1531-1542.

Barley, S. R. (1996). Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organizational Studies. *Administrative Science Quarterly*, *41*(3), 404-441.

Barley, S. R. (1986). Technology as an occasion for structuring: Evidence from observations of CT scanners and the social order of radiology departments. *Administrative Science Quarterly*, 78-108.

Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 671-732.

Barral, L. V., Pinet, F., Tacnet, J. M., & Jousselme, A. L. (Eds.) (2023). Combining UML profiles to design serious games dedicated to trace information in decision processes. In *Research Anthology on Game Design, Development, Usage, and Social Impact* (pp. 212-239). IGI Global.

Barrett, L. (2017). Reasonably suspicious algorithms: Predictive Policing at the United States border. *NYU Rev. L. & Soc. Change*, *41*, 327.

Barsky, R. F. (1998). *Noam Chomsky: A Life Of Dissent*. MIT Press.

Bartlett, J. (2018). *The People Vs Tech: How the Internet Is Killing Democracy (and How We Save It)*. Penguin.

Bartosik-Purgat, M., & Ratajczak-Mrożek, M. (2018). Big data analysis as a source of companies' competitive advantage: A review. *Entrepreneurial Business and Economics Review*, *6*(4), 197-215.

Bartz–Beielstein, T., Parsopoulos, K. E., & Vrahatis, M. N. (2004). Design and analysis of optimization algorithms using computational statistics. *Applied Numerical Analysis & Computational Mathematics*, *1*(2), 413-433.

Batistič, S., & van der Laken, P. (2019). History, evolution and future of big data and analytics: A bibliometric analysis of its relationship to performance in organizations. *British Journal of Management*, *30*(2), 229-251.

Bayamlıoğlu, E., & Leenes, R. (2018). The 'rule of law 'implications of data-driven decision-making: a techno-regulatory perspective. *Law, Innovation and Technology*, *10*(2), 295-313.

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Jama*, *319*(13), 1317-1318.

Beauchamp, T. L., & Childress, J. F. (2009). *Principles of Biomedical Ethics* (6th ed.). Oxford University Press.

Becker, H. S. (1963). *Outsiders* (Vol. 1973). New York: Free Press.

Beer, D. (2017). The Social Power of Algorithms. *Information, Communication & Society*, *20*(1), 1-13.

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious, *New Media & Society*, 11(6). 985-1002.

Beer, D., & Burrows, R. (2013). Popular culture, digital archives and the new social life of data. *Theory, Culture & Society*, 30(4), 47–71.

Beer, D., & Burrows, R. (2010). Consumption, prosumption and participatory web cultures: an introduction, *Journal of Consumer Culture*, 10 (1). 3-12.

Bendre, M. R., & Thool, V. R. (2016). Analytics, challenges and applications in big data environment: a survey. *Journal of Management Analytics*, *3*(3), 206-239.

Bennett, C. (2016). Voter Databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America?. *International Data Privacy Law*, *6*(4), 261-275.

Bennett, C. J., & Lyon, D. (2019). Data-driven elections: Implications and challenges for democratic societies. *Internet Policy Review*, *8*(4).

Bergamaschi F., Bianconi D. e Mattavelli A. (2023). *Business intelligence per le PMI. Manuale per Professionisti e Imprenditori*, Maggioli.

Berk, R. (2021). Artificial Intelligence, Predictive Policing, and Risk Assessment for Law Enforcement. *Annual Review of Criminology*, 4(1), 209-237.

Berman, E. (2019). Individualized suspicion in the age of Big Data. *Iowa L. Rev.*, *105*, 463.

Berman, E. (2018). A government of laws and not of machines. *Bul Rev.*, *98*, 1277.

Bertani, F., Raberto, M., & Teglio, A. (2020). The productivity and unemployment effects of the digital transformation: an empirical and modelling assessment. *Review of Evolutionary Political Economy*, *1*, 329-355.

Bertot, J. C., & Choi, H. (2013, June). Big data and e-government: issues, policies, and recommendations. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 1-10).

Bhatt, A., Sengar, V., & Pandey, R. (2022). Role of Big data Analysis and its Challenges. *Res Militaris*, *12*(5), 1431-1440.

Biggio, B., Rieck, K., Ariu, D., Wressnegger, C., Corona, I., Giacinto, G., & Roli, F. (2014, November). Poisoning behavioral malware clustering. In *Proceedings of the 2014 Workshop on Artificial Intelligence and Security Workshop* (pp. 27-36).

Bignami, F. (2007). European Versus American Liberty: A Comparative Privacy Analysis Of Antiterrorism Data Mining. *BCL Rev.*, *48*, 609-698.

Bikakis, N. (2018). Big Data visualization tools. *arXiv preprint arXiv:1801.08336*.

Bikakis, N., Papastefanatos, G., & Papaemmanouil, O. (2019). Big Data exploration, visualization and analytics. *Big Data Res*, *18*(10), 1016.

Biltgen, P., & Ryan, S. (2016). *Activity-Based Intelligence: Principles and Applications*. Artech House.

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018, April). 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems*, 1-14.

Birrer, F. A. (2005). Data mining to combat terrorism and the roots of privacy concerns. *Ethics and Information Technology*,*7* , 211-220.

Bishop, C.M. (2006). *Pattern Recognition And Machine Learning*. Springer, 2006

Black, E., Elzayn, H., Chouldechova, A., Goldin, J., & Ho, D. (2022, June). Algorithmic fairness and vertical equity: Income fairness with IRS tax audit models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1479-1503).

Blasch, E., Pham, T., Chong, C. Y., Koch, W., Leung, H., Braines, D., & Abdelzaher, T. (2021). Machine learning/artificial intelligence for sensor data fusion–opportunities and challenges. *IEEE Aerospace and Electronic Systems Magazine*, *36*(7), 80-93.

Boccia Artieri, G. (2014). La rete dopo l'overload informativo. La realtà dell'algoritmo da macchia cieca a bene comune. *Paradoxa*, *7*(2), 100-113.

Boccia Artieri, G., & Marinelli, A. (2018). Introduzione: piattaforme, algoritmi, formati. Come sta evolvendo l'informazione online. *Problemi dell'Informazione*, *43*(3), 349-368.

Bongiovi, J. R. (2019). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power by Shoshana Zuboff. *Social Forces*, *98*(2), e45.

Bonilla, D.N. (2014). Information Management Professionals Working For Intelligence Organizations: Ethics and Deontology Implications. *Security and Human Rights 24*(3-4). 264-279.

Bonneau, J., & Preibusch, S. (2010). The privacy jungle: On the market for data protection in social networks. In *Economics Of Information Security And Privacy* (pp. 121-167). Boston, MA: Springer US.

Bonner, S., Kureshi, I., Brennan, J., & Theodoropoulos, G. (2017). Exploring the evolution of big data technologies. In *Software Architecture For Big Data And The Cloud* (pp. 253-283). Morgan Kaufmann.

Borgesius, F. J. Z., Möller, J., Kruikemeier, S., Fathaigh, R. Ó., Irion, K., Dobber, T., ... & De Vreese, C. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82-96.

Borradaile, G., & Reeves, J. (2020). Sousveillance Capitalism. *Surveillance & Society*, *18*(2), 272-275.

Boulos, M. N. K., Sanfilippo, A. P., Corley, C. D., & Wheeler, S. (2010). Social Web mining and exploitation for serious applications: Technosocial Predictive Analytics and related technologies for public health, environmental and national security surveillance. *Computer Methods and Programs in Biomedicine*, *100*(1), 16-23.

Boustani, N., Emrouznejad, A., Gholami, R., Despic, O., & Ioannou, A. (2023). Improving the predictive accuracy of the cross-selling of consumer loans using deep learning networks. *Annals of Operations Research*, 1-18.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662-679.

Bragazzi, N. L., Dai, H., Damiani, G., Behzadifar, M., Martini, M., & Wu, J. (2020). How big data and artificial intelligence can help better manage the COVID-19 pandemic. *International Journal of Environmental Research And Public Health*, *17*(9), 3176.

Brayne, S. (2020). *Predict And Surveil: Data, Discretion, And The Future Of Policing*. Oxford University Press.

Brayne, S. (2017). Big Data Surveillance: The case of policing. *American Sociological Review*, *82*(5), 977-1008.

Brayne, S., & Christin, A. (2021). Technologies of Crime prediction: The Reception of Algorithms in Policing and Criminal Courts. *Social Problems*, *68*(3), 608-624.

Bresciani, S. (2016). *Le Innovazioni Dirompenti*. Giappichelli.

Brezina, P., Eberhartinger, E., & Zieser, M. (2021). *The Future of Tax Audits?: The Acceptance of Online-based, Automated Tax Audits and Their Effects on Trust and Power*. WU Vienna University of Economics and Business.

Bridgelall, R. (2022). Applying unsupervised machine learning to counterterrorism. *Journal of Computational Social Science*, *5*(2), 1099-1128.

Bright, J. (2016). Explaining the emergence of echo chambers on social media: the role of ideology and extremism. *arXiv preprint arXiv:1609.05003*.

Bruce, F., Malcolm, J., & O'Neill, S. (2017). Big data: Understanding how creative organisations create and sustain their networks. *The Design Journal*, *20*(sup1), S435-S443.

Bruns, A. (2017, September). Echo chamber? What echo chamber? Reviewing the evidence. In *6th Biennial Future of Journalism Conference (FOJ17)*.

Buchi, G., Cugno, M., & Castagnoli, R. (2019). Industry 4.0 and Internationalization: a systematic approach to the analysis of causal relationships. *Identità, innovazione e impatto dell'aziendalismo italiano. Dentro l'economia digitale. Atti del XXXIX Convegno Nazionale AIDEA*, 1-13.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 1-12.

Busch, L. (2014). A Dozen Ways To Get Lost In Translation: Inherent Challenges In Large Scale Data Sets. *International Journal of Communication*, *8*(18). 1727-1744.

Butt, M. F. U., Yaqub, R., Hammad, M., Ahsen, M., Ansir, M., & Zamir, N. (2019, April). Implementation of aquaponics within IoT framework. In *2019 SoutheastCon* (pp. 1-6). IEEE.

Cadena, J., Korkmaz, G., Kuhlman, C. J., Marathe, A., Ramakrishnan, N., & Vullikanti, A. (2015). Forecasting social unrest using activity cascades. *PloS one*, *10*(6), e0128879.

Cadamuro, J. (2023). Disruption vs. evolution in laboratory medicine. Current challenges and possible strategies, making laboratories and the laboratory specialist profession fit for the future. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *61*(4), 558-566.

Calo, R. (2017). Artificial Intelligence Policy: a primer and roadmap. *UCDL Rev.*, *51*, 399.

Campo, E., Martella, A., & Ciccarese, L. (2018). Gli algoritmi come costruzione sociale. Neutralità, potere e opacità. *The Lab's Quarterly*, *20*(4), 7-24.

Canter, D., & Larkin, P. (1993). The environmental range of serial rapists. *Journal of Environmental Psychology*, *13*(1), 63-69.

Cardon, D. (2016). Deconstructing the algorithm: four types of digital information calculations. In R. Seyfert & J. Roberge (Eds.) *Algorithmic Cultures*, (pag. 95-110). Routledge.

Carter, D. (2009). *Law Enforcement Intelligence: A Guide for State, Local, and Tribal Law Enforcement Agencies*. Michigan State University.

Castagnoli, R., Büchi, G., Coeurderoy, R., & Cugno, M. (2022). Evolution of industry 4.0 and international business: A systematic literature review and a research agenda. *European Management Journal*, *40*(4), 572-589.

Cate, F. H. (2008). Government Data Mining: The need for a legal framework. *Harv. CR-CLL Rev.*, *43*, 435.

Cerrone, C. R. (2022). Dati personali come moneta di scambio per le prestazioni rese nell'ambito dei mercati digitali. *Data Protection Law, 2*, 3-19.

Chadwick, A., & Stromer-Galley, J. (2016). Digital media, power, and democracy in parties and election campaigns: Party decline or party renewal?. *The International Journal of Press/Politics*, *21*(3), 283-293.

Chae, B. K. (2019). A General framework for studying the evolution of the digital innovation ecosystem: The case of big data. *International Journal of Information Management*, *45*, 83-94.

Chai, C., & Li, G. (2020). Human-in-the-loop Techniques in Machine Learning. *IEEE Data Eng. Bull.*, *43*(3), 37-52.

Chan, J. (2021). The future of AI in policing: Exploring the sociotechnical imaginaries. In J. McDaniel and K. Pease (Eds.) *Predictive Policing and Artificial Intelligence* (pp. 41-57). Routledge.

Chan, J., & Bennett Moses, L. (2016). Is big data challenging criminology?. *Theoretical Criminology*, *20*(1), 21-39.

Chan, J., Sanders, C., Bennett Moses, L., & Blackmore, H. (2022). Datafication and the practice of intelligence production. *Big Data & Society*, *9*(1), 1-13.

Chancellor, S. (2023). Toward Practices for Human-Centered Machine Learning. *Communications of the ACM*, *66*(3), 78-85.

Chancellor, S., Baumer, E. P., & De Choudhury, M. (2019, November). Who is the" human" in human-centered machine learning: The case of predicting mental health from social media. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 147:1-32.

Chapman, R. (2003). A Policy Mix for Environmentally Sustainable Development – Learning from the Dutch Experience. *New Zealand Journal of Environmental Law, 7*(1), 29-51.

Chaudhary, M., & Bansal, D. (2022, July 07). Open source intelligence extraction for terrorism-related information: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *12*(5), e1473.

Chen, N., Chen, Y., You, Y., Ling, H., Liang, P., & Zimmermann, R. (2016, April). Dynamic urban surveillance video stream processing using fog computing. In *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)* (pp. 105-112). IEEE.

Chen, Z. Y., Fan, Z. P., & Sun, M. (2023). Machine Learning Methods for Data-Driven Demand Estimation and Assortment Planning Considering Cross-Selling and Substitutions. *INFORMS Journal on Computing*, *35*(1), 158-177.

Chen, Y. C., & Hsieh, T. C. (2014). Big data for digital government: Opportunities, challenges, and strategies. *International Journal of Public Administration in the Digital Age (IJPADA)*, *1*(1), 1-14.

Chen, Y., & Ji, W. (2021). Enhancing situational assessment of critical infrastructure following disasters using social media. *Journal of Management in Engineering*, *37*(6), 04021058.

Chen, Y., Li, C., & Wang, H. (2022). Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000–2021). *Forecasting*, *4*(4), 767-786.

Chen, H., Reid, E., Sinai, J., Silke, A., & Ganor, B. (2008). *Terrorism Informatics: Knowledge Management And Data Mining For Homeland Security* (Vol. 18). Springer.

Chen, Y., Shi, Y., Wei, X., & Zhang, L. (2014). Domestic systemically important banks: A quantitative analysis for the Chinese banking system. *Mathematical Problems in Engineering*, *4,* 1-19.

Chen, Y., Wang, Q., & Ji, W. (2020). Rapid assessment of disaster impacts on highways using social media. *Journal of Management in Engineering*, *36*(5), 04020068.

Chen, D., Yan, Q., Wu, C., & Zhao, J. (2021). SQL injection attack detection and prevention techniques using deep learning. *Journal of Physics: Conference Series,* 1757(1), 012055.

Chen, C. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314-347.

Cheney-Lippold, J. (2017). *We Are Data*. New York University Press.

Cheney-Lippold, J. (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, *28*(6), 164-181.

Cheng, X., Guo, F., Chen, J., Li, K., Zhang, Y., & Gao, P. (2019). Exploring the trust influencing mechanism of robo-advisor service: A mixed method approach. *Sustainability*, *11*(18), 4917.

Chermak, S., Carter, J., Carter, D., McGarrell, E. F., & Drew, J. (2013). Law enforcement's information sharing infrastructure: A national assessment. *Police Quarterly*, *16*(2), 211-244.

Chern, C. C., Lei, W. U., Huang, K. L., & Chen, S. Y. (2021). A decision tree classifier for credit assessment problems in big data environments. *Information Systems and e-Business Management*, *19*, 363-386.

Chesney, R. M. (2002). Civil Liberties and the Terrorism Prevention Paradigm: The Guilt by Association Critique. *Mich. L. Rev.*, *101*, 1408.

Chi, M. (2017). *Big Data in National Security*. Australian Strategic Policy Institute.

Chiarvesio, M., & Romanello, R. (2018). Industry 4.0 technologies and internationalization: Insights from Italian companies. In *International business in the information and digital age* (Vol. 13, pp. 357-378). Emerald Publishing Limited.

Chinoracky, R., & Corejova, T. (2021). How to evaluate the digital economy scale and potential?. *Entrepreneurship and Sustainability Issues*, *8*(4), 536.

Choi, D., Chun, S., Oh, H., Han, J., & Kwon, T. T. (2020). Rumor propagation is amplified by echo chambers in social media. *Scientific Reports*, *10*(1), 310.

Christl, W., & Spiekermann, S. (2016). *Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy*. Facultas Verlags- und Buchhandels AG Facultas Universitätsverlag.

Chua, Y.T. (2023). Sale of private, confidential, and personal data. In D. Hummer, and J. Byrne (Eds.) *Handbook on Crime and Technology* (pp. 138-155). Edward Elgar.

Church, D. L., & Naugler, C. (2022). Using a systematic approach to strategic innovation in laboratory medicine to bring about change. *Critical Reviews in Clinical Laboratory Sciences*, *59*(3), 178-202.

Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2020). Echo chambers on social media: A comparative analysis. *arXiv preprint arXiv:2004.09603*.

Cobbe, J. (2019). Administrative law and the machines of government: judicial review of automated public-sector decision-making. *Legal Studies*, *39*(4), 636-655.

Coglianese, C., & Lehr, D. (2019). Transparency and algorithmic governance. *Administrative Law Review*, *71*(1), 1-56.

Cohen, S. (1985). *Visions of Social Control: Crime, Punishment and Classification.* Polity Press.

Colangelo, G. (2016). Big data, piattaforme digitali e antitrust. *Mercato Concorrenza Regole*, *18*(3), 425-460.

Colbary, K. (2021). Outsourcing the Police: How Reliance on the Private Sector for Law Enforcement Threatens Privacy Legislation Around the World. *Nw. J. Int'l L. & Bus.*, *41(*2), 213-224.

Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and other National Goals, National Research Council. (2008). *Protecting Individual Privacy In The Struggle Against Terrorists: A Framework For Program Assessment*, National Academies Press.

Conover, M. D., Gonçalves, B., Flammini, A., & Menczer, F. (2012). Partisan asymmetries in online political activity. *EPJ Data Science*, *1*(1), 1-19.

Corejova, T., & Chinoracky, R. (2021). Assessing the potential for digital transformation. *Sustainability*, *13*(19), 11040.

Corsi, A., de Souza, F. F., Pagani, R. N., & Kovaleski, J. L. (2021). Big data analytics as a tool for fighting pandemics: a systematic review of literature. *Journal of Ambient Intelligence and Humanized Computing*, *12*(10), 9163-9180.

Côrte-Real, N., Ruivo, P., Oliveira, T., & Popovič, A. (2019). Unlocking the drivers of big data analytics value in firms. *Journal of Business Research*, *97*, 160-173.

Couldry, N. (2015). The myth of 'us': digital networks, political change and the production of collectivity. *Information, Communication & Society*, *18*(6), 608-626.

Crawford, K., Gray, M.L., & K. Miltner. (2014). Critiquing Big Data: Politics, Ethics, Epistemology. Special Section Introduction. *International Journal of Communication 8*(10). 1663-1672.

Crawford, K., & Schultz, J. (2014). Big Data and due process: Toward a framework to redress predictive privacy harms. *BCL Rev.*, *55*, 93.

Criminal Intelligence Service Canada. (2007). *Strategic early warning for criminal intelligence: theoretical framework and sentinel methodology*. CISC, 2007(6).

Cuffaro V., D'Orazio, R. & Ricciuto, V. (Eds.) (2019) *I dati personali nel diritto europeo*, Giappichelli.

Cui, Y., Kara, S., & Chan, K. C. (2020). Manufacturing big data ecosystem: A systematic literature review. *Robotics and Computer-Integrated Manufacturing*, *62*, 101861.

Cui, Y., Koppol, P., Admoni, H., Niekum, S., Simmons, R., Steinfeld, A., & Fitzgerald, T. (2021, August). Understanding the relationship between interactions and outcomes in human-in-the-loop machine learning. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, Survey Track*. (pp. 4382-4391). IJCAI.

Custers, B. (2012). Technology in policing: Experiences, obstacles and police needs. *Comput. Law Secur. Rev., 28*, 62-68.

D'Acquisto, G., & Naldi, M. (2017). *Big Data e Privacy by Design* (Vol. 5). Giappichelli.

Das, S. R. (2016). Big Data's Big Muscle. *Finance & Development*, *53*(3), 26-27.

Das, S., Behera, R. K., & Rath, S. K. (2018). Real-time sentiment analysis of twitter streaming data for stock prediction. *Procedia Computer Science*, *132*, 956-964.

Deeks, A. S. (2018). Predicting enemies. *Virginia Law Review*, *104*(8), 1529-1592.

De Franceschi, A. (2017). *La circolazione dei dati personali tra privacy e contratto.* ESI.

De Giorgi, G., & De Masi, F. (2019). Big Data e Assicurazioni. Regolamentazione dei Mercati e Tutela Giuridica. *Ithaca: Viaggio nella Scienza*, *(1)*, 81-86.

Degli Esposti, S. (2014). When big data meets dataveillance: The hidden side of analytics. *Surveillance & Society* 12(2), 209- 225.

De Mauro, A. (2019). *Big Data Analytics: Analizzare e interpretare dati con il Machine Learning*. Apogeo.

De Minico, G. (2019). Big Data e la debole resistenza delle categorie giuridiche. Privacy e lex mercatoria. *Diritto Pubblico*, *25*(1), 89-116.

Dencik, L., Hintz, A., & Carey, Z. (2018). Prediction, pre-emption and limits to dissent: Social Media and Big Data uses for policing protests in the United Kingdom. *New Media & Society*, *20*(4), 1433-1450.

Deng, S., Wang, C., Fu, Z., & Wang, M. (2021). An intelligent system for insider trading identification in Chinese security market. *Computational Economics*, *57*, 593-616.

Deng, S., Wang, C., Li, J., Yu, H., Tian, H., Zhang, Y., ... & Yang, T. (2019). Identification of insider trading using extreme gradient boosting and multi-objective optimization. *Information*, *10*(12), 367.

Deng, S., Wang, C., Wang, M., & Sun, Z. (2019). A gradient boosting decision tree approach for insider trading identification: An empirical model evaluation of China stock market. *Applied Soft Computing*, *83*, 105652.

DeRosa, M. (2004). *Data Mining And Data Analysis For Counterterrorism*. CSIS Press.

De Rosa, R., & Aragona, B. (2017). Unpacking big data in education. A research framework. *Statistics, Politics and Policy*, *8*(2), 123-137.

Deskus, C. (2018). Fifth Amendment limitations on criminal algorithmic decision-making. *NYUJ Legis. & Pub. Pol'y*, *21*, 237.

De Stefani, F. (2018). *Le regole della privacy: guida pratica al nuovo GDPR*: Hoepli.

Dhingra, S., Dhingra, A. K., & Gupta, S. B. (2022). Smart farming: An IOT based automation. In *Ambient Communications and Computer Systems: Proceedings of RACCCS 2021* (pp. 79-88). Singapore: Springer Nature Singapore.

Diebold, Francis X. (September 21, 2012). *On the Origin(s) and Development of the Term 'Big Data'*. PIER Working Paper No. 12-037.

Dignum, V.(2018). Ethics in Artificial Intelligence: Introduction to the special issue. *Ethics Inf. Technol. 20,* 1-3.

Divya, K. S., Bhargavi, P., & Jyothi, S. (2018). Machine learning algorithms in big data analytics. *Int. J. Comput. Sci. Eng*, *6*(1), 63-70.

Dobbe, R., Dean, S., Gilbert, T., & Kohli, N. (2018). A broader view on bias in automated decision-making: Reflecting on epistemology and dynamics. *arXiv preprint arXiv:1807.00553.*

Dommett, K. (2019). Data-driven political campaigns in practice: Understanding and regulating diverse data-driven campaigns .*Internet Policy Review,8*(4), 1-18

Dommett, K., Barclay, A., & Gibson, R. (2023). Just what is data-driven campaigning? A systematic review. *Information, Communication & Society*, 1-22.

Dong, C., Akram, A., Andersson, D., Arnäs, P. O., & Stefansson, G. (2021). The impact of Emerging and Disruptive Technologies on freight transportation in the digital era: current state and future trends. *The International Journal of Logistics Management*, *32*(2), 386-412.

Dong, J., Wu, H., Zhou, D., Li, K., Zhang, Y., Ji, H., ... & Liu, Z. (2021). Application of big data and artificial intelligence in COVID-19 prevention, diagnosis, treatment and management decisions in China. *Journal of Medical Systems*, *45*(9), 84.

Donohue, L. K. (2005). Anglo-American Privacy and Surveillance. *J. Crim. L. & Criminology*, *96*, 1059.

Dourish, P. (2016). Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, *3*(2).

Downes, L., & Nunes, P. (2017). *Big Bang Disruption: L'era Dell'innovazione Devastante*. EGEA.

Doyle, A., Katz, G., Summers, K., Ackermann, C., Zavorin, I., Lim, Z., ... & Ramakrishnan, N. (2014, October). The EMBERS architecture for streaming predictive analytics. In *2014 IEEE International Conference on Big Data,* (pp. 11-13). IEEE.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management*, *48*, 63-71.

Durrant-Whyte, H. (1988). Sensor models and multisensor integration, I*nternational Journal of Robotics Research*, 7(6), 97-113.

Eckerson, W. W. (2007). Predictive analytics. Extending the Value of Your Data Warehousing Investment. *TDWI Best Practices Report*, *1*, 1-36.

Efthymiou, O. K., & Ponis, S. T. (2021). Industry 4.0 Technologies and their impact in contemporary logistics: a systematic literature review. *Sustainability*, *13*(21), 11643.

Egbert, S., Mann, M. (2021). Discrimination in Predictive Policing: The (Dangerous) Myth of Impartiality and the Need for STS Analysis. In A. Završnik, and V. Badalič (Eds.) *Automating Crime Prevention, Surveillance, and Military Operations* (pp. 25-46). Springer.

Eichler, R., Gröger, C., Hoos, E., Schwarz, H., & Mitschang, B. (2022, July). From data asset to data product–the role of the data provider in the enterprise data marketplace. In *Symposium and Summer School on Service-Oriented Computing* (pp. 119-138). Springer.

Eikermann, R., Look, M., Roth, A., Rumpe, B., & Wortmann, A. (2017). Architecting cloud services for the digital me in a privacy-aware environment. In I, Mistrik, R. Bahsoon, N. Ali, M.

Heisel, and B. Maxim (Eds.) *Software Architecture for Big Data and the Cloud* (pp. 207-226). Morgan Kaufmann.

Ellinger, E. W. (2020). Book Review: Shoshana Zuboff The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. *Organization Studies, 41*(11), 1577-1584.

Elzayn, H., Smith, E., Hertz, T., Ramesh, A., Goldin, J., Ho, D. E., & Fisher, R. (2023). *Measuring And Mitigating Racial Disparities In Tax Audits*. Stanford Institute for Economic Policy Research (SIEPR).

Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M. F. (2020, February). Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies. *NYU School of Law, Public Law Research Paper*, (20-54).

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018, January). Runaway feedback loops in Predictive Policing. In *The 1st Conference on Fairness, Accountability and Transparency*, 81, (pp. 160-171). PMLR.

Ericson, R. V., & Haggerty, K. D. (1997). *Policing the Risk Society*. University of Toronto Press.

Erwin, S. I. (2013). Defense, Intelligence Agencies Struggle to Unify Data Networks. *National Defense*, *97*(712), 30-31.

Espeland, W. N., & Vannebo, B. I. (2007). Accountability, Quantification, and Law. *Annu. Rev. Law Soc. Sci.*, *3*, 21-43.

Evangelista, R. (2019). Review of Zuboff's The Age of Surveillance Capitalism. S*urveillance & Society*, *17*(1/2), 246-251.

Evangelista, R., Guerrieri, P., & Meliciani, V. (2014). The economic impact of digital technologies in Europe. *Economics of Innovation and New Technology*, *23*(8), 802-824.

Fabiano, N. (2020). *GDPR & Privacy: consapevolezza e opportunità. L'approccio con il Data Protection and Privacy Relationships Model (DAPPREMO)*. goWare.

Fagan, F., & Levmore, S. (2019). The impact of artificial intelligence on rules, standards, and judicial discretion. *S. Cal. L. Rev.*, *93*(1), 1-36.

Fairclough, B. (2016). Privacy Piracy: The Shortcomings of the United States' Data Privacy Regime and How to Fix It. *J. Corp. L.*, *42*(1), 461-480.

Fan, C., Yao, W., Mostafavi, A., & Huang, R. (2019, January). A graph-based approach for detecting critical infrastructure disruptions on social media in disasters. In *the 52nd Hawaii International Conference on System Sciences*.

Fantin, S., & Vogiatzoglou, P. (2020). Purpose Limitation By Design As A Counter To Function Creep And System Insecurity In Police Artificial Intelligence. *United Nations Interregional Crime and Justice Research Institute (UNICRI)*.

Farooqui, N. A., & Ritika, A. S. (2019). Sentiment analysis of twitter accounts using natural language processing. *International Journal of Engineering and Advanced Technology*, *8*(3), 473-479.

Ferguson, A. G. (2017). The rise of big data policing. In *The Rise of Big Data Policing*. New York University Press.

Ferguson, A. G. (2015). Big Data and predictive reasonable suspicion. *University of Pennsylvania Law Review*, 327-410.

Ferooz, F., Hassan, M. T., Awan, M. J., Nobanee, H., Kamal, M., Yasin, A., & Zain, A. M. (2021). Suicide bomb attack identification and analytics through data mining techniques. *Electronics*, *10*(19), 2398.

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: impact on firm performance. *Management Decision*, *57*(8), 1923-1936.

Fiebrink, R., & Gillies, M. (2018). Introduction to the special issue on human-centered machine learning. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *8*(2), 1-7.

Fiske, J. (1998). Surveilling the city: Whiteness, the black man and democratic totalitarianism. *Theory, Culture & Society*, *15*(2), 67-88.

Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition*. McGraw-Hill.

Frattini, F. (1997). La disciplina del segreto di Stato. Normativa vigente, prassi applicative e profili di criticità, *Per Aspera ad Veritatem, III*, 9.

Fourcade, M., & Healy, K. (2017). Seeing like a market. *Socio-economic Review*, *15*(1), 9-29.

Fuchs, A., Passarella, A., & Conti, M. (2022). Modeling Human Behavior Part II - Cognitive Approaches and Uncertainty. *arXiv preprint arXiv:2205.06483*.

Fulda, J. S. (2000). Data Mining and Privacy. *Alb. LJ Sci. & Tech.*, *11*, 105.

Galeotti, M. (2018). Discriminazione e algoritmi. Incontri e scontri tra diverse idee di fairness. *The Lab's Quarterly*, *20*(4), 73-96.

Ganor, B. (2021). Artificial or human: A new era of counterterrorism intelligence?. *Studies in Conflict & Terrorism*, *44*(7), 605-624.

Garimella, K., De Francisci Morales, G., Gionis, A., & Mathioudakis, M. (2018, April). Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship. In *Proceedings of the 2018 World Wide Web Conference* (pp. 913-922).

Gee, H. (2020). Last Call for the Third-Party Doctrine in the Digital Age after Carpenter?. *BUJ Sci. & Tech. l.*, *26*, 286.

Gehl, R. W. (2011). The Archive and the Processor: The internal logic of Web 2.0. *New Media & Society*, *13*(8), 1228-1244.

Genovino, C., Caprino, R. M., & Salmista, O. (2020). Internazionalizzazione e innovazione: un sodalizio rinverdito nelle nuove catene globali del valore delle PMI. *Esperienze d'Impresa*, *1*(2), 27-57.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and Management. *Academy of Management Journal*, *57*(2), 321-326.

Gherghina, E. M., Paşa, A. T., & Onofrei, N. (2021). The effects of digitalization on economic growth. *Economic Convergence in European Union*, *131*.

Gilbert, A. S. (2018). Algorithmic culture and the colonization of life-worlds. *Thesis Eleven*, *146*(1), 87-96.

Gillespie, T. (2014). The relevance of algorithms. *Media technologies: Essays on Communication, Materiality, and Society*, 167.

Gillespie, T., & Seaver, N. (2016). Critical algorithm studies: A reading list. *Social Media Collective*, *15*.

Gitelman, L. (Ed.). (2013). *Raw Data is an Oxymoron*. MIT Press.

Giudici, P., Givens, G. H., & Mallick, B. K. (2013). *Wiley Series in Computational Statistics* (Vol. 596). Wiley Online.

Gkougkoudis, G., Pissanidis, D., & Demertzis, K. (2022). Intelligence-led policing and the new technologies adopted by the Hellenic Police. *Digital*, *2*(2), 143-163.

Goel, S., Shroff, R., Skeem, J., & Slobogin, C. (2021). The accuracy, equity, and jurisprudence of criminal risk assessment. In R. Vogl (Ed.) *Research Handbook on Big Data Law*. (pp. 9-28). Elgar.

Goffman, E. (2014). Stigma and social identity. In *Understanding Deviance* (pp. 256-265). Routledge.

Goffman, E. (1963). Embarrassment and Social Organization. In N. J. Smelser & W. T. Smelser (Eds.), *Personality and Social Systems* (pp. 541–548). John Wiley.

Gohar, F., Butt, W. H., & Qamar, U. (2014). Terrorist group prediction using data classification. *Work. Multi Relational Data Min. MRDM2003*, *10*, 199-208.

González, F., Yu, Y., Figueroa, A., López, C., & Aragon, C. (2019, May). Global reactions to the Cambridge Analytica scandal: A cross-language social media study. In *Companion Proceedings of the 2019 World Wide Web Conference* (pp. 799-806).

Goswami, S., Chakraborty, S., Ghosh, S., Chakrabarti, A., & Chakraborty, B. (2018). A review on application of data mining techniques to combat natural disasters. *Ain Shams Engineering Journal*, *9*(3), 365-378.

Govindarajan, S., Mustafa, M. A., Kiyosov, S., Duong, N. D., Raju, M. N., & Gola, K. K. (2023). An optimization based feature extraction and machine learning techniques for named entity identification. *Optik*, *272*, 170348.

Goyal, T., Saini, J. K., & Bansal, D. (2019). Analyzing behavior of ISIS and Al-Qaeda using association rule mining. In *Proceedings of 2nd International Conference on Communication, Computing and Networking: ICCCN 2018, NITTTR Chandigarh, India* (pp. 669-675). Springer.

Graham, S., & Wood, D. (2017). Digitizing Surveillance: Categorization, Space, Inequality. In D. Wilson, and C. Norris (Eds.) *Surveillance, Crime and Social Control* (pp. 537-558). Routledge.

Greaves, R. F., Bernardini, S., Ferrari, M., Fortina, P., Gouget, B., Gruson, D., ... & Kricka, L. J. (2019). Key questions about the future of laboratory medicine in the next decade of the 21st century: a report from the IFCC-Emerging Technologies Division. *Clinica Chimica Acta*, *495*, 570-589.

Greaves, R. F., Kricka, L., Gruson, D., Martin, H., Ferrari, M., & Bernardini, S. (2023). Emerging technology: a definition for laboratory medicine. *Clinical Chemistry and Laboratory Medicine (CCLM)*,*61*(1), 33-36.

Grigorescu, A., Pelinescu, E., Ion, A. E., & Dutcas, M. F. (2021). Human capital in digital economy: An empirical analysis of Central and Eastern European Countries from the European Union. *Sustainability*, *13*(4), 2020.

Gross, S., Possley, M. & Stephens, K. (2017). Race and Wrongful Convictions in the United States. *The National Registry of Exonerations*. Newkirk Center for Science and Society.

Gruber, H. (2017). Innovation, skills and investment: A digital industrial policy for Europe. *Economia e Politica Industriale*, *44*(3), 327-343.

Gundabathula, V. T., & Vaidhehi, V. (2018). An efficient modelling of terrorist groups in India using machine learning algorithms. *Indian J. Sci. Technol*, *11*(15), 1-10.

Gupta, S., Drave, V. A., Dwivedi, Y. K., Baabdullah, A. M., & Ismagilova, E. (2020). Achieving superior organizational performance via big data predictive analytics: A dynamic capability view. *Industrial Marketing Management*, *90*, 581-592.

Gupta, K., & Jiwani, N. (2021). A systematic Overview of Fundamentals and Methods of Business Intelligence. *International Journal of Sustainable Development in Computing Science*, *3*(3), 31-46.

Gupta, D., & Rani, R. (2018). Big data framework for zero-day malware detection. *Cybernetics and Systems*, *49*(2), 103-121.

Haafza, L. A., Awan, M. J., Abid, A., Yasin, A., Nobanee, H., & Farooq, M. S. (2021). Big data covid-19 systematic literature review: Pandemic crisis. *Electronics*, *10*(24), 3125.

Hacking, I. (1990). *The Taming of Chance,* (Vol. 17), Cambridge University Press.

Haggerty, K., & Ericson, V. (2000). The Surveillant Assemblage. *The British Journal of Sociology* *51*(4),605- 622.

Halevi, G., & Moed Dr, H. F. (2012). The evolution of big data as a research and scientific topic: Overview of the literature. *Research Trends*,*1*(30), 2.

Halterman, A., Schrodt, P. A., Beger, A., Bagozzi, B. E., & Scarborough, G. I. (2023). Creating Custom Event Data Without Dictionaries: A Bag-of-Tricks. *arXiv preprint arXiv:2304*.01331.

Hammond-Errey, M. (2022). *Big Data and National Security: A Guide for Australian Policymakers*. Lowy Institute for International Policy.

Hammouri, Q., Atobishi, T., Altememi, A., Al-Zagheer, H., & Khataybeh, H. (2022). The impact of investing in Big Data Analytics (BDA) in enhancing organizational agility and Performance. *Central European Management Journal*, *30*(4), 1090-1093.

Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.

Hardy, K., & Maurushat, A. (2017). Opening up government data for Big Data analysis and public benefit. *Computer Law & Security Review*, *33*(1), 30-37.

Hargreaves, E., Agosti, C., Menasché, D., Neglia, G., Reiffers-Masson, A., & Altman, E. (2019). Fairness in online social network timelines: Measurements, models and mechanism design. *ACM Sigmetrics Performance Evaluation Review*, *46*(3), 68-69.

Hassanien, A. E., Dey, N., & Elghamrawy, S. (Eds.). (2020). *Big Data Analytics and Artificial Intelligence against COVID-19: Innovation Vision and Approach* (Vol. 78). Springer.

Haxhiu, A. (2020). Third-Party Doctrine: The Threat of the Digital Age. *Trento Student Law Review*, *2*(1), 89-107.

Hayashi, T., & Ohsawa, Y. (2020). TEEDA: an interactive platform for matching data providers and users in the data marketplace. *Information*, *11*(4), 218.

Hayden, E. C. (2012). *A Broken Contract*. Nature Publishing Group.

He, S., He, Y., & Li, M. (2019, March). Classification of illegal activities on the dark web. In *Proceedings of the 2nd International Conference on Information Science and Systems* (pp. 73-78).

He, W., Wang, F. K., & Akula, V. (2017). Managing extracted knowledge from big social media data for business decision making. *Journal of Knowledge Management*, *21*(2), 275-294.

Heawood, J. (2018). Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity*, *23*(4), 429-434.

Henderson, S. E. (2017). A few criminal justice big data rules. *Ohio St. J. Crim. L.*, *15*, 527.

Hendricks, J. A., & Kaid, L. L. (Eds.). (2014). *Techno Politics in Presidential Campaigning: New Voices, New Technologies, and New Voters*. Routledge.

Henke, N., Bughin, J., Chui, M., Manyika, J., Saleh, T., Wiseman, B., & Sethupathy, G. (December 2016). *The Age of Analytics: Competing in a Data-driven World*, McKinsey Global Institute.

Hermann, E. (2023). Psychological targeting: nudge or boost to foster mindful and sustainable consumption?. *AI & Society*, *38*(2), 961-962.

Hinds, J., Williams, E. J., & Joinson, A. N. (2020). "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, *143*, 102498.

Hinton, G. (2018). Deep learning—a technology with the potential to transform health care. *Jama*, *320*(11), 1101-1102.

Hoerl, R. W., Snee, R. D., & De Veaux, R. D. (2014). Applying statistical thinking to 'Big Data' problems. *Wiley Interdisciplinary Reviews: Computational Statistics*, *6*(4), 222-232.

Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big data and consumer behavior: Imminent opportunities. *Journal of Consumer Marketing*, *33*(2), 89-97.

Hoffman, S., & Podgurski, A. (2013). Big Bad Data: Law, Public Health, and Biomedical Databases. *Journal of Law, Medicine & Ethics*, *41*(S1), 56-60.

Holt, T.J. (Ed.) (2016). Situating the problem of cybercrime in a multidisciplinary context. In *Cybercrime Through an Interdisciplinary Lens* (pp. 15-28). Routledge.

Hopster, J. (2021, November). What are socially disruptive technologies?. *Technology in Society*, *67*, 101750.

Hossain, K. S. M., Harutyunyan, H., Ning, Y., Kennedy, B., Ramakrishnan, N., & Galstyan, A. (2022). Identifying geopolitical event precursors using attention-based LSTMs. *Frontiers in Artificial Intelligence*, *5*, 893875.

Hosseini, R., Chen, A., Yang, K., Patra, S., Su, Y., Al Orjany, S. E., ... & Ahammad, P. (2022, August). Greykite: Deploying Flexible Forecasting at Scale at LinkedIn. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 3007-3017).

Houser, K., & Sanders, D. (February 2018). The use of big data analytics by the IRS: What tax practitioners need to know. *Journal of Taxation*, *128*(2).

Howlett M, Mukherjee I, Woo JJ. (2015). From Tools to Toolkits in Policy Design Studies: The New Design Orientation towards Policy Formulation Research. *Policy & Politics, 43*(2), 291-311.

Hoyer, C., Gunawan, I., & Reaiche, C. H. (2020). The implementation of industry 4.0–a systematic literature review of the key factors. *Systems Research and Behavioral Science*, *37*(4), 557-578.

Hoyer, W. D., MacInnis, D. J., & Pieters, R. (2016). *Consumer Behavior*. Cengage Learning.

Hua, T., Chen, F., Zhao, L., Lu, C. T., & Ramakrishnan, N. (2013, August). STED: semi-supervised targeted-interest event detection in in Twitter. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1466-1469).

Huamaní, E. L., Alva, M. A., & Roman-Gonzalez, A. (2020). Machine learning techniques to visualize and predict terrorist attacks worldwide using the global terrorism database. *International Journal of Advanced Computer Science and Applications*, *11*(4).

Huang, Y., Dong, H., Yesha, Y., & Zhou, S. (2014, May). A scalable system for community discovery in twitter during hurricane sandy. In *2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing* (pp. 893-899). IEEE.

Hung, J. C., & Chang, J. W. (2021). Multi-level transfer learning for improving the performance of deep neural networks: Theory and practice from the tasks of facial emotion recognition and named entity recognition. *Applied Soft Computing*, *109*, 107491.

Hunter, D. R., Krivitsky, P. N., & Schweinberger, M. (2012). Computational statistical methods for social network models. *Journal of Computational and Graphical Statistics*, *21*(4), 856-882.

Hurley, M., & Adebayo, J. (2016). Credit scoring in the era of Big Data. *Yale JL & Tech.*, *18*, 148.

Imran, M., Liu, X., Wang, R., Saud, S., Zhao, Y., & Khan, M. J. (2022). The Influence of Digital Economy and Society Index on Sustainable Development Indicators: The Case of European Union. *Sustainability*, *14*(18), 11130.

Innerarity, D. (2021). Making the black box society transparent. *AI & Society*, *36*, 975-981.

Innes, M., & Graef, R. (2012). 'The Anvil' in the Information Age: Police, Politics and Media. In T. Newburn, & J. Peay (Eds.), *Policing: Politics, Culture and Control Essays in Honour of Robert Reiner* (pp. 155-172). Hart Publishing.

Ioannoni Fiore, F. (2022). Applicazioni dell'intelligenza artificiale e nuovi strumenti di governo del rischio nei contratti assicurativi. *Rivista di Diritto dell'Impresa*, *1*, 95-118.

Islam, S. R., Ghafoor, S. K., & Eberle, W. (2018, December). Mining illegal insider trading of stocks: A proactive approach. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 1397-1406). IEEE.

Islam, Y.; Zahidul, A. (2015). Data Mining and Privacy of Social Network Sites' Users. Implications of the Data Mining Problem. *Sci. Eng. Ethics, 21,* 941–966.

Jacobi, T., & Stonecipher, D. (2021). A Solution for the third-party doctrine in a time of data sharing, contact tracing, and mass surveillance. *Notre Dame L. Rev.*, *97*, 823.

Jafari, N., Azarian, M., & Yu, H. (2022). Moving from Industry 4.0 to Industry 5.0: what are the implications for smart logistics?. *Logistics*, *6*(2), 26.

Jalilovich, T. M. (2021). The Reveal Of Artificial Intelligence In Information Society. *Conferencea*, 82-83.

Jani, K. P., & Soni, A. (2018). Promise and perils of big data science for intelligence Community. In M.E. Kosal (Ed.) *Technology and the Intelligence Community: Challenges and Advances for the 21st Century*, (pp. 183-203). Springer.

Jayagopal, V., & Basser, K. K. (2022). Data management and big data analytics: Data management in digital economy. In *Research Anthology on Big Data Analytics, Architectures, and Applications*, (pp. 1614-1633). IGI Global.

Javalgi, R. R. G., Martin, C. L., & Young, R. B. (2006). Marketing research, market orientation and customer relationship management: a framework and implications for service providers. *Journal of Services Marketing*, *20*(1), 12-23.

Jemal, I., Cheikhrouhou, O., Hamam, H., & Mahfoudhi, A. (2020). SQL injection attack detection and prevention techniques using machine learning. *International Journal of Applied Engineering Research*, *15*(6), 569-580.

Jia, Q., Guo, Y., Wang, G., & Barnes, S. J. (2020). Big data analytics in the fight against major public health incidents (Including COVID-19): a conceptual framework. *International Journal of Environmental Research and Public Health*, *17*(17), 6161.

Jirovsky, V., Jirovsky Jr., V. (2020). Impact of Disruptive Technologies on the Human Attitude. In Ahram, T., Taiar, R., Colson, S., Choplin, A. (Eds) *Human Interaction and Emerging Technologies*. IHIET 2019. Advances in Intelligent Systems and Computing, vol 1018. Springer, Cham.

Jo, J. H., Sharma, P. K., Sicato, J. C. S., & Park, J. H. (2019). Emerging Technologies for sustainable smart city network security: Issues, challenges, and countermeasures. *Journal of Information Processing Systems*, *15*(4), 765-784.

Joh, E. E. (2017). Feeding the Machine: Policing, Crime Data, & Algorithms. *Wm. & Mary Bill Rts. J.*, *26*, 287.

Johnson, B. R. (2007). A look at fusion centers: Working together to protect America. *FBI L. Enforcement Bull.*, *76*, 28.

Jørgensen, R. F. (2023). Data and rights in the digital welfare state: the case of Denmark. *Information, Communication & Society*, *26*(1), 123-138.

Joshi, A., & Geetha, V. (2014, July). SQL Injection detection using machine learning. In *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)* (pp. 1111-1115). IEEE.

Jovanović, M., Dlačić, J., & Okanović, M. (2018). Digitalization and society's sustainable development–Measures and implications. *Zbornik radova Ekonomskog fakulteta u Rijeci: časopis za ekonomsku teoriju i praksu*, *36*(2), 905-928.

Jung, D., Dorner, V., Weinhardt, C., & Pusmaz, H. (2018). Designing a robo-advisor for risk-averse, low-budget consumers. *Electronic Markets*, *28*, 367-380.

Kade, P. S., & Dhande, N. M. (2017, March-April). Web Data Segmentation for Terrorism Detection using Named Entity Recognition Technique. IJSRST 3(2), 217-222.

Kalantari, F., Mohd Tahir, O., Mahmoudi Lahijani, A., & Kalantari, S. (2017, November). A review of vertical farming technology: A guide for implementation of building integrated agriculture in cities. In *Advanced engineering forum* (Vol. 24, pp. 76-91). Trans Tech Publications Ltd.

Kaluarachchi, T., Reis, A., & Nanayakkara, S. (2021). A review of recent deep learning approaches in human-centered machine learning. *Sensors*, *21*(7), 2514.

Kamtuo, K., & Soomlek, C. (2016, December). Machine Learning for SQL injection prevention on server-side scripting. In *2016 International Computer Science and Engineering Conference (ICSEC)* (pp. 1-6). IEEE.

Kapadia, A. (2020). All that is solid melts into code: The age of surveillance capitalism: The fight for a human future at the new frontier of power, by Shoshan Zuboff, New York, NY, Public Affairs, 2019, e-book. *Economy and Society*, *49*(2), 329-344.

Karuna, P., Rana, M., & Purohit, H. (2017, May). Citizen helper: A streaming analytics system to mine citizen and web data for humanitarian organizations. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 11, No. 1, pp. 729-730).

Katz, R., Koutroumpis, P., & Martin Callorda, F. (2014). Using a digitization index to measure the economic and social impact of digital agendas. *Info*, *16*(1), 32-44.

Kazansky, B. (2021). 'It depends on your threat model': the anticipatory dimensions of resistance to data-driven surveillance. *Big Data & Society*, *8*(1), 205395172098555.

Kazansky, B., & Milan, S. (2021). "Bodies not templates": Contesting dominant algorithmic imaginaries. *New Media & Society*, *23*(2), 363-381.

Kefford, G. (2021). *Political Parties and Campaigning in Australia*. Springer.

Kefford, G., Dommett, K., Baldwin-Philippi, J., Bannerman, S., Dobber, T., Kruschinski, S., ... & Rzepecki, E. (2023). Data-driven campaigning and democratic disruption: Evidence from six advanced democracies. *Party Politics*, *29*(3), 448-462.

Keim, D., Qu, H., & Ma, K. L. (2013). Big-Data visualization. *IEEE Computer Graphics and Applications*, *33*(4), 20-21.

Khan, S., Nazir, S., García-Magariño, I., & Hussain, A. (2021). Deep learning-based urban big data fusion in smart cities: Towards traffic monitoring and flow-preserving fusion. *Computers & Electrical Engineering*, *89*, 106906.

Khanra, S., Dhir, A., & Mäntymäki, M. (2020). Big data analytics and enterprises: a bibliometric synthesis of the literature. *Enterprise Information Systems*, *14*(6), 737-768.

Khatab, Z., & Yousef, G. M. (2021). Disruptive innovations in the clinical laboratory: Catching the wave of precision diagnostics. *Critical Reviews In Clinical Laboratory Sciences*, *58*(8), 546-562.

Khatri, M. (2021). How digital marketing along with artificial intelligence is transforming consumer behaviour. *International Journal for Research in Applied Science and Engineering Technology*, *9*(VII), 523-527.

Khorshid, M., Abou-El-Enien, T. H., & Soliman, G. (2015, May). A comparison among support vector machine and other machine learning classification algorithms. *IPASJ International Journal of Computer Science (IIJCS)*, *3*(5), 25-35.

Khurana, H., Basney, J., Bakht, M., Freemon, M., Welch, V., & Butler, R. (2009, April). Palantir: a framework for collaborative incident response and investigation. In *Proceedings of the 8th Symposium on Identity and Trust on the Internet* (pp. 38-51).

Kiani Mavi, R., Kiani Mavi, N., Olaru, D., Biermann, S., & Chi, S. (2022). Innovations in freight transport: A systematic literature evaluation and COVID implications. *The International Journal of Logistics Management*, *33*(4), 1157-1195.

Kim, I., Pottenger, W. M., & Behe, V. (2018, October). Can a student outperform a teacher? Deep learning-based named entity recognition using automatic labeling of the global terrorism database. In *2018 IEEE International Symposium on Technologies for Homeland Security (HST)* (pp. 1-6). IEEE.

Kim, G. H., Trimi, S., & Chung, J. H. (2014). Big-data applications in the government sector. *Communications of the ACM*, *57*(3), 78-85.

Kireyev, K., Palen, L., & Anderson, K. (2009, December). Applications of topics models to analysis of disaster-related Twitter data. In *NIPS Workshop on Applications for Topic Models: Text and Beyond* (Vol. 1).

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, *20*(1), 14-29.

Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. SAGE.

Kling, R. (1991). Cooperation, coordination and control in computer-supported work. *Communications of the ACM*, *34*(12), 83-88.

Kohler-Hausmann, I. (2013). Misdemeanor justice: Control without conviction. *American Journal of Sociology*, *119*(2), 351-393.

Kocatepe, A., Ulak, M. B., Sriram, L. M. K., Pinzan, D., Ozguven, E. E., & Arghandeh, R. (2018, November). Co-resilience assessment of hurricane-induced power grid and roadway network disruptions: A case study in Florida with a focus on critical facilities. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 2759-2764). IEEE.

Kontokosta, C. E., & Malik, A. (2018). The Resilience to Emergencies and Disasters Index: Applying big data to benchmark and validate neighborhood resilience capacity. *Sustainable Cities & Society*, *36*, 272-285.

Kościelniak, H., & Puto, A. (2015). Big Data in decision making processes of enterprises. *Procedia Computer Science*, *65*, 1052-1058.

Koukaras, P., & Tjortjis, C. (2019). Social media analytics, types and methodology. *Machine Learning Paradigms: Applications of Learning and Analytics in Intelligent Systems*, 401-427.

Kreiss, D. (2017). Micro-targeting, the quantified persuasion. *Internet Policy Review,6*(4), 1-14.

Kreiss, D. (2016). *Prototype Politics: Technology-Intensive Campaigning and the Data of Democracy*. Oxford University Press.

Kreiss, D., & Jasinski, C. (2016). The tech industry meets presidential politics: Explaining the Democratic Party's technological advantage in electoral campaigning. *Political Communication, 33*(4), 544-562.

Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 US presidential cycle. *Political Communication, 35*(2), 155-177

Kreiss, D., & Saffer, A. J. (2017). Networks and innovation in the production of communication: Explaining innovations in US electoral campaigning from 2004 to 2012. *Journal of Communication*, *67*(4), 521-544.

Kris, D. S. (2006). The Rise and Fall of the FISA Wall. *Stan. L. & Pol'y Rev.*, *17*, 487.

Krishnan, A., & Swarna, S. (2020, October). Robotics, IoT, and AI in the automation of agricultural industry: a review. In *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)* (pp. 1-6). IEEE.

Krishnamurthy, R., & Desouza, K. C. (2014). Big data analytics: The case of the social security administration. *Information Polity*, *19*(3-4), 165-178.

Kruschinski, S., & Haller, A. (2017). Restrictions on data-driven political micro-targeting in Germany. *Internet Policy Review*, *6*(4), 1-23.

Kumar, S. (2020). Modern Advertising Strategies and Consumer Buying Behavior. In B.B. Tiwari, and B.W. Lyall (Eds.) *E-Business: Issues and Challenges of 21st Century*, (pp. 30-41). Allied Publishers.

Kumar, V., & Garg, M. L. (2018). Predictive analytics: a review of trends and techniques. *International Journal of Computer Applications*, *182*(1), 31-37.

Kumar, R., Jain, V., Yie, L. W., & Teyarachakul, S. (Eds.). (2023). *Convergence of IoT, Blockchain, and Computational Intelligence in Smart Cities*. CRC Press.

Kumar, S., Koolwal, V., & Mohbey, K. K. (2019). Sentiment analysis of electronic product tweets using big data framework. *Jordanian Journal of Computers and Information Technology*, *5*(1).

Kumar, V., Mazzara, M., Messina, A., & Lee, J. (2020). A conjoint application of data mining techniques for analysis of global terrorist attacks: prevention and prediction for combating terrorism. In *Proceedings of 6th International Conference in Software Engineering for Defence Applications: SEDA 2018 6* (pp. 146-158). Springer.

Kumar, V. N., & Shindgikar, P. (2018). *Modern Big Data Processing With Hadoop: Expert Techniques For Architecting End-To-End Big Data Solutions To Get Valuable Insights*. Packt Publishing Ltd.

Labib, N. M., Rizka, M. A., & Shokry, A. E. M. (2020). Survey of machine learning approaches of anti-money laundering techniques to counter terrorism finance. In *Internet of Things—Applications and Future: Proceedings of ITAF 2019* (pp. 73-87). Springer.

Lachapelle, G., & Maarek, P. (Eds.). (2015). *Political Parties In The Digital Age: The Impact Of New Technologies In Politics*. Walter de Gruyter GmbH & Co KG.

Lagorio, A., Zenezini, G., Mangano, G., & Pinto, R. (2022). A systematic literature review of innovative technologies adopted in logistics management. *International Journal of Logistics Research and Applications*, *25*(7), 1043-1066.

Lakshmi, B. N., & Raghunandhan, G. H. (2011, February). A conceptual overview of data mining. In *2011 National Conference on Innovations in Emerging Technology* (pp. 27-32). IEEE.

Lamardini, M. (2016). La vigilanza diretta dell'Esma. Un modello per il futuro?. *Giur. Comm. 4*, 450.

Lamdan, S. (2019). When Westlaw fuels ICE surveillance: Legal ethics in the era of big data policing. *NYU Rev. L. & Soc. Change*, *43*(2). 255-293.

Lampo, A., Mancarella, M., & Piga, A. (2020). (Non)-neutrality of science and algorithms: Machine Learning between fundamental physics and society. *arXiv preprint arXiv:2006.10745*.

Lander, S.F. (2004). International intelligence cooperation: an inside perspective. *Cambridge Review of International Affairs, 17*, 481-493.

Landon-Murray, M. (2016). Big data and intelligence: Applications, human capital, and education. *Journal of Strategic Security*, *9*(2), 92-121.

Larose, D. T. (2015). *Data Mining and Predictive Analytics*. John Wiley & Sons.

Laskov, P. and Lippmann, R. (2010) Machine Learning in Adversarial Environments. *Machine Learning Journal, 81*, 115-119.

Laub, J. H. (2014). *Understanding inequality and the justice system response: Charting a new way forward*. William T. Grant Foundation.

Lavorgna, A., & Ugwudike, P. (2021). The datafication revolution in criminal justice: An empirical exploration of frames portraying data-driven technologies for crime prevention and control. *Big Data & Society*, *8*(2), 1-15.

Le, T. M., & Liaw, S. Y. (2017). Effects of pros and cons of applying big data analytics to consumers' responses in an e-commerce context. *Sustainability*, *9*(5), 798.

Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, *60*(3), 293-303.

L'Heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *IEEE Access*, *5*, 7776-7797.

Li, J., Jia, M., Liu, Q., Fu, Y., & Chen, Q. (2022, December). Getting insights from news data mining: a case of BBC's counter-terrorism news text analysis. In *Second International Symposium on Computer Technology and Information Science (ISCTIS 2022)* (Vol. 12474, pp. 118-122). SPIE.

Li, D., Kong, Y., Zheng, Z., & Pan, J. (2022). Recent Advances in Big Data Analytics. *The Palgrave Handbook of Operations Research*, 805-834.

Li, J., Sun, A., Han, J., & Li, C. (2020). A survey on deep learning for named entity recognition. *IEEE Transactions on Knowledge and Data Engineering*, *34*(1), 50-70.

Li, T., Xie, N., Zeng, C., Zhou, W., Zheng, L., Jiang, Y., ... & Iyengar, S. S. (2017). Data-driven techniques in disaster information management. *ACM Computing Surveys (CSUR)*, *50*(1), 1-45.

Liang, R., Jiao, Z., & Liu, Z. (2020). An Empirical Study on the Influencing Factors of Customers' Acceptance Intention towards Online Behavioral Advertising. *Tehnički Vjesnik*, *27*(4), 1142-1149.

Liang, T. P., & Liu, Y. H. (2018). Research landscape of business intelligence and big data analytics: A bibliometrics study. *Expert Systems with Applications*, *111*, 2-10.

Liggett, R., Lee, J.R., Roddy, A.L., & Wallin, M.A. (2020). The Dark Web as a Platform for Crime: An Exploration of Illicit Drug, Firearm, CSAM, and Cybercrime Markets. In T. J. Holt, and A. M. Bossler (Eds.) *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (pp. 91-116). Palgrave Macmillan.

Linardos, V., Drakaki, M., Tzionas, P., & Karnavas, Y. L. (2022). Machine Learning in Disaster Management: recent developments in methods and applications. *Machine Learning and Knowledge Extraction*, *4*(2), 446-473.

Liu, J., Gao, L., Guo, S., Ding, R., Huang, X., Ye, L., ... & Thiruvady, D. (2021). A hybrid deep-learning approach for complex biochemical named entity recognition. *Knowledge-Based Systems*, *221*, 106958.

Liu, Y., Kohlberger, T., Norouzi, M., Dahl, G. E., Smith, J. L., Mohtashamian, A., ... & Stumpe, M. C. (2019). Artificial intelligence–based breast cancer nodal metastasis detection: Insights into the black box for pathologists. *Archives of Pathology & Laboratory Medicine*, *143*(7), 859-868.

Liu, X., Sun, R., Wang, S., & Wu, Y. J. (2020). The research landscape of big data: a bibliometric analysis. *Library Hi Tech*, *38*(2), 367-384.

Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *The Journal of Strategic Information Systems*, *24*(3), 149-157.

Löfgren, K., & Webster, C. W. R. (2020). The value of Big Data in government: The case of 'smart cities'. *Big Data & Society*, *7*(1).

Lowe, J., & Matthee, M. (2020). Requirements of data visualisation tools to analyse big data: A structured literature review. In *Responsible Design, Implementation and Use of Information and Communication Technology: 19th IFIP WG 6.11 Conference on E-business, E-services, and E-society, I3E 2020, Skukuza, South Africa, April 6–8, 2020, Proceedings, Part I 19* (pp. 469-480). Springer.

Lucchini Guastalla, E. (2019). Privacy e data Protection: principi generali. In E. Tosi (Ed.) *Privacy Digitale. Riservatezza e protezione dei dati personali tra GDPR e nuovo Codice Privacy* (pp. 55-92). Giuffré.

Luhn, H. P. (1958). A Business Intelligence System. *IBM Journal of Research and Development*, *2*(4), 314-319.

Lupton, D. (2014). The commodification of patient opinion: The digital patient experience economy in the age of big data. *Sociology of Health & Illness 36*(6). 856-869.

Lyon, D. (2003). *Surveillance after September 11, 11*. Polity.

Lyon, D. (1994). *The Electronic Eye: The Rise of Surveillance Society*. University of Minnesota Press.

Lyon, D., & Wood, D. M. (Eds.). (2020). *Big Data Surveillance and Security Intelligence: The Canadian Case*. UBC Press.

MacIntyre, A. (2007). *After Virtue: A Study in Moral Theory* (3rd ed.). Gerald Duckworth & Co Ltd.

Madison, M. C., Cowell, A. J., Butner, R. S., Fligg, K., Piatt, A. W., McGrath, L. R., & Ellis, P. C. (2012). Knowledge encapsulation framework for Technosocial Predictive Modeling. *Security Informatics*, *1*(1), 1-18.

Mageto, J. (2021). Big data analytics in sustainable supply chain management: A focus on manufacturing supply chains. *Sustainability*, *13*(13), 7101.

Maggiolino, M. (2016). Big data e prezzi personalizzati. *Concorrenza e Mercato*, *23*, 95-138.

Malone, E. L., Izaurralde, R. C., Thomson, A. M., & Morgan, L. G. (2009). Resilience, Climate Change, and Security: Modeling the Connections. In *AAAI Spring Symposium: Technosocial Predictive Analytics* (pp. 76-81).

Mandelli, A. (2017). *Big Data Marketing: Creare valore nella Platform Economy con Dati, Intelligenza Artificiale e IoT*. EGEA.

Manning, P. (1992). Information Technologies and the Police. *Crime Justice, 15*, 349-398.

Mantelero, A. (2015). Rilevanza e tutela della dimensione collettiva della protezione dei dati personali. *Contr. Impr. Europa, 1,* 141

Markopoulou, D., Papakonstantinou, V., & De Hert, P. (2019). The new EU cybersecurity framework: The NIS Directive, ENISA's role and the General Data Protection Regulation. *Computer Law & Security Review*, *35*(6), 105336.

Markowetz, A., K. Błaszkiewicz, C. Montag, C. Switala, and T.E. Schlaepfer. 2014. Psychoinformatics: Big data shaping modern Psychometrics. *Medical Hypotheses, 82*(4). 405–411.

Maroufkhani, P., Wagner, R., Wan Ismail, W. K., Baroto, M. B., & Nourani, M. (2019). Big data analytics and firm performance: A systematic review. *Information*, *10*(7), 226.

Mashiat, T., Gitiaux, X., Rangwala, H., & Das, S. (2023, May). Counterfactually Fair Dynamic Assignment: A Case Study on Policing. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems* (pp. 2526-2528). AAMAS.

Martin Jr, D., Prabhakaran, V., Kuhlberg, J., Smart, A., & Isaac, W.S. (2020). Participatory problem formulation for fairer machine learning through community based system dynamics. *arXiv preprint arXiv:2005.07572*.

Martis, R. J., Gurupur, V. P., Lin, H., Islam, A., & Fernandes, S. L. (2018). Recent advances in big data analytics, internet of things and machine learning. *Future Generation Computer Systems*, *88*, 696-698.

Marx, G. T. (2016). Surveillance and Surveys: the Soft Interview of the Future. *Society*, *53*(3), 301-306.

Marx, G. T. (2002). What's New About the" New Surveillance"? Classifying for Change and Continuity. *Surveillance & Society*, *1*(1), 9-29.

Marx, G. T. (1988). *Undercover: Police Surveillance in America*. University of California Press.

Mason, C. (2012). New police surveillance technologies and the good-faith exception: Warrantless GPS tracker evidence after United States v. Jones. *Nev. LJ*, *13*, 60.

Mastrelia, D. (2018). Gestione dei Big Data in una prospettiva orientata alla tutela della privacy degli individui. *Il Diritto Industriale*, *4*, 364-372.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, *114*(48), 12714-12719.

Matz, S. C., & Netzer, O. (2017). Using big data as a window into consumers' psychology. *Current Opinion in Behavioral Sciences*, *18*, 7-12.

Matzner, T. (2016). Beyond data as representation: The performativity of Big Data in surveillance. *Surveillance & Society*, *14*(2), 197-210.

Maxwell, T. A. (2005, January). Information policy, data mining, and national security: False positives and unidentified negatives. In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences* (pp. 134c-134c). IEEE.

Mazarr, M. J., Bauer, R. M., Casey, A., Heintz, S., & Matthews, L. J. (2019). *The Emerging Risk Of Virtual Societal Warfare. Social Manipulation in a Changing Information Environment.* RAND Corporation.

Mazzarisi, P., Ravagnani, A., Deriu, P., Lillo, F., Medda, F., & Russo, A. (2022). A machine learning approach to support decision in insider trading detection. *arXiv preprint arXiv:2212.05912*

Mazzei, M. J., & Noble, D. (2017). Big data dreams: A framework for corporate strategy. *Business Horizons*, *60*(3), 405-414.

Mazzotti, M. (2015). Per una sociologia degli algoritmi. *Rassegna Italiana di Sociologia, 56(3-4)*, 465-478.

McCarthy, M. T. (2002). USA Patriot Act. Harv. *Journal on Legis*. 39, 435.

McCue, C. (2006). Data mining and predictive analytics in public safety and security. *IT Professional Magazine*, *8*(4), 12.

McDermott, A. F., Veinott, E., Eusebi, L., Whitaker, E. T., Trewhitt, E. B., Mueller, S., ... & Guarino, S. (2021, July). Developing an adaptive framework to support intelligence analysis. In *International Conference on Human-Computer Interaction* (pp. 550-558) Springer International.

McGruddy, J. (2013). Multilateral Intelligence Collaboration and international oversight. *Journal of Strategic Security*, *6*(3), 214-220.

Mehta, N., & Shukla, S. (2022). Pandemic analytics: how countries are leveraging big data analytics and artificial intelligence to fight COVID-19?. *SN Computer Science*, *3*(1), 54.

Meissner, D., Gokhberg, L., & Saritas, O. (2019). *What Do Emerging Technologies Mean for Economic Development?* (pp. 1-10). Springer International Publishing.

Mell, A. (2012). *Reputation in the Market for Stolen Data*. University of Oxford.

Melnyk, L., Dehtyarova, I., Kubatko, O., Karintseva, O., & Derykolenko, A. (2019). Disruptive technologies for the transition of digital economies towards sustainability. *Економічний часопис (Economic Journal)-XXI*, (9-10), 22-30.

Meng, X., Nie, L., & Song, J. (2019). Big data-based prediction of terrorist attacks. *Computers & Electrical Engineering*, *77*, 120-127.

Merton, R. K. (1968). *Social Theory and Social Structure*. Simon and Schuster.

Messingschlager, T., & Holtz, P. (2020). Filter bubbles und echo chambers. Die Psychologie des Postfaktischen: Über Fake News. Lügenpresse. *Clickbait & Co.*, 91-102.

Mikalef, P., Boura, M., Lekakos, G., & Krogstie, J. (2019). Big data analytics and firm performance: Findings from a mixed-method approach. *Journal of Business Research*, *98*, 261-276.

Mikhailov, D. I. (2023). Optimizing National Security Strategies through LLM-Driven Artificial Intelligence Integration. *arXiv preprint arXiv:2305.13927*.

Miller, K. (2014). Total Surveillance, Big Data, and Predictive Crime Technology: Privacy's perfect storm. *J. Tech. L. & Pol'y*, *19*, 105.

Mingo, I., & Bracciale, R. (2016). Social Inequalities in Digital Skills. The European context and the Italian case. In *The Praxis of Social Inequalities, a global perspective* (Vol. 1, pp. 81-111). Lexington Books.

Mishra, D., Luo, Z., Hazen, B., Hassini, E., & Foropon, C. (2019). Organizational capabilities that enable big data and predictive analytics diffusion and organizational performance: A resource-based perspective. *Management Decision*, *57*(8), 1734-1755.

Mittelstadt, B.D., Floridi, L. (2016). The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. In B. Mittelstadt, and L. Floridi. (Eds.) *The Ethics of Biomedical Big Data* (pp. 445-480). Springer.

Mohammed, D. Y., & Karabatak, M. (2018, March). Terrorist attacks in Turkey: An evaluate of terrorist acts that occurred in 2016. In *2018 6th International Symposium on Digital Forensic and Security (ISDFS)* (pp. 1-3). IEEE.

Mollá, D., Van Zaanen, M., & Cassidy, S. (2007, December). Named entity recognition in question answering of speech data. In *Proceedings of the Australasian Language Technology Workshop 2007* (pp. 57-65).

Molnar, P. (2022). Territorial and digital borders and migrant vulnerability under a pandemic crisis. In A. Triandafyllidou, *Migration and Pandemics: Spaces of Solidarity and Spaces of Exception* (pp. 45-64), IMISCOE Research Series.

Molnar, P. (2021a). Robots and refugees: the human rights impacts of artificial intelligence and automated decision-making in migration. In M. McAuliffe (Ed.) *Research Handbook on International Migration and Digital Technology* (pp. 134-151). Elgar.

Molnar, P. (2021b). Surveillance sovereignty: Migration management technologies and the politics of privatization. *Migration, security, and resistance: global and local perspectives.* Routledge.

Molnar, P. (2019a). Technology on the margins: AI and global migration management from a human rights perspective. *Cambridge International Law Journal*, *8*(2), 305-330.

Molnar, P. (2019b). New technologies in migration: human rights impacts. *Forced Migration Review*, *61*, 7-9.

Monahan, T. (2009). The murky world of 'Fusion Centres' Torin Monahan critiques the emergence of data-sharing 'Fusion Centres' intended to reduce crime and prevent terrorism. *Criminal Justice Matters*, *75*(1), 20-21.

Monahan, T., & Palmer, N. A. (2009). The emerging politics of DHS fusion centers. *Security Dialogue*, *40*(6), 617-636.

Moore, P., Xhafa, F., Barolli, L., & Thomas, A. (2013, October). Monitoring and detection of agitation in dementia: Towards real-time and big-data solutions. In *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing* (pp. 128-135). IEEE.

Morelli, G., Musso, F., Murmura, F., & Bravi, L. (2022). Knowledge Analysis on the Industry 4.0 Diffusion in Italian Manufacturing: Opportunities and Threats. In *Digital Transformation in Industry: Digital Twins and New Business Models* (pp. 195-214). Cham: Springer International Publishing.

Morstatter, F., Kumar, S., Liu, H., & Maciejewski, R. (2013, August). Understanding twitter data with Tweetxplorer. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1482-1485).

Moşescu, A. I., Chivu, R. G., Căescu, Ş. C., Ionuţ-Claudiu, P., & Botezatu, F. (2020). Using big data in marketing and advertising: a case study. *Journal of Emerging Trends in Marketing and Management*, *1*(1), 259-264.

Mosqueira-Rey, E., Hernández-Pereira, E., Alonso-Ríos, D., Bobes-Bascarán, J., & Fernández-Leal, Á. (2023). Human-in-the-loop machine learning: A state of the art. *Artificial Intelligence Review*, *56*(4), 3005-3054.

Mudigonda, S., Ozbay, K., & Bartin, B. (2019). Evaluating the resilience and recovery of public transit system using big data: Case study from New Jersey. *Journal of Transportation Safety & Security*, *11*(5), 491-519.

Mugavero, R., & Thorossian, W. (2021). Intelligenza artificiale e machine learning: nuovi strumenti per il contrasto della conflittualità asimmetrica e per la gestione delle crisi-il caso di studio pandemia covid-19. *Rivista di Criminologia, Vittimologia e Sicurezza*, *15*(1-3), 66-76.

Munir, A., Kwon, J., Lee, J. H., Kong, J., Blasch, E., Aved, A. J., & Muhammad, K. (2021). FogSurv: A fog-assisted architecture for urban surveillance using artificial intelligence and data fusion. *IEEE Access*, *9*, 111938-111959.

Monarch, R., & Munro, R. (2021). *Human-in-the-Loop Machine Learning: Active Learning and Annotation for Human-centered AI*. Simon & Schuster.

Muthiah, S., Butler, P., Khandpur, R. P., Saraf, P., Self, N., Rozovskaya, A., ... & Ramakrishnan, N. (2016, August). EMBERS at 4 years: Experiences operating an open source indicators forecasting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery And Data Mining* (pp. 205-214).

Muthiah, S., Huang, B., Arredondo, J., Mares, D., Getoor, L., Katz, G., & Ramakrishnan, N. (2015, January). Planned protest modeling in news and social media. In *Proceedings of the AAAI Conference on Artificial Intelligence 29*(2), (pp. 3920-3927).

Myne, A. K., Leahy, K. J., & Soklaski, R. J. (2022). *Knowledge-Integrated Informed AI for National Security*. Massachusetts Institute of Technology.

Nagy, S., & Somosi, M. V. (2022). The relationship between social innovation and digital economy and society. *Regional Statistics, 12*(2), 3-29.

Nair, L. R., Shetty, S. D., & Shetty, S. D. (2017). Streaming big data analysis for real-time sentiment based targeted advertising. *International Journal of Electrical and Computer Engineering*, *7*(1), 402.

Nakamura, L. (2013). *Cybertypes: Race, Ethnicity, and Identity on the Internet*. Routledge.

Nam, S. H. (2015). Why disruptive innovations matter in laboratory diagnostics. *Clinical Chemistry*, *61*(7), 935-937.

Nasser, T., & Tariq, R. S. (2015). Big Data challenges. *J Comput Eng Inf Technol (4)*3, 1000135.

National Institute on Drug Abuse (2007). *Drug, Brains and Behavior: The Science of Addiction.* National Institutes of Health Publication No. 07-5605.

National Institutes of Health. (2008). *Innovative Computational and Statistical Methodologies for the Design and Analysis of Multilevel Studies on Childhood Obesity (R01). Request for Applications (RFA) Number: RFA-HD-08-023*. Department of Health and Human Services.

Naylor, C. D. (2018). On the prospects for a (deep) learning health care system. *Jama*, *320*(11), 1099-1100.

Neppalli, V. K., Caragea, C., Squicciarini, A., Tapia, A., & Stehle, S. (2017). Sentiment analysis during Hurricane Sandy in emergency response. *International Journal of Disaster Risk Reduction*, *21*, 213-222.

Ng, A. K., & Mahkeswaran, R. (2021, August). Emerging and Disruptive Technologies for Urban Farming: A review and assessment. In *Journal of Physics: Conference Series* (Vol. 2003, No. 1, p. 012008). IOP Publishing.

Ngiam, K. Y., & Khor, W. (2019). Big data and machine learning algorithms for health-care delivery. *The Lancet Oncology*, *20*(5), e262-e273.

Nicita, A. & Delmasto, M. (2019). *Big Data: come trasformano l'economia e la politica*. Il Mulino.

Nickerson, D. W., & Rogers, T. (2014). Political Campaigns and Big Data. *Journal of Economic Perspectives*, *28*(2), 51-74.

Niranga, M., Sedera, D., & Sorwar, G. (2022). Does IT Matter (Now)? A Global Panel Data Analysis of 7 Regions from 2018-2020 on Digitalization and Its Impact on Economic Growth. *arXiv preprint arXiv:2212.03071*.

Nosratabadi, S., Atobishi, T., & Hegedűs, S. (2023). Social sustainability of digital transformation: Empirical evidence from EU-27 countries. *Administrative Sciences*, *13*(5), 126.

Novikov, S. V. (2020). Data science and big data technologies role in the digital economy. *TEM Journal, 9*(2), 756-762.

Nti, I. K., Quarcoo, J. A., Aning, J., & Fosu, G. K. (2022). A mini review of machine learning in big data analytics: Applications, challenges, and prospects. *Big Data Mining and Analytics*, *5*(2), 81-97.

Nunn, S. (2001, January). Police technology in cities: changes and challenges. *Technology in Society*, *23*(1), 11-27.

Odabas, M., Holt, T. J., & Breiger, R. L. (2017). Governance in online stolen data markets. In J. Beckert, and M. Dewey (Eds.) *The Architecture of Illegal Markets: Towards an Economic Sociology of Illegality in The Economy* (pp. 87-107). Oxford University Press.

Odeniyi, O. A., Adeosun, M. E., & Ogundunmade, T. P. (2022). Prediction of terrorist activities in Nigeria using machine learning models. *Innovations*, *71*, 87-96.

Okidegbe, N. (2019). When they hear us: Race, algorithms and the practice of criminal law. *Kan. JL & Pub. Pol'y*, *29*, 329.

Olabanjo, O. A., Aribisala, B. S., Mazzara, M., & Wusu, A. S. (2021). An ensemble machine learning model for the prediction of danger zones: Towards a global counter-terrorism. *Soft Computing Letters*, *3*, 100020.

Oladejo, B. F., & Onyemenam, N. I. (2019, June). Development of an extended natural language processing (nlp)-based framework for knowledge discovery in terrorism-based communication. *UIJSLICTR, 3*(1), 60-71.

O'Donovan, P., Leahy, K., Bruton, K., & O'Sullivan, D. T. (2015). Big data in manufacturing: a systematic mapping study. *Journal of Big Data*, *2*, 1-22.

Oehler, A., Horn, M., & Wendt, S. (2022). Investor characteristics and their impact on the decision to use a robo-advisor. *Journal of Financial Services Research*, *62*(1-2), 91-125.

Ofori, K. S., Boakye, K., & Narteh, B. (2018). Factors influencing consumer loyalty towards 3G mobile data service providers: evidence from Ghana. *Total Quality Management & Business Excellence*, *29*(5-6), 580-598.

O'Hara, K., & Stevens, D. (2015). Echo chambers and online radicalism: Assessing the Internet's complicity in violent extremism. *Policy & Internet*, *7*(4), 401-422.

Okidegbe, N. (2019). When they hear us: Race, algorithms and the practice of criminal law. *Kan. JL & Pub. Pol'y*, *29*, 329.

Omand, D., & Phythian, M. (2018). *Principled Spying: The Ethics of Secret Intelligence*. Oxford University Press.

Onay, C., & Öztürk, E. (2018). A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance*, *26*(3), 382-405.

O'Neill, K. (2020). *Big Data. Weapons Of Mathematical Destruction. How Big Data Increases Inequality and Threatens Democracy* / trans. z anhl. O. Kalinina. K. Force Ukraine. Kyiv, Ukraine.

Ongsulee, P., Chotchaung, V., Bamrungsi, E., & Rodcheewit, T. (2018, November). Big data, predictive analytics and machine learning. In *2018 16th International Conference on ICT and Knowledge Engineering (ICT&KE)* (pp. 1-6). IEEE.

Orefice, M. (2016). I big data. Regole e concorrenza. *Politica del Diritto*, *47*(4), 697-744.

Ormerod, P. C., & Trautman, L. J. (2017). A Descriptive Analysis of the Fourth Amendment and the Third-Party Doctrine in the Digital Age. *Alb. LJ Sci. & Tech.*, *28*, 73.

OSCE Organizzazione per la Sicurezza e la Cooperazione in Europa. (2021). *Project Report: Intelligence-Led Policing (ILP) 2017–2020.*

Osoba, O., & Welser IV, W. (2017). *An Intelligence in Our Image*. RAND Corporation.

Oswald, M. (2018). Algorithm-assisted decision-making in the public sector: framing the issues using administrative law rules governing discretionary power. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *376*(2128), 20170359.

Pager, D. (2007). The use of field experiments for studies of employment discrimination: Contributions, critiques, and directions for the future. *The Annals of the American Academy of Political and Social Science*, *609*(1), 104-133.

Pan, X. (2021). Quantitative Analysis and Prediction of Global Terrorist Attacks Based on Machine Learning. *Scientific Programming*, *2021*, 1-15.

Panimalar, A., Shree, V., Kathrine, V. (2017). The 17 V's of big data. *Int. Res. J. Eng. Technol.*, *4*(9), 329-333.

Papachristos, A. V., Hureau, D. M., & Braga, A. A. (2013). The corner and the crew: The influence of geography and social networks on gang violence. *American Sociological Review*, *78*(3), 417-447.

Parenzo, B. (2021) Sull'importanza del dire le cose come stanno: ovvero, sul perché della necessità di riconoscere la natura patrimoniale dei dati personali e l'esistenza di uno scambio sotteso ai c.d. servizi digitali "gratuiti". *Diritto di Famiglia e delle Persone, (Il)3,* 1457.

Pariser, E. (2012). *O filtro invisível: o que a internet está escondendo de você*. Editora Schwarcz-Companhia das Letras.

Park, Y. E. (2022). Developing a COVID-19 crisis management strategy using news media and social media in big data analytics. *Social Science Computer Review*, *40*(6), 1358-1375.

Parker, A. K. (2011). Dragnet Law Enforcement: Prolonged Surveillance & the Fourth Amendment. *W. St. UL Rev.*, *39*, 23.

Paßmann, J., & Boersma, A. (2017). Unknowing Algorithms. In M.T. Schäfer, and K. van Es (Eds.) *The Datafied Society* (pp. 139-267). Amsterdam University Press.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money And Information*. Harvard University Press.

Patil, S. & Lokesha, V. (2022, May 25). Live Twitter Sentiment Analysis Using Streamlit Framework. *Proceedings of the International Conference on Innovative Computing & Communication (ICICC).*

Pavletic, J. (2018). The Fourth Amendment in the age of persistent aerial surveillance. *The Journal of Criminal Law and Criminology, 108*(1), 171-196.

Payne, J., Solomon, J., Sankar, R., & McGrew, B. (2008, October). Grand challenge award: Interactive visual analytics Palantir: The future of analysis. In *2008 IEEE Symposium on Visual Analytics Science and Technology* (pp. 201-202). IEEE.

Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., & Turini, F. (2019, July). Meaningful explanations of black box AI decision systems. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, No. 01, pp. 9780-9784).

Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Pappalardo, L., Ruggieri, S., & Turini, F. (2018). Open the black box data-driven explanation of black box decision systems. *arXiv preprint arXiv:1806.09936.*

Pelillo, M. & Scantamburlo, T. (Eds.) (2021). *Machines We Trust: Perspective on Dependable AI.* MIT Press.

Pellegrino, E.D., & Thomasma, D.C. (1993). *The Virtues in Medical Practice*. Oxford University Press.

Pencheva, I., Esteve, M., & Mikhaylov, S. J. (2020). Big Data and AI–A transformational shift for government: So, what next for research?. *Public Policy and Administration*, *35*(1), 24-44.

Perna, A., & Runfola, A. (2017). Relazioni business to business e cambiamenti tecnologici: una prospettiva di marketing industriale. *Relazioni Business to Business e Cambiamenti Tecnologici*, 1-108.

Perrucci, A. (2019). Dai Big Data all'ecosistema digitale. Dinamiche tecnologiche e di mercato e ruolo delle politiche pubbliche. *Analisi Giuridica dell'Economia*, *18*(1), 61-88.

Perry, W. L. (2013). *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Rand Corporation.

Peterson, D. (2022). AI and the surveillance state. In W. C. Hannas & H. Cheng (Eds.), *Chinese Power and Artificial Intelligence: Perspectives and Challenges* (pp. 189–215). Routledge.

Peterson, E. D. (2019). Machine learning, predictive analytics, and clinical practice: can the past inform the present?. *Jama*, *322*(23), 2283-2284.

Pettit, C. J., Zarpelon Leao, S., Lock, O., Ng, M., & Reades, J. (2022). Big data: the engine to future cities - a reflective case study in urban transport. *Sustainability*, *14*(3), 1727.

Pham, Q. V., Nguyen, D. C., Huynh-The, T., Hwang, W. J., & Pathirana, P. N. (2020). Artificial intelligence (AI) and big data for coronavirus (COVID-19) pandemic: a survey on the state-of-the-arts. *IEEE Access*, *8*, 130820-130839.

Picciano, A. (2012). The evolution of big data and learning analytics in american higher education. *Journal of Asynchronous Learning Networks, 16(3)*, 9-20.

Picozzi, M., & Zappalà, A. (2002). *Criminal Profiling. Dall'analisi della scena del delitto al profilo psicologico del criminale,* McGraw-Hill.

Piegorsch, W. W., Levine, R. A., Zhang, H. H., & Lee, T. C. (Eds.). (2022). *Computational Statistics in Data Science*. John Wiley & Sons.

Pitruzzella, G. (2016). Big data, competition and privacy: a look from the Antitrust perspective. *Concorrenza e Mercato*, *23*, 15-28.

Plesničar, M. M., Završnik, A., & Šarf, P. (2020). Fighting impunity with new tools: how big data, algorithms, machine learning and AI shape the new era of criminal justice. *The Fight Against Impunity in EU Law*. Hart Publishing.

Podesta, J., Pritzker, P., Moniz, E. J., Holdren, J., & Zients, J. (2014). Big data: Seizing opportunities. *Executive Office of the President of USA*, 1-79.

Popkova, E.G. (2022). Vertical Farms Based on Hydroponics, Deep Learning, and AI as Smart Innovation in Agriculture. In: Popkova, E.G., Sergi, B.S. (Eds) *Smart Innovation in Agriculture*. Smart Innovation, Systems and Technologies, vol. 264. Springer, Singapore.

Porcedda, M. G., & Wall, D. S. (2019, June). Cascade and chain effects in big data cybercrime: Lessons from the TalkTalk hack. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (pp. 443-452). IEEE.

Porrini, D. (2000). Asimmetrie informative e concorrenzialità nel mercato assicurativo italiano. *Mercato Concorrenza Regole*, *2*(3), 491-514.

Porter TM (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton University Press.

Pourebrahim, N., Sultana, S., Edwards, J., Gochanour, A., & Mohanty, S. (2019). Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy. *International Journal of Disaster Risk Reduction*, *37*, 101176.

Preis, T., Moat, H. S., Bishop, S. R., Treleaven, P., & Stanley, H. E. (2013). Quantifying the digital traces of Hurricane Sandy on Flickr. *Scientific Reports*, *3*(1), 3141.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, *1*(1), 51-59.

Prpić J, Taeihagh A, Melton J. (2015). The Fundamentals of Policy Crowdsourcing. *Policy & Internet, 7*(3), 340-361.

Puri, D., & Mohan, T. (2020). Behavioral advertising with big data: A consumer's perspective. *Int. J. Emerg. Technol*, *11*(3), 771-776.

Pyne, S., Rao, B. P., & Rao, S. B. (Eds.). (2016). *Big Data Analytics: Methods and Applications*. Springer.

Qi, E., & Deng, M. (2019). R&D investment enhance the financial performance of company driven by big data computing and analysis. *Computer Systems Science and Engineering*, *34*(4), 237-248.

Qiao, F., Li, P., Zhang, X., Ding, Z., Cheng, J., & Wang, H. (2017). Predicting social unrest events with hidden Markov models using GDELT. *Discrete Dynamics in Nature and Society*, *2017*.

Qu, J., Simes, R., & O'Mahony, J. (2017). How do digital technologies drive economic growth?. *Economic Record*, *93*, 57-69.

Quillian, L., & Pager, D. (2001). Black neighbors, higher crime? The role of racial stereotypes in evaluations of neighborhood crime. *American Journal of Sociology*, *107*(3), 717-767.

Raban, D. R., & Gordon, A. (2020). The evolution of data science and big data research: A bibliometric analysis. *Scientometrics*, *122*(3), 1563-1581.

Rackow, S. H. (2001). How the USA Patriot Act Wil Permit Governmental Infringement upon the Privacy of Americans in the Name of Intelligence Investigations. *U. Pa. L. Rev.*, *150*, 1651.

Ram, J., Zhang, C., & Koronios, A. (2016). The implications of big data analytics on business intelligence: A qualitative study in China. *Procedia Computer Science*, *87*, 221-226.

Ramadhan, M. I. (2017, July). An analysis of natural disaster data by using K-means and K-medoids algorithm of data mining techniques. In *2017 15th International Conference on Quality in Research (QiR): International Symposium on Electrical and Computer Engineering* (pp. 221-225). IEEE.

Ramakrishnan, N., Butler, P., Muthiah, S., Self, N., Khandpur, R., Saraf, P., ... & Mares, D. (2014, August). 'Beating the news' with EMBERS: forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1799-1808).

Ramasastry, A. (2005). Lost in Translation-Data Mining, National Security and the Adverse Inference Problem. *Santa Clara Computer & High Tech. LJ*, *22*(4), 757-796.

Ramos, G., Suh, J., Ghorashi, S., Meek, C., Banks, R., Amershi, S., ... & Bansal, G. (2019, May). Emerging perspectives in human-centered machine learning. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-8).

Rashid, A., & Khurshid, M. M. (2022, July). A descriptive literature review and classification of business intelligence and big data research. In *Science and Information Conference* (pp. 865-879). Cham: Springer International Publishing.

Ratcliffe, J. (2008). *Intelligence-Led Policing*. Willan.

Refonaa, J., Lakshmi, M., & Vivek, V. (2015, March). Analysis and prediction of natural disaster using spatial data mining technique. In *2015 International Conference on Circuits, Power and Computing Technologies (ICCPCT-2015)* (pp. 1-6). IEEE.

Reilly, B. C. (2015). Doing more with more: The efficacy of big data in the intelligence community. *American Intelligence Journal*, *32*(1), 18-24.

Ren, S. (2022). Optimization of enterprise financial management and decision-making systems based on big data. *Journal of Mathematics*, *2022*, 1-11.

Renan, D. (2016). The Fourth Amendment as administrative governance. *Stanford Law Review*, *68*(5), 16-38.

Renda, A., Arroyo, J., Fanni, R., Laurer, M., Sipiczki, A., Yeung, T., ... & de Pierrefeu, G. (2021). *Study to support an impact assessment of regulatory requirements for artificial intelligence in Europe*. European Commission: Brussels, Belgium.

Renke, W. N. (2005). Who controls the past now controls the future: counter-terrorism, data mining and privacy. *Alta. L. Rev.*, *43*, 779.

Rezzani, A. (2013). *Big Data: Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati.* Maggioli.

Ricci, O. (2018). Social media e attacchi terroristici. *Problemi dell'Informazione (2)*, 330-334.

Ricciuto, V. (2020). Il contratto ed i nuovi fenomeni patrimoniali: il caso della circolazione dei dati. *Rivista di Diritto Civile, 3*, 642.

Ricciuto, V. (2019). La patrimonializzazione dei dati personali. Contratto e mercato nella ricostruzione del fenomeno. In V. Cuffaro, R. D'Orazio, V. Ricciuto (Eds.), *I dati personali nel diritto europeo* (pp. 23-59). Giappichelli.

Rich, M. L. (2016). Machine learning, automated suspicion algorithms, and the Fourth Amendment. *University of Pennsylvania Law Review*, 871-929.

Richards, N. (2016). The Third Party Doctrine and the Future of the Cloud. *Wash. UL Rev.*, *94*, 1441.

Richardson, R. (2021). Defining and demystifying automated decision systems. *Md. L. Rev.*, *81*, 785.

Richelson, J. (2002). *The Wizards of Langley: Inside the CIA's Directorate of Science and Technology*. Westview Press.

Richey, M. K. (2015). From crowds to crystal balls: Hybrid analytic methods for anticipatory intelligence. *American Intelligence Journal*, *32*(1), 146-151.

Ridgeway, G. (2018). Policing in the era of Big Data. *Annual Review of Criminology*, *1*, 401-419.

Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, *1*(1), 33-36.

Riensche, R. M., & Whitney, P. D. (2012). Combining modeling and gaming for predictive analytics. *Security Informatics*, *1*(1), 1-7.

Rifai, N., Topol, E., Chan, E., Lo, Y. D., & Wittwer, C. T. (2015). Disruptive innovation in laboratory medicine. *Clinical Chemistry*, *61*(9), 1129-1132.

Rios, V. M. (2011). *Punished: Policing the Lives of Black and Latino Boys*. NYU Press.

Ritzer, G., Dean, P. & Jurgenson, N. (2012). The coming of age of the prosumer, *American Behavioral Scientist*, 56(4). 379-398.

Robertson, V. H. (2020). Excessive data collection: privacy considerations and abuse of dominance in the era of big data. *Common Market Law Review*, *57*(1).

Rocher, L., Hendrickx, J. M., & De Montjoye, Y. A. (2019). Estimating the success of re-identifications in incomplete datasets using generative models. *Nature Communications*, *10*(1), 1-9.

Rodotà, S. (1999). Conclusioni. In V. Cuffaro, V. Ricciuto, & Z. Zencovich (Eds.) *Trattamento dei dati e tutela della persona* (p. 695). Giuffré.

Roff, H. (2020). Uncomfortable ground truths: Predictive analytics and national security. *Brookings National Security Report*.

Rolli, R., & D'Ambrosio, M. (2022). Consenso e accountability: i poli del commercio dei dati personali online. *PA - Persona e Amministrazione 10(1)*, 783-800.

Rosenzweig, P. (2009). Privacy and counter-terrorism: The pervasiveness of data. *Case W. Res. J. Int'l L.*, *42*, 625.

Rosenzweig, P. (2006). Privacy and Consequences: Legal and Policy Structures for Implementing New Counter-Terrorism Technologies and Protecting Civil Liberty. In R.L. Popp & J. Yen (Eds.) *Emergent Information Technologies and Enabling Policies for Counter-Terrorism*, (pp. 421-438). Institute of Electrical & Electronics Engineers, Inc.

Ross Arguedas, A., Robertson, C., Fletcher, R., & Nielsen, R. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. Reuters Institute for the Study of Journalism.

Rossmo, D. K. (2012). Recent developments in Geographic Profiling. *Policing: A Journal of Policy and Practice*, *6*(2), 144-150.

Rossmo, D. K., & Baeza, J. J. (1998). The Upper East Side Rapist: A case study in Geographic Profiling. In *Meeting of the American Society of Criminology*, *Washington, DC.*

Rossmo, D. K., & Rombouts, S. (2016). Geographic Profiling. In R. Wortley and M. Townsley. (Eds.) *Environmental Criminology and Crime Analysis*. Routledge.

Rota, A. (2017). Rapporto di lavoro e Big Data Analytics: profili critici e risposte possibili. *Labour & Law Issues*, *3*(1), I-32.

Rothstein, M. A. (2013). Privacy and technology in the twenty-first century. *U. Louisville L. Rev.*,*52*, 333-344.

Roy, K. C., Hasan, S., & Mozumder, P. (2020). A multilabel classification approach to identify hurricane-induced infrastructure disruptions using social media data. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1387-1402.

Rubinstein, I. S., Lee, R. D., & Schwartz, P. M. (2008). Data mining and Internet profiling: Emerging regulatory and technological approaches. *U. Chi. L. Rev.*, *75*, 261.

Rule, J. B. (1974). *Private Lives and Public Surveillance: Social Control in the Computer Age*. Schocken Books.

Ruppert, E. (2012). The governmental topologies of database devices, *Theory, Culture & Society*, 29(4-5), 116-136.

Ruppert, E., Law, J., & Savage, M. (2013). Reassembling social science methods: the challenge of digital devices. *Theory, Culture & Society*, *30*(4), 22-46.

Russom, P. (2011). Big Data Analytics. *TDWI Best Practices Report, Fourth Quarter*, *19*(4), 1-34.

Sachan A, Roy D (2012) TGPM: terrorist group prediction model of counter terrorism. *Int J Comput Appl 44*(10), 49-52.

Saetnan, A. R., Schneider, I., & Green, N. (Eds.). (2018). *The Politics and Policies of Big Data: Big Data, Big Brother?*. Routledge

Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management*, *54*(5), 758-790.

Saini, J.K., & Bansal, D. (2023, June 03). Computational techniques to counter terrorism: a systematic survey. *Multimed Tools Appl*, 1-26.

Saini, J. K., & Bansal, D. (2019). A comparative study and automated detection of illegal weapon procurement over dark web. *Cybernetics and Systems*, *50*(5), 405-416.

Sampson, R. J., & Bartusch, D. J. (1998). Legal cynicism and (subcultural?) tolerance of deviance: The neighborhood context of racial differences. *Law and Society Review*, 777-804.

Samuel, O., Javaid, N., Alghamdi, T. A., & Kumar, N. (2022). Towards sustainable smart cities: A secure and scalable trading system for residential homes using blockchain and artificial intelligence. *Sustainable Cities and Society*, *76*, 103371.

Sandhu, A., & Fussey, P. (2021). The 'uberization of policing'? How police negotiate and operationalise predictive policing technology. *Policing and Society*, *31*(1), 66-81.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2016). Automation, algorithms, and politics| when the algorithm itself is a racist: Diagnosing ethical harm in the basic components of software. *International Journal of Communication*, *10*, 19.

Sanfilippo, A., Butner, S., Cowell, A., Dalton, A., Haack, J., Kreyling, S., ... & Whitney, P. (2011). Technosocial predictive analytics for illicit nuclear trafficking. In *Social Computing, Behavioral-Cultural Modeling and Prediction: 4th International Conference, SBP 2011, College Park, MD, USA, March 29-31, 2011. Proceedings 4* (pp. 374-381). Springer.

Sanfilippo, A., Cowell, A.J., Malone, L., Riensche, R., Thomas, J., Unwin, S., Whitney, P., Wong, P.C. (2009, June). Technosocial Predictive Analytics in support of naturalistic decision making. *Proceedings of NDM9, the Ninth International Conference on Naturalistic Decision Making, London*.

Sanfilippo, A., Gilbert, N., & Greaves, M. (2012). Technosocial Predictive Analytics for security informatics. *Security Informatics*, *1*, 1-3.

Sanfilippo, A. P., Riensche, R. M., Unwin, S. D., & Amaya, J. P. (2010). *Bridging the gap between human judgment and automated reasoning in Predictive Analytics* (No. PNNL-SA-71894). Pacific Northwest National Lab.(PNNL), Richland, WA.

Sardi, A., Sorano, E., Cantino, V., & Garengo, P. (2020). Big data and performance measurement research: trends, evolution and future opportunities. *Measuring Business Excellence*, Vol. ahead-of-print No. ahead-of-print.

Sawicki, A. (2016). The Internet of things. *World Scientific News*, *48*, 89-96.

Schaub, N. (2018). The role of data providers as information intermediaries. *Journal of Financial and Quantitative Analysis*, *53*(4), 1805-1838.

Schuelke-Leech, B. A. (2018). A model for understanding the orders of magnitude of disruptive technologies. *Technological Forecasting and Social Change*, *129*, 261-274.

Schwab, K. (2017). *The Fourth Industrial Revolution*. Currency.

Schwartz, P. M. (2011). Regulating Governmental Data Mining in the United States and Germany: Constitutional Courts, the State, and New Technology. *Wm. & Mary L. Rev.*, *53*, 351-387.

Scoblic, J. P. (2018). Beacon and Warning: Sherman Kent, Scientific Hubris, and the CIA's Office of National Estimates. *Texas National Security Review*, 1(4).

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, *4*(2).

Seaver, N. (2013, April). Knowing algorithms. *Media in Transition, 8.* Cambridge, MA.

Sedkaoui, S., & Khelfaoui, M. (2020). *Sharing Economy and Big Data Analytics*. John Wiley & Sons.

Seele, P. (2017, June 01). Predictive Sustainability Control: A review assessing the potential to transfer big data driven 'predictive policing' to corporate sustainability management. *J. Clean. Prod.*, 153, 673-686.

Seidel, S., Berente, N., Lindberg, A., Lyytinen, K., & Nickerson, J. V. (2018). Autonomous tools and design: a triple-loop approach to human-machine learning. *Communications of the ACM*, *62*(1), 50-57.

Seifert, J. W. (2004). Data mining and the search for security: Challenges for connecting the dots and databases. *Government Information Quarterly*, *21*(4), 461-480.

Selbst, A. D. (2017). Disparate impact in big data policing. *Ga. L. Rev.*, *52*, 109.

Senigaglia, R. (2020). La dimensione patrimoniale del diritto alla protezione dei dati personali. *Contratto e Impresa, 2*, 760.

Shabana, M., & Sharma, K. V. (2019). A study on Big data advancement and big data analytics. *Journal of App. Sc. Computations*, *6*(1).

Shabat, H. A., & Omar, N. (2015). Named entity recognition in crime news documents using classifiers combination. *Middle-East Journal of Scientific Research*, *23*(6), 1215-1221.

Shabat, H., Omar, N., & Rahem, K. (2014). Named entity recognition in crime using machine learning approach. In *Information Retrieval Technology: 10th Asia Information Retrieval Societies Conference, AIRS 2014, Kuching, Malaysia, December 3-5, 2014. Proceedings 10* (pp. 280-288). Springer International.

Shah, N. D., Steyerberg, E. W., & Kent, D. M. (2018). Big data and predictive analytics: recalibrating expectations. *Jama*, *320*(1), 27-28.

Shaikh, M. A., Wang, J., Liu, H., & Song, Y. (2007). Investigative data mining for counterterrorism. In *Advances in Hybrid Information Technology: First International Conference, ICHIT 2006, Jeju Island, Korea, November 9-11, 2006, Revised Selected Papers 1* (pp. 31-41). Springer.

Shapiro, A. (2019). Predictive Policing for reform? Indeterminacy and intervention in Big Data Policing. *Surveillance & Society*, *17*(3/4), 456-472.

Sharma, S., Dhanda, N., & Verma, R. (2023, January). Urban Vertical Farming: A Review. In *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 432-437). IEEE.

Sharma, A., Podoplelova, E., Shapovalov, G., Tselykh, A., & Tselykh, A. (2021). Sustainable smart cities: convergence of artificial intelligence and blockchain. *Sustainability*, *13*(23), 13076.

Shaw, C. R. (1929). *Delinquency Areas.* University of Chicago Press.

Shaw, C., & McKay, H. (1942). *Juvenile Delinquency and Urban Areas*. Chicago: University of Chicago Press.

Shelton, T., Poorthuis, A., Graham, M., & Zook, M. (2014). Mapping the data shadows of Hurricane Sandy: Uncovering the sociospatial dimensions of 'big data'. *Geoforum*, *52*, 167-179.

Shi, Y. (2022*). Advances in Big Data Analytics Theory, Algorithms and Practices*. Springer.

Shiau, W.L.; Chen, H.; Wang, Z., & Dwivedi, Y.K. (2023), Exploring core knowledge in business intelligence research, *Internet Research*, 33(3), 1179-1201.

Shilton, K. (2012). Participatory Personal Data: An emerging research challenge for the information sciences. *Journal of the American Society for Information Science and Technology 63*(10). 1905-1915.

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, *36*(6), 495-504.

Shuja, J., Alanazi, E., Alasmary, W., & Alashaikh, A. (2021). COVID-19 open source data sets: a comprehensive survey. *Applied Intelligence*, *51*, 1296-1325.

Sidorenko, E.L., Khisamova, Z.I. (2020). The Readiness of the Economy for Digitalization: Basic Methodological Approaches. In: Ashmarina, S., Vochozka, M., Mantulenko, V. (Eds) *Digital Age: Chances, Challenges and Future*. ISCDTE 2019. Lecture Notes in Networks and Systems, vol 84. Springer.

Simon, F. M. (2019). "We power democracy": Exploring the promises of the political data analytics industry. *The Information Society*, *35*(3), 158-169.

Simmons, R. (2018). Big data, machine judges, and the legitimacy of the criminal justice system. *UC Davis L. Rev.*, *52*, 1067.

Simmons, R. (2016). Quantifying criminal procedure: how to unlock the potential of big data in our criminal justice system. *Mich. St. L. Rev.*, 947.

Simran, K., Sriram, S., Vinayakumar, R., & Soman, K. P. (2020). Deep learning approach for intelligent named entity recognition of cyber security. In *Advances in Signal Processing and Intelligent Recognition Systems: 5th International Symposium, SIRS 2019, Trivandrum, India, December 18–21, 2019, Revised Selected Papers 5* (pp. 163-172). Springer.

Singh, K., Chaudhary, A. S., & Kaur, P. (2019, August). A machine learning approach for enhancing defence against global terrorism. In *2019 Twelfth International Conference on Contemporary Computing (IC3)*(pp. 1-5). IEEE.

Singh, S. K., & Del Giudice, M. (2019). Big Data Analytics, dynamic capabilities and firm performance. *Management Decision*, *57*(8), 1729-1733.

Singh, J., Sajid, M., Gupta, S. K., & Haidri, R. A. (2022). Artificial Intelligence and Blockchain Technologies for Smart City. *Intelligent Green Technologies for Sustainable Smart Cities*, 317-330.

Singh, S., Sharma, P. K., Yoon, B., Shojafar, M., Cho, G. H., & Ra, I. H. (2020). Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. *Sustainable Cities and Society*, *63*, 102364.

Skillicorn, D. B., & Vats, N. (2007). Novel information discovery for intelligence and counterterrorism. *Decision Support Systems*, *43*(4), 1375-1382.

Skoff, G., & Rollo, S. (2022). Techno-social Futures: Trapped or Transformative. In T. Ray, R. Pillai Rajagopalan, and P. Moha (Eds.) *Digital Debates: CyFy Journal 2022*. (pp. 67-73). Observer Research Foundation.

Slobogin, C. (2008). Government data mining and the Fourth Amendment. *The University of Chicago Law Review*, *75*(1), 317-341.

Smith, A. (2019). *Consumer Behaviour and Analytics*. Routledge.

Snook, B., Zito, M., Bennell, C., & Taylor, P. J. (2005). On the complexity and accuracy of geographic profiling strategies. *Journal of Quantitative Criminology*, *21*, 1-26.

Soliman, G. M., & Abou-El-Enien, T. H. (2019). Terrorism Prediction Using Artificial Neural Network. *Rev. d'Intelligence Artificial, 33*(2), 81-87.

Solove, D. J. (2008). Data mining and the security-liberty debate. *University of Chicago Law Review*, *75*(1), 343-362.

Solove, D. J. (2007). I've got nothing to hide and other misunderstandings of privacy. *San Diego L. Rev.*, *44*, 745.

Solove, D. J. (2001). Digital dossiers and the dissipation of Fourth Amendment privacy. *S. Cal. L. Rev.*, *75*, 1083.

Solow-Niederman, A. (2022). Information privacy and the inference economy. *Nw. UL Rev.*, *117*, 357.

Southerland, V. M. (2020). The intersection of race and algorithmic tools in the criminal legal system. *Md. L. Rev.*, *80*, 487.

Sperrle, F., El-Assady, M., Guo, G., Borgo, R., Chau, D. H., Endert, A., & Keim, D. (2021, June). A Survey of Human-Centered Evaluations in Human-Centered Machine Learning. In *Computer Graphics Forum*, 40(3), 543-568.

Spiegel, J. (2018). The Ethics of Virtual Reality Technology: Social Hazards and Public Policy Recommendations. *Sci. Eng. Ethics, 24,* 1537-1550.

Stalph, F., & Heravi, B. (2021). Exploring Data Visualisations: An Analytical Framework Based on Dimensional Components of Data Artefacts in Journalism. *Digital Journalism*, 1-23.

Strang, K. D., & Sun, Z. (2017). Analyzing relationships in terrorism big data using Hadoop and statistics. *Journal of Computer Information Systems*, *57*(1), 67-75.

Stanger, A. (2022). The Real Cost of Surveillance Capitalism: Digital Humanism in the United States and Europe. *Perspectives on Digital Humanism*, 33-40.

Strange, R., & Zucchella, A. (2017). Industry 4.0, global value chains and international business. *Multinational Business Review*, *25*(3), 174-184.

Stransky, G., Knight, T., & Zych, F. (January, 2023). California Consumer Privacy Act Enforcement and Preparing for 2023 Data Privacy Rules. *Pratt's Privacy & Cybersecurity Law Report 2023, 9(1),* 12-16.

Stromer-Galley, J. (2019). *Presidential campaigning in the Internet age*. Oxford University Press.

Stuart, F. (2016). Becoming "copwise": Policing, culture, and the collateral consequences of street-level criminalization. *Law & Society Review*, *50*(2), 279-313.

Stylianou, K., Dimitriou, L., & Abdel-Aty, M. (2019). Big data and road safety: A comprehensive review. *Mobility Patterns, Big Data And Transport Analytics*, 297-343.

Sufi, F. K. (2022). AI-Global Events: A Software for analyzing, identifying and explaining global events with Artificial Intelligence. *Software Impacts*, *11*, 100218.

Sun, C., Yang, Z., Wang, L., Zhang, Y., Lin, H., & Wang, J. (2021). Deep learning with language models improves named entity recognition for PharmaCoNER. *BMC Bioinformatics*, *22*(1), 1-16.

Sun, X., Yu, H., Solvang, W. D., Wang, Y., & Wang, K. (2021). The application of Industry 4.0 Technologies in sustainable logistics: A systematic literature review (2012–2020) to explore future research opportunities. *Environmental Science and Pollution Research*, 1-32.

Sun, Z., Sun, L., & Strang, K. (2018). Big data analytics services for enhancing business intelligence. *Journal of Computer Information Systems*, *58*(2), 162-169.

Sun, Z., Zou, H., & Strang, K. (2015). Big data analytics as a service for business intelligence. In *Open and Big Data Management and Innovation: 14th IFIP WG 6.11 Conference on e-Business, e-Services, and e-Society, I3E 2015, Delft, The Netherlands, October 13-15, 2015, Proceedings 14* (pp. 200-211). Springer.

Sunagar, P., Hanumantharaju, R., Siddesh, G. M., Kanavalli, A., & Srinivasa, K. G. (2020). Influence of big data in smart tourism. In S. Bhattacharyya, V. Snášel, D. Gupta, and A. Khanna (Eds.) *Hybrid Computational Intelligence. Challenges and Applications* (pp. 25-47). Academic Press.

Supriyadi, B., Windarto, A. P., & Soemartono, T. (2018). Classification of natural disaster prone areas in Indonesia using K-means. *International Journal of Grid and Distributed Computing*, *11*(8), 87-98.

Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of Emerging Disruptive Technologies. *Regulation & Governance*, *15*(4), 1009-1019.

Taeihagh, A., Givoni, M., & Bañares-Alcántara, R. (2013). Which policy first? A network-centric approach for the analysis and ranking of policy measures. *Environment and Planning B: Planning and Design*, *40*(4), 595-616.

Taipale, K. A. (2003). Data mining and domestic security: connecting the dots to make sense of data. *Colum. Sci. & Tech. L. Rev.*,*5*(1), 21-74.

Talapina E.V. (2022) Artificial Intelligence Processing and Risks of Discrimination. *Law. Journal of the Higher School of Economics, 15*(1), 4-27.

Tallman, E. F., Richardson, D., Rogow, T. M., Kendrick, D. C., & Dixon, B. E. (2023). Leveraging HIE to facilitate large-scale data analytics. In B. E. Dixon (Ed.) *Health Information Exchange*: *Navigating and Managing a Network of Health Information Systems* (pp. 399-421). Academic Press.

Talreja, D., Nagaraj, J., Varsha, N. J., & Mahesh, K. (2017, September). Terrorism analytics: Learning to predict the perpetrator. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1723-1726). IEEE.

Tan, K. H., Ji, G., Lim, C. P., & Tseng, M. L. (2017). Using big data to make better decisions in the digital economy. *International Journal of Production Research*, *55*(17), 4998-5000.

Tan, K. H., & Zhan, Y. (2017). Improving new product development using big data: a case study of an electronics company. *R&D Management*, *47*(4), 570-582.

Taylor, R. W., & Russell, A. L. (2012). The failure of police 'fusion'centers and the concept of a national intelligence sharing plan. *Police Practice and Research*, *13*(2), 184-200.

Terren, L., & Borge-Bravo, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, *9*, 99-118.

Terry, N. P. (2012). Protecting patient privacy in the age of big data. *UMKC L. Rev.*, *81*, 385.

Thobani, S. (2018). *Diritti della personalità e contratto. Dalle fattispecie più tradizionali al trattamento in massa dei dati personali.* Ledizioni.

Thomas, J. (2019). An overview of emerging disruptive technologies and key issues. *Development*, *62*(1-4), 5-12.

Thuraisingham, B. (2009, September). Data mining for malicious code detection and security applications. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology* (Vol. 2, pp. 6-7). IEEE.

Thuraisingham, B. (2004). Data mining for counter-terrorism. *Data Mining: Next Generation Challenges and Future Directions*, 157-183.

Thuraisingham, B. (2002, December). Data mining, national security, privacy and civil liberties. *ACM SIGKDD Explorations Newsletter*, *4*(2), 1-5.

Thuraisingham, B., Khan, L., Masud, M. M., & Hamlen, K. W. (2008, December). Data mining for security applications. In *2008 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing* (Vol. 2, pp. 585-589). IEEE.

Tien, L. (2004). Privacy, Technology and Data Mining. *Ohio NUL Rev.*, *30*, 389.

Tokic, D. (2018). BlackRock Robo-Advisor 4.0: When artificial intelligence replaces human discretion. *Strategic Change*, *27*(4), 285-290.

Tolan, G. M., & Soliman, O. S. (2015). An experimental study of classification algorithms for terrorism prediction. *International Journal of Knowledge Engineering-IACSIT*, *1*(2), 107-112.

Torre-Bastida, A. I., Del Ser, J., Laña, I., Ilardia, M., Bilbao, M. N., & Campos-Cordobés, S. (2018). Big Data for transportation and mobility: recent advances, trends and challenges. *IET Intelligent Transport Systems*, *12*(8), 742-755.

Tosi, E. (2019). *Riservatezza e protezione dei dati personali tra GDPR e nuovo Codice Privacy*. Giuffré.

Tranos, E., & Mack, E. (2019). Big data: A new opportunity for transport geography?. *Journal of Transport Geography*, *76*.

Treverton, G. F., & Gabbard, C. B. (2008). *Assessing the Tradecraft of Intelligence Analysis*. Rand Corporation.

Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., & Rajagopal, R. (2021). NeuralProphet: Explainable forecasting at scale. *arXiv preprint arXiv:2111*.15397.

Tudose, M. B., Georgescu, A., & Avasilcăi, S. (2023). Global Analysis Regarding the Impact of Digital Transformation on Macroeconomic Outcomes. *Sustainability*, *15*(5), 4583.

Tufekci, Z. (2015, February). Algorithms in our midst: Information, power and choice when software is everywhere. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 1918-1918).

Tullini, P. (2018). L'economia delle piattaforme e le sfide del diritto del lavoro, in *Economia e Società Regionale*, *1*, 36-51.

Turner Lee, N. (2018). Detecting racial bias in algorithms and machine learning. *Journal of Information, Communication and Ethics in Society*, 16(3), 252-260.

Tzu, S. (2017). *The Art of War.(The Oldest Military Treatise in the World)*. Bahribook.

Uche, S.O., Tsopze, N. & Ebem, D.U. (2020): Data Mining Approach to Counterterrorism. *Computing, Information Systems, Development Informatics & Allied Research Journal*. 11(2), 77-90.

Ugwudike, P. (2022). AI audits for assessing design logics and building ethical systems: the case of predictive policing algorithms. *AI & Ethics*, *2*(1), 199-208.

Umoja Noble, S. (2018). *Algorithms of Oppression*. New York University Press.

Unsworth, K. (2016). The social contract and big data. *Journal of Information Ethics*, *25*(1), 83-97.

Uwagbole, S. O., Buchanan, W. J., & Fan, L. (2017, May). Applied machine learning predictive analytics to SQL injection attack detection and prevention. In *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)* (pp. 1087-1090). IEEE.

Vajjhala, N. R., Strang, K. D., & Sun, Z. (2015, August). Statistical modeling and visualizing open big data using a terrorism case study. In *2015 3rd International Conference on Future Internet of Things and Cloud* (pp. 489-496). IEEE.

Valverde, M. (2014), Studying the Governance of Crime and Security: Space, Time and Jurisdiction. *Criminology & Criminal Justice, 14*, 379-391.

Van Ark, B., De Vries, K., & Erumban, A. (2021). How to not miss a productivity revival once again. *National Institute Economic Review*, *255*, 9-24.

Van Brakel, R. (2016). Pre-emptive big data surveillance and its (dis) empowering consequences: The case of predictive policing. In B. van der Sloot, D. Broeders & E. Schrijvers (Eds.), *Exploring the Boundaries of Big Data* (pp. 117–141). Amsterdam University Press.

Van Puyvelde, D., Coulthart, S., & Hossain, M. S. (2017). Beyond the buzzword: big data and national security decision-making. *International Affairs*, *93*(6), 1397-1416.

Vardan, A., Sofya, O., Nane, M., & Arpine, M. (2023). Philosophical Comprehension of the Psychological and Information Influence Technologies in the Modern World. *Bulletin of the Armenian State Economic University*, (1), 87-98.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting Horizons*, *29*(2), 381-396.

Vasilescu, M. D., Serban, A. C., Dimian, G. C., Aceleanu, M. I., & Picatoste, X. (2020). Digital divide, skills and perceptions on digitalisation in the European Union—Towards a smart labour market. *PloS one*, *15*(4), e0232032.

Vassakis, K., Petrakis, E., & Kopanakis, I. (2018). Big data analytics: Applications, prospects and challenges. In G. Skourletopoulos, G. Mastorakis, C. X. Mavromoustakis, C. Dobre, and E. Pallis (Eds.) *Mobile Big Data: A Roadmap from Models to Technologies* (pp. 3-20). Springer.

Verma, C., Malhotra, S., Verma, S., Verma, V., & Tirwa K. (2018). Predictive modeling of terrorist attacks using machine learning. *Int. J. Pure Appl. Math*, *119*, 49-61.

Virk, A. L., Noor, M. A., Fiaz, S., Hussain, S., Hussain, H. A., Rehman, M., ... & Ma, W. (2020). Smart farming: an overview. *Smart Village Technology: Concepts and Developments*, 191-201.

Visentin, C. (2018) Il potere razionale degli algoritmi tra burocrazia e nuovi idealtipi. *The Lab's Quarterly, 20*(4), 47-72.

von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, *34*(4), 1607-1622.

Vossen, G. (2014). Big Data as the new enabler in Business and other Intelligence. *Vietnam Journal of Computer Science*, *1*(1), 3-14

Wahyudi, M., Meilinda, V., & Khoirunisa, A. (2022). The Digital Economy's Use of Big Data. *International Transactions on Artificial Intelligence*, *1*(1), 62-70.

Wakefield, S., & Wildeman, C. (2013). *Children of the Prison Boom: Mass Incarceration and the Future of American Inequality*. Oxford University Press.

Wang, L., & Alexander, C. A. (2016). Machine learning in big data. *International Journal of Mathematical, Engineering and Management Sciences*, *1*(2), 52.

Wang, F., Ding, L., Yu, H., & Zhao, Y. (2020). Big data analytics on enterprise credit risk evaluation of e-Business platform. *Information Systems and e-Business Management*, *18*, 311-350.

Wang, M., Fu, W., He, X., Hao, S., & Wu, X. (2020). A survey on large-scale machine learning. *IEEE Transactions on Knowledge and Data Engineering*, *34*(6), 2574-2594.

Wang, Y., & Sarkis, J. (2021). Emerging digitalisation technologies in freight transport and logistics: Current trends and future directions. *Transportation Research Part E: Logistics and Transportation Review*, *148*, 102291.

Wang, B., & Zhuang, J. (2017). Crisis information distribution on Twitter: a content analysis of tweets during Hurricane Sandy. *Natural Hazards*, *89*, 161-181.

Watney, M. (2019, October). Law Enforcement Use of Artificial Intelligence for Domestic Security: Challenges and Risks. In *Proceedings of the European Conference of Artificial Intelligence and Robotics (ECIAR), Oxford, UK, November* (pp. 341-348).

Watson, H. J. (2019). Update tutorial: Big Data analytics: Concepts, Technology, and Applications. *Communications of the Association for Information Systems*, *44*(1), 21.

Waytowich, N. R., Goecks, V. G., & Lawhern, V. J. (2018). Cycle-of-learning for autonomous systems from human interaction. *arXiv preprint arXiv:1808.09572*.

Weiss, C. (2004). The Coming Technology of Knowledge Discovery: A Final Blow to Privacy Protection?. *U. Ill. JL Tech. & Pol'y*, 253.

Weiss, S. M., & Indurkhya, N. (1998). *Predictive Data Mining: A Practical Guide*. Morgan Kaufmann.

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, *3(*1), 1-40.

Welch, T. F., & Widita, A. (2019). Big data in public transportation: a review of sources and methods. *Transport Reviews*, *39*(6), 795-818.

Wen, C., Yang, J., Gan, L., & Pan, Y. (2021). Big Data driven Internet of Things for credit evaluation and early warning in finance. *Future Generation Computer Systems*, *124*, 295-307.

Western, B., & Pettit, B. (2005). Black-white wage inequality, employment rates, and incarceration. *American Journal of Sociology*, *111*(2), 553-578.

Wiegand, W. A., & Donald Jr, G. (2015). *Encyclopedia of Library History* (Vol. 503). Routledge.

Wiktorowicz, Q. (2004). *Islamic Activism: A Social Movement Theory Approach*, Indiana University Press.

Wiil, U. K., Memon, N., & Gniadek, J. (2011). Crimefighter: A toolbox for counterterrorism. In *Knowledge Discovery, Knowledge Engineering and Knowledge Management: First International Joint Conference, IC3K 2009, Funchal, Madeira, Portugal, October 6-8, 2009, Revised Selected Papers 1* (pp. 337-350). Springer.

Wilson, S., Steele, S., & Adeli, K. (2022). Innovative technological advancements in laboratory medicine: Predicting the lab of the future. *Biotechnology & Biotechnological Equipment*, *36*(sup1), S9-S21.

Wischmeyer, T. (2020). Artificial intelligence and transparency: opening the black box. *Regulating Artificial Intelligence*, 75-101.

Wong, P. C., Leung, L. R., Lu, N., Paget, M., Correia Jr, J., Jiang, W., ... & Sanfilippo, A. (2009, March). Predicting the Impact of Climate Change on US Power Grids and Its Wider Implications on National Security. In *AAAI Spring Symposium: Technosocial Predictive Analytics* (pp. 148-153).

Wood, M. A., & Warren, I. (2022). Analysing the multiple dimensions of predictive policing's techno-social harms. *Justice, Power and Resistance*, *5*(3), 208-226.

Woschank, M., Rauch, E., & Zsifkovits, H. (2020). A review of further directions for Artificial Intelligence, Machine Learning, and Deep Learning in smart logistics. *Sustainability*, *12*(9), 3760.

Wright, B., Payne, J., Steckman, M., & Stevson, S. (2009, October). Palantir: A visualization platform for real-world analysis. In *2009 IEEE Symposium on Visual Analytics Science and Technology* (pp. 249-250). IEEE.

Wu, J., Wang, J., Nicholas, S., Maitland, E., & Fan, Q. (2020). Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations. *Journal of Medical Internet Research*, *22*(10), e21980.

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, *135*, 364-381.

Wong, P. C., Leung, L. R., Lu, N., Paget, M., Correia Jr, J., Jiang, W., ... & Sanfilippo, A. (2009, March). Predicting the Impact of Climate Change on US Power Grids and Its Wider Implications on National Security. In *AAAI Spring Symposium: Technosocial Predictive Analytics* (pp. 148-153).

Wysokińska, Z. (2021). A Review of the Impact of the Digital Transformation on the Global and European Economy. *Comparative Economic Research. Central and Eastern Europe*, *24*(3), 75-92.

Xia, T., & Gu, Y. (2019, July). Building terrorist knowledge graph from Global Terrorism Database and Wikipedia. In *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)* (pp. 194-196). IEEE.

Xin, D., Ma, L., Liu, J., Macke, S., Song, S., & Parameswaran, A. (2018, June). Accelerating Human-in-the-loop Machine Learning: Challenges and Opportunities. In *Proceedings of the Second Workshop on Data Management for End-To-End Machine Learning, 9* (pp. 1-4).

Yadrovskaia, M., Porksheyan, M., Petrova, A., Dudukalova, D., & Bulygin, Y. (2023, March). About the attitude towards artificial intelligence technologies. In *E3S Web of Conferences* (Vol. 376, p. 05025).

Yang, H. L. (2022, December). Application of Big Data in Counter-Terrorism Intelligence Analysis and Early Warning. In *2022 International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2022) 2022 International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2022)* (pp. 1193-1200). Atlantis Press.

Ying, S., Sindakis, S., Aggarwal, S., Chen, C., & Su, J. (2021). Managing big data in the retail industry of Singapore: Examining the impact on customer satisfaction and organizational performance. *European Management Journal*, *39*(3), 390-400.

Yoo, I., & Yi, C. G. (2022). Economic innovation caused by digital transformation and impact on social systems. *Sustainability*, *14*(5), 2600.

Yu, M., Huang, Q., Qin, H., Scheele, C., & Yang, C. (2020). Deep learning for real-time social media text classification for situation awareness–using Hurricanes Sandy, Harvey, and Irma as case studies. In Z. Li, Q. Huang, and C.T. Emrich (Eds.) *Social Sensing and Big Data Computing for Disaster Management* (pp. 33-50). Routledge.

Yuan, F., Liu, R., Mao, L., & Li, M. (2021). Internet of people enabled framework for evaluating performance loss and resilience of urban critical infrastructures. *Safety Science*, *134*, 105079.

Yuvaraj, N., Praghash, K., Logeshwaran, J., Peter, G., & Stonier, A. A. (2023). An Artificial Intelligence Based Sustainable Approaches—IoT Systems for Smart Cities. In *AI Models for Blockchain-Based Intelligent Networks in IoT Systems: Concepts, Methodologies, Tools, and Applications* (pp. 105-120). Cham: Springer International Publishing.

Zagorecki, A. T., Johnson, D. E., & Ristvej, J. (2013). Data mining and machine learning in the context of disaster and crisis management. *International Journal of Emergency Management*, *9*(4), 351-365

Zamin, N. (2009, November). Information extraction for counter-terrorism: A survey. In *2009 Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns* (pp. 520-526). IEEE.

Zanzotto, F. M. (2019). Human-in-the-loop Artificial Intelligence. *Journal of Artificial Intelligence Research*, *64*, 243-252.

Završnik, A. (2020, March). Criminal justice, artificial intelligence systems, and human rights. In *ERA Forum* (Vol. 20, No. 4, pp. 567-583). Springer.

Zeitzoff, T., Kelly, J., & Lotan, G. (2015). Using social media to measure foreign policy dynamics: An empirical analysis of the Iranian–Israeli confrontation (2012–13). *Journal of Peace Research*, *52*(3), 368-383.

Zhang, L., Pan, Y., & Zhang, T. (2004, July). Focused named entity recognition using machine learning. In *Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 281-288).

Zhang, L., Pentina, I., & Fan, Y. (2021). Who do you choose? Comparing perceptions of human vs robo-advisor in the context of financial services. *Journal of Services Marketing*, *35*(5), 634-646.

Zhang, L., Priestley, J., DeMaio, J., Ni, S., & Tian, X. (2021). Measuring customer similarity and identifying cross-selling products by community detection. *Big Data*, *9*(2), 132-143.

Zhang, Y., Song, B., Du, X., & Guizani, M. (2018). Vehicle tracking using surveillance with multimodal data fusion. *IEEE Transactions on Intelligent Transportation Systems*, *19*(7), 2353-2361.

Zhang, J. Z., Srivastava, P. R., Sharma, D., & Eachempati, P. (2021). Big data analytics and machine learning: A retrospective overview and bibliometric analysis. *Expert Systems with Applications*, *184*, 115561.

Zhang, C., Yao, W., Yang, Y., Huang, R., & Mostafavi, A. (2020). Semiautomated social media analytics for sensing societal impacts due to community disruptions during disasters. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1331-1348.

Zhang, Y., Zhang, M., Li, J., Liu, G., Yang, M. M., & Liu, S. (2021). A bibliometric review of a decade of research: Big data in business research–Setting a research agenda. *Journal of Business Research*, *131*, 374-390.

Zhao, L. (2021). Event prediction in the big data era: A systematic survey. *ACM Computing Surveys (CSUR)*, *54*(5), 1-37.

Zheng, L., Wang, F., & Zheng, X. (2017, July). Complex network construction method to extract the nature disaster chain based on data mining. In *2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC)* (pp. 25-28). IEEE.

Zhou, L., Pan, S., Wang, J., & Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350-361.

Zhu, Y., Ozbay, K., Xie, K., & Yang, H. (2016). Using big data to study resilience of taxi and subway trips for hurricanes Sandy and Irene. *Transportation Research Record*, *2599*(1), 70-80.

Zhu, X., & Yang, Y. (2021). Big data analytics for improving financial performance and sustainability. *Journal of Systems Science and Information*, *9*(2), 175-191.

Ziewitz, M. (2016). Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*, *41*(1), 3-16.

Zolkover, A., Petrunenko, I., Iastremska, O., Stashkevych, O., & Mehdizade, M. M. (2022). Benefits and risks of digital business transformation: The example of eastern Europe countries. *Journal of Eastern European and Central Asian Research (JEECAR)*, *9*(2), 344-356.

Zuboff, S. (2022). Surveillance capitalism or democracy? The death match of institutional orders and the politics of knowledge in our information civilization. *Organization Theory*, *3*(3).

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The fight for a human future at the new frontier of power*. Profile Books Ltd.

Zuboff, S. (2019, January). Surveillance capitalism and the challenge of collective action. In *New Labor Forum* (Vol. 28, No. 1), 10-29. SAGE.

Zuboff, S. (2010). Creating value in the age of distributed capitalism. *McKinsey Quarterly*, *4*, 45-55.

Zuboff, S., Möllers, N., Wood, D. M., & Lyon, D. (2019). Surveillance Capitalism: An interview with Shoshana Zuboff. *Surveillance & Society*, *17*(1/2), 257-266.

Sitografia

AGCM, AGCOM & Autorità Garante per la Protezione dei Dati Personali. (July 2019). *Big Data. Indagine conoscitiva congiunta. Linee Guida e Raccomandazioni di Policy*. Retrived from agcm.it: https://www.agcm.it/dotcmsdoc/allegati-news/Big_Data_Lineeguida_Raccomandazioni_di_policy.pdf

AGCOM (2018). *News vs. Fake nel sistema dell'informazione*. Retrived from agcom.it: https://www.agcom.it/documents/10179/12791486/Pubblicazione+23-11-2018/93869b4f-0a8d-4380-aad2-c10a0e426d83?version=1.0

AIMC - Associazione Italiana Medicina delle Catastrofi. (2023, June 8-10). *Gli scenari del Terzo Millennio: Nuove Tecnologie e Antichi Saperi*. Retrieved from simlaweb.it: https://www.simlaweb.it/wp-content/uploads/2023/06/PROGRAMMA-XXI-Congresso-AIMC-2023-Roma.pdf

Airoldi, M. & Gambetta, D. (April 2019). *Sul mito della neutralità algoritmica*. Retrived from researchgate.net: https://www.researchgate.net/publication/332254603_Sul_mito_della_neutralita_algoritmica

Alrasheed, G. & Rigato, B. (2019, February 5). *Exploring the Dark Web: Where Terrorists Hide?*. Retrived from carleton.ca: https://carleton.ca/align/2019/illuminate-exploring-the-dark-web-where-terrorists-hide/

Antonielli, A. (2021, July 29). *Crittografia: come funziona e come sfruttarla per la sicurezza informatica*. Retrived from blog.osservatori.net: https://blog.osservatori.net/it_it/crittografia-cosa-si-intende-e-quali-sono-le-principali-applicazioni

Antwan, J. (2017, May 11). *Amazon Prime and the Economics of Race*. Retrieved from huffpost.com: https://www.huffpost.com/entry/amazon-prime-and-the-econ_b_9895716

Auschitzky, E., Hammer, M., & Rajagopaul, A. (2014). *How big data can improve manufacturing*. Retrived from mckinsey.com: https://www.mckinsey.com/capabilities/operations/our-insights/how-big-data-can-improve-manufacturing

Australian Competition & Consumer Commission. (2023, September 27). *The Consumer Data Right*. Retrieved from accc.gov.au: https://www.accc.gov.au/by-industry/banking-and-finance/the-consumer-data-right#:~:text=The%20Consumer%20Data%20Right%20allows%20consumers%20to%20safely,will%20come%20soon%2C%20with%20other%20sectors%20to%20follow.

Australian Government Productivity Commission. (2016). *Digital Disruption: What Do Governments Need to Do?*. Retrived from pc.gov.au: https://www.pc.gov.au/research/completed/digital-disruption

Australian Strategic Policy Institute (ASPI). (2017). *Big data in national security: online resource*. Retrieved from ad-aspi.s3.ap-southeast-2.amazonaws.com: https://ad-aspi.s3.ap-southeast-2.amazonaws.com/2017-08/Big%20data%20online%20resource.pdf?VersionId=ozBjJ5aHozFL4Hj47ZNJ.whWz_RDWIWl

Australian Strategic Policy Institute (ASPI). (2010). *Annual Report 2009-2010*. Retrieved from s3-ap-southeast-2.amazonaws.com: https://s3-ap-southeast-2.amazonaws.com/ad-aspi/2017-07/ASPI-AR_0910.pdf

Bailo, F. (2015). Le banche dati al servizio di piccole e medie imprese. *Riv. Elettron. Dir. Econ. Manag. 1*, 9-29. Retrived from clioedu.it: https://www.clioedu.it/documenti/eventi-live-ondemand/rivista-elettronica/Rivista-elettronica-1-2015_93.pdf

Bao En, T. (n.d.). *Swimming In Sensors, Drowning In Data - Big Data Analytics For Military Intelligence*. Retrieved from mindef.gov.sg: https://www.mindef.gov.sg/oms/safti/pointer/documents/pdf/V42N1_SwimmingInSensors.pdf

Barbaschow, A. (2017, November 26). *Australians will own their banking and internet data under new legislation*. Retrieved from zdnet.com: https://www.zdnet.com/article/australians-will-own-their-banking-and-internet-data-under-new-legislation/

Barocas, S., Hood, S.& Ziewitz, M, (2013, March 29).*Governing Algorithms: A Provocation Piece*. Retrived from ssrn.com: https://ssrn.com/abstract=2245322 or http://dx.doi.org/10.2139/ssrn.2245322

Behm, A. (2007, November). *The Australian Intelligence Community in 2020*. Retrieved from search.informit.org: https://search.informit.org/doi/abs/10.3316/agispt.20083068

Bellini, M. (2023, June 15). *Big data, cosa sono e come sono utili alle aziende: soluzioni ed esempi pratici*. Retrived from bigdata4innovation.it: https://www.bigdata4innovation.it/big-data/big-data-analytics-data-science-e-data-scientist-soluzioni-e-skill-della-data-driven-economy/

Best Jr, R. A. (2007, February 13*). Sharing law enforcement and intelligence information: The congressional role*. CRS Report for Congress. Retrived from sgp.fas.org: https://sgp.fas.org/crs/intel/RL33873.pdf

Bettiol, M., Capestro, M., De Marchi, V., & Di Maria, E. (2020). *Industry 4.0 Investments In Manufacturing Firms And Internationalization*. Retrived from economia.unipd.it: https://www.economia.unipd.it/en/sites/economia.unipd.it/files/20200245.pdf

Bougleux, E. (2021). *Un General Data Protection Regulation (GDPR) non molto general*. Retrieved from journals.openedition.org: https://journals.openedition.org/aam/4098

Burns, M. (2015, January 12). *Leaked Palantir Doc Reveals Uses, Specific Functions, and Key Clients*. Retrived from techcrunch.com: https://techcrunch.com/2015/01/11/leaked-palantir-doc-reveals-uses-specific-functions-and-key-clients/

Calzolaio, S. (2017, December 20). *Privacy by design. Principi, dinamiche, ambizioni del nuovo Reg. Ue 2016/679*. Retrieved from u-pad.unimc.it: https://u-pad.unimc.it/retrieve/de3e5026-e7c0-83cd-e053-3a05fe0a1d44/Calzolaio%202017%20-%20Federalismi%20reg.%20Ue%20679-16.pdf

Camera dei Deputati, Parlamento Italiano. (2020, February 12). *AGCOM: conclusa l'indagine conoscitiva sui Big data*. Retrieved from temi.camera.it: https://temi.camera.it/leg18/post/OCD15-13943/agcom-conclusa-l-indagineconoscitiva-sui-big-data.html

Caputo, C., & Ferorelli, G. (2023, May 12*). "Pseudonimizzazione" o "Anonimizzazione"? Sui dati, questo è il dilemma (e non solo)*. Retrived from agendadigitale.eu: https://www.agendadigitale.eu/sicurezza/privacy/pseudonimizzazione-o-anonimizzazione-sui-dati-questo-e-il-dilemma-e-non-solo/

Casalini, L. (2021, May 21). *La patrimonializzazione dei dati personali nella giurisprudenza del Consiglio di Stato*. Retrived from dirittodelrisparmio.it: https://www.dirittodelrisparmio.it/2021/05/21/la-patrimonializzazione-dei-dati-personali-nella-giurisprudenza-del-consiglio-di-stato/

Casey, R. (2021). *Left to their own devices: A technosocial ethnography of penal electronic monitoring in Scotland* (Doctoral dissertation, University of Glasgow). Retrived from theses.gla.ac.uk: https://theses.gla.ac.uk/82367/

Castigli, M. (2022, September 20). *Volume, varietà, velocità: le 3 V dei big data*. Retrieved from bigdata4innovation.it: https://www.bigdata4innovation.it/big-data/volume-varieta-velocita-le-3-v-dei-big-data/

Chakravorti, B. (2020). *Why It's So Hard for Users to Control Their Data*. Retrived from hbr.org: https://hbr.org/2020/01/why-companies-make-it-so-hard-for-users-to-control-their-data

Chihai, C. (2019). *Applicazione dell'art. 35 del Regolamento UE n. 2016/679-Data Protection Impact Assessment.* Retrieved from dspace.unive.it: http://dspace.unive.it/bitstream/handle/10579/15645/843121-1218112.pdf?sequence=2

Ciaramitaro, M. (2022). *Legami tra Industria 4.0 e internazionalizzazione delle piccole e medie imprese italiane.= Links between Industry 4.0 and Internationalisation of Italian SMEs* (Doctoral dissertation, Politecnico di Torino). Retrived from webthesis.biblio.polito.it: https://webthesis.biblio.polito.it/25845/

CMA Competition and Markets Authority. (2015, January 27). *Commercial use of consumer data.* Retrieved from gov.uk: https://www.gov.uk/cma-cases/commercial-use-of-consumer-data

Coglianese, C., & Lehr, D. (2017). Adjudicating by Algorithm, Regulating by Robot. *The Regulatory Review Opinion*. 798. Retrived from scholarship.law.upenn.edu: https://scholarship.law.upenn.edu/regreview-opinion/798

Committee on Homeland Security & Governmental Affairs. (2011, October 12). *Ten Years After 9/11: A Status Report on Information Sharing*. Retrieved from markle.org: https://www.markle.org/publications/1716-ten-years-after-911-status-report-information-sharing/

Crawford, K. (2013, April 01). *The Hidden Biases in Big Data.* Retrived from hbr.org: https://hbr.org/2013/04/the-hidden-biases-in-big-data

Data to Decisions Cooperative Research Centre. (2017, October). *Annual Report 2016/17*. Retrieved from uploads-ssl.webflow.com: https://uploads-ssl.webflow.com/5cd23e823ab9b1f01f815a54/5d0887e57acb80de62c7f691_DTD_023%202017%20Annual%20Report_A4_PUBLIC_FINAL%20LR.pdf

Davies, A. (2016, October 06). *Reviewing intelligence: send in the red team*. Retrieved from aspistrategist.org.au: https://www.aspistrategist.org.au/reviewing-intelligence-send-in-red/

De Chant, T. (2014, March 26). *The Inevitability of Predicting the Future*. Retrieved from pbs.org: https://www.pbs.org/wgbh/nova/article/predicting-the-future/

Deparday, V., Gevaert, C. M., Molinario, G., Soden, R., & Balog-Way, S. (2019). *Machine Learning for Disaster Risk Management*. World Bank. Retrived from documents.worldbank.org: http://documents.worldbank.org/curated/en/503591547666118137/Machine-Learning-for-Disaster-Risk-Management

Devlin H. (2016, December 19). *Discrimination by algorithm: scientists devise test to detect AI bias.* Retrived from theguardian.com:

https://www.theguardian.com/technology/2016/dec/19/discrimination-by-algorithm-scientists-devise-test-to-detect-ai-bias

Di Falco, D. (2022). *Digital business transformation: il ruolo di big data e privacy per l'innovazione*. Retrived from tesi.luiss.it: http://tesi.luiss.it/33589/

Dimitrakopoulos, V. (2017, March 23). *New Data Visualisation tools for the Department of Defence*. Retrieved from cooperativeresearch.org.au: https://cooperativeresearch.org.au/new-data-visualisation-tools-for-the-department-of-defence/

Di Stefano, M. (2020, February 14). *Ricerca investigativa: le metodologie OSINT/SOCMINT sulle fonti aperte*. Retrived from altalex.com: https://www.altalex.com/documents/news/2020/02/14/ricerca-investigativa-metodologie-osint-socmint-sulle-fonti-aperte

Di Porto, F. (2017, November 21). *Big data per i sevizi pubblici: tutti i vantaggi e i rischi*. Retrived from agendadigitale.eu: https://www.agendadigitale.eu/cittadinanza-digitale/big-data-per-i-servizi-pubblici-tutti-i-vantaggi-e-i-rischi/?__hstc=98839809.656156d732094db5bd6cb5604c6840f7.1652979544851.1652979544851.1652979544851.1&__hssc=98839809.2.1652979544852&__hsfp=4230998910&_ga=2.178499428.1903935846.1652979545-656034233.1652979545

Discovery Analytics Center. (2016, July 04). *EMBERS is a system for forecasting significant societal events from open source*. Retrieved from dac.cs.vt.edu: https://dac.cs.vt.edu/research-project/embers/

Discovery Analytics Center. (2014). *Case study: Forecasting the future: the EMBERS predictive analytics success*. Retrieved from basistech.com: https://www.basistech.com/wp-content/uploads/2017/09/EMBERS-Case_Study.pdf

Donato, A. (2020). *Big Data, Concorrenza, Privacy e loro interdipendenza.* Retrieved from webthesis.biblio.polito.it: https://webthesis.biblio.polito.it/16586/1/tesi.pdf

Duhigg, C. (2012, February 16). *How Companies Learn Your Secrets*. Retrieved from nytimes.com: https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=0

Easton, S. (2017, November 27). *Feds promise 'sector-by-sector' data rights, more data reforms in a few weeks*. Retrieved from themandarin.com.au: https://www.themandarin.com.au/86466-feds-promise-sector-by-sector-data-rights-more-data-reforms-in-a-few-weeks/

European Commission. (2015, July 20). *Special Eurobarometer 431: Data Protection.* Retrieved from data.europa.eu: https://data.europa.eu/data/datasets/s2075_83_1_431_eng?locale=en

European Commission Data Protection. (2017, October 27). *Guidelines on the right to "data portability" (wp242rev.01).* Retrieved from ec.europa.eu: https://ec.europa.eu/newsroom/article29/items/611233/en

European Data Protection Supervisor (2022). *AEPD-EPDS Joint Paper – 10 Misunderstandings about Machine Learning*. Retrived from edps.europa.eu: https://edps.europa.eu/data-protection/our-work/publications/papers/2022-09-20-aepd-edps-joint-paper-10-misunderstandings-about-machine-learning_en

Flynn, M., *Study on technical requirements for data spaces in law enforcement – Executive summary*, Publications Office, 2021. Retrived from data.europa.eu: https://data.europa.eu/doi/10.2837/853645

Facchini, A., & Termine, A. (2022, January 20). *Explainable AI: come andare oltre la black box degli algoritmi*. Retrieved from agendadigitale.eu: https://www.agendadigitale.eu/cultura-digitale/explainable-ai-come-andare-oltre-la-black-box-degli-algoritmi/

Favretto, G. (2020). *Big Data: tra criticità concorrenziali e prospettive di regolamentazione. Data brokers nel mirino dell'Antitrust*. Retrived from webthesis.biblio.polito.it: https://webthesis.biblio.polito.it/secure/13891/1/tesi.pdf

Federal Trade Commission. (2014, May 27). *Data Brokers: A Call For Transparency and Accountability: A Report of the Federal Trade Commission (May 2014).* Retrived from ftc.gov: https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014

Fingar, T. (2009). *Reducing Uncertainty: Intelligence and National Security Using Intelligence to Anticipate Opportunities and Shape the Future*. Lecture held at Stanford University, Stanford, CA. October 21, retrieved from http://iis-db.stanford.edu/evnts/5859/lecture_text.pdf

Firican, G. (2023). *The history of Big Data*. Retrieved from lightsondata.com: https://www.lightsondata.com/the-history-of-big-data/#:~:text=Some%20argue%20that%20it%20has,the%20O'Reilly%20Media%20group

Fisch, J. E.; Labouré, M., & Turner, J. A. (2018). The Emergence of the Robo-advisor. *Wharton Pension Research Council Working Papers. 10*. Retrived from repository.upenn.edu: https://repository.upenn.edu/prc_papers/10

Fischer, L. H., Wunderlich, N., & Baskerville, R. (2023).*Artificial Intelligence and Digital Work*. Paper presented at Hawaii International Conference on System Science 2023, Maui. Retrived from scholarspace.manoa.hawaii.edu: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/4fa6eece-c043-4ea4-bd59-8f49bd445878/content

Franca, S. (2021), *L'intreccio fra disciplina delle pratiche commerciali scorrette e normativa in tema di protezione dei dati personali: il caso Facebook approda al Consiglio di Stato*. Retrived from rivistadellaregolazionedeimercati.it/: http://www.rivistadellaregolazionedeimercati.it/Article/Archive/index_html?ida=236&idn=17&idi=-1&idu=-1

Franzini, A. (2019). *A framework for CBM implementation: from data acquisition to decision-making*. Retrived from politesi.polimi.it: https://www.politesi.polimi.it/handle/10589/145881

Franzini, M. (2019, June 02). *Il capitalismo della Sorveglianza secondo Shoshana Zuboff*. Retrived from eticaeconimia.it: https://eticaeconomia.it/il-capitalismo-della-sorveglianza-secondo-shoshana-zuboff/

Frazao, V., & Strachan, D. (2023, June 26). *A Busy Year for U.S. Privacy Laws*. Retrieved from verasafe.com: https://verasafe.com/blog/us-privacy-laws-coming-into-effect-in-2023/

Goodman, B., & Flaxman, S. (2016, June). EU regulations on algorithmic decision-making and a "right to explanation". In *ICML Workshop on Human Interpretability in Machine Learning (WHI 2016), New York, NY*. Retrived from metromemetics.net: http://metromemetics.net/wp-content/uploads/2016/07/1606.08813v1.pdf

Greggio, R. (2022). *Insuretech: le implicazioni e l'impatto dei Big Data nel mercato assicurativo*. Retrieved from thesis.unipd.it: https://thesis.unipd.it/handle/20.500.12608/35318

Goodfellow, I. (2015, July). *Deep Learning Adversarial Examples—Clarifying Misconceptions*. Retrived from kdnuggets.com: https://www.kdnuggets.com/2015/07/deep-learning-adversarial-examples-misconceptions.html

Grosser, B. (2014*). What do metrics want? How quantification prescribes social interaction on Facebook*. Retrived from computationalculture.net: http://computationalculture.net/what-do-metrics-want/

Haag, F., Hopf, K., Vasconcelos, P. M., & Staake, T. (2022). *Augmented cross-selling through explainable AI--a case from energy retailing*. Retrived from arXiv.org: https://arxiv.org/abs/2208.11404

Harford, T. (2014, March 28). *Big Data: Are We Making a Big Mistake?*. Retrived from ft.com: https://www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0

Hollin, R. (2015). *Drilling into the Big Data Gold Mine: Data Fusion and High-Performance Analytics for Intelligence Professionals*. Retrieved from oreilly.com: https://www.oreilly.com/library/view/application-of-big/9780128019672/XHTML/B9780128019672000021/B978012801967200021.xhtml

Home Affairs Select Committee. (2009). *The Macpherson Report - Ten Years On, Session 2008-09*, Retrived from publications.parliament.uk: https://publications.parliament.uk/pa/cm200809/cmselect/cmhaff/427/427.pdf

Iliadis, A., & Acker, A. (2022, August). *The seer and the seen: Surveying Palantir's surveillance platform*. Retrieved from researchgate.net: https://www.researchgate.net/publication/362882456_The_seer_and_the_seen_Surveying_Palantir%27s_surveillance_platform

International Organizations of Security Commissions. (2017, February). *IOSCO Research Report on Financial Technologies (Fintech).* Retrived from iosco.org.: https://www.iosco.org/library/pubdocs/pdf/IOSCOPD554.pdf

IAPP International Association of Privacy Professionals. (2023). *2023 US State Data Protection Laws – A Summary of Opt-Out Rights and Preference Signal Requirements*. Retrieved from

iapp.org: https://iapp.org/resources/article/thompson-hine-2023-state-data-laws-compliance-chart/

Iaselli, M. (2018, June 21). *Pseudonimizzazione*. Retrived from Altalex.com: https://www.altalex.com/documents/altalexpedia/2018/06/04/pseudonomizzazione

Jackson-Barnes, S. (2023, January 11). *The Evolution of Big Data*. Retrieved from orientsoftware.com: https://www.orientsoftware.com/blog/big-data-evolution/#1940s%20to%201989%20%E2%80%93%20Data%20Warehousing%20and%20Personal%20Desktop%20Computers

Jeffries, A. (2017, February 28). *J.C. Penney's troubles are reflected in satellite images of its parking lots*. Retrieved from theoutline.com: https://theoutline.com/post/1169/jc-penney-satellite-imaging

Jones, S. (2022, December 02). *Expanded CDR legislation to make online tasks safer and easier*. Retrieved from ministers.treasury.gov.au: https://ministers.treasury.gov.au/ministers/stephen-jones-2022/media-releases/expanded-cdr-legislation-make-online-tasks-safer-and

Katz, Y. (2017, November 27). *Manufacturing an Artificial Intelligence Revolution*. Retrived from ssrn.com: https://ssrn.com/abstract=3078224

Kearns, J. (2015, July 09). *Satellite Images Show Economies Growing and Shrinking in Real Time*. Retrieved from bloomberg.com: https://www.bloomberg.com/news/features/2015-07-08/satellite-images-show-economies-growing-and-shrinking-in-real-time

Kelly, E. (2022, September 2022). *Statutory Review of the Consumer Data Right - Report*. Retrieved from treasury.gov.au: https://treasury.gov.au/publication/p2022-314513

Kinsman, D., & Wong, T. A. (2023). *Proactive policing as reinforcement learning*. Retrived from openreview.net: https://openreview.net/forum?id=lmcPpHDa0B

Knight, W. (2023, March 06). *La Cina insidia il primato tecnologico degli Stati Uniti*. Retrieved from wired.it: https://www.wired.it/article/cina-dominio-tecnologico-stati-uniti/

Krebs, V. (2002). *Connecting the dots: tracking two identified terrorists*. Retrieved from orgnet.com: http://orgnet.com/tnet.html

Kulshrestha, S. (2016, January 22). *Big Data in Military Information & Intelligence*. Retrived from ssrn.com: https://ssrn.com/abstract=2765008

LeCates, R. (2018, October 17). *Intelligence-led Policing: Changing the Face of Crime Prevention*. *Police Chief*. Retrived from policechiefmagazine.org: https://www.policechiefmagazine.org/changing-the-face-crime-prevention/

Leonelli, S. (2018). *La ricerca scientifica nell'era dei Big Data*. Retrieved from philarchive.org: https://philarchive.org/archive/LEOLRS

Lev-Ram, M. (2016, March 09). *Palantir Connects the Dots with Big Data*. Retrieved from fortune.com: https://fortune.com/longform/palantir-big-data-analysis/

Lindsay, T.; Defence Science and Technology Organisation. (2016, December 07). *National Security and Intelligence, Surveillance and Reconnaissance Division*. Retrieved from dst.defence.gov.au: https://www.dst.defence.gov.au/sites/default/files/events/documents/National-Security-and-ISR-Division-presentation-PW2015.pdf

Lotan, G. (2014, August 04). *Israel, Gaza, War & Data*. Retrieved from medium.com: https://medium.com/i-data/israel-gaza-war-data-a54969aeb23e#.auy5s5t7s

Lundblad, C. T., Yang, Z., & Zhang, Q. (2022, September 23). *Detecting Insider Trading in the Era of Big Data and Machine Learning*. Retrived from ssrn.com: https://ssrn.com/abstract=4240205

Lyn, A. (2020, December 16) *Risky Business: Artificial Intelligence and Risk Assessments in Sentencing and Bail Procedures in the United States*. Retrived from ssrn.com: https://ssrn.com/abstract=3831441

Magnanelli, M. (2022). *La protezione dei dati personali in Europa. L'impatto del GDPR sull'azienda italiana TOD'S S.p.A.* Retrieved from tesi.univpm.it: https://tesi.univpm.it/handle/20.500.12075/10160

Malick, R. (2023, August 24). *90% of your data is unstructured — and it's full of untapped value*. Retrieved from blog.box.com: https://blog.box.com/90-your-data-unstructured-and-its-full-untapped-value

Marcelli, S. (2021, February, 11). *Il valore dei dati: prospettive e caratteristiche dei big data*. Retrived from iusinitinere.it: https://www.iusinitinere.it/il-valore-dei-dati-prospettive-e-caratteristiche-dei-big-data-35418

Martorana M., & Pinelli, L. (2021, June 08). *Dati personali: anonimizzazione e pseudonimizzazione*. Retrived from altalex.com: https://www.altalex.com/documents/news/2021/06/08/dati-personali-anonimizzazione-e-pseudonimizzazione

Matshazi, N. (2018, October 18). *Google AI's LYNA Better Than Humans In Detecting Advanced Breast Cancer*. Retrived from healthcareweekly.com: https://healthcareweekly.com/google-ais-lyna-better-than-humans-in-detecting-advanced-breast-cancer/

Mikalef, P., Pappas, I., Krogstie, J., & Pavlou, P. (2019*). Big data and business analytics: A research agenda for realizing business value*. Retrived from ntnuopen.ntnu.no: https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2644384/Mikalef.pdf?sequence=4

Miller, C. C. (2015, June 26). *When Algorithms Discriminate*. Retrived from nytimes.com: https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html

Miller, P. E. (2005). *How can we improve information sharing among local law enforcement agencies?* (Doctoral dissertation, Monterey California. Naval Postgraduate School). Retrived from apps.dtic.mil: https://apps.dtic.mil/sti/citations/ADA439576

Mischitelli, L. (2020, January 28). *Ricerca scientifica e protezione dati: le raccomandazioni dell'EDPS*. Retrieved from agendadigitale.eu: https://www.agendadigitale.eu/sicurezza/ricerca-scientifica-e-protezione-dati-le-raccomandazioni-delledps/

Moazed, A. (2019). *How GDPR is Helping Big Tech and Hurting the Competition*. Retrieved from applicoinc.com: https://www.applicoinc.com/blog/how-gdpr-is-helping-big-tech-and-hurting-the-competition/

Monahan, T., & Regan, P. (2014, June 27). *Zones of Opacity: Data Fusion in Post-9/11 Security Organizations*. Retrieved from cambridge.org: https://www.cambridge.org/core/journals/canadian-journal-of-law-and-society-la-revue-canadienne-droit-et-societe/article/abs/zones-of-opacity-data-fusion-in-post911-security-organizations/DE3A642B0DC42922D19BBACD35DD0B73

Morana, S., Gnewuch, U., Jung, D., & Granig, C. (2020). *The Effect of Anthropomorphism on Investment Decision-Making with Robo-Advisor Chatbots*. Retrived from researchgate.net: https://www.researchgate.net/publication/341277570_The_Effect_of_Anthropomorphism_on_Investment_Decision-Making_with_Robo-Advisor_Chatbots

Munn, L. (2017). Seeing with software: Palantir and the regulation of life. *Studies in Control Societies*, *2*(1). Retrieved from studiesincontrolsocieties.org: https://studiesincontrolsocieties.org/seeing-with-software/

Napoli, P. M. (2013). The algorithm as institution: Toward a theoretical framework for automated media production and consumption. *Fordham University Schools of Business Research Paper*. Retrived from papers.ssrn.com: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2260923

Naranjo, D., & Molnar, P. (2020, February 24). *The Privatization of Migration Control*. Retrieved from cigionline.org: https://www.cigionline.org/articles/privatization-migration-control/

National Commission on Terrorist Attacks upon the United States, Washington DC. (2004, July 22). *Final report of the National Commission on Terrorist Attacks upon the United States*. Retrieved from 9-11commission.gov: https://www.9-11commission.gov/report/911Report.pdf

Neslen, A. (2021). *Feature - Pushback against AI policing in Europe heats up over racism fears*. Retrived from reuters.com: https://www.reuters.com/article/europe-tech-police-idINL8N2R92HQ

New York State Intelligence Center . (2008, September 03). *New York State Law Enforcement Terrorism Indicators Reference Card*. Retrieved from publicintelligence.net: https://publicintelligence.net/new-york-state-law-enforcement-terrorism-indicators-reference-card/

Nie, S., & Sun, D (2016, March). *Research on counter-terrorism based on big data. 2016 IEEE International Conference on Big Data Analysis (ICBDA)*. Retrived from researchgate.net: https://www.researchgate.net/publication/305333459_Research_on_counter-terrorism_based_on_big_data

Nurkin, T., & Konaev, M. (2022, May 25). *Eye to eye in AI: Developing artificial intelligence for national security and defense*. Retrieved from atlanticcouncil.org: https://www.atlanticcouncil.org/in-depth-research-reports/report/eye-to-eye-in-ai/

Obolentsev, V. F., & Yushchenko, O. G. (2019). *Application of artificial intelligence methods in law*. Retrived from repository.kpi.kharkov.ua: https://repository.kpi.kharkov.ua/server/api/core/bitstreams/1b819b19-7d8a-424c-91f5-15aa338ee812/content

OECD Organisation for Economic Co-operation and Development. (2020, June 10-12). *Consumer Data Rights and Competition - Background note*. Retrieved from one.oecd.org: https://one.oecd.org/document/DAF/COMP(2020)1/en/pdf

OECD Organisation for Economic Cooperation and Development. (2019, November 26). *Enhancing Access to and Sharing of Data*. Retrieved from oecd.org: https://www.oecd.org/publications/enhancing-access-to-and-sharing-of-data-276aaca8-en.htm

OECD Organisation for Economic Co-operation and Development. (2013, July 11). *Recommendation of the Council concerning Guidelines governing the Protection of Privacy and Transborder Flows of Personal Data (2013)*. Retrieved from oecd.org: https://www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf

O'Hara, K., & Bergin, A. (2009, November 27). *Information Sharing in Australia's National Security Community*. Retrieved from jstor.org/stable: https://www.jstor.org/stable/resrep03940#metadata_info_tab_contents

Ottonelli, M. (2020). *GDPR: impatto, applicazione e prospettive di miglioramento, con l'analisi un business case*. Retrieved from webthesis.biblio.polito.it: https://webthesis.biblio.polito.it/13898/

Passani, L. (2018, April 08). *Cambridge Analytica: come ti "hackero" la Democrazia: L'intersezione di Social Network, Intelligenza Artificiale (IA) e psicologia comportamentale è perfettamente in grado di costruire la macchina acchiappa anime*. Retrived from lavocedinewyork.com: https://lavocedinewyork.com/news/primo-piano/2018/04/08/cambridge-analytica-come-ti-hackero-la-democrazia/

Patidar, A. (2023). *Data Privacy Laws for 2023: A Closer Look at 9 Key Regulations*. Retrieved from loginradius.com: https://www.loginradius.com/blog/identity/stay-compliant-with-data-privacy-laws-2023/

Pennasilico, A. (2018, November 16). *I data broker e il vero prezzo dei nostri dati: che c'è da sapere*. Retrived from agendadigitale.eu: https://www.agendadigitale.eu/sicurezza/privacy/i-data-broker-e-il-vero-prezzo-dei-nostri-dati-che-ce-da-sapere/

Peterson, M. (2005). *Intelligence-led Policing: The New Intelligence Architecture*. Washington, DC: Bureau of Justice Assistance. Retrived from ojp.gov: https://www.ojp.gov/ncjrs/virtual-library/abstracts/intelligence-led-policing-new-intelligence-architecture

Phillips, A. (2021, April 01). *A history and timeline of big data*. Retrieved from techtarget.com: https://www.techtarget.com/whatis/feature/A-history-and-timeline-of-big-data

Philipson, G. (2017, November 27). *Consumers to own their own data, with new bill*. Retrieved from governmentnews.com.au: https://www.governmentnews.com.au/consumers-data-new-bill/

Piretti, V. (2020, May 12). *Indagine conoscitiva sui Big Data: conclusioni e prospettive*. Retrieved from diritto.it: https://www.diritto.it/indagine-conoscitiva-sui-big-data-conclusioni-e-prospettive/

Porche III, I., Wilson, B., Johnson, E., Tierney, S., & Saltzman, E. (2014). *Data Flood: Helping the Navy Address the Rising Tide of Sensor Information*. Retrieved from rand.org: https://www.rand.org/pubs/research_reports/RR315.html

Power D.J. (2014) *What is data-driven decision making?*. Retrived from dssresources.com: http://dssresources.com/faq/index.php?action=artikel&id=314

Prat, A., & Valletti, T. (2021, May 25). *Attention Oligopoly*. Retrieved from papers.ssrn.com: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3197930

Prestipino, D. (2017, July 15*). Nuovi scenari di rischio per la protezione dei dati personali in contesti data intensive (New Scenarios of Risk for Personal Data Protection in Data-intensive Contexts)* Institutional Doctoral Theses Repository by AlmaDL University of Bologna Digital Library. Retrived from amsdottorato.unibo.it: https://amsdottorato.unibo.it/8248/

Privacy International. (2020, September). *All roads lead to Palantir: A review of how the data analytics company has embedded itself throughout the UK*. Retrieved from privacyinternational.org: https://privacyinternational.org/sites/default/files/2020-11/All%20roads%20lead%20to%20Palantir%20with%20Palantir%20response%20v3.pdf

Rabasca Roepe, L. (2019, July 17). *How Big Data Helped Chicago Improve Its Food Safety*. Retrived from dell.com: https://www.dell.com/en-us/perspectives/how-big-data-helped-chicago-improve-its-food-safety/

Radden Keefe, P. (2006, March 12). *Can network theory thwart terrorists?*. Retrived from nytimes.com: https://www.nytimes.com/2006/03/12/magazine/can-network-theory-thwart-terrorists.html

Rainie, L., & Anderson, J. (2012, July 20). *Big Data: Experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well*. Retrieved from academia.edu: https://www.academia.edu/9994874/Big_Data_Experts_say_new_forms_of_information_analysis_will_help_people_be_more_nimble_and_adaptive_but_worry_over_humans_capacity_to_understand_and_use_these_new_tools_well

Risen, J., & Lichtblau, E. (2013, June 08). *How the US uses technology to mine more data more quickly*. Retrived from nytimes.com: https://www.nytimes.com/2013/06/09/us/revelations-give-look-at-spy-agencys-wider-reach.html

Riservato, E., & Scaini, S. (s.d.). *OSINT - Intelligence da fonti aperte: un'incessante evouzione dalle origini ai giorni nostri*. Retrived from safetysecuritymagazine.com:

https://www.safetysecuritymagazine.com/articoli/osint-intelligence-da-fonti-aperte-unincessante-evoluzione-dalle-origini-ai-giorni-nostri/

Rodriguez, J., & Nayor, D. (2022). The Balance of Embracing Technological Advancements in Law Enforcement. *Contemporary Issues in LE & Public Safety Leadership*. Retrived from julierodriguez.org: https://www.julierodriguez.org/uploads/1/4/5/7/145767614/the_balance_of_embracing_technological_advancements_in_law_enforcement.pdf

Roff, H. (2020, November). Uncomfortable ground truths: Predictive analytics and national security. *Brookings National Security Report*. Retrived from brookings.edu: https://www.brookings.edu/articles/uncomfortable-ground-truths/

Russel, K. (2017, February 23). *SpaceKnow CEO: Why Economists Should Look to the Skies*. Retrieved from satellitetoday.com: https://www.satellitetoday.com/innovation/2017/02/23/spaceknow-ceo-economists-look-skies/

Saetta, B. (2018, September 07). *Le sanzioni in materia di protezione dei dati personali*. Retrieved from protezionedatipersonali.it: https://protezionedatipersonali.it/sanzioni-protezione-dati-personali

Saini, J.K. (2023, October 16). *LSTM based deep learning approach to detect online violent activities over dark web*. Retrived from link.springer.com: https://link.springer.com/article/10.1007/s11042-023-17222-8

Schneier, B. (2013, May 07). *Intelligence Analysis and the Connect-the-Dots Metaphor*. Retrieved from schneier.com: https://www.schneier.com/blog/archives/2013/05/intelligence_an.html

Schwartz, M. (2015, January 26). *The whole haystack*. Retrived from newyorker.com: https://www.newyorker.com/magazine/2015/01/26/whole-haystack

Società Italiana di Antropologia Culturale. (2021, February 04). *General Data Protection Regulation in Antropologia - Seminari 2020/2021*. Retrieved from siacantropologia.it: https://www.siacantropologia.it/appuntamenti/general-data-protection-regulation-in-antropologia/

Soomro, H. (2023, January 32). *Mastering the 10 Vs of Big Data* . Retrieved from datasciencedojo.com: https://datasciencedojo.com/blog/10-vs-of-big-data/

Stato Maggiore della Difesa. (2022). *L'impatto delle Emerging & Disruptive Technologies (EDTs) sulla Difesa*. Retrieved from difesa.it: https://www.difesa.it/SMD_/Staff/Sottocapo/UGID/Dottrina/Documents/Concetto_Impatto_delle_EDT_sulla_Difesa_Ed_2022.pdf

Street, L., Brady, M., & Moroney, K. (2008, March 14). *Street review: A Review of Interoperability Between the AFP and Its National Security Partners*. Retrieved from ojp.gov: https://www.ojp.gov/ncjrs/virtual-library/abstracts/street-review-review-interoperability-between-afp-and-its-national

Symon, S., & Tarapore, A. (2015). *Defense Intelligence Analysis in the Age of Big Data*. Retrieved from semanticscholar.org: https://www.semanticscholar.org/paper/Defense-Intelligence-Analysis-in-the-Age-of-Big-Symon-Tarapore/6ea57c140c27a09b4a5d90c2d7deaec07aae7fd8

Takatsuki, Y. (2019, March 28). *CCPA Blog Series, Part 2: Rethinking access and data portability rights*. Retrived from fieldfisher.com: https://www.fieldfisher.com/en/services/privacy-security-and-information/privacy-security-and-information-law-blog/ccpa-blog-series-part-2-rethinking-access-and-data-portability-rights

Technical Committee ISO/IEC JTC 1/SC 27. (2018, February). *ISO/IEC 27000:2018 - Information technology — Security techniques — Information security management systems — Overview and vocabulary*. Retrieved from iso.org: https://www.iso.org/obp/ui/en/#iso:std:iso-iec:27000:ed-5:v1:en

The White House, Executive Office of the President. (2016). Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights. Retrived from obamawhitehouse.archives.gov: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

The White House (2016). Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights. Executive Office of the President. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

The White House (2014). Big Data: Seizing opportunities, preserving values. Executive Office of the President. https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_5.1.14_final_print.pdf

Thompson, N., Evans, C., & Armbrust, D. (2023, February 28). *America's lead in advanced computing is almost gone*. Retrieved from gppreview.com: https://gppreview.com/2023/02/28/americas-lead-in-advanced-computing-is-almost-gonepart-1-systems-and-capabilities/

Vallone, D. (2022). *Big data e Privacy= Big data and Privacy*. Retrived from webthesis.biblio.polito.it: https://webthesis.biblio.polito.it/25363/

Vance, A., & Stone, B. (2011, November 22). *Palantir, the War on Terror's secret weapon*. Retrived from Bloomberg.com: https://www.bloomberg.com/news/articles/2011-11-22/palantir-the-war-on-terrors-secret-weapon#xj4y7vzkg

Visentin, C. (April 2019). *Il potere razionale degli algoritmi tra burocrazia e nuovi idealtipi*. Retrived from researchgate.net: https://www.researchgate.net/publication/332254383_Il_potere_razionale_degli_algoritmi_tra_burocrazia_e_nuovi_idealtipi

Wang, J. C., & Perkins, C. B. (2019). *How Magic a Bullet is Machine Learning For Credit Analysis? An Exploration with Fintech Lending Data.* CEAR/CenFIS Conference, Oct 30, 2019. Retrived from pdfs.semanticscholar.org: https://pdfs.semanticscholar.org/21a1/14f037fdd3a27d386734fc71a3744bcb3562.pdf

Wattenberg, M. Viégas, F. & Hardt, M. (2017). *Attacking discrimination with smarter machine learning*. Retrived from research.google.com: https://research.google.com/bigpicture/attacking-discrimination-in-ml/

Weisel, D. L. (2005). *Analyzing Repeat Victimization* (pp. 1-80). Washington, DC: US Department of Justice, Office of Community Oriented Policing Services. Retrived from portal.educoas.org: https://portal.educoas.org/sites/default/files/nw/wg/docs/Analyzing-Repeat-Victimization.pdf

Winston, A. (2018). *Palantir has secretly been using New Orleans to test its predictive policing technology*. Retrived from theverge.com: https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd

Zanasi, A. (2003). *Information Warfare, Business Intelligence, Text Mining*. Retrieved from reserchgate.net: https://www.researchgate.net/profile/Alessandro-Zanasi/publication/237680074_Information_Warfare_Business_Intelligence_Text_Mining1/links/5592ad3708aed7453d46364e/Information-Warfare-Business-Intelligence-Text-Mining1.pdf

Sommario

ISBN: 978-88-6086-253-2
Pubblicato in digitale da Aiep Editore, Repubblica di San Marino, Ottobre 2024