

PRIVACY E INNOVAZIONE

STUDI E RICERCHE SULLA PROTEZIONE
DEI DATI PERSONALI NELL'ERA DIGITALE

Direttore

Riccardo ACCIAI

Garante per la protezione dei dati personali

Comitato scientifico

César ALONSO-IRIARTE

Commissione europea

Sauro ANGELETTI

Presidenza del Consiglio dei Ministri

Luigi CANNADA-BARTOLI

Avvocato del Foro di Roma

Daniele DE PAOLI

Garante per la protezione dei dati personali

Federico FERRO-LUZZI

Università degli Studi di Sassari

Fabio GIGLIONI

Sapienza – Università di Roma

Sergio LARICCIA

Sapienza – Università di Roma

Stefano LEONARDI

Sapienza – Università di Roma

Daniele PERUCCHINI

Fondazione Ugo Bordoni

Marilena VENDITTELLI

Sapienza – Università di Roma

Andrea VITALETTI

Sapienza – Università di Roma

PRIVACY E INNOVAZIONE

STUDI E RICERCHE SULLA PROTEZIONE DEI DATI PERSONALI NELL'ERA DIGITALE



La collana ospita i risultati delle attività di studio e ricerca avviate o promosse dal Centro studi privacy e nuove tecnologie: monografie tematiche, paper e position paper che seguono l'evoluzione del tema della protezione dei dati personali in una prospettiva multidisciplinare (giuridica, politico-sociale, tecnologica) proponendo chiavi di lettura innovative. Sono inseriti in collana anche gli atti dei convegni organizzati dal Centro studi, gli interventi e i documenti presentati dai membri in occasione della partecipazione, in qualità di relatori, a conferenze e convegni. In "Privacy e Innovazione" sono pubblicate opere di alto livello scientifico, anche in lingua straniera, per facilitarne la diffusione internazionale.

Il direttore approva le opere e le sottopone alla revisione paritaria con il sistema del "doppio cieco" (*double blind peer review*) nel rispetto dell'anonimato sia dell'autore, sia dei due revisori che sceglie: l'uno da un elenco deliberato dal comitato scientifico, l'altro dallo stesso comitato in funzione di revisore interno. I revisori rivestono o devono aver rivestito la qualifica di professore universitario di prima fascia nelle università italiane o una qualifica equivalente nelle università straniere.

Ciascun revisore formulerà una delle seguenti valutazioni:

- pubblicabile senza modifiche;
- pubblicabile previo apporto di modifiche;
- da rivedere in maniera sostanziale;
- da rigettare;

tenendo conto della: a) significatività del tema nell'ambito disciplinare prescelto e originalità dell'opera; b) rilevanza scientifica nel panorama nazionale e internazionale; c) attenzione adeguata alla dottrina e all'apparato critico; d) adeguato aggiornamento normativo e giurisprudenziale; e) rigore metodologico; f) proprietà di linguaggio e fluidità del testo; g) uniformità dei criteri redazionali. Nel caso di giudizio discordante fra i due revisori, la decisione finale è assunta dal direttore, salvo casi particolari in cui questi provveda a nominare tempestivamente un terzo revisore a cui rimettere la valutazione dell'elaborato. Il termine per la valutazione non deve superare i venti giorni, decorsi i quali il direttore della collana, in assenza di osservazioni negative, ritiene approvata la proposta. Sono escluse dalla valutazione gli atti di convegno, le opere dei membri del comitato scientifico e le opere collettive di provenienza accademica. Il direttore, su sua responsabilità, può decidere di non assoggettare a revisione scritti pubblicati su invito o comunque di autori di particolare prestigio.

Paradisi artificiali

La nuova frontiera dei dati sintetici
tra diritto e tecnologia

a cura di

Sergio Aracu
Carlo Rossi Chauvenet
Lorenzo Cristofaro

Prefazione di

Luca Bolognini

Contributi di

Riccardo Acciai
Sergio Aracu
Luca Bolognini
Carlo Rossi Chauvenet
Giulio Coraggio
Lorenzo Cristofaro
Giuseppe D'Acquisto
Gabriele Franco
Raffaella Grisafi
Shalini Kurapati
Aldo Lamberti
Alessandra Nicolosi
Daniele Panfilo





ISBN
979-12-218-1455-2

PRIMA EDIZIONE
ROMA 30 SETTEMBRE 2024

- 11 *Prefazione*
Luca Bolognini
- 13 *Introduzione*
Riccardo Acciai
- 17 *Indice delle abbreviazioni*
- 19 Verso una definizione di dati sintetici
Lorenzo Cristofaro e Gabriele Franco
1. La rivoluzione dei dati sintetici, 19 – 2. Una definizione tecnica di dati sintetici, 23 – 3. Una definizione “giuridica” di dati sintetici, 29 – 3.1. Premessa, 29 – 3.2. I dati sintetici nel GDPR: tra dati personali e dati anonimi, 30 – 3.3. I dati sintetici nell’Artificial Intelligence Act, 33 – 4. Prospettive future, 35
- 39 Paradisi artificiali: stato dell’arte
Carlo Rossi Chauvenet
1. Sintetico quindi anonimo? Stato dell’arte sull’utilizzabilità di un nuovo paradigma di anonimizzazione e condivisione del dato personale, 39 – 2. Alcuni cenni in tema di classificazione dei dati sintetici e profili tecnici, 43 – 3. L’utilizzabilità dei dati sintetici ai sensi della attuale prassi interpretativa in tema di anonimizzazione, 46 – 4. Nuove norme europee in tema di dati sintetici ed utilizzabilità come PET, 51 – 5. Alcuni recenti interventi delle autorità garanti in tema di dati sintetici, 55 – 6. Nuovo paradigma di condivisione del dato nel contesto europeo e temi aperti, 57
- 63 Synthetic data and data protection laws
Giuseppe D’Acquisto
1. Synthetic data: what they are, 63 – 2. Why they are gaining attention, 64 – 3. How are synthetic data generated?, 67 – 4. Where can synthetic data be used?, 71 – 5. What kind of insight on real world can we get from synthetic data?, 73 – 6. Data protection considerations on the use of synthetic data: are synthetic data anonymous data?, 76 – 7. Possible compliance options, 79 – 8. A new regulatory urgency: dealing with stochasticity, 82

85 **Dati sintetici strutturati: metodologie di generazione e verifica della privacy***Daniele Panfilo*

1. Premessa, 85 – 2. Panoramica sulle diverse metodologie per la produzione di dati sintetici, 86 – 3. Modelli di generazione dei dati sintetici basati su simulazioni, 87 – 3.1. Tipi di modelli di dati sintetici basati sulla simulazione, 87 – 3.2. Discussione, 90 – 4. Modelli generativi basati sull'intelligenza artificiale, 92 – 4.1. Tipologie di modelli generativi basati sull'intelligenza artificiale, 92 – 4.2. Modelli basati su tecniche di apprendimento profondo, 94 – 4.3. Discussione, 98 – 5. Hybrid Synthetic Data Models, 102 – 6. Metriche di controllo e verifica della privacy, 102 – 7. Conclusioni, 108

111 **I dati sintetici per mettere il dato al centro dell'IA e per la compliance***Shalini Kurapati e Alessandra Nicolosi*

1. Nel gioco dell'IA vince chi ha i dati, 111 – 2. I dati sintetici sono anonimi?, 113 – 3. Non solo privacy, 115 – 4. Il dato sintetico è FAIR & Open, 117 – 5. Il Clearbox AI: il lavoro nel campo dei dati sintetici, 117

121 **Technical Standard on Synthetic Data***Aldo Lamberti e Giulio Coraggio*

1. The IEEE's Role in Setting the Global Technical Standard on Synthetic Data, 121 – 2. Technical Aspects of Synthetic Data, 123 – 2.1. Anonymization under GDPR, 123 – 2.2. Differentially Private Synthetic Data, 124 – 2.3. Privacy Metrics for Synthetic Data, 126 – 2.4. Recommendations, 127 – 3. Use Cases of Synthetic Data, 128 – 4. Future Projections, 135 – 4.1. Looking ahead, 135 – 4.2. Open Areas of Research and Ethical Concerns, 136

139 **Prime applicazioni pratiche e potenzialità***Sergio Aracu*

1. Inquadramento, 139 – 1.1. I possibili ambiti di utilizzo dei dati sintetici: il marketing, 139 – 1.2. L'attuale ruolo e importanza dei così detti Big Data, 141 – 1.3. La posizione dell'Unione Europea, 144 – 2. Come l'utilizzo di tecnologie in grado di generare dati sintetici può contribuire significativamente a riequilibrare il mercato, 146 – 3. Ulteriori vantaggi dell'utilizzo dei dati sintetici in ambito marketing per imprese e consumatori, 147 – 3.1. Profili privacyistici, 147 – 4. Conclusioni: i dati sintetici sono la soluzione?, 152

155 **Il dato sintetico nell'approccio ermeneutico funzionale del diritto dei consumatori***Raffaella Grisafi*

1. Profili introduttivi, 155 – 2. Dalla tutela del dato personale a quella del dato sintetico, 159 – 3. La relatività del dato sintetico, 164 – 4. Il dato sintetico nella prospettiva della disciplina consumerista. L'approccio ermeneutico funzionale, 167

171 L'arte giuridica dei dati sintetici

Luca Bolognini

1. Una relazione tra dati sintetici, arte e diritto, 171 – 2. I dati sintetici resi sufficientemente anonimi, a valle, 173 – 3. Il trattamento di dati personali propedeutico alla generazione dei dati sintetici, a monte, 174 – 4. Le basi giuridiche del trattamento di dati personali, a monte, 176 – 5. Sintetizzazione come pubblicazione o come intermediazione e condivisione dei valori-dati: la strada dei codici di condotta, 178

181 Bibliografia

(a cura di Daniele Panfilo)

185 Gli Autori

Prefazione

di LUCA BOLOGNINI*

Una “gemma sintetica”. Così definirei quest’opera, per due ragioni facilmente intuibili: essa affronta nel merito la materia dei dati sintetici, appunto, e lo fa con una forma prismatica, colorata ma trasparente, brillante, a tratti illuminante. Le Autrici e gli Autori indagano il medesimo macro-argomento – che rappresenta insieme un enigma e una scommessa sul piano giuridico-tecnico – e ne colgono aspetti salienti, seguendo sensibilità, esperienze e competenze originali.

La sensazione è quella di una plurale “variazione solista sul tema”, non di un coro conformista, cioè di un’operazione epistemologica, oggi, più che mai necessaria e di rara onestà dottrinale, ma anche foriera di valore estetico per il lettore. Ottiche e prospettive cangianti, mutevoli come le opportunità e le conseguenze della generazione di dati che non sono reali ma realistici, che non sono veri ma veritieri, che non sono umani ma ricalcati sull’umanità. Tutto ciò, mai scadendo in eccessi di ultra-relativismo dei diritti.

Molteplici le sfaccettate tematiche considerate in questo libro, grazie a illustri esperti provenienti dalla professione, dalla ricerca, dalle autorità, dal mercato: tra esse, le tecniche di sintesi, di pseudonimizzazione e di anonimizzazione dei dati, i modelli di valorizzazione dei dati percorribili per scopi pubblicistici o privatistici, le basi giuridiche e le condizioni di liceità dei trattamenti, la delicata questione dei *bias* (pregiudizi) e dei connessi correttivi. Dato sintetico come problema-

*Avvocato cassazionista del foro di Roma, socio fondatore dello studio legale internazionale ICT Legal Consulting, Presidente dell’Istituto italiano per la privacy e la valorizzazione dei dati (IIP).

inganno, da segnalare e smascherare (il *deep-fake* frutto dell'Intelligenza Artificiale) o, piuttosto, come strumento-soluzione per minimizzare l'invasione nella privacy degli individui e il trattamento dei dati personali, senza rinunciare a innovazione, sviluppo e condivisione della conoscenza. La certezza e l'incertezza del diritto rimangono – quasi alla stregua di avvertimenti di rischio per l'interprete – in filigrana e all'orizzonte, riga dopo riga, pagina dopo pagina.

Gli Autori e le Autrici ci trasmettono un sincero e umanissimo (non sintetico!) senso critico: non assumono supinamente le *relata* degli entusiasti o dei detrattori, ma setacciano, osservano – al microscopio legale – ogni elemento d'interesse per lo studioso teorico e per il pratico, e ne problematizzano la statica e la dinamica, talvolta risolvendone nodi o apparenti contraddizioni. Probabilmente, il fatto stesso che le idee qui espresse derivino da un pregresso lavoro di compartecipazione e discussione – nel corso del convegno “Paradisi Artificiali” (con l'allusivo, ironico richiamo dei *paradis artificiels* di Baudelaire) promosso nell'estate del 2023 dall'autorevole *Centro Studi Privacy e Nuove Tecnologie* – ha reso possibile un'analisi aperta, organica e robusta di un argomento tanto complesso, complicato e inesplorato.

Nel *presente* volume, per *sintetizzare*, troverete il *futuro* della sostenibilità *data-driven*. Come è ormai pacifico e diffuso il consenso scientifico e sociale – fondato e consapevole – con riferimento alla raccolta differenziata e al riciclo dei rifiuti, così la collezione, la trasformazione e la valorizzazione dei dati (pur in origine personali) rappresentano il destino info-ecologico dei nostri Paesi, delle nostre istituzioni e industrie. Un destino che può rivelarsi amico o nemico dei diritti e delle libertà fondamentali, a seconda dei bilanciamenti che verranno via via calcolati e scelti. Ragionare di dati sintetici ora significa, dopotutto, elaborare e indirizzare forme di “neo-ambientalismo digitale”, in grado di coniugare assieme etica, rispetto degli esseri umani, incentivo all'innovazione e alla ricerca. Questa pubblicazione è proprio una gemma pionieristica, dunque, e costituisce un *unicum* nel panorama editoriale giuridico italiano, di cui fare buon tesoro.

Introduzione

di RICCARDO ACCIAI¹

Questo volume, contiene – in forma parzialmente rielaborata ed aggiornata – gli interventi svolti in occasione del convegno “*Paradisi artificiali. La nuova frontiera dei dati sintetici tra diritto e tecnologia*”, svoltosi il 21 giugno 2023 nella splendida cornice di Palazzo Falletti a Roma.

L’ineccepibile organizzazione di quell’evento si deve ai nostri tre curatori del volume, Sergio Aracu, Carlo Rossi Chauvenet e Lorenzo Cristofaro, con il sostegno delle relative realtà di appartenenza, ossia Area Legale s.r.l., CRCLEX e Panetta Studio Legale.

OIC – Osservatorio Imprese e Consumatori e Centro Studi privacy e nuove tecnologie, avevano fornito il loro patrocinio gratuito all’iniziativa.

Quest’ultimo, in particolare, ha risposto con convinzione ed entusiasmo all’ipotesi prospettata dagli organizzatori di pubblicare gli atti del convegno e ha affidato a chi scrive, che già aveva svolto il ruolo di moderatore della prima parte dei lavori di giugno, il compito di curare queste brevi pagine introduttive.

In primo luogo, credo sia importante provare a riferire dell’atmosfera di grande attenzione ed interesse registrata in occasione del convegno: le relazioni sono state tutte – come si spera risulterà dalla lettura delle pagine che seguono – di alto spessore e seguite da un pubblico numeroso e competente che non ha fatto mancare interventi e domande.

¹ Dirigente del Garante per la protezione dei dati personali. Direttore del Centro studi privacy e nuove tecnologie. Ogni opinione è qui espressa a titolo personale e non impegna in alcun modo l’amministrazione di appartenenza.

Davvero una rara occasione di confronto su un tema, ancora tutto da esplorare e di cui, probabilmente, il convegno in questione ha costituito un primo importante banco di prova nel nostro Paese.

Una ragione in più per dar merito a chi lo ha pensato e poi organizzato.

Questi atti, anche per via del tentativo di amalgamare contributi provenienti da esperienze diverse, giungono con un po' di ritardo rispetto all'evento, ma questo consente di presentare dei testi aggiornati alla più recente evoluzione normativa, a partire dall'intervenuta approvazione e pubblicazione del Regolamento (UE) sull'intelligenza artificiale, che – come si vedrà – reca interessanti spunti per l'inquadramento dei dati sintetici, e che al momento del convegno era presa in considerazione solo come “proposta” della Commissione europea.

Il lettore troverà nel volume testi in italiano e in inglese: assecondando un *trend* in atto ormai da alcuni anni, si è preferito non uniformare il linguaggio al fine di lasciare l'originalità dell'espressione degli Autori e ponendo contestualmente le basi per una più ampia diffusione del lavoro.

Il testo si apre con il contributo di Lorenzo Cristoforo e Gabriele Franco, che mira a confrontare le varie definizioni sin qui elaborate sui dati sintetici al fine di verificarne la tenuta e valutare l'esigenza di ulteriori precisazioni.

Un tema questo propedeutico per molti degli altri approfondimenti, a partire da quello di Carlo Rossi Chauvenet, che indaga sull'attuale “stato dell'arte”, alla luce dell'evoluzione normativa e degli interventi delle autorità di protezione dati.

Analogo percorso compie Giuseppe D'Acquisto, ma con un taglio decisamente tecnologico, incentrato sugli aspetti di protezione dei dati e, in particolare, sulla equiparabilità dei dati sintetici alla nozione di dato anonimo, come derivante dalla lettura GDPR.

Sempre di carattere tecnico è l'analisi di Daniele Panfilo che, nel quarto capitolo, si sofferma sulle metodologie per la produzione di dati sintetici e sulle metriche di controllo e verifica relativamente agli “aspetti privacy”.

Aldo Lamberti e Giulio Coraggio, poi, si soffermano sul ruolo che possono svolgere gli standard tecnici per i dati sintetici e delineano alcune linee di sviluppo, alla luce anche delle preoccupazioni di carattere etico che queste possono comportare.

Forti della loro esperienza sul campo, Shalini Kurapati e Alessandra Nicolosi analizzano le interazioni tra intelligenza artificiale e i “nuovi dati”, evidenziando il valore strategico delle informazioni per lo sviluppo dell’A.I.

Negli ultimi capitoli vengono analizzati i primi possibili effetti pratici derivanti dall’uso dei dati sintetici: Sergio Aracu, in particolare, si sofferma sulla loro utilizzabilità nell’ambito del marketing e, più in generale, Raffaella Grisafi, con un’approfondita analisi, si interroga sui nuovi scenari che si possono realizzare per i consumatori.

Al volume si aggiunge anche il prezioso contributo di Luca Bolognini, che, impossibilitato a suo tempo a partecipare al convegno, aveva voluto far giungere qualche spunto di riflessione, che ora troviamo nell’ultimo capitolo del volume e nelle gentili parole di presentazione dello stesso.

Non resta quindi che ringraziare tutti coloro che hanno contribuito, dapprima all’ottima riuscita del convegno e, ora, alla pubblicazione di questi atti, augurando buona lettura a coloro che vorranno soffermarsi sulle pagine seguenti.

Roma, 5 Agosto 2024

Indice delle abbreviazioni

Vengono riportate qui di seguito le abbreviazioni più frequentemente utilizzate nel volume:

AgCom	Autorità per le garanzie nelle comunicazioni
AGCM	Autorità garante per la concorrenza ed il mercato
AI Act	Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale).
art./artt.	articolo / articoli
cap.	capitolo / capitoli
Cass. Civ.	Corte di Cassazione – Sezione civile
Cass. Pen.	Corte di Cassazione – Sezione penale
CGUE	Corte di Giustizia dell'Unione Europea
Corte cost.	Corte costituzionale
c.d.	cosiddetto/
cit.	citato
cfr.	confronta
Cons. Stato.	Consiglio di Stato
Cost.	Costituzione
Data Governance Act	Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022, relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724 (regolamento sulla governance dei dati)
Digital Market Act	Regolamento (UE) 2022/1925 del Parlamento europeo e del Consiglio del 14 settembre 2022, relativo a mercati equi e contendibili nel settore digitale e che modifica le direttive (UE) 2019/1937 e (UE) 2020/1828 (regolamento sui mercati digitali)

d.l.	decreto-legge
d.lgs.	decreto legislativo
d.m.	decreto ministeriale
dPCM	decreto del Presidente del Consiglio dei Ministri
DPIA	<i>Data Protection Impact Assessment</i> (la valutazione di impatto prevista dall'art. 35 del Regolamento)
DPO	<i>Data Protection Officer</i> (Responsabile della protezione dei dati)
dPR	decreto del Presidente della Repubblica
EDPB	<i>European Data Protection Board</i> (ex WP29) o Comitato per la protezione dei dati personali
Garante	Garante per la protezione dei dati personali
GPDP	Garante per la protezione dei dati personali
GDPR	<i>General Data Protection Regulation</i> (il Regolamento (UE) 2016/679)
G.U.	Gazzetta Ufficiale della Repubblica Italiana
G.U.C.E.	Gazzetta Ufficiale delle Comunità Europee
G.U.U.E.	Gazzetta Ufficiale dell'Unione Europea
l.	legge
lett.	lettera
ord.	ordinanza
p. / pp.	pagina / pagine
par. / parr.	paragrafo / paragrafi
provv. / provv.ti	provvedimento / provvedimenti
Regolamento	Regolamento (UE) 2016/679 (GDPR)
Reg.	Regolamento (UE) 2016/679 (GDPR)
RPD	Responsabile della Protezione dei dati
RPO	Registro pubblico delle opposizioni
sent.	sentenza
ss.mm.ii.	successive modifiche e integrazioni
Tar	Tribunale amministrativo regionale
TFUE	Trattato sul funzionamento dell'Unione Europea
t.u.	testo unico
TUE	Trattato sull'Unione europea
UE	Unione europea
v.	vedi / vedasi
WP29	Gruppo di lavoro Articolo 29 (ora EDPB)

Verso una definizione di dati sintetici

di LORENZO CRISTOFARO* E GABRIELE FRANCO**

1. La rivoluzione dei dati sintetici

Nel 1993 il professore di statistica di Harvard Donald B. Rubin pubblicava sul *Journal of Official Statistics* l'articolo "*Statistical Disclosure Limitation*". In un contesto in cui la tutela della riservatezza iniziava a interessare anche il mondo delle rilevazioni statistiche, Rubin suggeriva di «*non rilasciare microdati reali, ma solo microdati sintetici, realizzati mediante l'imputazione multipla, in modo da poter essere analizzati validamente con un software statistico standard*»¹.

Nell'approccio innovativo proposto dal docente americano, molti ritrovano la prima, embrionale concettualizzazione dei dati sintetici². Vent'anni dopo, questa tecnologia unica nel proprio genere, per impatti e applicazioni, ha raggiunto risultati sorprendenti. Nel 2022 il mercato globale dei *synthetic data* ha superato i 163 milioni di dollari (l'anno prima sfiorava "solo" i 123 milioni) ed entro il 2030 toccherà quota 1,79 miliardi³. Ri-

* Avvocato, *Partner* di Panetta Studio Legale.

** Avvocato, *Associate* di Panetta Studio Legale.

¹ D.B. RUBIN, *Statistical disclosure limitation*, in «*Journal of Official Statistics*», 9(2), 1993, pp. 461-468 (traduzione degli autori).

² C'è chi risale ancora prima nel tempo e riconduce al lavoro sui metodi di simulazione Monte Carlo svolto negli anni '40 da Stanislaw Ulam e John von Neumann, i primi casi di impiego di dati sintetici generati al computer per risolvere compiti particolari (J. JORDON, L. SZPRUCH, F. HOUSSIAU, M. BOTTARELLI, G. CHERUBIN, C. MAPLE, S. N. COHEN, A. WELLER, *Synthetic data: what, why and how?*, in «arXiv», arXiv:2205.03257, 2022, p. 5).

³ GRAND VIEW RESEARCH, *Synthetic Data Generation Market Size, Share & Trends Analysis Report By Data Type, By Modeling Type, By Offering, By*

viste come il MIT Technology Review⁴ e Forbes⁵ hanno riconosciuto la sintetizzazione tra le tecnologie a più rilevante impatto del 2022. Anche l'European Data Protection Supervisor (EDPS) ha dedicato particolare considerazione ai dati sintetici, inserendoli tra gli *emerging technology trends* sia per il 2021/2022⁶, che per il 2022/2023⁷.

Molto si potrebbe dire in merito alle ragioni che hanno determinato il successo di questa rivoluzione, ma qui ne vogliamo indicare solamente due, particolarmente significative sia per la loro stretta interconnessione, sia per la rilevanza ai fini dell'esercizio definitorio che ci si appresta a intraprendere. Sotto un primo angolo visuale, più squisitamente tecnico, la diffusione su scala globale dei sistemi di intelligenza artificiale ha determinato un inedito picco nella domanda di dati utili ai fini di addestramento degli algoritmi, non solo da un punto di vista quantitativo, ma anche – e soprattutto – in termini qualitativi. Si pensi al *training* dei sistemi di *machine learning* alla base del funzionamento delle *self-driving car*.

Tali veicoli a guida autonoma dovrebbero circolare in strada per centinaia di milioni di miglia, talvolta persino per centinaia di miliardi di miglia, prima di poter dimostrare che sono sicuri⁸.

Di fronte all'oggettiva difficoltà – al limite dell'impossibilità – di raccogliere un numero di dati in grado di rappresentare in modo significativo la grande variabilità del contesto stradale, i dati sintetici offrono una soluzione ad alta efficienza, permettendo di simulare un'esperienza di guida in tutto e per tutto ana-

Application, By End-use, By Region, And Segment Forecasts, 2023-2030, settembre 2022.

⁴ R. COURTLAND, M. HONAN, A. NORDRUM, M. REILLY, D. ROTMAN, *10 Breakthrough Technologies 2022*, in «MIT Technology Review», marzo/aprile 2022.

⁵ B. MARR, *The 5 Biggest Data Science Trends In 2022*, in «Forbes», ottobre 2021.

⁶ T. ZERDICK, S. LEUCCI (a cura di), *TechSonar reports*, dicembre 2021.

⁷ M. ATTORESI, S. LEUCCI (a cura di), *TechSonar reports*, novembre 2022.

⁸ N. KALRA, S.M. PADDOCK, *How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?*, RAND Corporation, 2016.