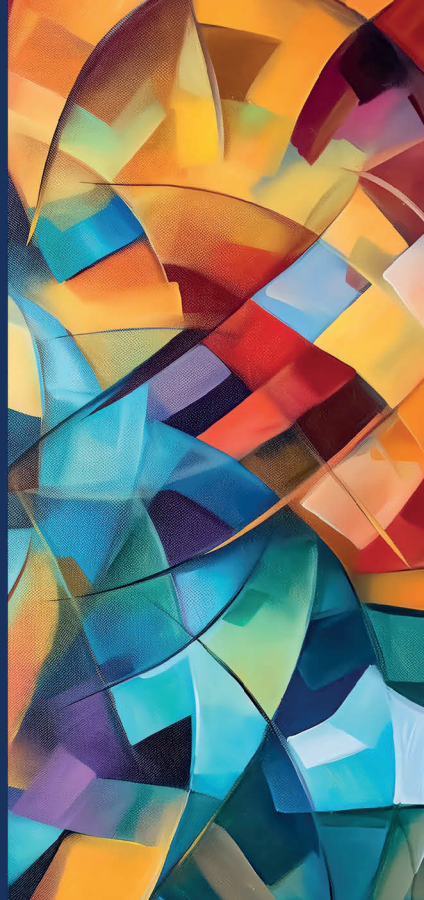


Silvia Scalzini

Banche di dati, sfruttamento dei dati digitali e concorrenza



Giappichelli

INTRODUZIONE

LE RAGIONI E GLI OBIETTIVI DEL LAVORO

La quarta rivoluzione industriale¹ ha il proprio punto focale nella digitalizzazione dell'informazione. I dati digitali costituiscono la rappresentazione di misure, fatti o informazioni in formato digitale. Grazie allo sviluppo di tecnologie capaci di misurare, combinare ed analizzare enormi volumi di dati provenienti da fonti eterogenee², nascono nuovi modi di organizzare e valorizzare l'informazione nei rapporti di produzione e di scambio di beni o servizi.

L'uso dei dati digitali, pertanto, oltre a costituire la nuova frontiera per l'avanzamento scientifico, culturale e sociale, è diventato uno strumento sempre più rilevante di concorrenza tra le imprese³. L'impulso innovativo ha spesso origine dalle capacità di raccolta, incrocio ed elaborazione di "big data" (o "macrodati")⁴. Il modo in cui le imprese accedono (ai) ed usano i dati si riflette sui mercati e sulle dinamiche competitive, quindi sulla allocazione delle risorse e sul

¹La quarta rivoluzione industriale è descritta come un fenomeno di radicale e rapidissimo mutamento di paradigma che, costruendo sulla rivoluzione digitale e combinando una serie di tecnologie, ha un impatto economico, sociale, commerciale ed industriale, trasformando l'intero sistema in modo trasversale. Per un inquadramento generale *cfr.*, in particolare, K. SCHWAB, *The Fourth Industrial Revolution*, Penguin Books, 2017; L. FLORIDI, *La quarta rivoluzione: come l'infosfera sta trasformando il mondo*, Raffaello Cortina Editore, Milano, 2017; F. PASQUALE, *Le nuove leggi della robotica. Difendere la competenza umana nell'era dell'intelligenza artificiale*, Luiss University Press, Roma, 2021. Parte della dottrina preferisce parlare a tal proposito di "società algoritmica". Si veda J. BALKIN, *The Three Laws of Robotics in the Age of Big Data*, in *Ohio State Law Journal*, 2017, Vol.78, 5, p. 1217 ss., spec. p. 1219 e, per la dottrina italiana, M. RICOLFI, *Il futuro della proprietà intellettuale nella società algoritmica*, in *Giur. It.* 2019, supplemento, p. 10 ss., 12.

²*Cfr.* EUROPEAN DATA PROTECTION SUPERVISOR, *Verso una nuova etica digitale. Dati, dignità, tecnologia*. Parere 4/2015, reperibile all'indirizzo https://edps.europa.eu/data-protection/our-work/publications/opinions/towards-new-digital-ethics-data-dignity-and_en.

³Il presente lavoro si concentra sullo sfruttamento economico dei dati digitali nell'attività di impresa in concorrenza con altre imprese. Non analizzerà nel dettaglio i pur fondamentali temi della circolazione dei dati tra impresa e settore pubblico e tra impresa e consumatori, se non nell'ambito del perimetro dell'analisi.

⁴*Cfr. ex multis*, V. MAYER-SCHÖNBERGER, K. KUKIER, *Big Data. A Revolution that Will Transform How we Live, Work and Think*, Houghton Mifflin Harcourt, Boston, 2013. Si parla, infatti di "data-driven innovation", o innovazione guidata dai dati. *Cfr.* OECD (2015), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Parigi, <https://doi.org/10.1787/9789264229358-en>.

benessere collettivo di lungo periodo⁵. L'equazione è fin troppo evidente: all'aumento della quantità (e qualità) di dati disponibili corrisponde una aumentata capacità computazionale di addivenire a risultati sempre migliori, a beni e servizi sempre più innovativi (e personalizzati), a previsioni sempre più accurate, che incidono sulle decisioni organizzative e commerciali delle imprese. I dati digitali, così, non solo costituiscono un importante vantaggio competitivo, ma acquistano anche un valore in sé grazie alla utilizzazione e riutilizzazione, potendo costituire oggetto di scambio nei mercati dell'economia digitale⁶. Secondo la ben nota (e forse abusata) metafora resa celebre dall'*Economist*⁷, i dati (o meglio la relativa utilizzazione e valorizzazione) rappresentano il “nuovo petrolio” dell'economia; essi divengono una risorsa imprenditoriale preziosa ed una variabile concorrenziale dalla cui considerazione non si può più prescindere per valutare il potere economico delle imprese nei nuovi mercati dell'economia digitale.

Tale rinnovata capacità di valorizzare i “dati digitali” ha fatto sorgere nuovi conflitti nei rapporti tra imprese, resi sempre più evidenti dalla casistica giurisprudenziale e dalle ricostruzioni dottrinali europee e nazionali su come interpretare ed applicare i principi e le norme del diritto d'impresa a questi nuovi fenomeni. Dalla prospettiva del diritto d'impresa⁸, si possono annoverare, da un lato, i conflitti tra gli interessi all'accumulo e uso esclusivo dei dati per beneficiare dei vantaggi competitivi e per recuperare gli investimenti profusi nelle attività di generazione, raccolta, elaborazione dei dati; dall'altro, gli interessi all'accesso ed all'uso dei dati da parte delle imprese che, invece, intendono entrare e svilupparsi nei nuovi mercati, con evidenti ricadute anche sulle dinamiche dell'innovazione e dell'interesse generale. Ciò solleva nuove questioni su quali siano i confini di liceità dei comportamenti innescati da tali conflitti e implica difficili bilanciamenti per l'applicazione dei rimedi disponibili nel quadro del diritto d'impresa, specialmente per quanto riguarda l'imposizione dell'obbligo di condivisione dei dati con altre imprese.

Al tempo stesso, lo sfruttamento economico dei dati digitali sta conducendo

⁵ Cfr. OECD (2022), *Data shaping firms and markets*, OECD Digital Economy Papers, No. 344, OECD Publishing, Parigi, <https://doi.org/10.1787/7b1a2d70-en>. Per una riflessione sull'inquadramento dell'attività di impresa nella “infosfera”, e sull'importanza della digitalizzazione per l'attività e l'organizzazione delle imprese, cfr. A. NUZZO, G. OLIVIERI (a cura di), *Algoritmi. Se li conosci, li regoli...*, in *Analisi Giuridica dell'Economia*, 1/2019; N. ABRIANI, G. SCHNEIDER, *Diritto delle imprese e intelligenza artificiale*, Il Mulino, Bologna, 2021; N. ABRIANI, R. COSTI, *Diritto societario, digitalizzazione e intelligenza artificiale*, Milano, 2023.

⁶ Cfr. A.M. GAMBINO, A. STAZI, *La circolazione dei dati. Titolarità, strumenti negoziali, diritti e tutele*, Pacini, Pisa, 2020; L. AMMANNATI, *La Circolazione dei Dati: dal Consumo alla Produzione*, reperibile all'indirizzo www.associazioneadde.it.

⁷ K. BHAGESHPUR, *The world's most valuable resource is no longer oil, but data*, *The Economist*, 6 maggio 2017, reperibile all'indirizzo www.theeconomist.com.

⁸ Cfr. oltre alle fonti che saranno menzionate nel prosieguo, V. FALCE, G. GHIDINI, G. OLIVIERI (a cura di), *Informazione e Big Data tra Innovazione e Concorrenza*, Giuffrè, Milano, 2018.

ad una rapida e corposa evoluzione normativa, sulla spinta delle istituzioni euro-unitarie, attraverso norme che disciplinano puntualmente alcune condotte delle imprese sui mercati dell'economia digitale. L'orientamento emergente di politica del diritto appare volto ad incentivare e consentire una maggior circolazione dei dati digitali, ritenuta essenziale per aumentare il benessere collettivo, stimolare l'innovazione e mantenere pressione concorrenziale, finanche prevedendo *ex ante* meccanismi di condivisione obbligatoria dei dati digitali tra imprese.

In questo quadro di importanti mutamenti tecnologici, economici e normativi non potevano non sorgere questioni interpretative e proposte di riforma – a circa un trentennio dalla relativa “armonizzazione”⁹ – degli strumenti che sembrano essere chiamati in causa per disciplinare la tutela del lavoro intellettuale¹⁰ e degli investimenti imprenditoriali nella attività di raccolta e sistematizzazione dei dati, e, correlativamente, la relativa libera utilizzazione: i diritti esclusivi sulle banche di dati. La tutela delle banche di dati non rappresenta l'unica (né forse la principale¹¹) fonte di disciplina della rete di rapporti economici sviluppata

⁹ Si fa riferimento più precisamente all'armonizzazione europea dei diritti esclusivi sulle banche di dati ad opera della Direttiva 96/9/CE del Parlamento europeo e del Consiglio, dell'11 marzo 1996, relativa alla tutela giuridica delle banche di dati, in Gazzetta ufficiale n. L 077 del 27/03/1996 pag. 0020 – 0028, recepita nel nostro ordinamento dal D.Lgs. 6 maggio 1999, n. 169, *Attuazione della direttiva 96/9/CE relativa alla tutela giuridica delle banche di dati*, che ha modificato la L. 22 aprile 1941, n. 633, *Protezione del diritto d'autore e di altri diritti connessi al suo esercizio* (di seguito “l.d.a.”).

¹⁰ Per l'opinione che collega i diritti di proprietà intellettuale alla tutela del lavoro intellettuale, in tutte le sue forme, “ivi inclusa quella che si realizza nell'ambito dell'impresa e che vede nell'azienda l'oggetto mediato della protezione offerta dall'ordinamento” *cfr.* G. OLIVIERI, S. SCALZINI, *La proprietà intellettuale*, in C. CAPORALE, L. MAFFEI, V. MARCHIS, J.C. DE MARTIN (a cura di), *Europa. Le sfide della Scienza*, Istituto della enciclopedia italiana fondata da Giovanni Treccani, p. 598-606, 600, condividendo la tesi di G. OPPO, *Creazione ed esclusiva nel diritto industriale*, in *Rivista del diritto commerciale*, 1964, 5-6, p. 187-203. In argomento *cfr.* anche M. LIBERTINI, *Tutela e promozione delle creazioni intellettuali e limiti funzionali della proprietà intellettuale*, in *AIDA*, 2014, p. 299 ss.

¹¹ Sulla rilevanza dei diritti esclusivi sulle banche di dati ai fini della regolazione dell'economia europea dei dati le visioni sono discordanti. Alcuni Autori sembrano evidenziare la centralità di tale disciplina, tra cui M. LEISNER, *Big Data and the Eu Database Directive 96/9/EC: Current Law and Potential for Reform*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Trading Data in the Digital Economy. Münster Colloquia on EU Law and the Digital Economy*, Nomos, Baden Baden, 2017, p.27 ss., 55; A. OTTOLIA, *Big Data e innovazione computazionale*, Giappichelli, Torino, 2017; G. NOTO LA DIEGA, *Artificial Intelligence and databases in the age of big machine data*, in *AIDA*, 2018, p. 93 ss. Ne sottolineano la rilevanza, in ottica decisamente più critica, E. DERCLAYE, M. HUSOVEC, *Sui Generis Database Protection 2.0: Judicial and Legislative Reform*, in *European Intellectual Property Review (EIPR)*, 2022, Vol.44(6), p.323-331. Sembrano ridimensionarne l'idoneità a fungere da parametro di riferimento per i problemi di cui si discute, e in particolare relativamente alla tutela dei *big data*, invece, M. RICOLFI, *Il futuro della proprietà intellettuale nella società algoritmica*, cit., p.12; P.B. HUGENHOLTZ, *Data Property in the System of Intellectual Property: Welcom Guest or Misfit?*, in S. LOHSSE, R. SCHULZE, D. STAUDENMAYER (a cura di), *Trading Data in the Digital Economy. Münster Colloquia on EU Law and the Digital Economy*, Hart Publishing, Nomos Verlagsgesellschaft, Baden-Baden, 2017, p. 75 ss., p. 83.

attorno allo sfruttamento commerciale dei dati digitali, considerata anche la contemporanea coesistenza di facoltà (di diritto e di fatto) di disposizione e uso degli stessi che incidono sulle vicende circolatorie¹².

D'altra parte, le innovative forme di raccolta, elaborazione e sfruttamento intensivo dei dati, rese possibili dall'avanzamento tecnologico e dal nuovo assetto dei mercati, (che – come si vedrà – non rendono irrilevanti le dimensioni ed il ruolo delle imprese coinvolte) e la crescente rilevanza degli interessi alla condivisione degli stessi per lo sviluppo dell'innovazione in molti settori industriali, rendono finalmente maturi i tempi per uno studio dedicato alla disciplina giuridica delle banche di dati alla luce delle mutate dinamiche concorrenziali legate allo sfruttamento dei dati digitali.

A sottolineare l'attualità di una riflessione su giustificazioni, contenuto e limiti dei diritti sulle banche di dati non è solo la necessità di soluzioni interpretative *de iure condito*, evidenziata dalla dottrina e dalla giurisprudenza¹³ per risolvere in modo coerente e controllabile le nuove tensioni tra tutela delle forme creative e degli investimenti nelle raccolte di dati e accesso e riutilizzazione degli stessi. Si assiste, infatti, anche alla proliferazione di nuove regole in via di costruzione nel panorama euro-unitario, oltre alle proposte dottrinali e istituzionali di riforma (radicale o più circoscritta) della disciplina, destinate ad avere importanti ripercussioni anche sulle peculiarità applicative della tutela in ambito nazionale, nonché sulla libertà di impresa.

I diritti esclusivi sulle banche di dati sono tipicamente esercitati da imprese, le quali sono i titolari “finali” dei relativi diritti¹⁴. L'attribuzione a titolo originario del diritto d'autore alla persona fisica che esprime la propria creatività nella selezione e disposizione dei dati nella banca di dati, infatti, viene temperata già a livello legislativo dalle norme che conferiscono l'esercizio dei diritti patrimoniali – salva diversa pattuizione – al datore di lavoro¹⁵. Il diritto *sui generis*, inoltre, nasce proprio per tutelare gli investimenti imprenditoriali nella costituzione, verifica, presentazione dei dati del costituente della banca di dati, che può essere

¹² Sulle ricostruzioni dei regimi che regolano l'appartenenza e la circolazione dei dati digitali, cfr., *ex multis*, J. DREXL, *Designing Competitive Markets for Industrial Data – Between Propertisation and Access*, in *JIPITEC*, 2017, 8, 257, § 1; A. OTTOLIA, *Big Data e innovazione computazionale*, cit.; M.L. MONTAGNANI, *Dati e proprietà intellettuale in Europa: dalla “proprietà” all'accesso*, in *Il diritto dell'economia*, n. 1, 2020, p. 539-569.

¹³ Riflette sull'evoluzione delle fonti del diritto commerciale M. SPERANZIN, *Un Nuovo Ordine delle Fonti del Diritto Commerciale*, in *Riv. Dir. Civ.*, 2019, p. 1174 ss.

¹⁴ Si vedano in proposito le considerazioni di G. GHIDINI, *Rethinking Intellectual Property. Balancing Conflicts of Interests in the Constitutional Paradigm*, Edward Elgar, Cheltenham-Northampton, 2018, p. 12, il quale ripercorre le ragioni storiche, filosofiche ed assiologiche dell'attribuzione a titolo originario dei diritti della proprietà intellettuale al creatore persona fisica e la relativa importanza per le finalità della disciplina, sottolineando altresì il ruolo centrale dell'impresa, anche in relazione ad alcuni espressi riconoscimenti da parte del legislatore, per l'esercizio dei diritti economici sul mercato come “*vérité de bon sens*”.

¹⁵ Cfr. art. 12-*bis* l.d.a. Sul punto si tornerà ampiamente al cap. 2, § 3.2.

identificato indifferentemente in una persona fisica o giuridica¹⁶. In altre parole, anche per i diritti sulle banche di dati, l'impresa è tipicamente il soggetto che esercita i relativi diritti esclusivi, i cui oggetti fanno parte dei beni aziendali di cui l'imprenditore si avvale, ai sensi dell'art. 2555 c.c., per operare sul mercato in concorrenza con altre imprese.

La prospettiva euristica del diritto d'impresa, infatti, assume come punto di partenza la teoria dell'impresa come organizzazione¹⁷, come tale tutelata dall'art. 41 della Costituzione nelle proprie scelte (e nei propri mezzi) di innovazione e sviluppo in un regime di concorrenza con altre imprese, nell'ambito di un orizzonte multi-livello di "economia sociale di mercato fortemente competitiva" e nello "sviluppo [economico] sostenibile" (art. 3 TUE).

Secondo la tesi qui difesa i regimi giuridici delle banche di dati consentono di risolvere alcune delle questioni sulla "riutilizzazione" dei dati tra imprese, attraverso interpretazioni che ne valorizzano la dimensione pro-concorrenziale che, peraltro, si ritiene ne abbia in parte costituito l'ispirazione originaria. Invece, per i diversi problemi relativi all'"accesso" ai dati rispetto a raccolte che, attraverso l'esercizio dei diritti, non sono rese disponibili a terzi la risoluzione dei relativi conflitti è demandata – ove ne sussistano i presupposti – a diversi rimedi derivanti da istituti esterni rispetto alla disciplina propria delle banche di dati, oppure a riforme normative volte a consentire un accesso pro-concorrenziale ai dati.

Se, infatti, l'orizzonte normativo positivo ha una matrice principalmente "privatistico-competitiva", gli interessi alla circolazione dell'informazione rappresentata dai dati digitali assumono una portata più generale, che non sempre sembra essere ben valorizzata dai bilanciamenti operati dal legislatore. In altri termini, le questioni da risolvere nell'ambito di conflitti inter-individuali si inquadrano in una visuale più ampia, di rango costituzionale¹⁸, là dove si configurano la possibilità di un controllo privatistico dei dati e dell'informazione, coinvolgendo la tutela di interessi generali.

¹⁶ Si veda *infra* cap. 2, § 4.3.

¹⁷ Cfr. in particolare M. LIBERTINI, *Impresa, proprietà intellettuale e Costituzione*, in *dircomm.it*, dicembre 2005, § 3-5. I beni immateriali possono, infatti, rientrare nel complesso di beni organizzati dall'imprenditore per l'esercizio dell'impresa di cui all'art. 2555 del codice civile. Pertanto, anch'essi, di riflesso, partecipano della garanzia costituzionale dell'art.41 Cost. e sono soggetti ai relativi limiti previsti dai commi 2 e 3. In particolare, l'esercizio dei relativi diritti, nell'ambito dell'iniziativa economica privata, non può svolgersi in contrasto con l'utilità sociale. Cfr. più ampiamente su tale profilo G. GHIDINI, *Profili evolutivi del diritto industriale. Innovazione, creatività, informazione, dinamiche conflittuali, esperienze di condivisione*, Giuffrè, Milano, 2015.

¹⁸ Valorizza la necessità di ancorare la risoluzione dei conflitti di interesse nell'ambito del quadro costituzionale, in particolare, G. GHIDINI, *Rethinking Intellectual Property. Balancing Conflicts of Interests in the Constitutional Paradigm*, cit., p. 1 ss. Per un approccio simile C. GEIGER, *The Social Function of Intellectual Property Rights, or How Ethics can Influence the Shape and Use of IP Law*, in G.B. DINWOODIE (a cura di), *Intellectual Property Law: Methods and Perspectives*, Edward Elgar, Cheltenham-Northampton, 2013, p. 153-176; A.B.L. BROWN, *IP, Human Rights and Competition: Access to Essential Innovation and Technology*, Edward Elgar, Cheltenham-Northampton, 2013, oltre alla dottrina citata a tal proposito nel capitolo 2.

Il primo capitolo, pertanto, analizza l'evoluzione dei mercati digitali derivate dal nuovo contesto economico e tecnologico, al fine di identificare con maggior nitidezza le questioni e le domande di ricerca che si intende affrontare. Preliminare a tale analisi è la necessità di fornire un primo inquadramento alla nozione di dati digitali, nell'ambito delle risorse che possono far parte del complesso dei beni aziendali organizzati dall'imprenditore e, dunque, funzionalmente preordinati a divenire strumento dell'esercizio dell'impresa. Sulla base di alcuni studi recenti¹⁹, sarà illustrato come lo sfruttamento economico dei dati digitali (e, in particolare, dei macrodati) si articoli in varie fasi che vanno dalla generazione, alla raccolta, alla interpretazione, passando dalla elaborazione dei dati. In ciascuna di tali fasi gli investimenti e le strategie imprenditoriali sono volte alla massimizzazione del valore che è possibile estrarre da tali attività, riflettendosi sulle dinamiche concorrenziali tra imprese che operano sugli stessi mercati o in mercati contigui. Il capitolo inquadra tali sviluppi, inoltre, nell'ambito dell'evoluzione normativa in corso, volta a riscrivere le regole sull'esercizio dell'attività imprenditoriale nei mercati digitali, senza tuttavia incidere radicalmente sulle norme sulla tutela esclusiva delle banche di dati.

Il secondo capitolo è dedicato all'analisi dei diritti esclusivi sulle banche di dati, d'autore e *sui generis*, al fine di delinearne funzioni e contenuto per poi affrontare le questioni del nuovo contesto economico e tecnologico poste in particolare dalle nuove utilizzazioni computazionali dei dati. La disciplina del diritto d'autore e del diritto *sui generis*, tuttavia, non consente sempre di arginare le forme abusive di esercizio dei diritti esclusivi volti a limitare l'accesso o l'uso dei dati, specialmente ove tale comportamento si leghi all'esercizio di potere di mercato.

Il terzo capitolo, pertanto, verifica se tali problemi possano essere affrontati attraverso l'applicazione di corpi normativi di recente introduzione volti a favorire un accesso pro-concorrenziale ai dati, tentando di risolvere i contrasti applicativi sollevati dall'interazione di tali fonti con i diritti sulle banche di dati o, alternativamente, dall'applicazione di rimedi *ex post* propri del diritto della concorrenza o – in generale – del diritto d'impresa.

Alla luce delle conclusioni raggiunte grazie alle diverse e complementari prospettive esaminate, infine, il lavoro discute le interpretazioni proposte e ipotizza alcuni scenari di riforma *de iure condendo* della disciplina sulla tutela delle banche di dati.

¹⁹Tra cui l'indagine conoscitiva congiunta sui Big Data dell'Autorità Garante della Concorrenza e del mercato (AGCM), dell'Autorità per le Garanzie nelle comunicazioni (AGCOM) e del Garante per la Protezione dei dati personali (GPDP), di seguito *Indagine Conoscitiva sui Big Data*, 2019.

CAPITOLO 1

SFRUTTAMENTO ECONOMICO
DEI DATI DIGITALI
E ATTIVITÀ D'IMPRESA

1. Il contesto: la nuova rivoluzione industriale ed il ruolo degli algoritmi

Premessa necessaria della ricerca è la presentazione dei fenomeni interessati dall'analisi che seguirà¹, al fine di mostrare – seppur con necessaria sintesi – il contesto in cui sorgono le domande giuridiche appena delineate in relazione alle incertezze ed alle questioni legate all'estensione ed all'esercizio dei diritti sulle banche di dati nei rapporti di concorrenza tra imprese, a fronte dei nuovi modi di utilizzazione dei dati digitali, contenuti in tali raccolte.

Nell'economia digitale lo sfruttamento dei dati ha un ruolo importante nell'ambito dell'attività economica di molte imprese, in quanto la capacità con cui le imprese generano, raccolgono e usano i dati si riflette sulla relativa innovazione e produttività². Come ha affermato la Commissione europea in una delle molteplici Comunicazioni dedicate all'economia digitale “è in atto una nuova rivoluzione industriale trainata dai dati digitali, dall'informatica e dall'automazione”³.

La rappresentazione di fatti, informazioni e comportamenti umani può essere, infatti, espressa o convertita in un linguaggio formato di combinazioni di zero ed uno interpretabile da una macchina⁴. L'informazione così codificata può

¹ Secondo l'insegnamento di T. ASCARELLI, *Norma giuridica e realtà sociale*, in *Diritto dell'economia*, 1955, p. 1179 ss. (poi in ID., *Problemi giuridici*, Giuffrè, Milano, 1959, p. 69-111).

² Cfr. da ultimo OECD, *Data shaping firms and markets*, cit., p. 12 ss., che peraltro sottolinea come “recent OECD analysis on microdata from France, Italy, the Netherlands and Sweden confirms that firms carrying out big data analytics are more likely to innovate”. Lo studio evidenzia anche come le grandi imprese siano sinora riuscite ad innovare maggiormente grazie alle nuove modalità di usare i dati rispetto alle piccole e medie imprese. Di rilievo per le evidenze empiriche contenute ed il livello di analisi anche OECD, *Big data: bringing competition policy to the digital era*, Parigi, 2016; cfr. OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, cit.

³ Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni, *Verso una florida economia basata sui dati* COM/2014/0442 final

⁴ Lo studio di OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, a p. 32,

essere più facilmente generata, raccolta, analizzata, archiviata e in seguito trasmessa attraverso mezzi più veloci e ad un costo inferiore, grazie a capacità connettive e computazionali sempre più avanzate. “Big data” è, infatti, una locuzione che indica lo sviluppo di tecnologie capaci di conservare, combinare ed analizzare enormi volumi di dati provenienti da fonti eterogenee⁵ ed ottenere dal trattamento di questi dati, grazie ad algoritmi sempre più sofisticati, “una quantità ancora più sterminata di nuovi dati, che consentano nuove conoscenze ed analisi relative ai fenomeni naturali ed ai comportamenti umani”⁶. Nonostante l’espressione evochi in primo luogo l’aspetto dimensionale del fenomeno⁷, la vera novità consiste nella “espansiva capacità di connettere, attraverso l’analisi algoritmica, *datasets* disparati, forgiando relazioni tra dati raccolti in differenti momenti e luoghi e per diverse finalità”⁸, con il fine ultimo di creare nuova conoscenza. Esempio è l’evoluzione del *marketing* d’impresa, dove le imprese hanno la possibilità di integrare le proprie banche di dati con i dati provenienti dai *social network* e dai dispositivi *smart* (si pensi ai dati di geolocalizzazione raccolti dagli *smartphone*), di profilare i consumatori e di conseguenza personalizzare le offerte commerciali, finanche ad anticipare le preferenze ed i comportamenti degli stessi⁹. Senza la possibilità di misurare i comportamenti dei con-

descrive il primo anello della catena del valore della innovazione guidata dai dati come “datafication and data collection – These refer to the activity of data generation through the digitisation of content, and monitoring of activities, including real world (offline) activities and phenomena, through sensors”.

⁵ Così EUROPEAN DATA PROTECTION SUPERVISOR, *Verso una nuova etica digitale. Dati, dignità, tecnologia*, cit. Sul fenomeno Big Data, la letteratura è sterminata. Ci si limiterà a ricordare, *ex multis*, V. MAYER-SCHÖNBERGER, K. KUKIER, *Big Data. A Revolution that Will Transform How we Live, Work and Think*, cit. R KITCHIN, *The Real Time City? Big Data and Smart Urbanism*, in *GeoJournal*, 2014, 79(1):1Z14; V. ZENO-ZENCOVICH, G. GIANNONE CODIGLIONE, *Ten Legal Perspectives on the “Big Data Revolution”*, in *Concorrenza e Mercato*, 2016, p. 29 ss.; V. FALCE, G. GHIDINI, G. OLIVIERI, *Informazione e Big Data tra innovazione e concorrenza*, cit. Per spunti di analisi economica del diritto M. DELMASTRO, A. NICITA, *Big Data. Come ci stanno trasformando*, Il Mulino, Bologna, 2019.

⁶ F.M. PIZZETTI, *Privacy e il Diritto Europeo alla Protezione dei Dati Personali. Dalla Direttiva 95/46 al nuovo Regolamento europeo*, Giappichelli, Torino, 2016.

⁷ Molte definizioni di “big data” si concentrano essenzialmente sulle c.d. “4V”: “il volume, con riferimento all’enorme dimensione dei dati generati e raccolti; la varietà, con riguardo alle numerose tipologie dei dati disponibili, tra i quali, oltre ai dati strutturati tradizionali, vi sono anche dati semi-strutturati e non strutturati come audio, video, pagine web e testi; la velocità delle operazioni di trattamento; il valore che i dati assumono allorquando vengono elaborati ed analizzati”. Cfr. AGCM, AGCOM, Garante Privacy, *Indagine Conoscitiva sui Big Data*, cit.; Gartner *IT Glossary: Big Data*, reperibile all’indirizzo <http://www.gartner.com/it-glossary/big-data/>. G. BRUZZONE, *Il mercato unico digitale, cybersecurity e big data*, Assonime, 31/2018.

⁸ J. METCALF, E. KELLER, F.D., BOYD, *Perspectives on Big Data, Ethics, and Society, White Paper of the Council for Big data, Ethics and Society*, 2016 reperibile all’indirizzo <http://bdes.data.society.net/>.

⁹ Cfr., F. CALABRÒ, *Capitale umano e digitale: come estrarre esperienza dai dati aziendali*, Agenda Digitale, 8 novembre 2021, reperibile all’indirizzo <https://www.agendadigitale.eu/documenti/capitale-umano-e-digitale-come-estrarre-esperienza-dai-dati-aziendali/>.

sumatori, organizzare (con mezzi per lo più automatici) cluster omogenei di dati ed incrociarli con dati tratti da altre fonti, tali risultati non sarebbero possibili.

Ai nuovi modi per generare, osservare e raccogliere dati si accompagna, infatti, una continua innovazione tecnologica sui modi di elaborazione, memorizzazione e trasmissione di tali dati¹⁰, tecniche che alimentano una forma di “innovazione computazionale”¹¹. La disponibilità di grandi masse di dati, e di *computer* sempre più potenti in grado di elaborare velocemente tali quantità di dati consente, inoltre, ai sistemi di intelligenza artificiale¹² di effettuare compiti

¹⁰ Cfr. in tema lo studio di IT Media Consulting, con il contributo scientifico del centro studi ASK dell'Università Commerciale Luigi Bocconi, *L'economia dei dati: tendenze di mercato e prospettive di policy*, 2018, reperibile all'indirizzo <https://www.itmedia-consulting.com>, p. 29 e 30, che descrive tale innovazione in termini di “potenza di calcolo dei microprocessori”, “disponibilità e capacità [di memoria]” e “capacità e velocità [di trasmissione] grazie allo sviluppo delle reti broadband e ultrabroadband”. Grazie allo sviluppo di tali operazioni hanno acquisito centralità, inoltre, non solo le tecniche analitiche dei Big Data, ma anche architetture di *Cloud Computing* che consentono di gestire “sia i servizi di rete primari per l'impresa che il traffico dati”. Cfr. anche il successivo studio *Big Data e Intelligenza Artificiale Condivisione e accesso ai dati tra mercato, concorrenza e regole*, 2020.

La digitalizzazione dell'informazione ha consentito anche la creazione di sistemi di trasmissione e registrazione digitale decentralizzata e trasparente dei dati digitali, attraverso le c.d. distributed ledger technologies ed i sistemi blockchain. In tema cfr. A. NUZZO (a cura di), *Blockchain e autonomia privata. Fondamenti Giuridici*, Luiss University Press, Roma, 2019-2020; F. SARZANA DI S. IPPOLITO, M. NICOTRA, *Diritto della Blockchain, Intelligenza Artificiale e IOT*, Ipsoa, 2018. Per il funzionamento della Blockchain cfr., in particolare, P. DE FILIPPI, A. WRIGHT, *Blockchain and the Law: the Rule of Code*, Harvard University Press, Cambridge, MA, 2018.

¹¹ A. OTTOLIA, *Big data e innovazione computazionale*, cit., *passim*. Cfr., altresì, A. NUZZO, G. OLIVIERI (a cura di), *Algoritmi. Se li conosci, li regoli ...*, cit.

¹² Cfr. la definizione elaborata dall'High-Level Expert Group on Artificial intelligence (AI) della Commissione europea secondo cui “AI refers to systems designed by humans that, given a complex goal, act in the physical or digital world by perceiving their environment, interpreting the collected structured or unstructured data, reasoning on the knowledge derived from this data and deciding the best action(s) to take (according to pre-defined parameters) to achieve the given goal. AI systems can also be designed to learn to adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)”. Si veda, in particolare il documento “A definition of AI: Main capabilities and scientific disciplines” (a cura dell'High-Level Expert Group on Artificial Intelligence), pubblicato il 18 dicembre 2018 e reperibile al seguente indirizzo <https://ec.europa.eu/digital-single-market/en/news/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines>.

Cfr. anche Commissione Europea, Proposta di regolamento del Parlamento europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione, 21 aprile 2021, COM (2021) 206 final. In tale atto, la definizione proposta di “sistema di intelligenza artificiale” è “un software sviluppato con una o più delle tecniche e degli approcci elencati nell'allegato I, che può, per una determinata se-

sempre più assimilabili a quelli che potrebbe compiere un essere umano¹³, se non finanche più complessi, innestando il tradizionale ciclo dei processi cognitivi¹⁴. Se sulla base delle informazioni elaborate (specialmente nella fase di “addestramento” degli algoritmi di *machine learning*¹⁵) tali sistemi sono capaci di compiere compiti, con un margine di “autonomia”¹⁶, in relazione a predeterminati obiettivi – come tradurre automaticamente testi, coadiuvare ed assistere l’utente nelle mansioni quotidiane o professionali, fino a (quasi) sostituire l’utente nella predisposizione di atti e documenti, o prevedere l’evoluzione di malattie, cambiamenti climatici, prevenire reati –, il relativo affinamento dipende dalla possibilità di accedere, selezionare e processare i macrodati.

Gli assi portanti di tale progresso tecnologico sono *hardware*, *software* (o algoritmi, in generale) e infrastrutture di connettività, che rendono possibile non solo le comunicazioni tra persone, ma altresì la connessione e comunicazione di oggetti e sensori via Internet. L’Internet delle cose (nella più nota versione anglofona, *Internet of Things*, abbreviato *IoT*) è definito, infatti, come l’evoluzione di beni e servizi da statici a “intelligenti, ossia in grado di connettersi ad altri dispositivi o reti, scambiare dati con essi e operare in una certa misura in modo interattivo e autonomo”¹⁷ ed è utilizzato nei più diversi settori di applicazione –

rie di obiettivi definiti dall’uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono”.

Sull’importanza del problema definitorio si è soffermato, in particolare, G. OLIVIERI, nella sua introduzione tenuta al Convegno “Artificial Intelligence and Regulation”, presso la Luiss Guido Carli, 2 marzo 2018. In tema *cfr.* anche R. LEENES, E. PALMERINI, B.-J. KOOPS., A. BERTOLINI, P. SALVINI, F. LUCIVERO, *Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues*, in *Law, Innovation and Technology*, 2017, 9:1, p. 1-44; G.F. ITALIANO, S. CIVITARESE MATTEUCCI, A. PERRUCCI, *L’intelligenza artificiale: dalla ricerca scientifica alle sue applicazioni. Una introduzione di contesto*, in A. PAJNO, F. DONATI, A. PERRUCCI, *Intelligenza Artificiale e Diritto: una rivoluzione?*, Vol. 1, Quaderni Astrid, Il Mulino, Bologna, 2022, p. 43 ss.

¹³ Per una approfondita riflessione filosofica su potenzialità e limiti dell’intelligenza artificiale il riferimento è d’obbligo all’opera di N. BOSTROM, *Superintelligence, Path, Dangers, Strategies*, 2014, Oxford University Press. In tema *cfr.* A. PAJNO, F. DONATI, A. PERRUCCI, *Intelligenza Artificiale e Diritto: una rivoluzione?*, cit.; A. D’ALOIA (a cura di), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo*, Franco Angeli, Milano, 2020; U. RUFFOLO (a cura di), *Intelligenza artificiale – Il diritto, i diritti, l’etica*, Giuffrè, Milano, 2020; A. PUNZI, *Il diritto e i nuovi orizzonti dell’intelligenza umana*, in *Analisi Giuridica dell’Economia*, 1/2019, p. 21 ss.

¹⁴ *Cfr.* in tema i vari modelli di descrizione delle relazioni tra dati, informazione, conoscenza H. ZECH, *Information as Property*, 6 (2015) JIPITEC, 192; G. COMANDÈ, *The Rotting Meat Error: From Galileo to Aristotle in Data Mining?*, in *European Data Protection Law Review*, 2018, 4, 3, p. 270 ss.

¹⁵ Sul funzionamento dei sistemi di intelligenza artificiale, tra cui il *machine learning*, si rinvia al saggio di G. ITALIANO, *Le sfide interdisciplinari dell’intelligenza artificiale* in A. NUZZO, G. OLIVIERI (a cura di) *Algoritmi. Se li conosci, li regoli...*, in *Analisi Giuridica dell’Economia*, 1/2019, p. 9 ss.

¹⁶ Per la differenza tra sistemi autonomi ed automatici *cfr.* G. COMANDÈ, *Responsabilità ed accountability nell’era dell’Intelligenza Artificiale*, in F. DI CIOMMO, O. TROIANO (a cura di), *Giurisprudenza e autorità indipendenti nell’epoca del diritto liquido. Studi in onore di Roberto Pardolesi*, 2018, La Tribuna, Piacenza, 1001-1015.

¹⁷ *Cfr.* da ultimo la Relazione della Commissione al Consiglio e al Parlamento Europeo, *Rela-*

industriale, commerciale e finanze, se non soprattutto, domestico e personale –. Si pensi agli elettrodomestici intelligenti, capaci di anticipare le scelte di consumo delle persone sulla base dell'osservazione delle abitudini dell'utente del bene fisico¹⁸, ai veicoli a guida autonoma che funzionano tramite l'integrazione in tempo reale dei dati captati dai propri sensori interni ed esterni, con i dati sul traffico dei servizi di mappe interattive, sulla mobilità urbana forniti dalle pubbliche amministrazioni, alla manutenzione predittiva delle macchine industriali o agricole basate in gran parte sui dati di funzionamento dei macchinari, o ancora agli assistenti virtuali ed ai dispositivi indossabili dai consumatori¹⁹, per la cui evoluzione è centrale la capacità di generazione e registrazione automatica dei dati (“dati generati dalle macchine”) e la successiva organizzazione e valorizzazione. L'utilizzazione di tali dispositivi da parte dei consumatori e degli utenti professionali, infatti, genera dati, che sono nel controllo del produttore del dispositivo (o del fornitore del servizio), che ha quindi a disposizione una vasta informazione da sfruttare commercialmente sul mercato. L'evoluzione dell'economia dei dati, tuttavia, mostra come molti altri soggetti possono avere un interesse qualificato all'uso – anche in tempo reale – di tali dati (consumatori, concorrenti, la pubblica amministrazione) sia per la manutenzione ed il buon funzionamento dei dispositivi o delle reti interconnesse²⁰, sia per competere in mercati collegati (ad esempio per offrire nuovi servizi a valore aggiunto) o per altre finalità (ad esempio, i dati raccolti da un'automobile sulle abitudini del conducente ai fini del perfezionamento dei servizi dell'auto potrebbero essere utili altresì ad una società assicurativa). Il carattere innovativo dello sfruttamento economico dei dati digitali consiste proprio nella capacità di incrociare, attraverso algoritmi, grandi insiemi di dati per rispondere a molte domande sulla base di mere correlazioni tra dati di natura, fonte, tipo differente²¹.

È evidente come in tale scenario i paradigmi di circolazione ed accesso alla conoscenza, come le tecniche di produzione e circolazione della ricchezza, appaiano diversi da quelli classici, definendo nuovi mercati e richiedendo un ripensamento sugli ordinamenti che ne regolano le dinamiche²². È altresì evidente

zione finale – Indagine settoriale sull'internet delle cose per i consumatori, del 20 gennaio 2022, COM(2022) 19 final. L'uso dei prodotti dell'Internet delle cose sta diventando sempre più rilevante per i consumatori (e.g. dispositivi domestici intelligenti, assistenti vocali, servizi, dispositivi indossabili). L'Internet delle cose, inoltre, è diffuso anche nei settori industriali.

¹⁸Per un ampio studio in tema cfr. G. NOTO LA DIEGA, *Internet of Things and the Law. Legal Strategies for Consumer-Centric Smart Technologies*, Routledge, Abingdon, New York, 2023; si veda anche T. FARKAS, *Data Created by the Internet of Things: The New Gold without Ownership*, in *Revista La Propiedad Inmaterial*, 2017, 23, p. 8.

¹⁹Per una riflessione in tema A. M. MATWYSHYN, *The Internet of Bodies*, in *Wm. & Mary L. Rev.*, 2019, 61(1), p. 77 ss.

²⁰Si pensi, ad esempio, alle auto a guida autonoma nell'ambito delle *smart cities*.

²¹J. DREXL, *Designing Competitive Markets for Industrial Data – Between Propertisation and Access*, cit., p. 262.

²²Per l'opzione metodologica di “ragionare di mercati, e dunque di leggi (o di sistemi norma-

come tale inarrestabile – ed ormai ineludibile – circolazione ed utilizzazione di dati, se da un lato consente applicazioni che offrono immense possibilità di sviluppo, dall'altro debba trovare forme di regolazione efficaci per prevenire ed eliminare possibili abusi.

L'interrogativo che si pone davanti a tali innovative tecniche di sfruttamento commerciale dei dati digitali concerne l'*an* ed il *quantum* di controllo del relativo flusso. Questo implica una riflessione sulla effettività e sui limiti della tutela delle banche di dati quali strumenti che attribuiscono diritti esclusivi sulle raccolte di dati in ragione rispettivamente dell'originalità della struttura, tramite il diritto d'autore, e della rilevanza dell'investimento in termini quantitativi e qualitativi per "il conseguimento, la verifica e la presentazione" dei contenuti, tramite il diritto *sui generis*. La funzione di incentivo allo sviluppo di sistemi di gestione e memorizzazione dei dati, espressa nella Direttiva 96/9/CE, che ha armonizzato la tutela autorale in Europa e dato i natali alla tutela *sui generis*, infatti, necessita di confermare la propria validità ed utilità davanti al mutamento delle determinanti economiche e tecnologiche descritte, nonché all'interferenza con le dinamiche competitive relative all'accumulo, accesso ed uso dei dati. La stessa Commissione europea, in sede di valutazione della Direttiva ha sostenuto come l'interazione della disciplina dei diritti esclusivi sulle banche di dati con l'economia dei dati non "è completamente chiara allo stato attuale e necessita di essere ulteriormente monitorata"²³.

Al fine di discutere le sfide giuridiche poste dal nuovo contesto, è necessario, tuttavia, tentare di definire la dimensione dei problemi che emergono, identificando l'oggetto di cui si discute e cogliendo le mutate dinamiche economiche sottese alla rivoluzione tecnologica in atto. Preliminare all'analisi giuridica è la necessità di comprendere più da vicino come e da chi è generato il valore economico nell'economia digitale tramite lo sfruttamento dei dati digitali.

2. Dati digitali e mercati dei dati digitali: un punto di partenza

Benché vi sia stata per lungo tempo – e ancora invero sussista – notevole incertezza sulla definizione giuridica di "informazione"²⁴, di "dati"²⁵, poi "di-

tivi) che li governano, partendo dai beni (in senso economico) che in essi vengono prodotti e scambiati" *cf.* G. OLIVIERI, *Dal mercato delle cose al mercato delle idee. Metamorfosi di un istituto*, in *Rivista delle società*, 2017, 4, p. 815-24.

²³ Queste le conclusioni della valutazione della Direttiva da parte della Commissione europea del 25 aprile 2018 (traduzione a cura dell'autrice, nella versione originale del testo: "The interaction of the *sui generis* right with the broader data economy is not fully clear at this stage and would need to be further monitored"). *Cfr.* Commission Staff Working Document, Evaluation of Directive 96/9/EC on the legal protection of databases, SWD(2018) 146 final.

²⁴ Occorre richiamare i tentativi civilistici dell'identificazione di uno (o più) "statuto (o statuti) dell'informazione". *Cfr.*, *ex multis* S. SCHAFF, *La nozione di informazione e la sua rilevanza giuridica*,

gitali”²⁶, il minimo comun denominatore dei più recenti atti normativi internazionali, che può essere assunto come punto di partenza sembra essere quello di “rappresentazione digitale di atti, fatti o informazioni, anche sotto forma di registrazione sonora, visiva o audiovisiva”²⁷ (seppur nella necessaria consape-

in *Dir. Inf.*, 1987; R. PARDOLESI, C. MOTTI, *L'informazione come bene*, in G. DE NOVA (a cura di), *Dalle res alla new properties*, Milano, 1991, p. 37 ss.; V. ZENO-ZENCOVICH, *Informazione (Profili Civiltistici)*, in *Digesto*, IV Edizione, Vol. IX Civile, Utet, 1993; p. 446 ss.; P. PERLINGIERI, *L'informazione come bene giuridico*, in P. PERLINGIERI, *Il diritto dei contratti fra persona e mercato*, Napoli, Edizioni Scientifiche Italiane, 2003, p. 353; S. ORLANDO, *Le informazioni*, Cedam, Padova, 2012; G. ALPA, C. SCOGNAMIGLIO, *Le informazioni sul regime dei diritti*, in *AIDA*, 2002, p. 272; A. DE FRANCESCHI, M. LEHMANN, *Data As Tradable Commodity and New Measures for Their Protection*, in *The Italian Law Journal*, Vol. 1, 2015, n. 1; in tema cfr. anche N. IRTI, *Il tessitore di Goethe (per la decisione robotica)*, in A. CARLEO (a cura di) *Decisione Robotica*, Il Mulino, Bologna, 2019, p. 19.

²⁵ Da un punto di vista etimologico, il termine “datum” richiama un elemento dalla cui osservazione, misurazione, ed elaborazione deriva un’informazione e, ancora, ad un livello di elaborazione e contestualizzazione più avanzato, una conoscenza. Molti Autori criticano finanche l’uso del termine “datum”, in quanto sarebbe più opportuno usare il termine “captum”: è ciò che è catturato, preso dalla natura, selezionato per una certa finalità, in altri termini, ciò che rileva, non ciò che è “dato”. Cfr. R. KITCHIN, *The Data Revolution. Big Data, Open Data, Data Infrastructures and Their Consequences*, Sage, 2014, § 1. Si veda anche la relazione di S. ORLANDO, “Data as Information and Data as Assets” alla Summer School *Digital Innovation and Sustainability*, Università di Napoli “L’Orientale”, 6-7 giugno 2022.

²⁶ Cfr. in tema l’ampio e ambizioso progetto dell’American Law Institute assieme all’European Law Institute. L.J. THOMAS, S. WEISE, N.B. COHEN, C. WENDEHORST, *ALI-ELI Principles for a Data Economy* (ELI Final Council Draft) (2021). Brooklyn Law School, Legal Studies Paper No. 719, reperibile su SSRN: <https://ssrn.com/abstract=4074993>.

²⁷ Tale definizione è fatta propria dai Regolamenti europei emanati nell’ambito della strategia sul mercato unico digitale, *infra*. Cfr., in particolare, Regolamento Ue 2022/868 del Parlamento Europeo e del Consiglio relativo alla governance europea dei dati del 25 novembre 2020 COM (2020) 767 final (Atto sulla governance dei dati, o *Data Governance Act*) all’art. 2 definisce i “dati”: “qualsiasi rappresentazione digitale di atti, fatti o informazioni e qualsiasi raccolta di tali atti, fatti o informazioni, anche sotto forma di registrazione sonora, visiva o audiovisiva”. La stessa definizione è fatta propria dall’art. 2 del 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, o nella più nota versione inglese *Digital Markets Act*); dalla Proposta di Regolamento del Parlamento Europeo e del Consiglio riguardante norme armonizzate sull’accesso equo ai dati e sul loro utilizzo, COM(2022) 68 final, del 23 febbraio 2022 (di seguito, Proposta di Data Act).

È interessante notare come tali definizioni includano anche le “raccolte di tali atti, fatti o informazioni”.

Cfr. anche Raccomandazione dell’OECD, *Recommendation of the Council on Enhancing Access to and Sharing of Data*, C/MIN(2021)20/FINAL, approvata il 6 ottobre 2021, dove vi è un accordo – per i fini della Raccomandazione – sulla seguente definizione: “‘Data refers to recorded information in structured or unstructured formats, including text, images, sound, and video’”. La Raccomandazione definisce anche le nozioni di “accesso ai dati”, “‘Data access’ or ‘access to data’ refers to the act of querying or retrieving data for its potential use, subject to applicable technical, financial, legal, or organizational access requirements. ‘Data sharing’ or ‘sharing of data’ refers to the act of providing data access for use by others, subject to applicable technical, financial, legal, or organisational use requirements”.

volezza del significato funzionale di ogni definizione giuridica²⁸).

In termini pratici, il dato digitale individualmente preso può quindi essere pensato come rappresentazione di una misura (di qualsiasi entità apprensibile) espressa in un formato leggibile da una macchina (secondo l'espressione di Herbert Zech, "*machine readable encoded information*"²⁹), che si può descrivere come una matrice composta da stringhe di numeri. Tale stringa, inoltre, è resa osservabile da un essere umano normalmente attraverso interfacce.

Al fine di meglio orientare la qualificazione giuridica di tale oggetto, alcuni Autori³⁰ suggeriscono di partire dalla semiotica nel distinguere tre livelli: il livello semantico dell'informazione (ovvero il significato o il messaggio³¹), il livello sintattico dell'informazione (la rappresentazione dell'informazione in particolari segni, il vettore dell'informazione³²) e il livello strutturale (il canale di comunicazione, il *medium*³³). Tale distinzione è utile per ricostruire le dinamiche del relativo sfruttamento ed identificare i soggetti capaci di accedere, disporre e utilizzare tali risorse. Ognuno dei tre livelli, infatti, da un punto di vista giuridico o economico, costituisce "*independent possibilities to define a certain amount of information*"³⁴. Inoltre, potrebbero essere diversi i soggetti "che detengono le chiavi di accesso a questi livelli (livello sintattico – titolarità del *software* di lettura, livello semantico – titolarità dell'informazione ricavata, livello del supporto fisico, proprietà del *physical carrier*)"³⁵, con la conseguenza che sui dati potrebbe sussistere un fascio di regimi proprietari e di circolazione interferenti, dalle finalità anche divergenti, facenti capo a una pluralità di soggetti. Altri Autori³⁶ descrivono in termini fattuali il "dato digitale" come entità immateriale, distinta dal dato come entità corporale formato da elettroni, soffermandosi sulla diffe-

²⁸ Cfr. P. SPADA, *Impresa*, in *Digesto delle discipline privatistiche, Sezione commerciale*, 7° Vol., Torino, 1992, p. 32-76.

²⁹ H. ZECH., *A legal framework for a data economy in the European Digital Single Market: rights to use data*, in *Journal of Intellectual Property Law & Practice*, 2016. Cfr. anche ID., *Data as tradeable commodity*, in A. DE FRANCESCHI (a cura di), *European Contract Law and the Digital Single Market*, Insentia, 2016, 51 ss.

³⁰ Cfr. ancora l'importante opera di H. ZECH, *Information as Property*, cit., p. 192; richiamata ed illustrata anche da J. DREXL, *Designing Competitive Markets for Industrial Data*, cit., p. 263 ss.

³¹ Tale qualificazione rileva ad esempio per l'applicabilità di regimi giuridici che si riferiscono alla natura dell'informazione rappresentata, come la tutela dei dati personali o la tutela del segreto commerciale.

³² In questo caso si tratta di informazioni o misure astratte dal significato semantico, come – per riprendere la definizione normativa – registrazioni sonore, visive o audiovisive.

³³ Il livello strutturale è il *corpus mechanicum* che contiene l'informazione (un libro, un cd, una memoria, un *cloud*).

³⁴ H. ZECH, *Information as Property*, cit., p. 194.

³⁵ Cfr. A. GALIANO, A. LEOGRANDE, S.F. MASSARI, A. MASSARO, *I dati non personali: la natura e il valore*, in *Rivista Italiana di Informatica e Diritto*, 1/2020, p. 63.

³⁶ A. OTTOLIA, *Big data e innovazione computazionale*, cit., p.15.

renza tra fruizione espressiva e fruizione computazionale dello stesso. In tal senso, il dato sarebbe “qualsiasi entità osservabile”³⁷ dalla macchina, con il corollario che – secondo tali opinioni – la relativa qualificazione possa solo identificarsi *ex post*³⁸.

Sebbene questa sia una premessa fondamentale per capire l'oggetto dello sfruttamento in parola, è bene rammentare sin da subito come ai fini della tutela delle banche di dati la disciplina giuridica è neutra rispetto al contenuto della raccolta (che potrebbe consistere di “opere, dati o altri elementi indipendenti”³⁹, a patto che siano “sistematicamente o metodicamente disposti ed individualmente accessibili grazie a mezzi elettronici o in altro modo”⁴⁰) e non si estende a tutelare i singoli elementi contenuti⁴¹. Ciò che rileva, da questo punto di vista è, invece, l'incentivo e la tutela alla costruzione e configurazione di banche di dati che contengono tale informazione, da un lato, e gli usi che possono essere fatti della raccolta e di parti sostanziali del relativo contenuto (in termini di riproduzione dell'intera banca di dati, modalità di estrazione dell'informazione ivi contenuta, comunicazione e reimpiego della stessa), dall'altro, in relazione alla creatività ed agli investimenti profusi per l'organizzazione e la disposizione di tali dati. Non saranno trattati, pertanto, i profili generali relativi ai regimi di appropriazione e circolazione dei dati digitali, se non in quanto funzionali a tale perimetro d'indagine. Come si vedrà, le nuove possibilità di analisi computazionale dei dati e lo sfruttamento che può derivarne delineano nuove dinamiche di concorrenza tra imprese e nuove collisioni di interessi, che appare opportuno cogliere preliminarmente all'analisi giuridica rispetto all'impatto sull'interpretazione ed applicazione della disciplina sui diritti esclusivi sulle banche di dati ed alle interferenze con altri corpi normativi.

2.1. La “filiera” dei macrodati digitali

Come descritto nel primo paragrafo, la raccolta, gestione ed elaborazione computazionale di grandi quantità di dati rappresenta la caratteristica più preminente ed il motore della “innovazione guidata dai dati”.

L'ampia indagine conoscitiva sui Big Data, svolta congiuntamente da AGCM, AGCOM e GPDP⁴², ha identificato la “filiera” o catena del valore dei “Big Data”⁴³ distinguendo le seguenti attività: “i) la raccolta, che a sua volta si

³⁷ A. OTTOLIA, *Big data e innovazione computazionale*, cit., p. 16.

³⁸ A. OTTOLIA, *Big data e innovazione computazionale*, cit., p. 16.

³⁹ Cfr. art. 1 Direttiva 96/9/CE.

⁴⁰ Cfr. art. 1 Direttiva 96/9/CE.

⁴¹ Cfr. art. 3 Direttiva 96/9/CE.

⁴² AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit.

⁴³ Anche la Raccomandazione Raccomandazione dell'OECD, *Recommendation of the Council on Enhancing Access to and Sharing of Data*, cit., p. 4, definisce la catena del valore nei seguenti

articola in generazione, acquisizione e memorizzazione, ii) l'elaborazione, che coinvolge attività di estrazione, integrazione e analisi, iii) l'interpretazione e l'utilizzo"⁴⁴. Tale indagine ampia e trasversale, i cui risultati spesso convergono e si integrano con quelli di altre indagini nazionali o internazionali⁴⁵, rappresenta una preziosa fonte per descrivere le attività che caratterizzano lo sfruttamento economico dei dati e per comprendere come tale sfruttamento abbia un impatto sull'interpretazione, l'applicazione, se non, addirittura, la riforma della disciplina giuridica delle banche di dati. Appare opportuno, pertanto, descrivere sinteticamente le fasi dello sfruttamento economico dei dati.

2.1.1. *La fase della raccolta dei dati*

La prima fase della filiera attiene alla raccolta dei dati, che potrebbero essere direttamente prodotti dall'impresa o raccolti presso altre fonti, se già esistenti, o generati *aliunde*.

I dati potrebbero essere, infatti, "generati"⁴⁶ direttamente dall'impresa attraverso misurazioni tramite sensori (ad esempio, i dati prodotti dal funzionamento dei macchinari relativi all'usura della macchina o a misurazioni di fattori esterni, come temperatura, stato del suolo nel caso dei macchinari agricoli), oppure attraverso un processo di elaborazione di altri dati (ad esempio, i dati relativi alla profilazione dei consumatori o alla manutenzione predittiva di un macchinario).

Anche l'osservazione e la misurazione delle attività degli utenti di servizi *online* e *offline*, o dei dispositivi intelligenti consente di generare una grande quantità di dati. Alcuni dati sono generati direttamente dagli utenti (intesi come consumatori o imprese), che li forniscono volontariamente ai servizi di cui fanno uso (si pensi ai contenuti resi accessibili sui *social network* o su altre *app* o anche

termini: "Data value cycle' refers to data-related processes through which value is created with data, including data creation, collection, validation, verification, storage, curation, enrichment, processing and analysis, access and sharing, and deletion".

⁴⁴ AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 8.

⁴⁵ Cfr., *ex multis*, AUTORITÉ DE LA CONCURRENCE, BUNDESKARTELLAMT, *Competition Law and Data*, 10 maggio 2016, reperibile all'indirizzo <https://www.autoritedelaconurrence.fr/sites/default/files/Big%20Data%20Papier.pdf>; European Commission Directorate General For Communications Networks, Content & Technology (Dg Connect), *European strategy on the data value chain*, novembre 2013; J. CRÉMER, Y.-A. DE MONTJOYE, H. SCHWEITZER, *Competition policy for the digital era – final report*, Lussemburgo, Publications Office of the European Union, 2019.

Cfr. altresì Investigation of Competition in Digital Markets (Majority Staff Report and Recommendations), of the U.S. Congress Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, Report on the Digital Economy, edited by the Global Antitrust Institute of the George Mason University; J. FURMAN, D. COYLE, A. FLETCHER, D. MCAULEY, P. MARSDEN, *Unlocking digital competition*, 2019; H. GILBERT MILLER, P. MORK, *From Data to Decisions: A Value Chain for Big Data*, in 15(1) *IT Professional*, 2013, p. 57-59; M.S. GAL, D.L. RUBINFELD, *Access Barriers to Big Data*, 59 *Arizona Law Review* 339 (2016); M. MAGGIOLINO, *I big data ed il diritto antitrust*, Egea, Milano, 2018.

⁴⁶ Si tornerà *infra*, cap. 2, § 4.1, sulla qualificazione dei "dati generati", al momento usato solo a fini descrittivi.

ai dati comunicati da imprese che si servono di servizi di intermediazione *online*), oppure semplicemente tramite l'interazione con un dispositivo (come nel caso della geolocalizzazione, dei dati ambientali o logistici)⁴⁷.

Ancora, i dati potrebbero essere raccolti o estratti da fonti esterne. Ad esempio, molti dati possono essere raccolti da fonti disponibili in Internet o nell'ambito dei rapporti tra imprese ed enti del settore pubblico (come nel caso dei settori interconnessi delle *smart cities*⁴⁸). Gli insiemi di dati sono anche tipicamente trasferiti nell'ambito di rapporti tra imprese⁴⁹, attraverso una pluralità di modelli⁵⁰, dai trasferimenti bilaterali ad assetti plurilaterali di condivisione e ge-

⁴⁷ Un'altra tassonomia in uso è quella che distingue i dati forniti volontariamente, i dati osservati ed i dati derivati o inferiti da altre analisi. Cfr. WORLD ECONOMIC FORUM, *Personal Data: The Emergence of a New Asset Class*, Gennaio 2011.

⁴⁸ Cfr. AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 12, che offre come esempio "una Smart City, in cui i cittadini attraverso un'applicazione presente nei propri smartphone hanno accesso in tempo reale ai dati sul traffico, sui parcheggi disponibili, sulla qualità dell'aria, sui tempi di attesa dei mezzi pubblici, sulle farmacie di turno aperte, sul numero di pazienti presenti nei pronto soccorsi. Tutto ciò grazie a sensori interconnessi, i quali trasmettono le proprie rilevazioni ad un server centrale che elabora e rende disponibili le informazioni ai propri utenti".

⁴⁹ I risultati della consultazione della Commissione europea, Directorate-General for Communications Networks, Content and Technology, condotta tra il 2018 ed il 2019, *SME panel consultation on B2B data-sharing principles and guidance, Report on the results*, p. 1, dimostra, ad esempio, come "a significant proportion of SMEs actively acquire data from other companies (33% of respondents – in line with results of other surveys) and are using (or plan to use) connected ('Internet of Things' – 'IoT') objects such as smart, industrial robots (30% of respondents)." La sempre più stretta interconnessione dei dati offre alle imprese l'opportunità (e talvolta la necessità) di utilizzare risorse raccolte o generate al di fuori della propria organizzazione imprenditoriale, secondo un paradigma emergente di "innovazione aperta e collaborativa" (sul tema cfr., in particolare, il fondamentale contributo di H.W. CHESBROUGH, *Open innovation: The new imperative for creating and profiting from technology*, Boston 2006, Harvard Business School Press; da una prospettiva giuridica, cfr. anche M. RICOLFI, *The New Paradigm of Creativity and Innovation and Its Corollaries For The Law Of Obligations*, in P. DRAHOS, G. GHIDINI, H. ULLRICH (a cura di), *Kritika: Essays on Intellectual Property*, Vol. I, Edward Elgar, Celtenham, 2015, p. 134 ss.).

⁵⁰ Alcuni recenti studi condotti dalla (e in favore della) Commissione europea propongono una classificazione abbastanza consolidata da una prospettiva economica dei possibili modelli commerciali di condivisione dei dati digitali, che ben evidenziano lo spettro degli interessi delle parti e che sembra opportuno brevemente richiamare. Cfr. Documento di Lavoro dei Servizi della Commissione, *Orientamenti sulla condivisione dei dati del settore privato nell'economia europea dei dati che accompagna il documento Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni "Verso uno spazio comune europeo dei dati" SWD/2018/125 final*; cfr. altresì lo studio European Commission, Directorate-General for Communications Networks, Content and Technology, E. SCARIA, A. BERGHMANS, M. PONT ET AL., *Study on data sharing between companies in Europe: final report*, Publications Office, 2018, <https://data.europa.eu/doi/10.2759/354943>. Si tratta di uno studio corposo, che ha considerato imprese di dimensioni diverse, operanti in sei diversi settori ("data-generating driving (i.e. automotive, transport and logistics), smart agriculture, smart manufacturing, telecom operators, smart living environments (i.e. home automation, sensors, robotics, or wearable technology), and smart grids & meters") in 31 Paesi. I risultati di tale studio sono stati ripresi anche dal successivo

stione collettiva dei dati⁵¹ (come i *data pools*⁵²), anche mediante l'intervento di intermediari⁵³.

studio in favore della Commissione europea, Directorate-General for Justice and Consumers, Autori, *Study on model contract terms and fairness control in data sharing and in cloud contracts and on data access rights, Final Report* JUST/2020/RCON/FW/CIVIL/0098, Aprile 2022. Cfr. anche il progetto “Data Governance” dell’Università von Humbolt di Berlino (<https://www.hiig.de/en/project/data-governance/>).

⁵¹Cfr. Documento di Lavoro dei Servizi della Commissione, *Orientamenti sulla condivisione dei dati del settore privato nell’economia europea dei dati che accompagna il documento Comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle regioni “Verso uno spazio comune europeo dei dati”*, cit., p. 5 ss. Ove vi sia un forte interesse ad un ampio riutilizzo dei dati può essere adottato un “approccio fondato su dati aperti”, in cui il soggetto “detentore” dei dati li mette a disposizione ad una “gamma aperta di (ri)utilizzatori con restrizioni minime, a titolo gratuito oppure dietro remunerazione minima”. Benché tale approccio non sia il più diffuso nell’ambito della condivisione dei dati tra imprese, potrebbe condurre al vantaggio di poter “usufruire di un ecosistema di sviluppatori di applicazioni di terze parti al fine di raggiungere i clienti finali”, oppure, come altri studi segnalano, potrebbe essere supportato da finalità di interesse generale, nell’ambito di modelli di responsabilità sociale di impresa. All’altro estremo dello spettro si collocano approcci volti alla “monetizzazione dei dati su un mercato di dati”, in cui le imprese condividono i dati con altre imprese (normalmente attraverso rapporti bilaterali a titolo oneroso) per ottenere una remunerazione ulteriore da tale condivisione. Dal lato della controparte, l’acquisto dei dati potrebbe essere utile come input per sviluppare beni o servizi da offrire sul mercato, per migliorare l’efficienza interna o, ancora, per rivendere i dati in formato aggregato. Tali scambi possono anche avvenire tramite “l’intermediazione di un mercato di dati” che può essere di interesse “per imprese che non conoscono potenziali riutilizzatori dei propri dati e che intendono intraprendere un’attività di monetizzazione dei dati una tantum. Tale meccanismo risulta idoneo nel caso in cui 1) sussistono rischi limitati di utilizzo illecito dei dati in questione, 2) il fornitore dei dati ha buoni motivi per fidarsi del (ri)utilizzatore, oppure 3) il fornitore dei dati dispone di meccanismi tecnici per prevenire o individuare gli usi illeciti”. Invece di avvenire in un mercato più aperto, lo scambio dei dati può anche configurarsi “su una piattaforma chiusa istituita da un attore chiave in un ambiente di condivisione dei dati oppure da un intermediario indipendente”. In questo caso i dati sono condivisi in cambio di remunerazione oppure di servizi a valore aggiunto erogati ad esempio all’interno della piattaforma. Tale meccanismo sembra essere particolarmente interessante per collaborazioni stabili tra imprese (si pensi ad imprese comuni, contratti di rete, associazioni di imprese, consorzi), che consentono anche maggiori meccanismi di controllo sulle finalità ed i modi di utilizzo dei dati). Tali modelli possono naturalmente combinarsi tra loro in relazione all’assetto di interessi delle parti, che si rispecchierà sia negli accordi contrattuali stipulati sia nelle soluzioni tecniche prescelte per la condivisione.

⁵²Gli *Ali-Eli Principles For A Data Economy – Data Transactions And Data Rights* (ELI Final Council Draft), p. 6, forniscono alcuni esempi di trasferimento o condivisione dei dati che possono verificarsi nella prassi.

Gli interessi delle parti potrebbero, infatti, dar vita a molteplici assetti. Ad esempio, sono diffusi i contratti in cui una parte concede all’altra l’accesso ai dati in uno spazio sicuro sul proprio *server* per far compiere agli algoritmi alcune operazioni. La prestazione dedotta in contratto potrebbe anche consistere nel consentire un accesso continuo e in tempo reale al flusso di dati raccolti da una parte, o – diversamente – nell’assunzione dell’obbligo di raccogliere dati automaticamente per trasmetterli alla controparte in uno specifico formato (si pensi ad esempio alla raccolta dati sugli *exit poll*). Frequente è anche l’obbligazione assunta da una parte di concedere accesso ai

Tali dati, raccolti da fonti variegatae, presentano le caratteristiche tipiche dei macrodati, ovvero “l’eterogeneità, la varietà, l’assenza di una struttura, la forte relazione spazio/tempo e la rapida crescita”⁵⁴. Come si vedrà in dettaglio nei successivi capitoli, comprendere se i dati siano generati nell’ambito dell’impresa, forniti da terzi, osservati o acquisiti *aliunde* è di importanza cruciale per valutazioni differenti, attinenti sia alla determinazione della sussistenza o meno della tutela esclusiva della banca di dati, sia alla valutazione se la banca dati rappresenti l’unica fonte di accesso ai dati, eventualmente essenziali per competere in un dato mercato⁵⁵.

Al fine di operare tali valutazioni, inoltre, occorre vagliare l’accessibilità giuridica e fattuale agli stessi, dal momento che i regimi incidenti sulla circolazione e sull’utilizzo di tali dati sono molteplici e dipendono dal tipo e dalla natura dei dati in questione⁵⁶. Tali fattori entrano in gioco, di volta in volta, ai fini del riconoscimento della tutela (ad esempio, per valutare gli investimenti necessari per la costituzione della banca di dati), ai fini della verifica della liceità o meno

dati ad un’altra parte che fornisce un servizio di elaborazione dei dati (come potrebbe avvenire, ad esempio, per le infrastrutture cloud, o per esternalizzare servizi di *big data analysis*).

Particolarmente rilevante è la tendenza al “data pooling”, ispirato a modelli di gestione collettiva dei diritti sui beni immateriali, come la gestione collettiva dei diritti d’autore e dei diritti connessi o i *patent pools*). Una descrizione di “data pool” è stata recentemente fornita nell’ambito di un procedimento antitrust della Commissione europea nei confronti di un ente assicurativo irlandese, poi conclusosi con impegni (Press Release, *Antitrust: Commission opens investigation into Insurance Ireland data pooling system*, 14 maggio 2019, reperibile all’indirizzo https://ec.europa.eu/commission/presscorner/detail/en/IP_19_2509): “a data pool is a sharing system between companies which involves an element of reciprocity, whereby at least some companies contribute to data”. Cfr. Commission Decision of 30.6.2022 relating to a proceeding under Article 101 of the Treaty on the Functioning of the European Union (TFEU) and Article 53 of the EEA Agreement. Case AT.40511 – *Insurance Ireland: Insurance claims database and conditions of access*.

La differenza rispetto ad altri schemi contrattuali consisterebbe, dunque, nell’elemento della reciprocità tra tutti i partecipanti al *pool* rispetto alla condivisione ed all’utilizzo di giacimenti di dati. Cfr. anche OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, 2018, p. 46 ss.

Secondo la dottrina che ha catalogato il fenomeno con riferimento all’ordinamento nazionale (A. OTTOLIA, *Big Data e Innovazione computazionale*, cit., p. 272 s.), è possibile distinguere il “data pool contratto”, che non costituisce alcun soggetto giuridico nuovo, da altri assetti organizzativi, come il “data pool soggetto”, volto a costituire una nuova società o un consorzio, o al “data pool patrimonio” (e.g. patrimonio destinato o “trust”). In tema, con specifico riferimento alle forme di *data pooling* di dati sanitari cfr. G. SCHNEIDER, *Health Data Pools Under European Data Protection and Competition Law, Health as a Digital Business*, Springer, Berlino, 2022.

⁵³ Cfr. AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 14, la quale fa riferimento ai “data broker”, i quali raccolgono e, soprattutto, “aggregano dati da diverse fonti (principalmente siti internet) e li organizzano per metterli a disposizione di soggetti terzi”, alimentando – tuttavia – “un mercato poco trasparente soprattutto per gli utenti finali”.

⁵⁴ AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 12.

⁵⁵ Cfr. J. CRÉMER, Y.A. DE MONTJOYE, H. SCHWEITZER, *Competition Policy for the Digital Era*, cit.

⁵⁶ Cfr. ad esempio la tassonomia proposta da OECD, *Enhancing Access to and Sharing of Data: Reconciling Risks and Benefits for Data Re-use across Societies*, cit., § 2.

delle varie utilizzazioni, nonché ai fini della valutazione circa l'esistenza di eventuali preclusioni concorrenziali.

Se l'informazione veicolata da tali dati rientra nell'ambito di applicazione della tutela dei dati personali, ad esempio, la relativa disciplina ne conforma la circolazione a tutela della persona e l'impresa potrà accedere e trattare i dati solo osservando le misure ivi previste. A fini di semplificazione, pertanto, si suole distinguere i "dati personali" dai "dati non personali", dove la definizione di dato personale si rinviene – quantomeno per l'ordinamento europeo – dal Regolamento generale sulla protezione dei dati personali⁵⁷ (RGPD), e dalle discipline speciali per le attività *online*⁵⁸, il quale definisce il dato personale "qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»)"⁵⁹. La categoria dei dati definiti "non personali" è, invece, residuale e sostanzialmente identificata in negativo, come conferma il Regolamento 2018/1807 sulla libera circolazione dei dati non personali⁶⁰.

Un'altra tassonomia rilevante ai fini del regime di accessibilità consiste nella distinzione tra "dati del settore pubblico" e "dati del settore privato". Per i primi, l'interesse generale alla circolazione ed alla fruizione da parte della collettività è prevalente rispetto ad altri interessi e dunque i relativi regimi di accesso e riutilizzo tendono ad essere più aperti⁶¹. Per i dati nella disponibilità del-

⁵⁷ Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati, di seguito RGPD).

⁵⁸ Cfr. Direttiva 2002/58/CE relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche. Quest'ultima è in corso di revisione, cfr. Proposta di Regolamento del Parlamento Europeo e del Consiglio relativo al rispetto della vita privata e alla tutela dei dati personali nelle comunicazioni elettroniche e che abroga la direttiva 2002/58/CE (regolamento sulla vita privata e le comunicazioni elettroniche), COM/2017/010 final.

⁵⁹ Cfr., *infra*, cap. 3, § 4.1. Si considera identificabile "la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale"; cfr. art.4 (1) RGPD.

⁶⁰ Regolamento (UE) 2018/1807 del Parlamento Europeo e del Consiglio del 14 novembre 2018 relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea. Per la definizione dei dati non personali cfr. art. 3. 1: "i dati diversi dai dati personali definiti all'articolo 4, punto 1, del regolamento (UE) 2016/679". Per una articolata critica a tale scelta definitoria cfr. in particolare M.L. MONTAGNANI, *La libera circolazione dei dati al bivio. Tra tutela dei dati personali e promozione dell'intelligenza artificiale europea*, in *Mercato Concorrenza Regole*, 2, 2019, p. 239. La distinzione tra "dato personale" e "dato non personale", tuttavia, non è agevole, in quanto un dato non personale potrebbe divenire personale in relazione al contesto ed alla finalità in cui è utilizzato (ad esempio, un dato non personale incrociato con un altro dato in uno specifico contesto, potrebbe comunque riferirsi ad una persona) e data la prassi di analizzare grandi volumi di dati, agli insiemi di dati misti finisce per applicarsi comunque la disciplina a tutela dei dati personali.

⁶¹ Ad esempio, alla circolazione dei dati del settore pubblico si applicano le regole sulla Diret-

le imprese, invece, potrebbero insistere altri diritti che ne limitano la circolazione: l'informazione ivi veicolata potrebbe essere tutelata dal segreto oppure i dati potrebbero essere contenuti in banche di dati e soggetti ai diritti esclusivi ivi disciplinati, che potrebbero limitare l'accesso, l'estrazione o il reimpiego dei dati (*infra*, cap. 2).

Anche i tempi ed i modi dell'accesso sono particolarmente rilevanti, in quanto, ad esempio, la possibilità di accedere in tempo reale, dinamico, o sistematico ai dati sono fattori che possono incidere sul valore degli stessi e sulle valutazioni giuridiche da operare sui regimi giuridici di appartenenza e circolazione. Determinanti sono, inoltre, le finalità di accesso ed uso dei dati, in quanto la relativa precisa identificazione consente di determinare le dinamiche competitive in gioco. A tal proposito alcuni principali studi in ambito europeo⁶² distinguono, anzitutto, i dati utilizzati a livello individuale, ovvero riferiti ad un determinato individuo o ad un determinato bene. Tali dati possono essere usati in modo non anonimo, quando l'uso è finalizzato a fornire direttamente un servizio alla persona o all'impresa a cui i dati si riferiscono⁶³, o anonimo, quando l'accesso a livello individuale serve per una finalità non necessariamente connessa al bene o al soggetto a cui i dati si riferiscono⁶⁴. In tale ultimo caso, l'accesso a grandi insiemi di dati utilizzati a livello individuale anonimo (e dunque un fascio di dati a livello individuale) potrebbe essere essenziale per una concorrenza effettiva. I dati, inoltre, potrebbero essere utilizzati in modo (irreversibilmente) aggregato, ovvero standardizzati a tal punto che l'accesso al livello individuale non sia più necessario. Esempi di tal sorta possono essere i dati sulle vendite, o i dati statistici. Infine, i dati "contestuali" non derivano da una raccolta a livello individuale, e sono utilizzati come tali, come i dati satellitari, le mappe, i dati stradali.

tiva sull'apertura dei dati e sulla riutilizzazione dell'informazione del settore pubblico (Direttiva (UE) 2019/1024 del Parlamento Europeo e del Consiglio del 20 giugno 2019 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico (*infra*, cap. 3, § 3), mentre per la circolazione dei dati del settore privato si riepande l'autonomia privata conformata ai vari regimi giuridici di volta in volta applicabili.

⁶²Per la classificazione dei diversi scenari di utilizzazione *cf.* J. CRÉMER, Y.-A. DE MONTJOYE, H. SCHWEITZER, *Competition Policy for the Digital Era*, cit., p. 24 ss. La stessa classificazione è condivisa recentemente da un importante studio della Commissione europea in materia di tutela dei segreti commerciali nell'economia dei dati. *Cfr.* A. RADAUER, M. BADER, T. APLIN, U. KONOPKA, N. SEARLE, R. ALTENBURGER, C. BACHNER, *Study On The Legal Protection Of Trade Secrets In The Context Of The Data Economy (Gro/Sme/20/F/206)*, Final Report, luglio 2022.

⁶³Ad esempio, l'utilizzazione dei dati di un determinato individuo per finalità di raccomandazione pubblicitaria o anche a fini di portabilità dei dati personali da un servizio all'altro, o l'utilizzazione dei dati riferiti ad un determinato dispositivo per offrire servizi di monitoraggio e manutenzione. *Cfr. infra*, cap. 3, § 4.1.

⁶⁴Un esempio potrebbe consistere nell'uso di dati raccolti per finalità diverse ed utilizzati a livello individuale per addestrare un algoritmo di *machine learning* di diagnosi medica, per cui rileva non tanto – o non solo – il fatto che i dati si riferiscano ad un determinato individuo, quanto la possibilità di analizzare grandi volumi di dati dello stesso tipo per inferire correlazioni e sviluppare l'algoritmo in relazione al risultato atteso.

Da tali distinzioni emerge come i modi, i tempi ed il tipo di accesso ai dati può dipendere dalle (infinite) finalità di utilizzazione degli stessi. Per i fini che qui rilevano, l'accesso ai dati tra imprese potrebbe avvenire (i) per competere in mercati contigui⁶⁵ (ad esempio per offrire beni o servizi complementari), per cui l'accesso continuo a dati (non anonimi o anonimi⁶⁶) a livello individuale, e magari in tempo reale, è essenziale per porre in essere l'attività, come potrebbe avvenire per servizi a valore aggiunto, connessi all'uso di un dispositivo o di un servizio di *social network* da parte di un consumatore, o servizi di manutenzione per un macchinario; (ii) per competere negli stessi mercati, e dunque in una situazione di concorrenza diretta, seppur potenziale (si pensi, ad esempio, a un concorrente entrante nel mercato dei motori di ricerca, che necessita l'accesso ai dati dell'*incumbent* ai fini di una concorrenza effettiva); o (iii) per finalità scollegate rispetto all'ambito di attività dell'impresa che controlla i dati, come può avvenire per l'accesso a dati individuali (anonimi o non anonimi), aggregati o contestuali per finalità di addestramento di algoritmi di *machine learning*, che necessitano di apprendere sulla base del trattamento di grandi volumi di dati selezionati a tal fine, o per scopi di ricerca.

L'importanza di tali distinzioni concettuali – sebbene semplificate – aiuta a comprendere come l'accesso e l'uso dei dati, collocabile nella prima fase della filiera del relativo sfruttamento, rilevino per le imprese per innovare e competere. Tali distinzioni saranno utili per discutere l'ambito della tutela dei diritti esclusivi sulle banche di dati rispetto agli interessi all'accesso ed alla riutilizzazione dei dati da parte dei terzi (utenti o concorrenti).

In tale scenario, infine, è essenziale comprendere quali siano gli sforzi e gli investimenti profusi dalle imprese in tale fase di “formazione” dei macrodati e quali di queste attività siano rilevanti per la tutela giuridica delle banche di dati. Alcuni studi dell'OECD sono molto utili a comprendere le tendenze di mercato, sia a livello globale che europeo. Anzitutto, dal momento che nell'attuale scenario il valore dei dati cresce in relazione alla aggregazione ed all'incrocio di dati di diverso tipo, “le imprese hanno incentivi a raccogliere dati anche in assenza di piani immediati sul relativo utilizzo”⁶⁷ ed anche se l'oggetto dell'impresa non si basa su attività guidate dai dati⁶⁸. Anche i costi relativamente più bassi rispetto al passato di tecnologie di raccolta (dispositivi connessi e *social media*), me-

⁶⁵ Per una riflessione sull'uso dei dati nei mercati a valle, o negli *aftermarkets*, cfr. *infra*, cap. 3, spec. § 2.

⁶⁶ Il Report a cura di J. CRÉMER, Y.-A. DE MONTJOYE, H. SCHWEITZER, *Competition Policy for the Digital Era*, cit., p. 75, menziona l'esempio della concorrenza nella manutenzione di aerei. I dati individuali sul singolo aeroplano possono essere essenziali per offrire il servizio relativo a quello specifico aeroplano, ma i dati aggregati o anonimi sull'insieme degli aerei dello stesso tipo potrebbe servire per offrire servizi a valore aggiunto di manutenzione predittiva, in quanto aumentano la conoscenza dei possibili guasti che potrebbero verificarsi.

⁶⁷ Lo conferma lo studio dell'OECD (2022), *Data shaping firms and markets*, cit., p. 9.

⁶⁸ OECD (2022), *Data shaping firms and markets*, cit., p. 10.

morizzazione (*cloud*) e risorse computazionali (algoritmi di intelligenza artificiale ed analisi *big data*) facilitano in teoria le attività raccolta dei dati e le successive fasi dello sfruttamento. Tuttavia, i dati digitali sono fonte di valore una volta “ripuliti, strutturati ed elaborati” e sono largamente complementari rispetto all’investimento in tecnologia effettuato dall’impresa. Gli studi più recenti evidenziano come a fronte di ingenti costi fissi ed investimenti iniziali, i costi marginali associati all’uso dei dati sono molto bassi. Anzi, come si vedrà *infra*, le caratteristiche economiche dei dati e dell’assetto dei mercati denotano come le imprese che per prime hanno investito in raccolta dei dati e in tecnologie di analisi – e che su tali risorse abbiano costruito nel tempo “ecosistemi” sempre più ramificati – beneficino di effetti che tendono ad un accumulo maggiore di dati ed alla capacità di trarne profitti, con tendenze di mercato volte alla concentrazione⁶⁹. Occorre, tuttavia, procedere con ordine nel ricostruire le fasi della filiera.

Una volta generati ed acquisiti nei tempi e nei modi descritti, i dati digitali sono memorizzati (e dunque trasferiti dagli strumenti di acquisizione a quelli di memorizzazione) nelle memorie a breve o lungo termine dei sistemi di elaborazione, o su *cloud*, per consentire la successiva attività di elaborazione. In tali fasi assumono rilevanza gli investimenti in integrità e sicurezza dei dati⁷⁰.

2.1.2. La fase della elaborazione

La fase dell’elaborazione dei dati consiste in una primaria fase di organizzazione dei dati grezzi attraverso attività di “estrazione durante la quale i dati vengono reperiti dalle diverse fonti disponibili, selezionati e caricati nella memoria del sistema di elaborazione”⁷¹ e di “integrazione di tutte le informazioni che si riferiscono agli stessi elementi o domini applicativi”⁷² per consentirne la successiva “analisi” tramite algoritmi al fine di produrre risultati interpretabili per la successiva utilizzazione⁷³. In tale fase potrebbero avvenire operazioni di copia

⁶⁹ OECD (2022), *Data shaping firms and markets*, cit., ne fa discendere peraltro un forte iato tra le grandi imprese e le piccole e medie imprese in termini di capacità di trasformare i dati in produttività. Cfr., in tema anche alcuni studi di diritto della concorrenza dedicati alle fasi di formazione dei Big data, M.S. GAL, D.L. RUBINFELD, *Access Barriers to Big Data*, cit.; M. MAGGIOLINO, *I big data e il diritto antitrust*, cit., p. 179.

⁷⁰ AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 14.

⁷¹ Tale attività ha di recente trovato anche una definizione normativa. L’utilizzazione algoritmica dei dati è stata recentemente definita nel linguaggio legislativo come *Text And Data Mining*, ovvero “qualsiasi tecnica di analisi automatizzata volta ad analizzare testi e dati in formato digitale avente lo scopo di generare informazioni inclusi, a titolo non esaustivo, modelli, tendenze e correlazioni”. Questa è la definizione di Text e data mining contenuta all’art. 2.2 della Direttiva (UE) 2019/790 sul diritto d’autore nel mercato unico digitale (*infra*, cap. 2, § 5).

⁷² AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 15.

⁷³ Sui profili rivoluzionari (e anche problematici) di tali metodi di produzione della conoscenza cfr. G. COMANDÈ, *The Rotting Meat Error: From Galileo to Aristotle in Data Mining?*, cit., p. 270 e ss. Sia consentito anche il riferimento a S. SCALZINI, *Big Data e Integrità nella Ricerca: Un Punto Di Partenza*, in *The Future of Science and Ethics*, 1, 2, 2016, p. 18 ss.

ed estrazione dei dati, ivi inclusa una fase preliminare di “*pre-processing*” ossia di “normalizzazione” dei formati dei dati, di ripulitura dei dati non ritenuti rilevanti, di categorizzazione dei dati, del loro eventuale caricamento su una piattaforma unica e così via⁷⁴. La fase di elaborazione vera e propria dei dati (*data mining/pattern discovery*) per raggiungere l’*output* desiderato (in relazione al sistema utilizzato⁷⁵), inoltre, può altresì includere una fase di ricombinazione dei dati per ottenere e visualizzare i risultati dell’elaborazione.

Anche tale attività di “sistematizzazione”, pertanto, non può dirsi “neutra o casuale”⁷⁶, riflettendo spesso scelte (e investimenti) a monte “di struttura parziale e selettiva”⁷⁷ dei dati che vengono analizzati, al fine di renderli omogenei ed idonei alla successiva analisi automatizzata⁷⁸, per cui è necessario riflettere sulla sussistenza e sul rilievo – come si vedrà nel dettaglio *infra* – della tutela accordata dalla disciplina sulle banche di dati. È necessario aggiungere, inoltre, che anche tali fasi preparatorie sono tipicamente compiute tramite algoritmi impostati per la selezione e l’organizzazione dei dati.

⁷⁴ Cfr. per tale ricostruzione al fine di analizzare l’utilizzazione computazionale dei dati in relazione al diritto d’autore ed ai diritti connessi J.-P. TRIAILLE, J. DE MEEÛS D’ARGENTEUIL, DE FRANQUEN, *Study on the Legal Framework on Text and Data Mining* (TDM), 2014, European Commission, reperibile all’indirizzo <https://publications.europa.eu/en/publication-detail/-/publication/074ddf78-01e9-4a1d-9895-65290705e2a5/language-en>; M.L. MONTAGNANI, G. AIME, *Il text and data mining e il diritto d’autore*, in AIDA, 2017, p. 376-394; C. GEIGER, G. FROSIO, O. BULAYENKO, *Text and Data Mining in the Proposed Copyright Reform: making EU Ready for an Age of Big Data?*, in IIC 49, 2018, p. 814-844; E. ROSATI, *The Exception for Text and Data Mining (TDM) in the Proposed Directive on Copyright in the Digital Single Market – Technical Aspects, Briefing per il Policy Department for Citizens’ Rights and Constitutional Affairs del Parlamento europeo* (PE604.942), 2018.

⁷⁵ Si distinguono approcci solo statistici capaci di trovare correlazioni statistiche tra dati di diversa natura da sistemi di *machine learning*. In tali casi, ad esempio, vi è una ulteriore fase successiva all’elaborazione che consiste nella creazione di modelli di addestramento derivanti dalle correlazioni rinvenute dall’elaborazione precedente che vengono poi applicati nuovamente per testare altri dati, si veda, in particolare, G.F. ITALIANO, *Le sfide interdisciplinari dell’intelligenza artificiale*, cit., p. 13.

⁷⁶ Così V. FALCE, *L’insostenibile Leggerezza” delle Regole sulle Banche Dati nell’Unione dell’innovazione*, in *Rivista di Diritto Industriale*, 4, 1, 2018, p. 377, § 2; simili considerazioni sono espresse anche in V. FALCE, *Uses and abuses of database rights*, in P. DRAHOS, G. GHIDINI, H. ULLRICH (a cura di), *Kritika: Essays on Intellectual Property*, Edward Elgar, 2020, p. 180-202.

⁷⁷ Così V. FALCE, *L’insostenibile Leggerezza” delle Regole sulle Banche Dati nell’Unione dell’innovazione*, cit., § 2.

⁷⁸ Ad esempio, nell’ambito della agricoltura digitale “che ha ad oggetto la raccolta di dati e informazioni sulle aziende agricole allo scopo di fornire a queste ultime servizi personalizzati o aggregati (come servizi di consulenza sui metodi di coltivazione)”, la concentrazione *Bayer-Monsanto* esaminata dalla Commissione europea ha reso evidente come “la disponibilità di dati in quanto tale non è da sola sufficiente per poter fornire prescrizioni agronomiche digitali: occorre infatti elaborare un processo per “ripulire” i dati grezzi, un algoritmo per modellare e prevedere un dato fenomeno e un software in grado di combinare i dati e fornire quindi un servizio maggiormente competitivo”. Cfr. AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 74. In tema, *infra*, cap. 3, § 2.7.

Gli algoritmi utilizzati per l'analisi possono essere tipicamente di "interrogazione" ad una precisa domanda dell'utente o di "apprendimento", tramite le menzionate sofisticate tecniche di intelligenza artificiale e di *machine learning*, che consentono di inferire correlazioni inaspettate, rivelare tendenze ed estrapolare nuova conoscenza. Tali algoritmi sono generalmente proprietari o segreti⁷⁹, in quanto sviluppati secondo le necessità delle varie imprese, sebbene sulla base di linguaggi generalmente pubblicamente disponibili. Essi sono sempre più raffinati e precisi nei risultati perseguiti quanto più sono allenati con *dataset* di qualità e con dati raccolti da fonti diversificate⁸⁰, per cui la possibilità di accedere ai dati digitali si conferma, di nuovo, di primaria importanza anche per tale fase di sfruttamento. È stato, inoltre, sottolineato come la disponibilità di dati in grande quantità, varietà e qualità sia spesso difficilmente replicabile da parte della generalità delle imprese e come sia proprio la disponibilità dei dati e la relativa qualità (*i.e.* anche il modo in cui sono raccolti, selezionati, verificati, presentati), piuttosto che l'innovazione dei modelli interpretativi, a costituire spesso il vero vantaggio competitivo. "A partire da una grande e variegata mole di dati, gli algoritmi di intelligenza artificiale sono in grado di individuare complessi schemi di relazioni che possono sfuggire ai ricercatori (umani)"⁸¹. Attraverso l'ausilio dei dati è altresì possibile monitorare ed analizzare le interpretazioni e decisioni adottate sulla base dell'analisi computazionale⁸².

⁷⁹ Riflettono su tale profilo M. RICOLFI, *Il futuro della proprietà intellettuale nella società algoritmica*, cit., p. 24; N. ABRIANI, G. SCHNEIDER, *Diritto delle imprese e intelligenza artificiale*, cit., cap. 2, § 5; sebbene nel contesto dei prodotti inventivi dell'intelligenza artificiale, M. LIBERTINI, *Prodotti inventivi dell'intelligenza artificiale*, in A. PAJNO, F. DONATI, A. PERRUCCI, *Intelligenza Artificiale e Diritto: una rivoluzione?*, cit., Vol. 3, p. 89.

⁸⁰ AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 17, sulla base delle audizioni delle imprese nel settore ICT.

⁸¹ AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 18. *Cfr.* anche le considerazioni a p. 17.

⁸² L'indagine conoscitiva congiunta di AGCM, AGCOM, GPDP, *Indagine Conoscitiva sui Big Data*, cit., p. 18 ha sottolineato come i grandi operatori dell'ICT godono al momento di un vantaggio dovuto al fatto che sono da tempo sul mercato. Per le imprese tradizionali l'investimento richiesto per *hardware* e *software* non è particolarmente proibitivo, in quanto esse possono altresì avvalersi di servizi e soluzioni informatiche mediante contratti di appalto o di servizi, sebbene, come sottolinea l'analisi conoscitiva congiunta "i soggetti dai quali è allo stato possibile acquisire i suddetti servizi sono proprio i grandi operatori dell'ICT" (p.18). Un aspetto che non viene, invece, considerato dalla IC, né da altri studi in materia (per uno studio che, invero, da tempo ha osservato tale profilo *cfr.* A. MANTELETO, *Il costo della privacy tra valore della persona e ragione d'impresa*, Giuffrè, Milano, 2007), sono i costi aggiuntivi per le imprese derivanti dalla *compliance* normativa per la gestione dei dati. Sebbene alcune attività computazionali siano gestite in *outsourcing*, infatti, nessuna impresa può esimersi dal dotarsi di una esperienza interdisciplinare interna (o esterna) per prevenire rischi giuridici e responsabilità nella gestione dei dati (personali e non personali) nell'ambito della propria attività di impresa. Tale attività è particolarmente complessa e può anche essere onerosa, in quanto necessita di un'attività di *compliance* integrata che consideri differenti regimi che necessariamente interagiscono ed interferiscono creando problemi interpretativi ed incertezza nelle strategie commerciali.