LUISS

Dipartimento di GIURISPRUDENZA

Cattedra di DIRITTO AMMINISTRATIVO 1

Circolazione ed organizzazione dei dati pubblici: tra banche dati e *Big Data*

Chir.mo Prof. Aldo Sandulli

RELATORE

Chiar.mo Prof. Aristide Police

CORRELATORE

Antonio Tolentino
Matr. 153853

CANDIDATO

Anno accademico 2022/2023

INDICE

INTRODUZIONE

"*In ancient Greece, when anyone needed to make a big decision on life's most important questions, they all consulted the oracle …. well, we have a new oracle, and it's name is big data*" [1]

La pubblica amministrazione è in grado di sfruttare le potenzialità correlate ai *Big Data*? È questa la domanda cui si è sostanzialmente cercato di dare una risposta nella presente analisi.

Prima di illustrare l'iter logico che è stato seguito nell'elaborato, occorre motivare le ragioni per cui si è voluta focalizzare l'analisi sulla tematica dei *Big Data* e sulle problematiche connesse alla loro gestione e inquadramento a livello normativo, adottando, per tutto il corpo della tesi, una prospettiva di carattere principalmente pubblicistico.

Lo studio parte dalla presa di coscienza che ormai facciamo parte della società dell'informazione, in cui ognuno di noi è connesso all'infosfera digitale ed ogni aspetto delle nostre vite è suscettibile di essere "*datizzato*"[2] a grandi velocità.

In tal modo chiunque, consapevolmente o inconsapevolmente, è produttore e consumatore di dati e contribuisce ad alimentare il "mercato dei dati".

Tutti questi dati costituiscono un valore immenso per la società dell'informazione tanto che nell'aprile del 2017 la copertina dell'Economist affermava che "*la risorsa più preziosa al mondo non è il petrolio, sono i dati*"[3].

---

[1] Con queste parole Tricia Wang ha aperto il suo Ted Talk nel 2016 a Cambridge, si veda https://www.ted.com/talks/tricia_wang_the_human_insights_missing_from_big_data?language=it.

[2] "Datizzare" un fenomeno significa convertirlo in forma quantitativa in modo da poterlo tabulare e analizzare. Tra i primi a parlare del fenomeno V.M. Schönberger, K. Cukier, *Big Data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, V ed., Milano, Garzanti, 2013. Tale processo ed il numero della quantità di dati sono destinati ad aumentare, anche con l'avvento di nuove tecnologie, come l'*Internet of Things* o il 5G.

[3] A. Nicita e M. Delmastro, *Big data. Come stanno cambiando il nostro mondo*, Bologna, Il Mulino, 2019, p.7. Intendendosi con ciò che i dati sono oramai il nuovo "carburante" dell'economia. Per

Per gestire una tale quantità è necessario prima di tutto poter organizzare tali dati e, a questo fine, l'implementazione di banche dati risulta una risposta efficace.

Il patrimonio informativo gestito dalle pubbliche amministrazioni, espresso in termini di basi di dati e *dataset*, ha assunto nel tempo un ruolo rilevante e significativo[4]. Esso, per ciò solo, costituisce un elemento fondamentale per i cittadini, le imprese e le pubbliche amministrazioni. Una caratteristica tanto importante da indurre a pensare che sia nata una nuova funzione amministrativa: "*la funzione amministrativa dei dati, volta alla organizzazione, gestione e fruizione degli stessi*"[5].

Ma come si gestisce e come si estrae valore da una realtà datizzata, composta da *Big Data*?

A tal fine sarà necessario implementare delle soluzioni normative ed organizzative (di immagazzinamento e di analisi dei dati - c.d. *Big Data Analytics* -), in grado di sfruttare le potenzialità dei Big Data. Per orientarsi in questo mare di informazioni è infatti necessario un lavoro sinergico tra banche dati ed algoritmi, poiché questa collaborazione permette di trasformare grandi quantità di dati in informazioni utili, guidando decisioni efficaci in vari settori, dall'economia alla salute, dalla sicurezza pubblica all'istruzione. Tale lavoro sinergico permette lo sviluppo di una società effettivamente capace di sfruttare le potenzialità dei dati a vantaggio del potere conoscitivo, la c.d. *data driven society*. Il tutto agevolato ed incrementato grazie all'aumento della capacità computazionale ed alle tecniche di immagazzinamento dei dati. Sicché alla metafora dei dati intesi come il nuovo petrolio del ventunesimo secolo, occorre aggiungere la

---

il riferimento al *The Economist* si può vedere K. Bhageshpur, *The world's most valuable resource is no longer oil, but data*, The Economist, 6 maggio 2017.

[4] Nel 2017, il valore totale delle ISP (informazioni del settore pubblico) nell'Unione Europea ha toccato un massimo di 52 miliardi di euro, con previsioni che indicano un possibile aumento fino a 194 miliardi di euro entro il 2030; in tal senso O. Borgogno, *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I.*, in *Il diritto dell'informazione e dell'informatica*, n. 3/2019, pp.700-701.

[5] G. Carullo, *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa*, Torino, Giappichelli Editore, 2018, p.36.

precisazione che è cambiato sia il motore della macchina (capacità computazionale), sia il serbatoio del carburante (metodologie per acquisire ed immagazzinare i dati)[6].

A fronte di quest'ingente quantità di dati e delle relative tecniche di analisi, si torna alla domanda iniziale: è in grado la pubblica amministrazione di sfruttare i *Big Data*? È in grado di orientarsi nel mare di dati?

Lo sfruttamento di tali potenzialità permetterebbe alla pubblica amministrazione di incrementare il potere conoscitivo nell'ottica di un modello di pubblica amministrazione fruitrice dei dati, e non solamente di intermediazione di questi.

La presente analisi si è focalizzata in particolare sui presupposti che permettono di creare un "ecosistema" favorevole alle moderne tecniche di analisi dei dati (come, ad esempio, l'intelligenza artificiale, il *machine learning*). Ed in quest'ecosistema un ruolo rilevante è ricoperto dai *Big Data*, perché la quantità di informazioni è necessaria per lo sviluppo di queste tecniche di analisi. Quindi nel lavoro di ricerca si ragionerà sulla possibilità per le pubbliche amministrazioni di sviluppare tale ecosistema, che richiede una centralizzazione ed un'ampia circolazione di dati. Di contro, le tecniche di analisi dei dati saranno studiate solo in quanto connesse ai dati.

Prima di descrivere l'impostazione adottata nella presente analisi, va fatta un'ultima considerazione riguardante l'attuale contesto tecnologico nazionale.

In quello che è stato definito uno "Stato digitale minimo"[7], in cui appare assolutamente necessaria una preliminare alfabetizzazione digitale, un potenziamento della connettività e della digitalizzazione in generale, può apparire prematura un'analisi volta ad esaminare la capacità di sfruttamento dei *Big Data*. Tuttavia, nel compiere il primo

---

[6] Metafora liberamente adatta a partire da U. Pagallo, *Big data, open data e black box society*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, p.49.

[7] Espressione rinvenibile in M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e Big Data*, Napoli, Editoriale scientifica, 2023, p.118. L'arretratezza del sistema italiano rispetto alle tecnologie innovative è sottolineata diverse volte in dottrina, ad esempio, in F. Merloni, *Data analysis e capacità conoscitive delle pubbliche amministrazioni*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p. 115, viene affermato che "*non ci si può occupare di questo profilo, cioè di come costruiamo le conoscenze tecniche informatiche delle amministrazioni del futuro, senza rimontare lo svantaggio che abbiamo accumulato*".

passo verso la digitalizzazione, non si può non guardare al tragitto che si sta percorrendo: pena la possibilità di incorrere in soluzioni organizzative che in seguito si rivelino inadeguate rispetto agli obiettivi futuri (si pensi all'architettura deconcentrata dei *data center* – c.d. a *data silos* – impostata a partire dagli anni Settanta, con l'obiettivo di garantire una maggiore tutela dei dati personali e che adesso pare essere così in contrasto con gli obiettivi attuali, incentrati sulla razionalizzazione dei *data center* e sull'uniformazione del patrimonio informativo). Poiché lo sfruttamento delle potenzialità dei *Big Data* appare essere uno dei temi più *disruptive* dal punto di vista tecnologico, si è scelto di analizzare tale tematica. Per verificare se il tragitto attualmente intrapreso a livello europeo abbia come destinazione anche lo sfruttamento dei *Big Data*, saranno indagate non solo le soluzioni organizzative su cui l'Europa e l'Italia stanno puntando, ma anche quelle necessarie allo sfruttamento dei *Big Data*.

Dopo quest'ulteriore premessa, è possibile illustrare l'*iter* seguito nell'elaborato.

Analizzare la questione relativa alla domanda iniziale, comporta indagare due aspetti, che, di tale domanda, ne rappresentano il *prius* ed il *posterius* logici.

Il primo aspetto è relativo ad un'indagine della normativa sulla circolazione dei dati, per verificare se essa consenta l'acquisizione e la circolazione di grandi quantità di dati (personali e non) per e nella pubblica amministrazione. La seconda questione è volta ad analizzare se la pubblica amministrazione disponga di soluzioni organizzative in grado di immagazzinare ed analizzare l'ampia quantità di dati acquisita in tal modo.

Per esaminare questi due aspetti, nel primo e nel secondo capitolo si analizza lo stato dell'arte relativo alla circolazione dei dati ed alle modifiche organizzative dei dati e dei *data center*. Più nello specifico, dopo aver brevemente esposto la Strategia europea per i dati, l'attenzione si focalizzerà sulla tutela dei dati personali. Successivamente, in base al programma NG-EU ed al PNRR, saranno individuate tra le principali trasformazioni organizzative in atto quelle concernenti l'interoperabilità ed il *cloud computing*. Per entrambe le soluzioni tecnologiche, dopo aver cercato di delinearne la nozione, verrà analizzato il quadro europeo e quello nazionale. Nel primo caso si farà quindi riferimento all'*European interoperability framework* (EIF), all'*Interoperable Europe Act* ed alla Piattaforma Digitale Nazionale Dati; nel secondo caso alla comunicazione *Bussola*

Digitale 2030 ed alla Strategia *cloud* Italia. L'analisi dello stato dell'arte in tema di interoperabilità e di *cloud* sarà l'occasione per svolgere anche le prime riflessioni sul modello di amministrazione che tale assetto organizzativo vuole restituire.

Nel terzo capitolo l'attenzione si sposterà sull'importanza dei *Big Data* e sul collegamento tra tale fenomeno e le relative tecniche di analisi (compresa l'intelligenza artificiale, il *machine learning* ed il *deep learning*), cercando successivamente di indagare le correlazioni ed in che misura la normativa sulla circolazione dei dati ed il *framework* organizzativo precedentemente descritto sia allineato con il nuovo paradigma dei *Big Data*.

Il quarto capitolo è orientato ad esaminare nel concreto l'utilizzo delle banche dati da parte dell'amministrazione finanziaria. Dopo aver individuato nell'analisi del rischio fiscale un terreno d'elezione per lo sfruttamento delle banche dati e delle relative tecniche di analisi algoritmica, si analizzeranno gli atti direttivi impartiti dal Ministro dell'economia e delle Finanze nei confronti dell'amministrazione finanziaria, per esaminare su quali soluzioni tecnologiche si sta puntando per efficientare la fase di selezione dei contribuenti da sottoporre a controllo. In seguito, si esamineranno i profili problematici tra la necessità di sfruttare tali soluzioni tecnologiche e la disciplina relativa alla tutela dei dati personali.

All'esito dell'analisi così effettuata si tenterà di delineare in che misura le pubbliche amministrazioni si trovano oggi nella condizione di sfruttare le potenzialità dei *Big Data* e le relative tecniche di analisi.

CAPITOLO I

BANCHE DATI PUBBLICHE: NOZIONE, TUTELA, ACQUISIZIONE E CIRCOLAZIONE DEI DATI

1. Banche dati pubbliche: nozione e tutela

Nel nostro ordinamento è presente una definizione di banca dati all'art. 9, co. 1, n. 9 della legge del 22 aprile 1941, n. 633 (c.d. legge sul diritto d'autore – l.d.a. – rubricata "*Protezione del diritto d'autore e di altri diritti connessi al suo esercizio*"), così come modificata dal decreto legislativo. 6 maggio 1999, n. 169 in attuazione della Direttiva 96/9/CE, relativa alla tutela giuridica delle banche di dati[8].

Secondo l'art. 1, paragrafo 2, della Direttiva, per banca dati[9] s'intende "*una raccolta di opere, dati o altri elementi indipendenti sistematicamente o metodicamente disposti ed individualmente accessibili grazie a mezzi elettronici o in altro modo*"[10].

[8] Data la coincidenza della definizione contenuta all'interno della legge sul diritto d'autore con quella europea, si prenderà in esame solamente quest'ultima disciplina.

[9] La versione tradotta in italiano della direttiva utilizza il termine "*banca dati*", mentre la versione inglese utilizza il termine "*database*". Ai fini di tale normativa, i due termini sono quindi equivalenti.

[10] In precedenza, vi era anche un'altra definizione contenuta nel Codice della Privacy (decreto legislativo 30 giugno 2003, n. 196), che è stata abrogata dal decreto legislativo 10 agosto 2018, n. 101, in attuazione del Regolamento (UE) 2016/679, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati (GDPR). La precedente definizione, di cui all'art. 4 lett. p), decreto legislativo n. 196/2003 intendeva per banca dati "*qualsiasi complesso organizzato di dati personali, ripartito in una o più unità dislocate in uno o più siti*". Si noti che tale definizione è sostanzialmente concorde con quella contenuta nella Direttiva 96/9/CE , specie per quanto riguarda l'elemento organizzativo dei dati.

Nonostante la disciplina si applichi anche alle banche dati non elettroniche[11], in questa analisi si concentrerà l'attenzione solamente su quelle elettroniche: sono queste ultime, infatti, che presentano i vantaggi connessi alle nuove tecnologie.

Dalla definizione appare chiaro che il nucleo essenziale di una banca dati è quello di essere una "raccolta sistematica o metodica di elementi individualmente accessibili"[12], il che porta a compiere due importanti osservazioni: la banca dati è un concetto indipendente sia dal *software*[13] che permette di interagire con la stessa, sia dall'*hardware* sulla quale è memorizzata. Queste tre componenti, infatti, lavorando in modo coordinato, vanno a comporre un centro elaborazione dati (CED o *data center*)[14]. Quindi all'interno del CED devono trovarsi queste tre componenti chiave: la base di dati, che costituisce l'organizzazione dei dati (che dal punto di vista tecnico possono essere strutturati, non strutturati o semi – strutturati; dal punto di vista normativo possono essere personali o non personali)[15]; vari tipi di *software* utilizzati per la gestione,

[11] Come si evince dalla definizione e dal Considerando 14 della Direttiva 96/9/CE, ai sensi del quale "*occorre estendere la tutela concessa dalla presente direttiva alle banche di dati non elettroniche*".

[12] Nel linguaggio tecnico per *database* s'intende principalmente lo schema di organizzazione dei dati. Tale schema descrive la conoscenza *ex- ante* che il progettista ha delle relazioni tra i dati e della struttura stessa del *database*; infatti, lo schema "*deve essere definito al momento della progettazione del database prima di iniziare a memorizzare i dati e definisce l'organizzazione logica dei dati in esso contenuti*". In tal senso M. Aldinucci, *La pubblica amministrazione con i Big Data,* in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021.

[13] Come specificato anche dal Considerato 23 che afferma "*il termine «banca di dati» non deve applicarsi ai programmi per elaboratore utilizzati per la costituzione o per il funzionamento di una banca di dati*"; tali programmi hanno infatti una tutela differente.

[14] In tal senso anche la Corte dei Conti nella sua Deliberazione 14 luglio 2022, n. 33/2022/G sulle infrastrutture digitali, ove afferma che «*per "infrastruttura digitale" deve intendersi un "centro elaborazione dati" (CED o "data center"),ossia un vero e proprio "quartier generale" informatico in cui sono allocati, manutenuti, custoditi, protetti e costantemente monitorati tutti i dispositivi elettronici, gli strumenti di elaborazione e di connettività, gli archivi digitali e quanto serve a far funzionare l'intera architettura informatica, a sua volta in grado di ospitare un numero indefinito di applicazioni, siti internet, portali e software*». Per consultare la delibera, si veda https://www.corteconti.it/Download?id=fb85e5e8-4ad8-4544-a13c-d239170b7f47.

[15] I dati, secondo criteri informatici, possono distinguersi in strutturati, non strutturati e semi-strutturati. I primi sono acquisiti, immagazzinati e categorizzati secondo schemi e regole precise e definite *ex-ante* (quali quelli contenuti in *database* relazionali ed ordinati in colonne e tabelle),

l'elaborazione, l'analisi dei dati e altri necessari per il funzionamento e la manutenzione del CED; un'infrastruttura *hardware*, che include *server*, *router*, alimentatori e altri componenti fisici necessari per l'elaborazione e la memorizzazione dei dati.

Da queste osservazioni si può comprendere che una banca dati può esistere anche senza il *software* e che vi possono essere diversi *software* che permettono di accedere e di utilizzare i suoi contenuti. Tale distinzione offre la flessibilità di scegliere il programma più adatto alle specifiche necessità di ogni situazione, tema rilevante soprattutto per quanto riguarda la possibilità di installare determinate API[16]. Inoltre, è possibile duplicare le banche dati innumerevoli volte su qualsiasi sistema capace di contenerle.

Occorre tuttavia chiarire un'ulteriore caratteristica delle banche dati intese quali raccolta sistematica di elementi. La Corte di Giustizia Europea, infatti, nella causa C-304/07, Directmedia, ha fatto emergere che gli elementi contenuti nella banca dati devono essere certamente connessi, ma non per forza secondo criteri di ordine formale, tecnico o materiale[17]. Ne risulta che la nozione di banca dati, *latu sensu* interpretata, non è più

---

in modo tale che possano essere facilmente analizzabili ed identificabili. I secondi sono tutti quei dati ricavati senza classificazioni e schemi e che richiedono tecniche sofisticate per essere trattati (ad esempio immagini, foto, testi, ecc.). Gli ultimi sono invece dati non strutturati accompagnati da metadati che ne facilitano l'elaborazione e l'analisi (ad esempio in una email, il cui corpo è costituito da dati non strutturati, mentre i dati che rientrano nelle categorie del mittente, destinatario o della data di invio costituiscono dati identificabili e più strutturati). Sul punto si veda M. Crovara, *Informazioni non strutturate*, in *Atti del convegno-L'innovazione tecnologica e metodologica al servizio del mondo del lavoro-primo seminario-Roma, Centro Congressi Villa Eur*, Tipolitografia INAIL, Milano, 2009, pp. 51 ss.; M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., p.10 e 26; D. Iacovelli, M. Fontana, *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale: set di training, algoritmi e qualificazione dei dati. Profili critici*, in *Il diritto dell'economia*, STEM Mucchi Editore, n. 3/2022, p.128.

[16] Acronimo di "*Application programming interface*", esse sono protocolli utilizzati come interfaccia di comunicazione tra componenti *software*.

[17] E questo in virtù dell'interpretazione data al pt.38 del Considerando 21. Quest'ultimo afferma "*che non è necessario che tali materie siano state memorizzate fisicamente in forma organizzata*", dove il termine "materie" è un termine generico che comprende "*testi, suoni, immagini, numeri, fatti, dati o combinazioni di questi*" (per questa definizione si veda la Relazione alla proposta di direttiva presentata dalla Commissione delle Comunità europee il 15 maggio 1992, G.U.C.E. 23 giugno 1992, reperibile al seguente link https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=OJ:C:1992:156:FULL&from=EL; in tale relazione viene inoltre chiarito

una “raccolta sistematica di elementi”, ma solamente una “raccolta di elementi (ossia di dati)”.

Alcuni autori sottolineano un altro aspetto importante che caratterizza la banca dati, che permette di distinguerle dai meri archivi (digitali o cartacei), vale a dire il metodo in esse adottato per consentire e facilitare la ricerca[18]. Solitamente, infatti, i dati presenti nel *database* sono “individualmente accessibili” per mezzo di una ricerca libera, effettuata con qualsiasi dato presente in banca dati, a differenza degli archivi, nei quali la ricerca è condizionata da chiavi d’accesso predeterminate (e.g. ricerca per indici, per lettere alfabetiche). Tuttavia, questa caratteristica non è separata dall’organizzazione e dalla

---

che con il termine “*banca dati*” si intende anche la raccolta di “*materie*”). Ma la Corte di Giustizia Europea lascia trasparire tale interpretazione anche in altri punti della sentenza. Al pt.39, in particolare, interpretando sempre il Considerando 38 della direttiva afferma che “*un’operazione di copiatura non autorizzata, accompagnata da un adattamento del contenuto della banca di dati copiata, rientra nelle operazioni contro le quali la suddetta direttiva intende tutelare il costitutore di tale banca mediante la creazione del diritto sui generis*”. Da ciò si evince che estrarre il contenuto da una banca dati e modificare la disposizione degli elementi rispetto alla banca dati originaria (adattamento del contenuto), costituisce comunque un’operazione di estrazione, quindi illecita. Non può essere quindi la disposizione degli elementi a caratterizzare la banca dati, poiché la sua tutela va intesa *latu sensu*, ricomprendendo anche la semplice connessione degli elementi, il semplice fatto che determinati elementi si trovino in uno stesso *database*. Ma l’oggetto la direttiva tutela anche le banche dati che non siano “originali” per via della scelta degli elementi. Ne risulta che, ai fini della Direttiva 96/9/CE, la nozione normativa di banca dati che è possibile estrapolare consiste semplicemente in una “raccolta di dati” che, a seconda di alcune caratteristiche, attribuirà posizioni giuridiche soggettive differenti ai soggetti coinvolti. In tal senso si veda anche E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, Aggiornamento V – 2001, Giuffrè, p.130.

[18] R. Borruso, S. Russo, C. Tiberi, *L’informatica per il giurista. Dal Bit a internet*, III ed., Milano, Giuffrè Editore, 2009, p. 291-294. A tal vengono esposti cinque principi fondamentali che dovrebbero caratterizzare la ricerca e l’elaborazione dei dati in banca dati: libertà e casualità della ricerca (possibilità di utilizzare ogni dato presente in banca dati come “chiave d’accesso” per selezionare un elemento); libera combinazione degli operatori booleani *and*, *or* e *not* (vale a dire che deve essere possibile cercare più dati cumulativamente, alternativamente o con l’esclusione di uno di essi); libera mascherabilità di un dato non conosciuto con un carattere convenzionale; opportunità di ricavare un riassunto conciso dai documenti contenuti nel database; capacità di comunicare con il computer eseguendo una ricerca attraverso una sequenza di comandi e informazioni successive

disposizione degli elementi: infatti l'efficienza della ricerca dipende essenzialmente dall'organizzazione degli elementi e dal *software* che utilizza la banca dati[19].

Alla luce delle considerazioni finora svolte possiamo affermare che la banca dati rappresenta un bene immateriale, consistente in una raccolta di dati individualmente accessibili.

Occorre adesso chiarire quando una banca dati si possa classificare come pubblica. A causa dell'assenza di una definizione normativa di banca dati pubblica[20], si deve ricorrere ad un criterio soggettivo: sono banche dati pubbliche quelle detenute da un soggetto pubblico. Cosa debba intendersi per soggetto pubblico, non è specificato, perché non esiste una disciplina specifica per le banche dati, proprio perché esse, essendo uno strumento con le quali viene esercita un'attività, coinvolgono numerose discipline: il diritto d'autore, l'accesso ai documenti amministrativi, concorrenza, *privacy*, diffusione e riutilizzo delle informazioni e così via. Molto spesso sono un elemento organizzativo, strumentale rispetto ad una determinata attività. Per questo motivo, si può dedurre che la disciplina relativa alle banche dati potrà seguire il regime pubblicistico in determinati settori e seguire regimi privatistici in altri. Infatti il Consiglio di Stato ha chiarito di come "*uno stesso soggetto possa avere la natura pubblica a certi fini e possa, invece, non averla ad altri fini, consentendo rispetto ad altri istituti regimi di natura privatistica*"[21]. E

[19] P. Atzeni, S. Ceri, S. Paraboschi, R. Torlone, *Basi di dati. Modelli e linguaggi di interrogazione,* II ed., Milano, McGraw-Hill, 2006, p. 5, ove si afferma che la semplicità per svolgere le operazioni da parte degli utenti e la capacità di rendere produttive le attività degli utenti dipende dalla bontà della realizzazione della base di dati da parte dei suoi progettisti e dalle applicazioni che la utilizzano.

[20] Infatti, benché l'art. 25 della legge n. 340/2000 sia rubricato 'accesso alle banche dati pubbliche', esso risulta inadeguato a perimetrarne la nozione, perché il primo comma disciplina la facoltà di dare in uso gratuito un *software*, di cui sia titolare una determinata pubblica amministrazione, ad un'altra pubblica amministrazione. Il secondo comma disciplina invece l'accesso delle pubbliche amministrazioni a dati contenuti in pubblici elenchi, che siano conoscibili da chiunque. Per un approfondimento sull'inidoneità dell'art. 25 a descrivere la nozione di banca dati pubblica si veda F. Cardarelli*, Le banche dati pubbliche: una definizione*, in *Dir. informaz*., 2002, vol. 18, n. 2, pp. 321-341.

[21] Cons.St., sez. VI, 11 luglio 2016, n. 3043.

lo stesso, ad avviso del Consiglio di Stato, vale anche per gli enti pubblici, che in tal modo assumono una "*nozione funzionale e cangiante*"[22].

Quindi una banca dati è pubblica se detenuta da un soggetto pubblico, ma uno stesso soggetto può essere ricompreso nella nozione di pubblica amministrazione (o di ente pubblico) per determinate attività, mentre per altre può non ricadere nella nozione di soggetto pubblico. E la banca dati dovrebbe seguire la stessa logica.

Di conseguenza, se, per esempio, si prende ad esame la disciplina sull'accesso (che per espressa menzione della Direttiva 96/9 resta impregiudicata)[23], tale disciplina si dovrebbe applicare, ai sensi dell'art. 23 della legge 7 agosto 1990, n. 241 "*nei confronti delle pubbliche amministrazioni, delle aziende autonome e speciali, degli enti pubblici e dei gestori di pubblici servizi*". L'art. 22 della stessa legge invece precisa che per pubblica amministrazione s'intendono "*tutti i soggetti di diritto pubblico e i soggetti di diritto privato limitatamente alla loro attività di pubblico interesse disciplinata dal diritto nazionale o comunitario*". Quindi, rispetto alla disciplina dell'accesso, saranno banche dati pubbliche anche le banche dati detenute dai gestori di pubblici servizi.

Dopo aver cercato di individuare la nozione ed in cosa consiste una banca dati, si analizzerà adesso la tutela delle banche dati prevista dalla Direttiva 96/9/CE, che, oltre a costituire una delle discipline più rilevanti in materia, è in grado di influenzare anche il regime di circolazione dei dati. Essa, infatti, attribuisce due diritti differenti in capo a due soggetti diversi: il diritto d'autore in capo all' "*autore della banca dati*"[24] e il diritto *sui*

---

[22] Cons.St., sez. VI, 26 maggio 2015, n. 2660. Secondo il Consiglio di Stato infatti "*il perimetro del concetto di ente pubblico non è sempre uguale a se stesso, ma muta a seconda dell'istituto o del regime normativo che deve essere applicato e della ratio ad esso sottesa*".

[23] Secondo l'articolo 13 della Direttiva 96/9 CE, rubricato '*Mantenimento di altre disposizioni*': "*La presente direttiva non osta all'applicazione delle disposizioni concernenti segnatamente il diritto d'autore, i diritti connessi o altri diritti od obblighi preesistenti su dati, opere o altri elementi inseriti in una banca di dati, brevetti, marchi commerciali, disegni e modelli industriali, la protezione dei beni appartenenti al patrimonio nazionale, le norme sulle intese e sulla concorrenza sleale, il segreto industriale, la sicurezza, la riservatezza, la tutela dei dati di carattere personale ed il rispetto della vita privata, l'accesso ai documenti pubblici o il diritto dei contratti*".

[24] Disciplinato dall'art. 3 all'art. 6 della Direttiva 96/9/CE.

generis in capo al "*costitutore*"[25] della banca dati. Tali posizioni giuridiche, benché molto differenti, attribuiscono dei diritti di esclusiva sulle banche dati che consentono di esercitare un controllo "*anche sull'accesso e sull'uso del contenuto delle stesse e, di conseguenza, sulla circolazione dei dati digitali*"[26]. Quindi tale disciplina va letta anche alla luce della tensione tra, da un lato, i diritti riconosciuti in capo all'autore ed al costitutore della banca dati, nell'ottica di promuovere lo sviluppo di tali tecnologie[27], e, dall'altro lato, dalle esigenze, sempre più sentite, di garantire un'ampia circolazione ed accesso ai dati[28].

Nel proseguo si cercherà quindi di delineare i due diritti previsti dalla Direttiva 96/9/CE, per poi cercare di analizzare le relazioni con altre discipline e il rapporto con l'attuale contesto economico e tecnologico.

Il diritto d'autore tutela le opere dell'ingegno di carattere creativo, ossia quelle opere che sono nuove e che possiedono un'originalità della creazione tale da tradursi in un'impronta individuale dell'opera[29].

L'originalità non deve tradursi necessariamente nella creazione di un'opera *ex novo*, ma può consistere anche in una raccolta delle opere dell'ingegno che, per la particolare

---

[25] Disciplinato dall'art. 7 all'art. 11 della Direttiva 96/9/CE.

[26] Si perviene a tale conclusione in S. Scalzini, M. Maggiolino, *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma?*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, volume III, Bologna, Il Mulino, 2022, p.200. Le autrici, infatti, affermano che determinati regimi giuridici possono interferire con la circolazione, accessibilità e riutilizzazione dei dati, sebbene "*non sussista allo stato un diritto esclusivo avente ad oggetto specificatamente i dati digitali*" (op.ult.cit., p.175).

[27] L'equazione alla base dell'attribuzione di tali diritti era "*più proprietà intellettuale = più innovazione*", così come affermato in R. Caso, *Open data, ricerca scientifica e privatizzazione della conoscenza*, *Trento law and technology research group*, *research paper n. 48*, Università di Trento, 2022, p.9.

[28] In merito al bilanciamento tra diritti di esclusiva sulle banche dati ed esigenze di circolazione dati, promosse specialmente attraverso un'evoluzione legislativa europea *in fieri* (come la Direttiva *Open Data* ed il *Data Governance Act*), si veda S. Scalzini, *Banche di dati, sfruttamento dei dati digitali e concorrenza*, Giappichelli, Torino, 2023, pp.151-153 e pp. 211-217.

[29] T. Ascarelli, *Teoria della concorrenza e dei beni immateriali: lezioni di diritto industriale*, II ed., Milano, Giuffrè, 1957, 282.

scelta o disposizione delle opere, assume un'autonoma efficacia rappresentativa[30]. Queste caratteristiche possono essere soddisfatte anche da una raccolta o pluralità di dati, nonostante il dato singolarmente considerato non costituisca oggetto di tale tutela[31]. A tal fine l'art. 3 della Direttiva 96/9/CE dispone che per stabilire se una banca dati costituisca una creazione dell'ingegno (quindi protetta dal diritto d'autore) si ha riguardo alla "*scelta*" o alla "*disposizione*" del materiale, impedendo l'applicazione di altri criteri per stabilire se ad una banca dati vada riconosciuta la tutela autoriale[32].
L'autore di tale banca dati sarà titolare di una posizione giuridica equivalente, ma non identica, a quella che deriva dal diritto d'autore[33]. Tale diritto attribuisce il potere

[30] *Ibidem*.

[31] E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, cit., p.132.

[32] La tutela concerne, oltre la raccolta di dati, anche "*sistemi elettronici necessari per utilizzarla, come, ad esempio, il sistema di presentazione delle informazioni, i sistemi di indici e i thesauri necessari per la ricerca; tutto ciò , insomma, che è necessario per costituire una banca di dati. Non comprende, invece, il sistema di ricerca, ossia il complesso dei programmi utilizzati per la costituzione e il funzionamento della banca*". In tal senso, E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, cit., p.133. La tutela autoriale non si estende invece al contenuto della banca dati, in quanto l'oggetto della tutela è la struttura-architettura della banca dati dovuta alla specifica selezione o disposizione dei dati. Infatti, la nozione di banca dati, come visto in precedenza, rappresenta un concetto distinto rispetto al contenuto della stessa. Invero, i diritti sul contenuto della banca dati sono diritti differenti sia dal diritto d'autore che *sui generis*, i quali insistono più propriamente sul *database*. In tal senso dispongono gli artt. 3 e 7 della Direttiva 96/9/CE. Nel merito si veda S. Scalzini, M. Maggiolino, *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma?*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., p. 179 e M. L. Montagnani, *Dati e proprietà intellettuale in Europa: dalla "proprietà" all'"accesso"*, in *Il diritto dell'economia*, STEM Mucchi Editore, n. 1/2020, p.555. Per la tutela attraverso il diritto *sui generis* conferita unicamente alla specifica conformazione dei dati nel *database* (e non ai dati in sé e per sé considerati), si veda A. Amidei, M. Maggiolino, *Intelligenza artificiale, dati digitali e proprietà intellettuale*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., p. 76.

[33] Il diritto d'autore, diritti patrimoniali compresi, dovrebbe spettare infatti esclusivamente all'autore - persona fisica - dell'opera ai sensi dell'art. 12 della l.d.a., tuttavia l'art. 12-*bis* l.d.a. dispone che sia il datore di lavoro, salvo patto contrario, ad essere titolare del diritto esclusivo di utilizzazione economica della banca dati del lavoratore dipendente nell'esecuzione delle sue mansioni o su istruzioni impartite dallo stesso datore di lavoro. L'autore della banca dati, benché sia titolare originario del diritto, è come se trasferisse *ope legis* e salvo patto contrario, il diritto

esclusivo di eseguire o autorizzare la riproduzione, la modificazione, la distribuzione, la comunicazione al pubblico[34]. Si tratta, invero, di un diritto avente un contenuto persino più ampio del diritto d'autore tradizionale[35].

Vi sono invece casi in cui le banche dati consistono in raccolte prive dei caratteri di originalità e creatività, nonostante abbiano richiesto un considerevole investimento economico per la loro costituzione. A tal fine la scelta del legislatore europeo è stata quella di attribuire un diritto *sui generis* al costitutore della banca dati[36], infatti l'art. 7, comma 1, della Direttiva 96/9/CE attribuisce al costitutore il diritto di "*vietare operazioni di estrazione e/o reimpiego della totalità o di una parte sostanziale del contenuto della stessa, valutata in termini qualitativi o quantitativi, qualora il conseguimento, la verifica*

al datore di lavoro. In quest'ultimo senso L. Chimenti, *Banche di dati e diritto d'autore*, Milano, Giuffrè, 1999, p.29. Si precisa che, secondo il diritto nazionale, solo una persona fisica può essere autore di una banca dati. Per disposizione del contenuto si deve intendere una scelta creativa nell'organizzazione e nel posizionamento del dato rispetto al resto della banca dati, escludendo criteri ovvi, necessitati o convenzionali. In quest'ultimo senso si veda C. Di Cocco, *Tutela delle banche di dati: patrimonio culturale e mercato unico digitale*, in *Aedon-Rivista di arti e di diritto on line*, 3/2020.

[34] Si veda l'art. 64-*quinquies* l.d.a. È da sottolineare inoltre che tali diritti si aggiungono a quelli generali previsti dal diritto d'autore, come il diritto esclusivo di pubblicare ed utilizzare economicamente l'opera previsto all'art. 12 l.d.a.

[35] E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, cit., p.134.

[36] L'introduzione di tale diritto costituiva una "scommessa" per incentivare lo sviluppo di banche dati e per superare le economie concorrenti rispetto a quella del continente europeo, in particolare gli Stati Uniti, che hanno optato di tutelare le banche dati solamente con una tutela di tipo autoriale e, per quelle sopravviste del carattere dell'originalità, non con un diritto equipollente a quello *sui generis* - peculiarità europea - , ma attraverso le norme sulla concorrenza sleale (*tort of misappropiation*). In merito si veda E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, cit., p.140; J.H. Reichman, *La guerra delle banche dati. Riflessioni sulla situazione americana*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997; S. Scalzini, M. Maggiolino, *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma?*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., p.177. È da segnalare inoltre che tale diritto ha suscitato diverse critiche sin dalla sua introduzione per il rischio che una tutela eccessiva potesse configurare un "*monopolio dell'informazione*", si veda nel merito op. ult. cit., p.180 e G. Ghidini, *Rethinking intellectual property: balancing conflicts of interest in the constitutional paradigm*, Cheltenham, Northampton, E. Elgar, 2018.

e la presentazione di tale contenuto attestino un investimento rilevante sotto il profilo qualitativo o quantitativo".

Il presupposto di tale diritto è differente da quello del diritto d'autore: non è l'originalità della banca dati, ma lo sforzo organizzativo necessario per la raccolta dei dati e, dalla differenza del presupposto, discendono diversità in relazione alla titolarità, durata e contenuto del diritto rispetto al diritto d'autore[37]. Infatti, tale diritto può spettare anche ad imprese o società che hanno costituito la banca dati[38]. Inoltre, ha una durata di quindici anni a partire dalla costituzione della banca dati; una durata che, benché molto inferiore ai classici settant'anni del diritto d'autore, è suscettibile di essere protratta per un tempo indeterminato[39], poiché il legislatore ha previsto che ogniqualvolta intervenga una "*modifica sostanziale*"[40] comportante un "*nuovo investimento sostanziale*"[41] decorre un nuovo termine di quindici anni a partire dalla modifica. Il contenuto del diritto è incentrato sulla possibilità di vietare le operazioni di estrazione e reimpiego[42], potendo il costitutore decidere se mettere la banca dati a disposizione del pubblico, a titolo gratuito o a pagamento.

---

[37] In tal senso viene sottolineato in E. Giannantonio, *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, cit., pp.137-138. In tal caso infatti il diritto è volto a tutelare interessi industriali (la remunerazione dell'investimento) e non il diritto morale e patrimoniale dell'autore. Si veda in merito P. Spada, *Banche dati e diritto d'autore (il "genere" del diritto d'autore sulle banche di dati)*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, pp.5 ss.
Tale diritto è stato definito come il primo esempio di una tutela immediata e diretta del puro investimento; in tal senso V. Di Cataldo, *Banche-dati e diritto sui generis: la fattispecie costitutiva*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, p. 27.

[38] Art. 11 Direttiva 96/9/CE.

[39] Sul tema si veda M. C., Cardarelli, *banche dati e direttiva comunitaria: il diritto sui generis. La durata*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, pp. 64-85.

[40] Art. 10 Direttiva 96/9/CE.

[41] *Ibidem*.

[42] Ai sensi dell'art. 7 della Direttiva 96/9/CE per estrazione s'intende "*il trasferimento permanente o temporaneo della totalità o di una parte sostanziale del contenuto di una banca di dati su un altro supporto con qualsiasi mezzo o in qualsivoglia forma*", mentre per reimpiego s'intende qualsiasi forma di "*messa a disposizione del pubblico della totalità o di una parte sostanziale del contenuto della banca di dati mediante distribuzione di copie, noleggio, trasmissione in linea o in altre forme*".

Identificare l'investimento sotteso alla banca dati rappresenta un passaggio cruciale al fine dell'attribuzione di tale diritto. Infatti, l'investimento rilevante ai fini della configurazione del diritto *sui generis*, non è quello di "*produzione-rilevazione*", ma è solamente quello "*di conseguimento*"[43], ossia l'investimento diretto alla realizzazione del *database*, quale soluzione tecnologica per conservare, organizzare e trattare dati già esistenti, non già l'investimento diretto alla creazione e generazione di nuovi dati.

Nel tempo il contesto tecnologico ed economico è mutato di molto rispetto a quando fu emanata la Direttiva 96/9/CE, la quale resta ancora una disciplina rilevante in materia di banche dati e di accesso a queste ultime. Per analizzare gli effetti che la Direttiva ha avuto nel corso del tempo allo scopo di ipotizzare eventuali modifiche, la Commissione europea si è adoperata a partire dal 2005 attraverso due valutazioni della Direttiva. Gli ultimi studi si sono compendiati nella "seconda valutazione della direttiva sulle banche dati"[44], pubblicata il 25 Aprile 2018. È sostanzialmente emerso che i tre obiettivi principali della Direttiva 96/9/CE (tutelare e di conseguenza stimolare gli investimenti in banche dati; corretto equilibrio tra diritto del titolare della banca dati ed utente; armonizzazione della protezione delle basi di dati) sono ancora ben bilanciati dalla Direttiva[45]. È stato evidenziato tuttavia che risulta difficile collegare la Direttiva 96/9/CE

[43] La distinzione di nomenclatura tra i differenti investimenti è contenuta in A. Ottolia, *Big Data e innovazione computazionale*, Torino, Giappichelli, 2017, pp. 76-84. Tale differenza è dovuta ad un'interpretazione restrittiva adoperata dalla CGUE in diverse sentenze (tra le quali, a titolo esemplificativo, C-46/02; C-444/02), che si giustificano in virtù del fatto che il diritto *sui generis* ha come *ratio* gli investimenti e gli sforzi effettuati per ottenere una particolare organizzazione o forma espressiva dei dati. Questa interpretazione permette anche di evitare che vi sia un'unica fonte di produzione delle informazioni, inaccessibile al mercato (c.d. *sole source data destination*). Si veda per un approfondimento S. Scalzini, M. Maggiolino, *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma?*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., pp.180-183; A. Amidei, M. Maggiolino, *Intelligenza artificiale, dati digitali e proprietà intellettuale*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., pp. 69-74; C. Galli, M. Bogni, *I requisiti per la tutela IP dei Big Data*, in V. Falce, G. Ghidini (a cura di), *Informazione e Big Data tra innovazione e concorrenza*, Milano, 2018, pp. 97 ss.

[44] "*Evaluation of Directive 96/9/EC on the legal protection of databases*", SWD(2018) 146 final.

[45] Per il diritto *sui generis*, tuttavia, non c'è una prova che abbia avuto un grande impatto sulla produzione di banche dati. in tal senso si esprime la Commissione nella sua seconda valutazione.

con la disciplina dell'informazione nel settore pubblico (c.d. PSI, ossia la *public sector information*)[46], problema che è stato risolto con l'emanazione della Direttiva n. 1024/2019 (c.d. *Open Data Directive*), relativa all'apertura dei dati e al riutilizzo[47]

Le maggiori proposte avanzate dalla dottrina circa un intervento sulla disciplina della tutela giuridica delle banche dati riguardano proprio il diritto *sui generis*, che non si è rivelato così proficuo per il progresso tecnologico così come si sperava e rappresenta "*il profilo più controverso della Direttiva europea*". In quest'ultimo senso e per un approfondimento sulle due valutazioni della Commissione sulla Direttiva 96/9/CE, si veda, per tutti, S. Scalzini, *Banche di dati, sfruttamento dei dati digitali e concorrenza*, Giappichelli, Torino, 2023, pp.39-51.

[46] Viene infatti riportato nell'*Executive Summary* della valutazione della Direttiva 96/9/CE (SWD(2018) 147 final) che "*There are no major inconsistencies between the Database Directive and other EU legislation. However, a clarification of how it interacts with the Public Sector Information Directive is needed*". La difficoltà nel coordinare le diverse discipline è sottolineata anche in M. Van Eechoud, *A Serpent Eating Its Tail: The Database Directive Meets the Open Data Directive*, in *IIC - International Review of Intellectual Property and Competition Law*, Vol.52, 04/2021, pp. 375-378. Le interferenze tra le due normative derivano dal fatto che alcune PSI potrebbero far parte di banche dati protette dal diritto *sui generis* o d'autore, il cui titolare può essere lo stesso ente pubblico o un terzo. Ed infatti il diritto *sui generis* ed il diritto d'autore sono attribuiti anche alle pubbliche amministrazioni. In merito al diritto d'autore occorre fare riferimento all'art. 11 della L.633/41 (legge sul diritto d'autore), il quale sancisce che: "*Alle Amministrazioni dello Stato, …, alle Provincie ed ai Comuni, spetta il diritto di autore sulle opere create e pubblicate sotto il loro nome ed a loro conto e spese. Lo stesso diritto spetta agli enti privati che non perseguano scopi di lucro, salvo diverso accordo con gli autori delle opere pubblicate, nonché alle accademie e agli altri enti pubblici culturali sulla raccolta dei loro atti e sulle loro pubblicazioni*". Si dovrebbe pervenire alla medesima conclusione anche rispetto al diritto *sui generis*, volto a tutelare quell'investimento significativo che esorbita il normale svolgimento dell'attività amministrativa.

[47] Per riutilizzo si intende l'utilizzo di dati in possesso di pubbliche amministrazioni "*da parte di persone fisiche o giuridiche a fini commerciali o non commerciali diversi dallo scopo iniziale nell'ambito dei compiti di servizio pubblico per i quali i dati sono stati prodotti, fatta eccezione per lo scambio di dati tra enti pubblici esclusivamente in adempimento dei loro compiti di servizio pubblico*". Questa è la definizione dall'art. 2, par. 1, n. 2 del *Data Governace Act* (Regolamento (UE) 2022/868), che è sostanzialmente conforme a quella utilizzata dall'art. 2, par. 1, n. 11, dell'*Open Data Directive*. Il riutilizzo dell'informazione pubblica (sia dei dati personali che non personali) riveste una rilevante importanza per l'innovazione tecnologica e per la *data-driven* economy; se si pensa che circa l'ottantacinque per cento dei dati raccolti in Europa non viene riutilizzato, si comprende l'importanza degli interventi normativi; si veda a tal proposito G. Resta, Pubblico, privato, collettivo nel sistema europeo di governo dei dati, in Rivista trimestrale di diritto pubblico, 4/2022, p.976 e T. Ramge, V. Mayer-Schönberger, *Fuori i dati! Rompere i monopoli sulle informazioni per rilanciare il progresso*, Egea, Milano, 2021.

dell'informazione del settore pubblico, e di rifusione delle precedenti direttive[48]. In ogni caso, nel programma di lavoro 2021[49] , nella sezione relativa ad "*Un'Europa pronta per l'era digitale*" si prevede una revisione della direttiva sulle banche dati entro il 2030.

La correlazione con la disciplina dell'informazione nel settore pubblico è particolarmente significativa, anche perché mostra di come la Direttiva 96/9/CE influisca sulla circolazione e sul riutilizzo dei dati. L'*Open Data Directive* assume inoltre particolare rilevanza in quanto è un atto che rientra nella c.d. Strategia europea per i dati[50]. La direttiva da ultimo citata costituisce il culmine di un percorso legislativo volto a valorizzare il vasto patrimonio informativo del settore pubblico[51] attraverso l'introduzione di un regime di riutilizzabilità dei documenti contenenti dati pubblici da applicarsi compatibilmente la disciplina sulla proprietà intellettuale. L'*Open Data Directive* si applica ai dati nella disponibilità degli enti pubblici[52] e sono compresi i "*dati della ricerca*"[53] e le "*serie di dati*

[48] Le precedenti direttive rifuse sono: Direttiva 2003/98/CE, relativa al riutilizzo dell'informazione del settore pubblico (cd Direttiva PSI) e Direttiva 2013/37/UE (c.d. Direttiva PSI 2.0). La Direttiva 1024/2019 è stata recepita nel nostro ordinamento con il decreto legislativo 8 novembre 2021, n. 200, che ha modificato il decreto legislativo 24 gennaio 2006, n. 36.

[49]Bruxelles, 19/10/2020 COM(2020) 690 final, https://eur-lex.europa.eu/legal content/IT/TXT/HTML/?uri=CELEX:52020DC0690&from=DA

[50] Così come riportato in M. Van Eechoud, *Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research*, European Commission – Directorate general for research and innovation, 2022, p.5. Sulla strategia europea in materia di dati si parlerà più diffusamente nel paragrafo seguente.

[51] Sulla rilevanza per l'economia e la società circa la riutilizzazione delle PSI si veda M. Maggiolino, *Il riutilizzo dell'informazione detenuta dal settore pubblico: alcune riflessioni di politica e diritto della concorrenza*, in *Concorrenza e mercato*, 2012, p. 765. Tra le preoccupazioni vi è quella di garantire la disponibilità delle informazioni contenute in banche dati disponibili solamente ad un soggetto, in virtù del ruolo pubblico ricoperto.

[52] In cui sono ricompresi gli organismi di diritto pubblico, le imprese pubbliche e private, i gestori di pubblici servizi in relazione ai servizi di pubblico interesse (art. 2, n. 11, Direttiva 2019/1024).

[53] Definiti quali i "*documenti in formato digitale, diversi dalle pubblicazioni scientifiche, raccolti o prodotti nel corso della ricerca scientifica e utilizzati come elementi di prova nel processo di ricerca, o comunemente accettati nella comunità di ricerca come necessari per convalidare le conclusioni e i risultati della ricerca*" (art. 2, Direttiva *Open Data*). Sull'importanza del riutilizzo dei dati nell'ambito della ricerca si veda M. Shabani, *The Data Governance Act and the EU's move towards facilitating data sharing*, in *Molecular Systems Biology*, 17, 2021, pp.1-3.

di elevato valore"[54]. Peraltro, rimangono impregiudicati i limiti all'accesso ai dati in caso di diritti di proprietà intellettuale (salvo alcune eccezioni)[55], specialmente quando alcuni documenti nella disponibilità delle pubbliche amministrazioni contengano banche dati protette dal diritto d'autore o *sui generis* facenti capo a soggetti privati[56]. In ogni caso, tale disciplina è volta anche a controbilanciare i diritti di esclusiva sulle banche dati[57].

Al termine di questo paragrafo si è visto, dopo aver inquadrato la nozione di banca dati, di come i diritti previsti dalla Direttiva 96/9/CE costituiscano una disciplina che interferisce e limita la circolazione dei dati. Si sono visti inoltre i rapporti e limitazioni di tali diritti con riguardo al riutilizzo delle ISP, esponendo alcuni profili della *Open Data Directive*[58]. Quest'ultima rappresenta uno dei diversi atti che fanno parte della Strategia

---

[54] Definiti all'art. 2 della Direttiva quali i "*documenti il cui riutilizzo è associato a importanti benefici per la società, l'ambiente e l'economia, sin particolare in considerazione della loro idoneità per la creazione di servizi, applicazioni a valore aggiunto e nuovi posti di lavoro dignitosi e di alta qualità, nonché del numero dei potenziali beneficiari dei servizi e delle applicazioni a valore aggiunto basati su tali serie di dati*".

[55] Infatti l'art. 3 della Direttiva dispone che "*Gli Stati membri provvedono affinché i documenti i cui diritti di proprietà intellettuale sono detenuti da biblioteche, comprese le biblioteche universitarie, musei e archivi, e i documenti in possesso delle imprese pubbliche siano riutilizzabili a fini commerciali o non commerciali, qualora il loro riutilizzo sia autorizzato*".

[56] Sul rapporto tra accesso e riutilizzazione si veda S. Gobbato, *Verso l'attuazione della direttiva (UE) 2019/1024 sul riutilizzo degli open data della PA: nuove opportunità per le imprese*, in *MediaLaws*, 2/2020, pp.247-261. In ogni caso, un'ulteriore spinta verso una maggiore circolazione dei dati pubblici, cercando di restringere le limitazioni derivanti dal diritto *sui generis*, proviene anche dal *Data Governance Act*. Si veda nel merito S. Scalzini, *Banche di dati, sfruttamento dei dati digitali e concorrenza*, cit., pp.216-217.

[57] Ad esempio, l'art. 1, par. 6, della Direttiva sancisce che "*Il diritto del costitutore di una banca di dati di cui all'articolo 7, paragrafo 1, della direttiva 96/9/CE non è esercitato dagli enti pubblici al fine di impedire il riutilizzo di documenti o di limitare il riutilizzo oltre i limiti stabiliti dalla presente direttiva*". Si veda R. Caso, *Open data, ricerca scientifica e privatizzazione della conoscenza*, cit., p.10.

[58] Si è sostenuto che il potenziamento dei diritti di esclusiva rappresentano una contraddizione con il paradigma dell'*open data* e della *open science*. Per questo, bisognerebbe effettuare una decisione su quale valore privilegiare: l'Unione europea, infatti, nonostante i recenti atti volti ad incentivare la circolazione dei dati, continua a potenziare anche la normativa sulla proprietà intellettuale (si veda la Direttiva 2019/790 – c.d. Direttiva *copyright*-). Sul rapporto tra *open science* e proprietà intellettuale si veda R. Caso, *Open data, ricerca scientifica e privatizzazione della conoscenza*, cit., passim.

europea per i dati, che, in base alle ultime normative, sembra voler incentivare una maggiore circolazione e riutilizzo dei dati. Per questo motivo nel prossimo paragrafo, dopo aver meglio esaminato la Strategia europea sui dati, ci si concentrerà sulla normativa che ne costituisce il caposaldo e che prevede le più nette limitazioni in materia di circolazione dei dati: il Regolamento generale sulla protezione dei dati (GDPR)[59].

2. Acquisizione e circolazione dei dati

2.1. Strategia europea per i dati

"*La strategia europea per i dati mira a far acquisire all'UE una posizione di leadership nella società basata sui dati. La creazione di un mercato unico consentirà ai dati di circolare liberamente all'interno dell'UE e in tutti i settori a vantaggio delle imprese, dei ricercatori e delle pubbliche amministrazioni*"[60] e grazie a tale mercato unico l'UE potrà "*divenire un modello di riferimento per una società che, grazie ai dati, dispone di strumenti per adottare decisioni migliori, a livello sia di imprese sia di settore pubblico*"[61]. La Strategia europea per i dati, consolidatasi con la comunicazione della Commissione

[59] Regolamento (UE) n. 2016/679.

[60] *Slogan* della strategia europea in materia di dati reperibile al portale della Commissione europea, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_it#:~:text=La%20creazione%20di%20un%20mercato,ricercatori%20e%20delle%20pubbliche%20amministrazioni.

[61] Comunicazione della Commissione europea, *Una strategia europea per i dati*, COM(2020) 66 final, 19 febbraio 2020.

del 2020[62], rappresenta un punto di svolta[63] rispetto al tradizionale approccio, che si basa in particolare sulla protezione e sulla "difesa" del dato[64].

Tale strategia s'inserisce nella più ampia strategia digitale dell'UE[65] ed è volta specificatamente a regolare il mercato dei dati. Allo stato, tra i più importanti atti della

[62] L'intento di creare un mercato interno dei dati si era già manifestato con la prima Direttiva PSI (2003/98/CE) ed era proseguita con la comunicazione della Commissione agli Stati membri "*Verso una florida economia basata sui dati*" (COM(2014) 442 final), con lo scopo di stabilire un mercato unico dei *Big Data* e del *cloud*.

[63] Come si può evincere negli estratti sopra riportati, la Strategia europea per i dati mira ad incentivare la circolazione dei dati e a rafforzare il potere conoscitivo per fini decisionali. Come si evince dalla comunicazione stessa la Strategia si basa su quattro pilastri: un quadro di *governance* intersettoriale per l'accesso e l'utilizzo dei dati; investimenti nei dati e rafforzamento delle infrastrutture e delle capacità europee per l'*hosting*, l'elaborazione e l'utilizzo dei dati, l'interoperabilità; Competenze: fornire strumenti alle persone, investire nelle competenze e nelle PMI; Spazi comuni europei di dati in settori strategici e ambiti di interesse pubblico. Particolare rilevanza, all'interno dello "spazio unico europeo dei dati", possono assumere spazi settoriali dei dati, come lo "Spazio europeo dei dati sanitari (EDHS)", il primo proposto dalla Commissione; sullo sviluppo delle infrastrutture in tale ambito si rimanda a J. Reichel, *The European Strategy for Data and Trust in EU Governance. The Case of Access to Publicly Held Data*, in *Ceridap*, 4/2023, pp. 129-158.

[64] Le riflessioni in merito ad un'eventuale evoluzione in merito all'approccio ai dati verranno svolte al termine del capitolo.

[65] Si veda il portale della Commissione https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_it. La strategia digitale dell'UE è più ampia e ricomprende tutti gli atti e le normative che hanno un impatto nell'ambito del digitale, come quelle relative ai semiconduttori, ai mercati digitali, all'intelligenza artificiale o all'identità digitale. Nel cercare di effettuare una tassonomia al fine di individuare le normative e gli atti che più specificatamente rientrano nella "Strategia per i dati" sono stati esclusi il *Digital Markets Act* (Regolamento (UE) 2022/1925), il *Digital Services Act* (Regolamento (UE) 2022/2065) e la proposta di Regolamento europeo per l'intelligenza artificiale, il c.d. *AI Act* (COM(2021) 206 final). Questi ultimi tre atti incidono in modo più diretto sul mercato unico, in quanto il primo, disciplinando i mercati, si occupa dei rapporti fra le piattaforme e i fornitori di servizi; il secondo ha ad oggetto principalmente i rapporti fra gli erogatori di servizi e gli utenti; il terzo si occupa di intelligenza artificiale, anche se in realtà ha un oggetto molto ampio. Questi atti rientrano più propriamente nella "*Strategia per il mercato unico digitale*". Per il riferimento a quest'ultima strategia, il cui processo regolatorio è stato avviato a partire dal 2015 con la relativa Comunicazione della Commissione, si veda L. Torchia, *Lo Stato digitale*, cit., pp.67-69 e 84. Vi è da segnalare che ai fini dell'attuazione della Strategia per il mercato unico digitale è stato istituito un apposito fondo attraverso il Programma europeo "Europa digitale", che inciderà sul quadro finanziario pluriennale 2021-2027, si veda in merito M. Cattari, *Il nuovo Programma europeo*

Strategia, vi rientrano: il GDPR; il Regolamento per la libera circolazione dei dati non personali (Regolamento (UE) 2018/1807, c.d. RDNP o Regolamento 1807); la Direttiva *Open Data*; il *Data Governance Act* (Regolamento (UE) 2022/868, c.d. DGA) ed il *Data Act* (Regolamento (UE) 2023/2854)[66].

"Europa digitale"(2021-2027) Proposta della Commissione Europea e documentazione, in *DigItalia - Rivista del digitale nei beni culturali*, Roma, ICCU, Anno XV, Numero 1, 2020, pp.125-130. Vi è da considerare che tali interventi, benché distinti dalla strategia per i dati, hanno forti connessioni con essa e si influenzano a vicenda; e lo stesso dicasi per altre discipline che indirettamente influiscono sul *framework* dei dati, come, si è visto in precedenza, la Direttiva 96/9/CE. Per le connessioni con quest'ultima normativa si veda C. Sganga, *Ventisei anni di Direttiva database alla prova della nuova strategia europea peri dati: evoluzioni giurisprudenziali e percorsi di riforma*, in *Il diritto dell'informazione e dell'informatica*, 3/2022, pp.651-704. In tale materia quindi, benché si stia procedendo ad una tassonomia, non si può non considerare che il successo di un'efficiente strategia digitale debba essere in grado di coordinare numerose normative, ciascuna delle quali deve essere in grado di inserirsi in un complesso contesto. Il "*filo rosso*" capace di ricollegare le differenti normative che si intrecciano potrebbe essere ricondotto, come affermato in G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, 4/2022, p.972, nella sequenza "*dati-algoritmi-piattaforme: i dati come risorsa fondamentale dell'economia e della società digitale; gli algoritmi come strumenti capaci di estrarre valore da tali dati (anche per finalità analitiche, predittive e decisionali); le piattaforme come luoghi virtuali nei quali avvengono — intermediati in via algoritmica e con costante profusione di dati, riutilizzati per aumentare il potere delle piattaforme medesime — buona parte degli scambi e delle interazioni sociali contemporanee*". In tale contesto, il DGA ed il *Data Act* inciderebbero sul primo nodo (i dati) in modo da promuoverne la libera circolazione, favorirne accesso e riuso. Per una disamina sul ruolo delle piattaforme e sui nuovi approcci regolatori che esse richiedono si rimanda a J. E. Cohen, *Law for the Platform Economy*, 51 U.C. Davis L. Rev, 2017. Per una necessità ed una panoramica (Europa, Stati Uniti, Cina e Russia) delle regolamentazioni delle piattaforme si veda G. Buttarelli, *La regolazione delle piattaforme digitali: il ruolo delle istituzioni pubbliche*, in *Giornale di diritto amministrativo*, 1/2023, pp. 116-127. Per il *Digital Markets Act* ed *Digital Services Act* si rinvia ad M. Libertini, *Il Regolamento europeo sui mercati digitali e le norme generali in materia di concorrenza*, in *Rivista trimestrale di diritto pubblico*, 4/2022 ed a F. Di Porto, T. Grote, G. Volpi, R. Invernizzi, *Talking at Cross Purposes? A computational analysis of the debate on informational duties in the digital services and the digital markets acts*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. I, Roma, Roma TrE-Press, 2023, pp. 299-352.

[66] L'elenco di tali strumenti rientranti nella Strategia per i dati è stato preso da M. Van Eechoud, *Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research*, European Commission – Directorate general for research and innovation, cit., p.5. Del resto, ciascuno di questi atti è citato, direttamente o indirettamente, nella comunicazione della Commissione "*Una strategia europea per i dati*". Non v'è dubbio che il DGA e il *Data Act* rientrino

Per delineare alcuni profili relativi alla circolazione dei dati, si potrebbero descrivere le caratteristiche salienti di tali atti, focalizzando successivamente l'attenzione sul GDPR (e sulla normativa complementare, il RDNP), che, benché cronologicamente anteriore rispetto agli altri atti, costituisce ancora il cardine della disciplina per quanto riguarda i dati personali[67].

Il *Data Governance Act* si articola lungo tre linee direttrici principali: il riutilizzo dei dati detenuti dagli enti pubblici; la disciplina dei servizi di intermediazione dei dati e la disciplina relativa all'altruismo dei dati[68].

Il primo ambito di intervento, che si pone in continuità con la Direttiva *Open Data*, è volto ampliare il *secondary use* dei dati, personali e non personali, raccolti dalla pubblica amministrazione, la quale, a sua volta, potrà pretendere delle tariffe necessarie a coprire

---

tra i più importanti tasselli della strategia sui dati, volti a regolare circolazione e fruizione dei dati, si veda in merito, A. Cataleta, *Innovazione della PA, un anno di fermento normativo tecnologico*, in *FPA-Annual Report 2023*, Edizioni Forum PA, 2023, pp.140-142.

[67] Benché le normative siano molteplici e vi possano essere contraddizioni, viene sottolineato che "*The GDPR [omissis] has priority over all three* (DGA; ODD; Data Act)". In tal senso M. Van Eechoud, *Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research*, European Commission – Directorate general for research and innovation, cit., p.9. Ad esempio, l'art. 3, par. 3, del DGA, sancisce che "*In caso di conflitto tra il presente regolamento e il diritto dell'Unione in materia di protezione dei dati personali* [omissis] *prevale il pertinente diritto dell'Unione o nazionale in materia di protezione dei dati personali. Il presente regolamento non crea una base giuridica per il trattamento dei dati personali e non influisce sui diritti e sugli obblighi di cui ai regolamenti (UE) 2016/679 e (UE) 2018/1725* [omissis]". Lo stesso dicasi, come si vedrà, con riferimento al RDNP. Ulteriori riflessioni saranno fatte al termine del capitolo. Nel proseguo verrà prima esposta la disciplina del DGA e del *Data Act*; dell'ODD si è invece già discusso in precedenza; in seguito verranno esposte le discipline del RDNP e del GDPR.

[68] In tal senso M. Van Eechoud, *Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research*, European Commission – Directorate general for research and innovation, cit., p.26. Del resto, la rilevanza di queste tre direttrici principali emerge dalla stessa sistematica del DGA; infatti il capo II, rubricato "*Riutilizzo di determinate categorie di dati protetti detenuti da enti pubblici*" (artt. da 3 a 9), disciplina il la prima linea direttrice, mentre la seconda e la terza sono disciplinate, rispettivamente, dal capo III (rubricato "*Requisiti applicabili ai servizi di intermediazione dei dati*", artt. da 10 a 15) e dal capo IV (rubricato "*Altruismo dei dati*", artt.da 16 a 25). Vi è da sottolineare inoltre che il capo VI ("*Comitato europeo per l'innovazione in materia di dati*", artt.29-30), prevede l'istituzione di un comitato europeo per l'innovazione in materia di dati sotto forma di un gruppo di esperti, volto ad implementare il Regolamento.

i costi dovuti dalla messa a disposizione dei dati (ad esempio, i costi relativi alla diffusione dei dati, al mantenimento di un ambiente di trattamento sicuro e così via)[69]; una novità da sottolineare in quest'ambito è che il DGA consente il riutilizzo di categorie di dati pubblici, ulteriori rispetto a quelli che rientrano nell'ambito di applicazione della *Open Data Directive*, soggetti a restrizioni relative alla circolazione, quali la personalità del dato, i diritti di proprietà intellettuale o i segreti commerciali[70].

Il secondo ambito di intervento mira invece a regolamentare aspetti più privatistici relativi ai servizi di intermediazione dei dati, distinguendo in tal modo, all'interno del mercato dei dati, fornitura, intermediazione ed utilizzo dei dati. Si consente agli intermediari di mettere a disposizione un'infrastruttura tecnica, al fine di interconnettere i "titolari dei dati" e gli "utenti dei dati"[71]. Nel riconoscere[72] e disciplinare tale mercato la prospettiva della Commissione è "*to create or to sustain new*

---

[69] Si veda nel merito F. Bravo, *Il principio di solidarietà tra data protection e data governance*, in *Il diritto dell'informazione e dell'informatica*, 3/2023, pp.494-495.

[70] Si veda nel merito l'art. 3 del DGA. Come misura volta a contrastare i diritti di esclusiva, rileva anche l'art. 4, rubricato "*Divieto di accordi di esclusiva*".

[71] Si veda il Considerando n. 32 del DGA.

[72] Gli intermediari di dati potrebbero essere assimilati ai c.d. *data brokers* (*information brokers* o *information resellers*), ossia quei soggetti che acquisiscono dati digitali, recuperati da diverse fonti (generalmente da fonti pubbliche ed *open data*, ma possono anche essere acquisiti direttamente dai titolari dei dati) allo scopo di rivenderli sul mercato, dopo averli aggregati, elaborati ed analizzati. Essi operano solitamente nel mercato secondario, cioè raggruppano ed organizzano dati che sono già in circolazione ed il valore aggiunto del loro lavoro consiste proprio nell'elaborazione ed organizzazione dei dati (per, ad esempio, incasellare gli individui in una categoria di riferimento per gli annunci pubblicitari). In E. Parsier, , *The Filter Bubble: what the Internet is Hidining from you*, Londra, Viking, 2011, è affermato che una società di *brokeraggio*, la *Acxiom* detiene le informazioni del novantasei per cento delle famiglie statunitensi. In tal senso si veda M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., pp.39-42. Per approfondire l'argomento del ruolo dei *data brokers* nel *framework* normativo europeo, si veda H. Ruschemeier, *Data Brokers and European Digital Legislation*, in *European Data Protection Law Review*, 1/2023, pp.27-38, mentre per la loro importanza in una società che deve affrontare un contesto *Big Data*, si veda C.L. Yeh, *Pursuing consumer empowerment in the age of big data: A comprehensive regulatory framework for data brokers*, in *Telecommunications policy*, 05/2018, Volume 42, Fascicolo 4, pp. 282-292. Si noti che la figura degli intermediari dei dati che operano nel mercato secondario, se ben regolamentata, potrebbe contrastare l'opacità che caratterizza tale mercato; a proposito si veda V. Zeno-Zencovich, *Do "data markets" exist?*, in *MediaLaws*, 2/2019, pp. 22 ss.

business opportunities for European enterprises and advantages for Public Institutions, as well"[73], considerando che questa profonda innovazione dovrà rispettare la normativa già esistente, coordinandosi in particolare con il GDPR[74]. Quindi tali servizi mirano ad instaurare "*rapporti commerciali ai fini della condivisione dei dati tra un numero indeterminato di interessati e di titolari dei dati, da un lato, e gli utenti dei dati, dall'altro*"[75]; tali soggetti, agevolando lo scambio bilaterale ed una messa in comune dei dati, hanno il potenziale di rendere più florido il mercato dei dati e di contrastare gli operatori che detengono un grado significativo di potere di mercato[76].

---

[73] In tal senso F. Bravo, *Intermediazione di dati personali e servizi di data sharing dal GDPR al Data Governance Act*, in *Contratto e impresa Europa*, 1/2021, p.200. Vi è inoltre da sottolineare che l'attenzione al mercato dei dati personali e non personali era già stata avvertita nel GDPR, specialmente con riferimento alla portabilità dei dati, che consente all'interessato di chiedere al titolare dei dati di trasferire questi ultimi ad altro titolare; come affermato in D. C. Cravo, *How to Make Data Portability Right More Meaningful for Data Subjects?,* in *European Data Protection Law Review*, 1/2022, p.53, la portabilità è un "*user-centered tool, which enables the data subject to play an active role in the data ecosystem*" e abilita al c.d. "*multihoming*" delle piattaforme, ossia la possibilità dell'interessato di passare dai servizi di un titolare dei dati ad un altro, in conseguenza al trasferimento dei dati e senza la necessità che quest'ultimo debba riacquisirli. Tale diritto ha una portata innovativa, in quanto viene previsto per la prima volta nel GDPR (non era disciplinato dalla Direttiva 95/46/CE), tuttavia, secondo l'autrice da ultimo citata, manca ancora una piena conoscenza di tale diritto da parte degli interessati. Tale diritto, in ambito concorrenziale, potrebbe essere rilevante anche per contrastare il potere di mercato delle poche – grandi – società che detengono un grande quantitativo di dati personali. Per approfondimenti al diritto sulla portabilità si rinvia a S. Troiano, *Il diritto alla portabilità dei dati*, in N. Zorzi Galgano (a cura di), *Persona e mercato dei dati: riflessioni sul GDPR*, Milano, Wolters Kluwer, 2019, pp. 195 ss ed a B. Engels, *Data portability among online platforms*, in *Internet policy review-Journal on internet regulation*, vol.5, issue 2, pp.1-17.

[74] *Ibidem*.

[75] Art. 2, par. 1, n. 11, DGA.

[76] F. Bravo, *Intermediazione di dati personali e servizi di data sharing dal GDPR al Data Governance Act*, in *Contratto e impresa Europa*, cit., pp.238-239. La strada intrapresa in quest'ambito di intervento è quella di potenziare circolazione e mercato dei dati, sebbene non manchino preoccupazioni di trattamenti GDPR *compliant*, sottolineate anche nell'*Opinion* n. 3/2020 dell'*European Data Protection Supervisor* (EDPS). Per quest'ultima tematica e per una disamina più approfondita dell'ambito di intervento relativo all'intermediazione dei dati si rimanda a F. Bravo, *Intermediazione di dati personali e servizi di data sharing dal GDPR al Data Governance Act*, in *Contratto e impresa Europa*, cit., pp. 199-256.

Il terzo ambito di intervento riguarda invece l'altruismo dei dati, che consiste nella condivisione volontaria ed a titolo gratuito di dati (salvo il rimborso per la messa a disposizione dei dati), personali o non personali, per finalità di interesse generale, in virtù del consenso espresso dagli interessati in base agli artt. 6 e 9 del GDPR[77]. Il DGA prevede l'istituzione di un registro in cui si devono iscrivere tali organizzazioni, che devono possedere specifici requisiti e rispettare specifici obblighi di trasparenza; sono previste inoltre misure di monitoraggio e di vigilanza ad opera di apposite autorità.

Il *Data Act*, misura complementare al DGA, mira a promuovere la libera circolazione dei dati, per favorirne accesso e riuso[78]; tuttavia, invece di regolare il flusso normativo *Administration to Business*, mira a potenziare il flusso *Business to Administration*,

[77] F. Bravo, *Il principio di solidarietà tra data protection e data governance*, in *Il diritto dell'informazione e dell'informatica*, cit., pp. 495-496. L'altruismo dei dati viene definito all'art. 2, par. 1, n. 16 del DGA come "*la condivisione volontaria di dati sulla base del consenso accordato dagli interessati al trattamento dei dati personali che li riguardano, o sulle autorizzazioni di altri titolari dei dati volte a consentire l'uso dei loro dati non personali, senza la richiesta o la ricezione di un compenso che vada oltre la compensazione dei costi sostenuti per mettere a disposizione i propri dati, per obiettivi di interesse generale, stabiliti nel diritto nazionale, ove applicabile, quali l'assistenza sanitaria, la lotta ai cambiamenti climatici, il miglioramento della mobilità, l'agevolazione dell'elaborazione, della produzione e della divulgazione di statistiche ufficiali, il miglioramento della fornitura dei servizi pubblici, l'elaborazione delle politiche pubbliche o la ricerca scientifica nell'interesse generale*". Sul concetto di altruismo dei dati si veda M. Taddeo, Data Philantropy and Individual Rights, in *27 Minds and Machines*, Vol.27, 2017, pp.1-5; sull'importanza dell'altrusimo dei dati per contrastare il potere degli oligopoli si veda B. Prainsack, *Data Donation: How to resist the iLeviathan*, in J. Krutzinna, L. Floridi (a cura di), *The Ethics of Medical Data Donation*, Cham, Springer, pp. 9-22.

[78] Tale atto risulta essere uno strumento orizzontale ed è articolato attorno a cinque obiettivi principali "*(i) facilitare l'accesso e l'utilizzo dei dati da parte di consumatori e attori di mercato, (ii) consentire al settore pubblico l'utilizzo, in casi eccezionali, di dati detenuti dalle principali imprese e piattaforme, (iii) rendere più agevole il passaggio tra diversi servizi cloud ed edge, (iv) offrire strumenti di salvaguardia contro trasferimenti di dati illegittimi e senza preventiva notifica da parte di cloud providers verso governi o altri enti di Stati terzi e (v) imporre standard di interoperabilità per i dati*", così come affermato in C. Sganga, *Ventisei anni di Direttiva database alla prova della nuova strategia europea peri dati: evoluzioni giurisprudenziali e percorsi di riforma*, in *Il diritto dell'informazione e dell'informatica*, cit., p.698.

cercando di codificare il modello della c.d. "*reverse PSI*", imponendo delle regole al trasferimento dei dati dal settore privato a quello pubblico[79].

Le altre due normative complementari sono il Regolamento per libera circolazione dei dati non personali ed il Regolamento per la protezione dei dati personali; dal punto di vista normativo, infatti, le banche dati, possono contenere due tipologie di dati, personali e non personali, a cui si ricollegano discipline differenti. Il RDNP è un altro atto-sintomo, che la questione relativa alla circolazione dei dati sta diventando sempre più un'esigenza trasversale e rilevante, perciò costituisce un atto volto a completare il "*quadro globale per uno spazio comune europeo dei dati e per la libera circolazione di tutti i dati all'interno dell'Unione europea*"[80].

---

[79] G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, 4/2022, pp. 979-980. Sui meccanismi di acquisizione dell'informazione B2A rileva in particolare il capo V del *Data Act*; l'art. 14 infatti consente ad un ente pubblico che dimostri una "*necessità eccezionale, di cui all'articolo 15, di utilizzare i dati*" per svolgere le proprie funzioni statutarie nell'interesse pubblico di ottenere i dati da privati su richiesta motivata. A norma dell'art. 15 tra le necessità eccezionali rientra la mancanza di dati, che impedisce all'ente pubblico di svolgere un compito specifico nell'interesse pubblico previsto dalla legge (quali la redazione di statistiche ufficiali, la mitigazione o la ripresa dopo un'emergenza pubblica); tale disciplina potrebbe costituire un tassello importante per rafforzare il potere conoscitivo delle pubbliche amministrazioni, sebbene sia ancora "*at an embryonic stage*" secondo quanto affermato (e a cui si rinvia per approfondimenti sul *data sharing* B2A) in A. Vigorito, *Government Access to Privately-Held Data: Business-to-Government Data Sharing. Voluntary and Mandatory Models*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. II, Roma, Roma TrE-Press, 2023, p. 718.

In dottrina, tuttavia, vengono sottolineate le possibili criticità del *Data Act*, dovute agli appesantimenti burocratici in assenza di agevolazioni di altro tipo, si veda W. Kerber, *Governance of IoT Data: Why the EU Data Act Will not Fulfill Its Objectives*, in *GRUR International*, 2/2023, pp. 120-135; più in generale, per il rischio che una *compliance* troppo elevata (come potrebbe essere quella imposta ultimamente dall'UE) possa andare a detrimento dello sviluppo tecnologico, si veda T. E. Frosini, *L'ordine giuridico del digitale*, in Ceridap, 2/2023, pp.36-65.

[80] Comunicazione della Commissione, *Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union*, COM (2019) 250 final, p. 2. L'intento della Commissione è proprio quello di garantire una circolazione maggiore, più chiara e più sicura dei dati, come affermato anche in altre sedi: "*Il regolamento sulla libera circolazione dei dati non personali, insieme al regolamento generale sulla protezione dei dati (GDPR), ha stabilito la circolazione illimitata di tutti i dati in tutta Europa*" (in tal senso, si veda *Cloud and Edge Computing: a different way of using IT — Brochure*). Inoltre, la necessità di tale atto emerge anche se si considera che l'ipotesi più frequente è quella in cui vengano trattati dati non aventi

L'analisi sarà focalizzata in particolare sui dati personali[81], mentre rispetto alla disciplina dei dati non personali, ci si limiterà ad osservare che "*tanto la definizione quanto la disciplina del dato non personale del Regolamento 1807 sono state plasmate in maniera tale da soddisfare al meglio le esigenze di libera circolazione all'interno dell'Unione*"[82].

carattere personale, così come affermato nell'indagine conoscitiva sui *Big data* effettuata congiuntamente dalle tre autorità indipendenti (AGCOM, AGCM e Garante della *Privacy*) e conclusasi nel 2019.In aggiunta, anche il Considerando n. 9 dell'appena menzionato Regolamento chiarisce che i dati non personali sono destinati a moltiplicarsi: "*L'espansione dell'Internet degli oggetti, l'intelligenza artificiale e l'apprendimento automatico rappresentano fonti importanti di dati non personali, ad esempio a seguito del loro utilizzo in processi automatizzati di produzione industriale*".

[81] In quanto è tale disciplina, caposaldo normativo dell'Unione, che sta attraversando una crisi dovuta all'emergere delle nuove tecnologie. Per esaminare tale crisi in questo paragrafo si esporrà la disciplina dello stato dell'arte, mentre nel terzo capitolo si vedranno gli aspetti critici.

[82] S. Torregiani, *Il dato non personale alla luce del Regolamento (UE) 2018/1807: tra anonimizzazione, ownership e Data by Design*, in federalismi.it, n. 18/2020, p.319. Tale circostanza può essere dedotta anche dal fatto che le basi giuridiche dei Regolamenti sono differenti, mentre per il GDPR è l'art. 16 TFUE (che sancisce il diritto di ogni persona alla protezione dei dati di carattere personale), per il RDNP è l'art. 114 TFUE (il quale attribuisce al Parlamento ed al Consiglio la competenza di armonizzazione al fine di ravvicinare le norme nazionali che rappresentano un ostacolo alla libera circolazione). Inoltre, lo stesso articolo 1 del Regolamento chiarisce che: "*Il presente regolamento mira a garantire la libera circolazione dei dati diversi dai dati personali all'interno dell'Unione stabilendo disposizioni relative agli obblighi di localizzazione dei dati, alla messa a disposizione dei dati alle autorità competenti e alla portabilità dei dati per gli utenti professionali*". Un aspetto interessante, legato ad una maggiore circolazione dei dati, è ad esempio rappresentato dall'articolo 4 del RDNP il quale sancisce un divieto generalizzato di imporre obblighi di localizzazione dei dati, salvo che siano giustificati da motivi di sicurezza pubblica. Secondo l'articolo 3, comma 5, gli obblighi di localizzazione sono definiti come disposizioni di legge, orientamenti o pratiche amministrative che, in modo diretto o indiretto, richiedono che il trattamento dei dati non personali avvenga all'interno di uno specifico Stato membro, o che ne impediscono il trattamento in un altro Stato.
Tuttavia, come osserva l'autore, l'assetto complessivo della circolazione dei dati, in virtù della residualità del RDNP rispetto al GDPR, rappresenta la volontà del legislatore europeo di attribuire maggiore importanza alla protezione dei dati rispetto che alla circolazione. Tra gli strumenti volti ad aumentare la certezza giuridica ed il livello di fiducia negli scambi relativi a tali dati vi sono, oltre alla rimozione degli obblighi di localizzazione dei dati, la portabilità dei dati non personali e l'autoregolamentazione affidata a codici di condotta.

2.2. La tutela dei dati personali

Per quanto riguarda, invece, la disciplina volta a tutelare i dati personali, essa riveste un ruolo di grande impatto sulle banche dati. Infatti, è in grado di influire sia su questioni organizzative della PA, sia sull'attività amministrativa di acquisizione e trattamento dei dati[83].

Per questi motivi, una tutela troppo rigida dei dati personali potrebbe essere di ostacolo all'interoperabilità e alle tecniche di analisi algoritmica, invece una tutela troppo

[83] La tutela dei dati personali è infatti uno dei fattori che ha contribuito a determinare l'architettura distribuita delle banche dati. Viene affermato in B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, I ed., Milano, FrancoAngeli, 2023, p.129ss, dove viene chiarito che "*le esigenze di tutela dei dati personali – oggi così centrali e impattanti rispetto alla modalità di acquisizione e trattamento – sono anche uno dei fattori che hanno contribuito a determinare – storicamente – l'architettura distribuita e deconcentrata dei data set, all'interno del sistema pubblico complessivamente considerato. Infatti, già agli albori dei processi di informatizzazione è maturata la consapevolezza che la concentrazione di informazioni disponibili a un medesimo soggetto abilitasse (sia in termini attuali, allora, sia in termini di potenzialità ancora inespresse) un potere conoscitivo idoneo a incidere in modo significativo sulle libertà e sui diritti dei cittadini, con rilevantissimi rischi di loro limitazione, condizionamento, manipolazione, negazione. Dunque, già negli anni '70 del secolo scorso veniva posta all'ordine del giorno la necessità di provvedere affinché fosse impedito allo Stato (al sistema pubblico considerato nel suo insieme) di realizzare una concentrazione dei dati personali raccolti e trattati per l'esercizio delle funzioni amministrative, mediante una illimitata integrazione delle basi di dati ovvero una loro illimitata/non presidiata circolazione, proprio a protezione e a garanzia della libertà dei consociati*". Dunque, l'organizzazione delle banche dati decentrate (c.d. a "*silos*") era volta al fine di ottenere una maggiore garanzia di tutela dei dati personali. Ma le modifiche a livello organizzativo (che riguardano sia l'organizzazione del titolare nel suo complesso, sia le caratteristiche dell'infrastruttura su cui poggiano le banche dati) sono anche dovute, direttamente ed indirettamente, dalle disposizioni del Regolamento del parlamento europeo e del consiglio del 27 aprile 2016 n. 679 (GDPR). Si pensi a determinati obblighi di sicurezza dal punto di vista tecnico, alla nomina obbligatorio del DPO, ai principi di *privacy by design e by default* e così via. Per un ulteriore esame sui rapporti tra protezione dei dati personali ed azione ed organizzazione amministrativa si rimanda a S. Franca, *La semplificazione nelle modalità di trattamento dei dati personali da parte della pubblica amministrazione*, in *Diritto pubblico*, 2/2021, pp. 635-665.

permissiva andrebbe a ledere un diritto, che è riconosciuto da molto prima che i dati costituissero il "*nuovo petrolio*"[84].

[84] In tal senso L. Torchia, *Lo Stato digitale. Una introduzione*, Bologna, Il Mulino, 2023, p.49, dove viene sottolineato che tale diritto storicamente nasce come il "*right to be let alone*" teorizzato da due giuristi americani in un saggio (D.Warren, L.D. Brandeis, *The right to privacy*, in *Harvard law review*, Vol.VI, n. 5, 1890) a fine '800 ed inteso come libertà negativa, cioè il diritto di essere lasciato solo, di escludere gli altri dalla propria sfera privata. Successivamente viene inteso anche nella sua accezione di libertà positiva, ossia la capacità dell'individuo di controllare l'uso che altri soggetti fanno dei dati che a lui si riferiscono. Tale accezione era contenuta nell'art. 8 della Carta dei diritti fondamentali dell'Unione Europea proclamata per la prima volta il 7 dicembre 2000 a Nizza ed è ugualmente contenuta (essendo rimasto invariato l'articolo 8) nella seconda versione pubblicata il 7 giugno 2016 (inoltre è da sottolineare che, con l'entrata in vigore del Trattato di Lisbona il primo dicembre 2009, la Carta è diventata giuridicamente vincolante). Il primo comma dell'art. 8 prevede che "*Ogni persona ha diritto alla protezione dei dati di carattere personale che la riguardano*", il secondo comma in aggiunta dispone che "*Tali dati devono essere trattati secondo il principio di lealtà, per finalità determinate e in base al consenso della persona interessata o a un altro fondamento legittimo previsto dalla legge. Ogni persona ha il diritto di accedere ai dati raccolti che la riguardano e di ottenerne la rettifica*". Tale diritto è mutato molto con l'avvento della società digitale, che ha fatto avvertire l'esigenza di una nuova disciplina. L'art. 8 appena menzionato e la giurisprudenza della Corte di giustizia hanno posto le basi al GDPR. Si pensi, solo per fare qualche esempio relativo alla giurisprudenza, alla sentenza del 13 maggio 2014 (cd "*Google Spain*" – Causa C-131/12), dove la Corte ha elaborato il "diritto all'oblio" ed ha esteso il diritto dell'Unione al titolare del trattamento, indipendentemente dalla sua collocazione territoriale. O, ancora, a quando la Corte il 6 ottobre 2015 (C-362/14), è giunta all'annullamento del *safe harbour* (decisione di adeguatezza della Commissione europea del 26 luglio 2000, n. 2000/520/CE), con il quale la Commissione aveva autorizzato il trasferimento dei dati personali nei confronti deli USA, poiché aveva ritenuto inidoneo il livello di tutela garantito dall'ordinamento statunitense, a differenza di quanto aveva ritenuto la Commissione. Tale giurisprudenza, che fa sempre riferimento anche all'art. 8 della CDFUE, ha posto le basi al Regolamento n. 679/2016. In ogni caso un ulteriore presidio alla tutela dei dati personali, come afferma B. Cortese, *Commento all'art. 286 TCE*, in A. Tizzano (a cura di), *Trattati dell'Unione e della Comunità Europea*, , Milano, Giuffrè, 2004, pp.1284-1287 , era già previsto dall'art. 286 della versione consolidata del 1997 del trattato che istituisce la comunità europea (versione del trattato che istituisce la comunità economica europea modificata dal trattato di Amsterdam nel 1997 - https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:11997E/TXT -) e dall'art. 16 TFUE (ex art. 286 del TCE), il quale sancisce che "*1. Ogni persona ha diritto alla protezione dei dati di carattere personale che la riguardano. 2. Il Parlamento europeo e il Consiglio, deliberando secondo la procedura legislativa ordinaria, stabiliscono le norme relative alla protezione delle persone fisiche con riguardo al trattamento dei dati di carattere personale da parte delle istituzioni, degli organi e degli organismi dell'Unione, nonché da parte degli Stati membri nell'esercizio di attività che rientrano nel campo di applicazione del diritto dell'Unione, e le norme*

Con l'avvento della digitalizzazione e delle banche dati, tuttavia, questo diritto subisce una prima minaccia, consistente in una circolazione ed acquisizione dei dati facilitata. Una seconda minaccia è invece rappresentata dalle recenti tecniche di elaborazione del dato attraverso procedure automatizzate ed algoritmiche[85].

Minacce che sono state opportunamente prese in considerazione dal legislatore comunitario e nazionale in base al progredire della tecnologia. La materia risulta attualmente disciplinata, a livello europeo, dal Regolamento generale sulla protezione dei dati – GDPR – e, a livello nazionale, dal decreto legislativo 30 giugno 2003, n. 139 (Codice in materia di protezione dei dati personali), così come recentemente modificato dal decreto legge 8 ottobre 2021, n. 139 (cd decreto "capienze"), convertito nella legge 3 dicembre 2021, n. 205[86].

Per comprendere come viene risolto il bilanciamento tra, da un lato, tutela dei dati personali e, dall'altro, le esigenze relative alla circolazione e all'elaborazione dei dati, è necessario, in primo luogo, analizzare le norme e i principi relativi al GDPR, in secondo luogo, individuare quali sono gli ambiti "scoperti", in cui il GDPR ha permesso al

---

relative alla libera circolazione di tali dati. Il rispetto di tali norme è soggetto al controllo di autorità indipendenti. Le norme adottate sulla base del presente articolo fanno salve le norme specifiche di cui all'articolo 39 del trattato sull'Unione europea". Per una ricostruzione della disciplina sulla tutela dei dati personali si rinvia a S. Rodotà, *Tecnologie e diritti*, Bologna, Il Mulino, 2021 e per l'importanza e la supremazia della protezione dei dati rispetto alla circolazione e riutilizzo di essi si veda T. Streinz, *The Evolution of European Data Law*, in P. Craig e G. de Búrca (a cura di), *The Evolution of EU Law*, Oxford, Oxford University Press, 2021, pp.902 ss.

[85] L. Torchia, *Lo Stato digitale*, cit., p. 25, dove si afferma che "*la doppia capacità, conoscitiva e predittiva, degli algoritmi pone numerosi problemi giuridici inediti [omissis]. La disciplina di tutela della privacy, la disciplina antitrust e la disciplina di tutela dei consumatori sono i tre principali plessi giuridici che vanno oggi modificandosi, almeno nell'ordinamento europeo, proprio per trovare una risposta ed una soluzione a quei problemi*".

[86] In realtà le ultime modifiche sono da attribuirsi al decreto legislativo 10 Marzo 2023, n. 24, ma non sono rilevanti ai nostri fini, in quanto le modifiche sono state apportate per fini attuativi della c.d. "direttiva *Whistleblowing*" (direttiva (UE) 2019/1937 del Parlamento europeo e del Consiglio, del 23 ottobre 2019), a tutela delle persone che segnalano violazioni di disposizioni normative nazionali o dell'Unione europea che ledono l'interesse pubblico o l'integrità dell'amministrazione pubblica o dell'ente privato, di cui siano venute a conoscenza in un contesto lavorativo pubblico o privato.

legislatore nazionale di modulare alcune norme del Regolamento (c.d. "margine di manovra") ed, in terzo luogo, esaminare la normativa nazionale in tali ambiti lasciati scoperti[87].

2.2.1. La normativa europea

Il regolamento ha l'obiettivo di garantire un livello adeguato di protezione dei dati personali delle persone fisiche, assicurando al contempo la libera circolazione di tali dati nell'ambito dell'Unione[88].

---

[87] Tale impostazione è stata ripresa da B. Ponti, *attività amministrativa e trattamento dei dati personali*, cit., *passim*. Nel terzo capitolo si affronteranno invece alcune questioni legate alle nuove sfide poste dai *Big Data* e dall'elaborazione algoritmica dei dati personali.

[88] Così come affermato in L. Durst, *Oggetto e finalità: un nuovo statuto giuridico dei dati personali,* in R. Panetta (a cura di), *Ciroclazione e protezione dei dati personali, tra libertà e regole del mercato. Commentario al Regolamento UE n. 679/2016 e al d.lgs n. 101/2018,* Milano, Giuffrè, 2019, p. 43, in cui viene evidenziato che la *ratio* del nuovo impianto di tutela è delineato all'art. 1 del regolamento, che al primo comma precisa che "*il presente Regolamento stabilisce norme relative alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché norme relative alla libera circolazione di tali dati*", mentre al secondo ed al terzo comma esplicita le finalità che riguardano ciascuno dei due ambiti indicati nel primo comma. Il secondo comma statuisce infatti "*il presente regolamento protegge i diritti e le libertà fondamentali delle persone fisiche, in particolare il diritto alla protezione dei dati personali*", mentre il terzo comma afferma che "*la libera circolazione dei dati personali nell'Unione non può essere limitata né vietata per motivi attinenti alla protezione delle persone fisiche con riguardo al trattamento dei dati personali*". Tale disposizione si pone in piena aderenza con l'art. 8 della Carta dei diritti fondamentali UE e con l'art. 16 del TFUE. Inoltre, come osserva F. Tigano, *Protezione dei dati personali e pubblica amministrazione: alcuni spunti di riflessione,* in *Diritto e società*, n. 2, 2022, pp.418-419, benché la protezione dei dati personali rappresenti un diritto fondamentale, ciò non significa che essi non possano essere limitati in virtù di preminenti interessi pubblici, come dimostra la fase pandemica e alcuni Considerando del GDPR, tra cui segnaliamo il quarto ed il sedicesimo. Il quarto considerando prevede che *"Il trattamento dei dati personali dovrebbe essere al servizio dell'uomo. Il diritto alla protezione dei dati di carattere personale non è una prerogativa assoluta, ma va considerato alla luce della sua funzione sociale e va contemperato con altri diritti fondamentali, in ossequio al principio di proporzionalità"*, il sedicesimo prevede che il Regolamento "*non si applica a questioni di tutela dei diritti e delle libertà fondamentali o di libera circolazione dei dati personali riferite ad attività che non rientrano nell'ambito di applicazione del diritto dell'Unione, quali le attività riguardanti la sicurezza nazionale* [omissis]". In altre parole, la formula adottata all'art. 1 GDPR, non

rappresenta un contrasto, ma la necessità di effettuare di volta in volta un bilanciamento tra i diversi principi, senza affermare direttamente la supremazia dell'uno o dell'altro; si veda in tal senso F. Bravo, *Sul bilanciamento proporzionale dei diritti e delle libertà "fondamentali", tra mercato e persona: nuovi assetti nell'ordinamento europeo?*, in *Contratto e impresa*, 1/2018, pp. 190-216. Sulla "funzione sociale" del trattamento si rimanda a F. Bravo, *Il principio di solidarietà tra data protection e data governance*, in *Il diritto dell'informazione e dell'informatica*, cit., pp.504-510 e a A. Ricci, *Sulla "funzione sociale" del diritto alla protezione dei dati personali*, in *Contratto e impresa*, 2/2017, pp.586 ss.

Il GDPR detta alcune regole generali, valevoli per qualsiasi titolare[89] del trattamento[90]; altre sono specifiche per il caso in cui il titolare sia un soggetto di diritto pubblico[91].

[89] Per "titolare del trattamento" si intende ai sensi dell'art. 4 GDPR: "*la persona fisica o giuridica, l'autorità pubblica, il servizio o altro organismo che, singolarmente o insieme ad altri, determina le finalità e i mezzi del trattamento di dati personali; quando le finalità e i mezzi di tale trattamento sono determinati dal diritto dell'Unione o degli Stati membri, il titolare del trattamento o i criteri specifici applicabili alla sua designazione possono essere stabiliti dal diritto dell'Unione o degli Stati membri*". Si noti che il GDPR si applica a qualsiasi titolare del trattamento, persona fisica o giuridica, mentre tutela solo le persone fisiche, in quanto i dati personali si possono riferire solamente a quest'ultime. Ed infatti per "dato personale", sempre ai sensi dello stesso articolo, s'intende: *"qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); 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*".
All'interno dei dati personali, inoltre, si distinguono i dati comuni e particolari. Questi ultimi (definiti sempre all'art. 4) ricomprendono i dati genetici, quelli biometrici e quelli relativi alla salute. Per dati genetici s'intende: "*dati personali relativi alle caratteristiche genetiche ereditarie o acquisite di una persona fisica che forniscono informazioni univoche sulla fisiologia o sulla salute di detta persona fisica, e che risultano in particolare dall'analisi di un campione biologico della persona fisica in questione*". Per dati biometrici s'intendono "*i dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica che ne consentono o confermano l'identificazione univoca, quali l'immagine facciale o i dati dattiloscopici".* Per dati relativi alla salute s'intendono "*i dati personali attinenti alla salute fisica o mentale di una persona fisica, compresa la prestazione di servizi di assistenza sanitaria, che rivelano informazioni relative al suo stato di salute*". La nozione di dato personale è importante anche ai fini della distinzione con la nozione di dato non personale, in quanto quest'ultima definizione è residuale rispetto alla prima, ed infatti sono dati non personali: "*dati diversi dai dati personali definiti all'articolo 4, punto 1, del regolamento (UE) 2016/679*" (L'articolo 3, punto 1, RDNP).
Tale nozione di dato, c.d. "*semantica*", in virtù del fatto che la nozione di dato si colora di significato in funzione del collegamento con la persona, è rimasta la definizione principale su cui concepire il dato a livello normativo; con l'avvento del DGA, per la prima volta, viene definito il dato a livello sintattico come "*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*" (art. 2, par. 1, n. 1). In merito a tale ricostruzione si veda G. Resta, *Towards a unified regime of data-rights? Rapport de synthèse*, in T. Pertot (a cura di), *Rechte an Daten*, 2020, pp.231-248 e G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, cit., pp.973-975.

In via generale qualsiasi titolare del trattamento, a cui si applica il Regolamento[92], può acquisire i dati solamente in virtù di una base giuridica che ne legittima l'acquisizione, secondo quanto dispone l'art. 6 GDPR. Una volta effettuata l'acquisizione, andranno rispettati alcuni principi generali[93] ed andranno rispettati degli obblighi organizzativi ed

[90] Per "trattamento" si intende ai sensi dell'art. 4 GDPR: "*qualsiasi operazione o insieme di operazioni, compiute con o senza l'ausilio di processi automatizzati e applicate a dati personali o insiemi di dati personali, come la raccolta, la registrazione, l'organizzazione, la strutturazione, la conservazione, l'adattamento o la modifica, l'estrazione, la consultazione, l'uso, la comunicazione mediante trasmissione, diffusione o qualsiasi altra forma di messa a disposizione, il raffronto o l'interconnessione, la limitazione, la cancellazione o la distruzione*". Si noti l'ampiezza di tale nozione che è idonea a ricomprendere ogni operazione che coinvolge la "*catena del valore del dato*" – espressione di M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., p.27 ss. – che comprende le fasi di: acquisizione; conservazione; modellazione; immagazzinamento. Inoltre, nella nozione sono ricomprese anche eventuali comunicazioni e diffusioni dei dati. in questo modo l'oggetto del GDPR, in virtù delle nozioni ampie di "dato personale" e di "trattamento", risulta molto esteso. A ciò si aggiunge che, stando alla nozione di trattamento, il Regolamento risulta applicabile ad ogni tipologia di trattamento, non solo digitale, ma anche cartaceo o orale.

[91] In dottrina viene osservato infatti che ai soggetti pubblici vengono applicate delle norme speciali, in aggiunta a tutti gli obblighi generali cui sono assoggettati i titolari di trattamento privati, che rimangono sempre applicabili; S. M. A. D'Ancona, *Trattamento e scambio di dati e documenti tra pubbliche amministrazioni, utilizzo delle nuove tecnologie e tutela della riservatezza tra diritto nazionale e diritto europeo*, in *Riv. it. dir. pubbl. com.*, 3, 2018, par. 4.

[92] Anche l'ambito di applicazione risulta molto ampio. Comprende un ambito di applicazione materiale (art. 2 del GDPR) e territoriale (art. 3 del GDPR). Secondo quanto dispone l'art. 2 la disciplina si applica "*al trattamento interamente o parzialmente automatizzato di dati personali e al trattamento non automatizzato di dati personali contenuti in un archivio o destinati a figurarvi*", anche se il secondo comma prevede delle eccezioni e la più rilevante riguarda i trattamenti "e*ffettuati da una persona fisica per l'esercizio di attività a carattere esclusivamente personale o domestico*". L'ambito di applicazione territoriale invece permette l'applicazione delle disposizioni del Regolamento sia in base al principio dello stabilimento (applicazione della disciplina nei confronti dei titolari e dei responsabili che abbiano il loro stabilimento nell'Unione, indipendentemente dal fatto che il trattamento sia effettuato o meno nell'Unione), sia in base al principio del *target*. Quest'ultimo (già affermato dalla Corte di Giustizia nel caso *Google Spain*), insieme ad un'impostazione strutturale europea più ampia, mira a comporre la "sovranità digitale europea" – in questo senso L. Torchia, *Lo Stato digitale*, cit., p.51 – , in quanto dispone che il Regolamento venga applicato al trattamento di dati personali di interessati che si trovano nell'Unione ed effettuato da un titolare o da un responsabile del trattamento che non è stabilito nell'Unione.

[93] Tra i più importanti vi sono quelli sanciti all'art. 5 ed all'art. 25 del GDPR.

informativi, pena la possibilità di incorrere in sanzioni amministrative, civili o penali (queste ultime, solo se introdotte dallo Stato membro).

Quindi il presupposto fondamentale per effettuare un trattamento è la liceità, cioè essere legittimati in virtù di una base giuridica individuata nell'art. 6 GDPR. Tale articolo, infatti, al primo comma dispone che:

"Il trattamento è lecito solo se e nella misura in cui ricorre almeno una delle seguenti condizioni:

a) *l'interessato ha espresso il consenso al trattamento dei propri dati personali per una o più specifiche finalità;*
b) *il trattamento è necessario all'esecuzione di un contratto di cui l'interessato è parte o all'esecuzione di misure precontrattuali adottate su richiesta dello stesso;*
c) *il trattamento è necessario per adempiere un obbligo legale al quale è soggetto il titolare del trattamento;*
d) *il trattamento è necessario per la salvaguardia degli interessi vitali dell'interessato o di un'altra persona fisica;*
e) *il trattamento è necessario per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento;*
f) *il trattamento è necessario per il perseguimento del legittimo interesse del titolare del trattamento o di terzi, a condizione che non prevalgano gli interessi o i diritti e le libertà fondamentali dell'interessato che richiedono la protezione dei dati personali, in particolare se l'interessato è un minore.*

La lettera f) del primo comma non si applica al trattamento di dati effettuato dalle autorità pubbliche nell'esecuzione dei loro compiti".

La base giuridica per eccellenza di un soggetto pubblico è individuata alla lettera e)[94], secondo lo schema della "*necessary clause*"[95] (clausola di necessarietà), secondo la quale il trattamento è lecito perché necessario al raggiungimento del determinato fine stabilito nella rispettiva lettera (schema con cui sono individuate tutte le basi giuridiche ad eccezione della lettera a), che permette di effettuare il trattamento in presenza del consenso).

Tale schema di liceità è previsto anche per i dati particolari, seppur in termini diversi[96]. L'articolo 9 GDPR elenca infatti alcune finalità che autorizzano il trattamento, tra cui figurano: la prestazione del consenso (salvo divieto imposto dal diritto dell'Unione o dello Stato membro); la tutela di un interesse vitale dell'interessato; motivi di interesse pubblico rilevante sulla base del diritto dell'Unione o degli Stati membri; fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici in conformità dell'articolo 89, paragrafo 1, sulla base del diritto dell'Unione o nazionale.

Si rileva che la base giuridica della lettera e) dell'art. 6 legittima il trattamento operato da un soggetto pubblico in due casi: da un lato, nel caso in cui sia necessario all'esecuzione di un compito di interesse pubblico e, dall'altro, nel caso in cui sia

[94] Il soggetto pubblico potrà infatti eseguire il trattamento in virtù di ciascuna base giuridica (ad eccezione della lett.f) se si tratta di una pubblica autorità). Ad esempio, potrà richiedere il consenso per offrire un determinato servizio. In altri casi la libertà di autodeterminazione viene meno in quanto il trattamento è imposto *ex lege*. Si pensi ai dati concernenti la posizione fiscale di un contribuente, la cui acquisizione è imposta dalla legge e che sono raccolti grazie alle dichiarazioni annuali o per mezzo di richieste ad altri enti. Tuttavia, il trattamento "fisiologico" per una pubblica amministrazione avviene in virtù della lett.e). Infatti il GDPR prevede dei limiti per l'utilizzo delle altre basi giuridiche da parte delle pubbliche amministrazioni; in merito a quest'ultimo punto si veda S. Franca, *I dati personali nell'amministrazione pubblica. Attività di trattamento e tutela del privato*, Università degli Studi di Trento, Collana della facoltà di giurisprudenza, 44, 2023, pp. 87-89.

[95] Come viene precisato in L. Torchia, *Lo Stato digitale*, cit., p. 57, la condizione di liceità può essere soddisfatta attraverso la prestazione del consenso o quando il trattamento è necessario (per questo motivo, clausola di necessarietà) per alcuni fini determinati all'art. 6. In quest'ultimo caso è lo stesso Regolamento che effettua un bilanciamento fra il diritto alla protezione dei dati ed altri interessi, come l'esecuzione di un compito di interesse pubblico.

[96] L'art. 9 del GDPR dispone come regola generale al primo paragrafo che è vietato trattare tali categorie di dati, tuttavia le numerose deroghe previste al secondo paragrafo dello stesso articolo, finiscono per inquadrare tale modello nello schema della *necessary clause*.

connesso all'esercizio di pubblici poteri. In ogni caso ciò che rileva è l'attribuzione specifica in capo al soggetto o di un determinato compito (di interesse pubblico) o dell'esercizio pubblici poteri, andando quindi a delineare una nozione oggettiva di funzione pubblica, poiché in entrambi i casi ciò che rileva è che l'attività svolta sia oggettivamente di pubblico interesse, a prescindere dalla natura formalmente pubblica o privata del soggetto[97].

Tale disposizione quindi, che ha efficacia diretta in quanto incorporata in un Regolamento, ha l'effetto di autorizzare il trattamento da parte di un'amministrazione per il sol fatto che questa risulta investita di un compito di interesse pubblico o

[97] La nozione oggettiva di funzione pubblica, così come la distinzione tra i due casi, viene evidenziata in B. Ponti, *Attività amministrativa e trattamento dei dati personali*, cit., p.46. Per quanto riguarda la differenza principale tra le due sotto-categorie, essa sembra consistere nella presenza o meno di un pubblico potere associato al compito di interesse pubblico, tuttavia in entrambi i casi "*vale la connessione di strumentalità necessaria quale requisito di legittimazione del trattamento di dati personali*", per questo motivo ciò che rileva in via principale è la "*strumentalità*" del trattamento rispetto all'attività di interesse pubblico, a prescindere dall'attribuzione dei pubblici poteri.

Tuttavia in determinati casi il Regolamento utilizza una nozione soggettiva di pubblica amministrazione, ad esempio sempre nel primo comma dell'art. 6 viene precisato che "*La lettera f) del primo comma non si applica al trattamento di dati effettuato dalle autorità pubbliche nell'esecuzione dei loro compiti*". In questo caso, per escludere la liceità del trattamento basato sul legittimo interesse del titolare, viene utilizzata la locuzione "*autorità pubbliche*" (non rientrando nella nozione quindi i soggetti privati, investiti di compiti di interesse pubblico o investiti di pubblici poteri – B. Ponti, *Attività amministrativa e trattamento dei dati personali*, cit., p.53). O ancora, tale locuzione è utilizzata anche all'art. 4 quando viene data una definizione di titolare del trattamento e si noti che in tale sede non viene specificata l'attività svolta e, quindi, il soggetto viene identificato in base ad un criterio prettamente soggettivo, come nota anche G. Carullo, *trattamento dei dati personali da parte delle pubbliche amministrazioni e natura del rapporto giuridico con l'interessato*, in *Rivista italiana di diritto pubblico comunitario*, n. 1-2, 2020, pp.134-137. L'autore si muove in senso sostanzialmente conforme rispetto alla distinzione tra nozione oggettiva e soggettiva, in quanto afferma che i soggetti titolari di una potestà pubblica (che possono essere soggetti privati o pubblici) possono esercitare il trattamento nell'esercizio di poteri autoritativi (ai sensi della lettera e) dell'art. 6, criterio oggettivo) o in quanto "autorità pubbliche" (criterio soggettivo).

Nel proseguo faremo riferimento generico ai "soggetti pubblici" o "amministrazione", per indicare sia i soggetti che effettuano il trattamento ai sensi della lettera e) (quindi investiti di una pubblica funzione, sia che essi siano soggetti pubblici o formalmente privati), sia le autorità pubbliche (nozione soggettiva).

dell'esercizio di un pubblico potere. Di conseguenza si pone il problema di stabilire quando un'amministrazione possa dirsi effettivamente investita di tale pubblica funzione, che la legittima ad effettuare il trattamento. La risposta si trova sempre nell'art. 6, al terzo comma:

"La base su cui si fonda il trattamento dei dati di cui al paragrafo 1, lettere c) ed e), deve essere stabilita:

a) dal diritto dell'Unione; o

b) dal diritto dello Stato membro cui è soggetto il titolare del trattamento.

La finalità del trattamento è determinata in tale base giuridica o, per quanto riguarda il trattamento di cui al paragrafo 1, lettera e), è necessaria per l'esecuzione di un compito svolto nel pubblico interesse o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento [omissis]. Il diritto dell'Unione o degli Stati membri persegue un obiettivo di interesse pubblico ed è proporzionato all'obiettivo legittimo perseguito".

Dal testo della norma possiamo evincere due osservazioni rilevanti.

Una prima osservazione è che la base giuridica (vale a dire l'attribuzione in capo al soggetto della pubblica funzione) che legittima il trattamento ai fini della lettera e) del primo comma dell'articolo 6 deve essere stabilita dal diritto dell'Unione o dello Stato Membro e che tale base giuridica, alla luce dei Considerando 41 e 47 GDPR , deve consistere in un atto legislativo[98].

[98] Sebbene il Considerando 41 GDPR affermi che per "*base giuridica*" non debba intendersi necessariamente un atto legislativo parlamentare ("*Qualora il presente regolamento faccia riferimento a una base giuridica o a una misura legislativa, ciò non richiede necessariamente l'adozione di un atto legislativo da parte di un parlamento, fatte salve le prescrizioni dell'ordinamento costituzionale dello Stato membro interessato*" - e quindi si potrebbe argomentare che la base giuridica, da cui deriva il potere di trattare i dati personali non debba necessariamente consistere in un atto legislativo), il Considerando 47 chiarisce che la base giuridica, da cui deriva il potere per l'amministrazione di trattare i dati personali, deve consistere in una legge ("…*spetta al legislatore prevedere per legge la base giuridica che autorizza le autorità pubbliche a trattare i dati personali*…"). Ed è proprio in virtù di quest'ultima osservazione che, il Considerando 47, in coerenza con quanto dispone il primo comma dell'art. 6, afferma che "l*a base giuridica per un legittimo interesse del titolare del trattamento non dovrebbe valere per il trattamento effettuato dalle autorità pubbliche nell'esecuzione dei loro compiti*". Infatti tale base giuridica (lett. f) dell'art. 6) connota maggiormente il settore privato, essendo quello pubblico invece fisiologicamente caratterizzato dalla lett. e) e dall'indissolubile

Una seconda osservazione ha come punto focale invece la finalità del trattamento, che non deve essere imprescindibilmente stabilita e contenuta nella base giuridica, ma deve semplicemente essere strumentale e necessaria allo svolgimento di quella pubblica funzione menzionata in precedenza. Per questo la base giuridica stabilita dall'ordinamento positivo, che attribuisce la *mission de service public,* costituisce il faro da cui ricavare la finalità del trattamento. In altre parole il trattamento ai sensi della lettera e) deve essere "*necessario per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento"*, ma tale pubblica funzione è attribuita dalla base giuridica che investe l'amministrazione di una specifica *mission*, per questo motivo il trattamento deve essere necessario e strumentale a quella *mission* che il diritto dello Stato Membro o dell'Unione ha attribuito a quel soggetto. In tal modo si ottiene una coincidenza tra la finalità attribuita dal diritto positivo e la finalità del trattamento (almeno in termini generali)[99].

---

collegamento tra *mission* stabilita dalla norma attributiva del potere e finalità del trattamento. In modo coerente con questa linea il Considerando 45 chiarisce che non è necessario un atto legislativo per ogni trattamento, ma sarà comunque richiesto "*per più trattamenti effettuati conformemente a un obbligo legale cui è soggetto il titolare del trattamento o se il trattamento è necessario per l'esecuzione di un compito svolto nel pubblico interesse o per l'esercizio di pubblici poteri*". Circa la conclusione relativa alla necessità di una legge quale base giuridica del trattamento, afferma in senso conforme G. Carullo, *Big Data e Pubblica Amministrazione nell'era delle banche dati interconnesse*, in *Concorrenza e mercato*, 23, 2016, p.197.

[99] Così come evidenziato in M. Bombardelli, *Dati personali (Tutela dei)*, in B.G. Mattarella e M. Ramajoli, *Funzioni amministrative – Enciclopedia del diritto - I tematici*, III volume, Milano, Giuffrè, 2020, p.360. L'autore sottolinea che il conferimento sulla base del diritto dell'dell'Unione o dello Stato membro di un compito di interesse pubblico o dell'esercizio di pubblici poteri, di per sé indica pure la finalità generale del trattamento ed attribuisce anche la possibilità di effettuare i trattamenti di dati necessari per l'esecuzione di tale compito, senza la necessità che ogni trattamento sia preventivamente individuato come adempimento di un obbligo legale. Tali considerazioni sottolineano il fatto che con la sola base giuridica stabilita dal diritto positivo: 1. La PA ha la facoltà (e quindi non deve sussistere necessariamente un obbligo legale – lettera c) di eseguire il trattamento; 2. Le finalità generali del trattamento sono ricondotte alle *mission* della PA. La norma attributiva del potere è il faro su cui si basa la legittimità del trattamento della pubblica amministrazione, inteso non solo come acquisizione del dato, ma anche comunicazione ed elaborazione. In tal senso G. Carullo, *Big Data e Pubblica Amministrazione nell'era delle banche dati interconnesse*, cit., p.200 e p.202: "*Il processo di digitalizzazione delle pubbliche amministrazioni consente dunque di esplicitare una caratteristica fondamentale delle banche dati pubbliche: il collegamento tra i dati in possesso*

Rilevare che la finalità del trattamento vada estrapolata dalla *mission* che il diritto positivo (cioè la base giuridica, quindi la legge) ha attribuito all'amministrazione, ha diverse conseguenze, in quanto il principio della finalità del trattamento è un principio cardine del GDPR, che permea trasversalmente tutta la disciplina[100].

Opera infatti il principio di "limitazione delle finalità", enunciato al primo paragrafo, lettera b) dell'art. 5 del Regolamento, secondo cui i dati personali sono "*raccolti per finalità determinate, esplicite e legittime, e successivamente trattati in modo che non sia incompatibile con tali finalità*".

Tale principio ha due ordini di conseguenze. Il primo, derivante dal fatto che i dati sono raccolti per "*finalità determinate, esplicite e legittime*", comporta che ogni trattamento,

---

dell'amministrazione e l'interesse pubblico sotteso al potere in relazione al quale detti dati possono essere utilizzati", ed ancora "*considerato che solo a fronte di un espresso potere è possibile l'accesso — e solo nei limiti necessari per l'espletamento di una specifica funzione —, si può altresì concludere che i dati sono funzionali al raggiungimento dell'interesse pubblico tutelato dalla norma attributiva del potere, e che solo in tale misura essi possono essere gestiti ed elaborati dall'Autorità fruitrice*". E tale disciplina non potrebbe essere diversa, se letta alla luce del principio di legalità, da cui deriva un'importanza necessaria della norma attributiva del potere.

[100] Esso è definito come la pietra angolare (*cornerstone*) della disciplina sulla tutela dei dati personali. Tale definizione è stata data dal gruppo di lavoro ex articolo 29 (WP29 – istituito dalla Direttiva 95/46/CE, che si è occupato delle questioni relative alla tutela della *privacy* e dei dati personali fino al 25 maggio 2018, data di entrata in vigore del GDPR) nel parere n. 03/2013 sulla limitazione delle finalità, dove a p.4 si legge "*When we share personal data with others, we usually have an expectation about the purposes for which the data will be used. There is a value in honouring these expectations and preserving trust and legal certainty, which is why purpose limitation is such an important safeguard, a cornerstone of data protection. Indeed, the principle of purpose limitation inhibits 'mission creep', which could otherwise give rise to the usage of the available personal data beyond the purposes for which they were initially collected*". Inoltre, secondo tale parere il principio di limitazione della finalità impone che la finalità sia individuata in modo esplicito (espressa in forma intellegibile) e specifica (finalità generiche e vaghe – es generiche finalità di *marketing* – non risultano adeguate, in quanto devono delimitare con precisione l'ambito delle operazioni di trattamento). Orientamento confermato anche dall'*European Data Protection Board* (EDPB – organismo indipendente dell'UE che sostituisce il WP29) in "*principles of data protection*" - https://edpb.europa.eu/system/files/2023-06/Principles%20Infographic.pdf -. Riallacciandoci al discorso precedente si noti che una base giuridica che non disciplina il trattamento (ma che attribuisce solamente la *mission*), comunque non dovrebbe risultare idonea ad individuare un trattamento lecito, in quanto carente del requisito di specificità. Ecco perché tale base giuridica andrebbe integrata.

ogni dato, deve recare con sé la finalità determinata per cui è trattato. E ciò con riferimento ad ogni base giuridica del trattamento, sia nel caso del consenso, sia in tutti gli altri casi di trattamento necessario. La differenza tra queste ultime due ipotesi è che nel primo caso l'interessato ha prestato il consenso rispetto ad un trattamento con una finalità specifica ed individuata dal titolare (il consenso deve essere ottenuto in merito ad ogni trattamento, che reca con sé una diversa finalità), nel secondo caso invece le finalità sono predeterminate dal legislatore unionale (esecuzione di un contratto; adempimento di un obbligo legale da parte del titolare; salvaguardia di interessi vitali; esecuzione di un compito di interesse pubblico; legittimo interesse del titolare).

Il secondo ordine di conseguenze deriva dal fatto che i dati personali devono essere "*successivamente trattati in modo che non sia incompatibile con tali finalità* (originarie)". Questo vincolo concerne la circolazione dei dati all'interno della PA e l'interoperabilità delle banche dati, in quanto viene disposto che la finalità del trattamento successivo (comprendente anche la circolazione del dato da una pubblica amministrazione ad un'altra) deve essere compatibile con la finalità del trattamento iniziale (di raccolta del dato presso l'interessato). Il vincolo di compatibilità, secondo quanto dispone il quarto paragrafo dell'articolo 6[101], può essere superato solo nel caso della prestazione del consenso da parte dell'interessato o sulla base di "*un atto legislativo dell'Unione o degli Stati membri che costituisca una misura necessaria e proporzionata in una società democratica per la salvaguardia degli obiettivi di cui all'articolo 23, paragrafo 1"* [102].

[101] Il quarto paragrafo prevede dei parametri di riferimento per analizzare se la finalità successiva sia compatibile con quella originaria, nel caso in cui la finalità successiva non sia stata autorizzata dall'interessato o da un atto legislativo. Tra i parametri esemplificativi (esemplificativi, in virtù della locuzione "*tra l'altro*") rientrano ad esempio: il nesso tra le finalità per cui i dati personali sono stati raccolti e le finalità dell'ulteriore trattamento previsto; del contesto in cui i dati personali sono stati raccolti, in particolare relativamente alla relazione tra l'interessato e il titolare del trattamento; delle possibili conseguenze dell'ulteriore trattamento previsto per gli interessati; dell'esistenza di garanzie adeguate, che possono comprendere la cifratura o la pseudonimizzazione; della natura dei dati personali (per esempio dati particolari).

[102] L'articolo 23 GDPR (rubricato "*limitazioni*" ed inserito nel capo III relativo ai diritti dell'interessato) consente al diritto dell'Unione o dello Stato membro mediante "*misure legislative*" di limitare i diritti dell'interessato o di alleggerire gli obblighi del titolare del trattamento, qualora tale misura sia necessaria e proporzionata per salvaguardare determinate finalità indicate nell'articolo (tra cui: la sicurezza nazionale; importanti obiettivi di interesse

Oltre a queste due deroghe, ve ne sono altre, come ad esempio alla lettera b) dell'art. 5, il quale prevede che *"un ulteriore trattamento dei dati personali a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici non è, conformemente all'articolo 89, paragrafo 1*[103]*, considerato incompatibile con le finalità iniziali".*

pubblico generale dell'Unione o di uno Stato membro, in particolare un rilevante interesse economico o finanziario dell'Unione o di uno Stato membro, anche in materia monetaria, di bilancio e tributaria, di sanità pubblica e sicurezza sociale). Si ricordi, come precedentemente affermato, che Il Considerando 41 GDPR chiarisce che per "*misura legislativa*" non debba intendersi necessariamente "*atto legislativo*". Sebbene il quarto paragrafo dell'art. 6 GDPR richiami il primo paragrafo dell'art. 23, le due norme restano distinte e l'effetto del combinato disposto, come evidenziato dalla Corte di Giustizia nella sentenza C-268/21 del 2 marzo 2023 al paragrafo n. 33, è quello per cui il trattamento eseguito per una differente finalità deve essere autorizzato da una legge e deve essere volto al perseguimento degli interessi indicati al primo paragrafo dell'art. 23 (quindi l'art. 23 riguarda solamente la limitazione di determinati obblighi e diritti, non prevedendo la possibilità di derogare al vincolo di compatibilità, possibilità prevista dal quarto paragrafo dell'art. 6). Questo, del resto, è chiarito anche nel Considerando 50: "*Ove l'interessato abbia prestato il suo consenso o il trattamento si basi sul diritto dell'Unione o degli Stati membri che costituisce una misura necessaria e proporzionata in una società democratica per salvaguardare, in particolare, importanti obiettivi di interesse pubblico generale, il titolare del trattamento dovrebbe poter sottoporre i dati personali a ulteriore trattamento a prescindere dalla compatibilità delle finalità. In ogni caso, dovrebbe essere garantita l'applicazione dei principi stabiliti dal presente regolamento, in particolare l'obbligo di informare l'interessato di tali altre finalità e dei suoi diritti, compreso il diritto di opporsi*".

Vi è inoltre da considerare che le due deroghe (consenso e atto legislativo) sono previste in modo coerente con i Considerando 111 e 112 GDPR. Il primo afferma che "*È opportuno prevedere la possibilità di trasferire dati in alcune circostanze se l'interessato ha esplicitamente acconsentito, se il trasferimento è occasionale e necessario in relazione a un contratto o un'azione legale, che sia in sede giudiziale, amministrativa o stragiudiziale, compresi i procedimenti dinanzi alle autorità di regolamentazione. È altresì opportuno prevedere la possibilità di trasferire dati se sussistono motivi di rilevante interesse pubblico previsti dal diritto dell'Unione o degli Stati membri o se i dati sono trasferiti da un registro stabilito per legge e destinato a essere consultato dal pubblico o dalle persone aventi un legittimo interesse*". Il secondo precisa che tali deroghe dovrebbero valere in particolare "*per i trasferimenti di dati richiesti e necessari per importanti motivi di interesse pubblico*", prevedendo quindi garanzie non di poco conto per la circolazione dei dati per motivi di interesse genarle previsti dal diritto positivo.

[103] Tale articolo prevede più nel dettaglio quali sono le garanzie da attuare in tal caso e quali sono le deroghe previste nel caso di trattamento a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici.

Quindi la circolazione e la comunicazione dei dati all'interno del sistema amministrativo dipenderanno anche dalla finalità originaria del trattamento. Si noti che la comunicazione non è altro che una tipologia di trattamento, perciò potrà avvenire se necessaria per l'espletamento della pubblica funzione del richiedente, salvo l'ulteriore limite derivante dal vincolo di compatibilità. A tal proposito sussiste una differenza in base alla finalità originaria di raccolta del dato: un conto è infatti che tale funzione consista in ottenere determinati dati per renderli disponibili ad altre amministrazioni affinché svolgano le loro funzioni, altro conto è che un'amministrazione abbia raccolto i dati presso l'interessato per eseguire delle funzioni pubbliche specifiche e circostanziate[104].

La finalità è correlata anche ai principi di limitazione della conservazione (lettera e) dell'articolo 5 del GDPR) e di minimizzazione dei dati (lettera c) dell'articolo 5 del GDPR). Secondo il primo, il dato è conservato per un arco di tempo non superiore al conseguimento della finalità originaria, fatta eccezione per i trattamenti effettuati esclusivamente a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici (conformemente al primo paragrafo dell'articolo 89), nel qual caso sarà possibile prevedere periodi più lunghi. Secondo quanto dispone il secondo

[104] Si pensi, ad esempio, ad un'amministrazione che operi nel campo dell'assistenza sociale e che, quindi, possa acquisire determinati dati correlati a tale finalità del trattamento, quali il consumo di gas o di luce che legittimano il beneficiario ad ottenere un determinato regime. Ebbene in questo caso la finalità originaria appare circoscritta e potranno sussistere problematiche legate alla comunicazione di tali dati, in quanto la comunicazione deve rispettare la finalità originaria.

Si faccia l'esempio adesso, invece, di una base di dati di interesse nazionale, quale l'ANPR (Anagrafe nazionale della popolazione residente). La funzione originaria (e la finalità del trattamento di acquisizione del dato) di tale banca dati centralizzata – presso il ministero dell'interno – è quella di rendere disponibili i dati ai comuni per l'esercizio delle funzioni istituzionali del sindaco e di renderli disponibili anche delle altre amministrazioni. Una funzione che è determinata, ma molto più ampia rispetto a quella dell'esempio precedente e che ci consente di comprendere come il principio di limitazione della finalità (nella misura in cui impone la compatibilità con la finalità originaria), sia diversamente modulato a seconda dell'ampiezza di tale finalità originaria.

Questi esempi sono stati sviluppati a partire da B. Ponti, *attività amministrativa e trattamento dei dati personali*, cit., rispettivamente a p.21 (il primo) ed a p.60 (il secondo).

principio i dati devono essere "*adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati*".

Una volta illustrato che il trattamento deve essere necessario e finalizzato ad una pubblica funzione, occorre chiarire quanto debba essere intenso il rapporto di strumentalità tra il trattamento e la pubblica funzione. Affermare infatti che l'amministrazione possa eseguire un trattamento, per il semplice fatto che questo sia connesso in qualche modo ai suoi compiti di interesse pubblico o all'esercizio dei pubblici poteri di cui risulta investita, garantirebbe una discrezionalità sull'*an* e sul *quomodo* del trattamento troppo ampi

Per questo motivo la Corte di Giustizia ha chiarito che "*il requisito di necessità non è soddisfatto quando l'obiettivo di interesse generale considerato può ragionevolmente essere raggiunto in modo altrettanto efficace mediante altri mezzi meno pregiudizievoli per i diritti fondamentali degli interessati"* [105], in tal modo affermando anche in questo contesto il principio di proporzionalità (fatto proprio dalla Corte di Giustizia in diverse materie ed ormai principio generale dell'ordinamento Europeo), che si declina nel rispetto dei test di idoneità, necessarietà ed adeguatezza[106].

Inoltre, è importante sottolineare che alla luce della giurisprudenza europea i requisiti idonei ad integrare il presupposto della liceità (e quindi le caratteristiche della base giuridica) possono variare a seconda della tipologia del trattamento.

Infatti nelle ipotesi che non comportano particolari incisioni nel diritto alla tutela dei dati personali dell'interessato, la condizione di liceità è soddisfatta attraverso una base giuridica che si limiti ad attribuire alla pubblica amministrazione una *mission* nei termini precedentemente menzionati, quindi saranno leciti tutti i trattamenti necessari all'espletamento della pubblica funzione.

---

[105] Cfr. C-439/19, punto 110.

[106] M. Clarich, *Manuale di diritto amministrativo*, IV ed., Bologna, Il Mulino, 2019, p.160, dove viene chiarito che l'idoneità mette in relazione il mezzo adoperato con l'obiettivo da perseguire; la necessarietà configura la "*regola del mezzo più mite*", cioè il mezzo deve comportare il minor sacrificio possibile degli interessi incisi da tale mezzo; l'adeguatezza consiste invece nel verificare che la restrizione nella sfera giuridica dei destinatari sia tollerabile e che il mezzo più mite sia effettivamente in grado di soddisfare il fine.

Quando l'incisione nel diritto alla tutela dei dati personali diventa maggiore (in virtù della tipologia dei dati oggetto del trattamento; della quantità degli stessi; delle caratteristiche del trattamento), sarà necessaria una base giuridica che ricomprenda le finalità del trattamento ed il tipo dei dati trattati[107]. In questo caso quindi la finalità del trattamento non potrà essere desunta in modo implicito dalla *mission* attribuita dal diritto positivo alla pubblica amministrazione, ma dovrà essere menzionata nella base giuridica in modo esplicito, chiaro e specifico. Tuttavia, in questo caso la determinazione delle finalità delle modalità del trattamento non dovranno essere necessariamente previste in un atto legislativo, ma l'amministrazione potrà integrare la base giuridica mediante propri atti, l'importante è che tali atti indichino in via preliminare le finalità dei trattamenti e i dati che ne sono oggetto (e quindi l'esistenza di una base giuridica che predetermini le caratteristiche del trattamento rappresenta il *discrimen* con il caso precedente).

L'incisione maggiore del diritto alla tutela dei dati personali si configura invece quando la tipologia del trattamento consiste nella comunicazione al pubblico, pubblicazione o diffusione via internet dei dati, ossia tutti quei trattamenti che comportano un'accessibilità generalizzata dei dati ed un possibile riutilizzo di essi da parte di altri soggetti (comunemente invece la disciplina dell'accesso ai dati da parte di altri soggetti, specialmente privati, è condizionato dalla tutela alla riservatezza della persona a cui si riferiscono i dati). In tali tipologie di trattamento, la base giuridica non avrà solamente dei requisiti a livello di contenuto (dovrà individuare le finalità e prevedere esplicitamente la tipologia del trattamento), ma anche a livello di forma, in quanto dovrà consistere in un atto legislativo del Parlamento. In tal modo viene configurato un vincolo di stretta indispensabilità tra il trattamento dei dati e la finalità ad esso sottesa[108].

---

[107] Cfr. C- 175/20, par. 68-69.

[108] A tal proposito si può richiamare la sentenza della Corte di Giustizia del 9 novembre 2010 resa nei procedimenti riuniti C-92/09 e C-93/09, punto 66: "*In primo luogo, è pacifico che l'ingerenza derivante dalla pubblicazione su un sito Internet di dati nominativi relativi ai beneficiari interessati deve essere considerata «prevista dalla legge» ai sensi dell'art. 52, n. 1, della Carta (*dei diritti fondamentali dell'UE*)."* In altre parole, quando la gravità dell'incisione sul diritto alla tutela dei dati personali diventa più penetrante, la giurisprudenza chiama in causa criteri di bilanciamento ulteriori rispetto alla lettera e) dell'art. 6 del GDPR, ossia il criterio di bilanciamento di cui all'art. 52 CDFUE, che costituisce un presidio dalle ingerenze ai diritti

In conclusione si può affermare che per soddisfare il requisito della liceità ai sensi dell'art. 6 paragrafo 1 lettera e) la pubblica amministrazione dovrà essere investita di una funzione pubblica (esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri), attribuita da una base giuridica stabilita dal diritto positivo dello Stato membro o dell'Unione (atto legislativo) e che tale base giuridica avrà dei requisiti di contenuto in dipendenza della tipologia del trattamento in questione, andando in tal modo ad incidere sull'individuazione della finalità del trattamento e sulla discrezionalità amministrativa, sia per quanto riguarda l'*an* (in virtù di quali finalità è possibile svolgere il trattamento e, quindi, quando è possibile svolgerlo), sia per quanto riguarda il *quomodo* (quali dati trattare). In via generale la base giuridica non dovrà indicare le finalità del trattamento e le modalità (quali dati trattare), precisazioni che potranno essere fatte dalla stessa amministrazione con propri atti[109], salvi i casi di trattamenti particolarmente incisivi della sfera giuridica dell'interessato, quali la pubblicazione o la diffusione via internet.

Questi appena visti sono i requisiti di liceità per effettuare il trattamento da parte della pubblica amministrazione alla luce del GDPR; tuttavia il Regolamento riconosce la possibilità per gli Stati membri di specificare o adattare determinate regole alla luce degli

---

fondamentali. L'effetto di questa giurisprudenza è quello di richiedere un atto legislativo ogni qual volta il trattamento consista in un'accessibilità generalizzata del dato personale.

[109] Le caratteristiche del trattamento effettuato dalla PA ai sensi della lettera e) del Regolamento ha indotto in dottrina a parlare anche di poteri impliciti – B. Ponti, *attività amministrativa e trattamento dei dati personali*, cit., p.43- 47 –, in quanto è sufficiente che la norma attributiva del potere investa di una *mission* l'amministrazione, senza attribuire nello specifico il potere di trattamento del dato personale e senza specificare le relative modalità dell'esercizio del potere di trattamento, potere che viene inquadrato in un'ingerenza da parte della PA nella sfera giuridica dell'interessato e, quindi, in quanto tale, dovrebbe sottostare al principio di legalità ed ai suoi corollari, quali il principio di nominatività e di tipicità. Il fatto che lo specifico potere di trattamento viene desunto in via implicita dalla *mission* ha indotto a parlare di poteri impliciti. Poteri che generano attrito con il principio di legalità del diritto interno, ma che non lo generano con l'ordinamento comunitario, perché, prosegue l'autore – ult.cit., p.188 –, nello spazio giuridico europeo non ha esplicito riconoscimento tale principio. Ma, senza dilungarci troppo in queste (diverse) questioni e che meriterebbero tutt'altro approfondimento, accogliamo una delle teorie prospettate dall'autore, ossia che nel diritto interno (in seguito alle modifiche introdotte con il decreto capienze) anche in quest'ambito si assiste in parte al processo di crisi della legalità, già in atto in diversi ambiti.

obiettivi di interesse pubblico generale, di politica nazionale e delle peculiarità del paese di riferimento.

Il margine di manovra nell'ambito di un trattamento effettuato da una pubblica amministrazione è disciplinato all'art. 6, al secondo ed al terzo paragrafo[110].

Il secondo paragrafo prevede la facoltà per gli Stati membri di "*introdurre o mantenere*[111] *disposizioni più specifiche per adeguare l'applicazione delle norme del presente regolamento con riguardo al trattamento, in conformità del paragrafo 1, lettere c) ed e), determinando con maggiore precisione requisiti specifici per il trattamento e altre misure atte a garantire un trattamento lecito e corretto*", mentre il terzo paragrafo, dopo aver stabilito che la base giuridica delle lettere c) ed e) deve essere stabilita dal diritto

---

[110] Tali disposizioni esplicitano quanto affermato nel considerando 10 GDPR, il quale stabilisce che "… *Per quanto riguarda il trattamento dei dati personali per l'adempimento di un obbligo legale, per l'esecuzione di un compito di interesse pubblico o connesso all'esercizio di pubblici poteri di cui è investito il titolare del trattamento, gli Stati membri dovrebbero rimanere liberi di mantenere o introdurre norme nazionali al fine di specificare ulteriormente l'applicazione delle norme del presente regolamento. [omissis] Il presente regolamento prevede anche un margine di manovra degli Stati membri per precisarne le norme, anche con riguardo al trattamento di categorie particolari di dati personali («dati sensibili»). In tal senso, il presente regolamento non esclude che il diritto degli Stati membri stabilisca le condizioni per specifiche situazioni di trattamento, anche determinando con maggiore precisione le condizioni alle quali il trattamento di dati personali è lecito*".

[111] La norma utilizza il termine "*mantenere*", perché tiene conto che tale materia era stata già normata nell'ordinamento dagli Stati membri. Infatti in precedenza diversi interventi normativi avevano disciplinato tale materia. A tal proposito, il legislatore comunitario è inizialmente intervenuto con la Direttiva 95/46/CE del 24 ottobre 1995 relativa alla "*tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati*", che è stata recepita nell'ordinamento nazionale ad opera della legge 31 dicembre 1996, n. 675, le cui norme sono state poi trasposte nel d.lgs. 30 giugno 2003, n. 196 (Codice in materia di protezione dei dati personali). Successivamente, quando il legislatore comunitario è intervenuto con il Regolamento 2016/679, il legislatore nazionale ha recepito la normativa dapprima attraverso il d.lgs. 10 agosto 2018, n. 101 ed, in seconda battuta dalla legge n. 205/2021.
La cronologia degli interventi normativi fa emergere la circostanza per cui il GDPR è intervenuto in un settore dove gli Stati membri avevano già legiferato per implementare la Direttiva precedente. Tutti gli Stati membri disponevano quindi di una regolamentazione che, seppur armonizzata dalla Direttiva, rispecchiava le scelte politiche di ogni ordinamento. In tale contesto la scelta del Regolamento è stata quella di lasciare un margine di manovra per permettere agli Stati membri di mantenere in parte le scelte normative già effettuate. In tal senso B. Ponti, *attività amministrativa e trattamento dei dati personali*, cit., pp.48-90.

dell'Unione o dello Stato membro, dispone che "*Tale base giuridica potrebbe contenere disposizioni specifiche per adeguare l'applicazione delle norme del presente regolamento, tra cui: le condizioni generali relative alla liceità del trattamento da parte del titolare del trattamento; le tipologie di dati oggetto del trattamento; gli interessati; i soggetti cui possono essere comunicati i dati personali e le finalità per cui sono comunicati; le limitazioni della finalità, i periodi di conservazione e le operazioni e procedure di trattamento, comprese le misure atte a garantire un trattamento lecito e corretto*". Tali previsioni sono specifiche per le lettere c) ed e), quindi, rientrano i trattamenti necessari per adempiere un obbligo legale ed i trattamenti necessari all'esecuzione di una pubblica funzione. Il settore pubblico è infatti uno di quelli maggiormente interessati dal margine di manovra concesso dal Regolamento, sicché in tale ambito alcuni autori parlano di "*dual legality standard*"[112], per affermare che la competenza legislativa di tale materia risulta effettivamente distribuita tra legislatore europeo e nazionale, data l'ampia facoltà per gli Stati membri di integrare o di adeguare le disposizioni del GDPR[113]. Tuttavia, se per molti Stati questo margine di discrezionalità, ha rappresentato la possibilità di mantenere o introdurre una tutela rafforzata rispetto allo standard di *default* previsto dal Regolamento, non è neanche esclusa la possibilità

[112] B. Ponti, *attività amministrativa e trattamento dei dati personali*, cit., *passim*, dove vengono ricercate anche le ragioni dell'attribuzione di questo margine di manovra. Gran parte delle motivazioni (op. ult. cit., p.30 e 49) sono ricondotte alla c.d. "*deference*", una sorta di riconoscimento da parte dell'Unione europea di nei confronti della legislazione amministrativa nazionale. Infatti l'autore, pur riconoscendo che i sistemi amministrativi sono armonizzati in misura significativa, ribadisce che non vi è uniformazione, anche in coerenza del metodo comunitario o funzionalista, secondo il quale l'Unione tende a riconoscere la sovranità degli Stati membri e a riconoscere in ambito amministrativo un nucleo essenziale non intaccabile. In realtà, discostandosi dal pensiero dell'autore, una ragione potrebbe consistere nel fatto che entrambi i trattamenti (lett.c) ed e)) abbisognano di un intervento da parte del diritto positivo per essere autorizzati, quindi anche in virtù di questa maggiore garanzia preliminare, è stato previsto un margine di manovra maggiore in tali ambiti.

[113] Inoltre, il margine di manovra è previsto anche con riferimento ai dati particolari, infatti il quarto paragrafo dell'art. 9 dispone che "*Gli Stati membri possono mantenere o introdurre ulteriori condizioni, comprese limitazioni, con riguardo al trattamento di dati genetici, dati biometrici o dati relativi alla salute*".

di introdurre una tutela meno garantista[114] (per i diritti dell'interessato) per allineare la disciplina agli obiettivi di politica nazionale.

2.2.2. La normativa nazionale

Esaminiamo quindi come il legislatore italiano abbia sfruttato i margini di discrezionalità previsti da tali disposizioni, in modo da incidere sulla disciplina relativa all'acquisizione e alla circolazione dei dati personali da parte delle pubbliche amministrazioni.

Il primo intervento di adeguamento alle disposizioni del GDPR è rappresentato dalla decreto legislativo n. 101/2018, che aveva previsto numerose abrogazioni al codice *privacy* per consentire la diretta applicazione del Regolamento e che aveva strutturato un assetto incentrato sulla '*strict legality rule'*, irrigidendo i presupposti (rispetto al Regolamento) per effettuare un trattamento in esecuzione di una pubblica funzione. Il sistema prevedeva infatti che la base giuridica prevista alla lettera b) del paragrafo terzo dell'art. 6 (ossia la base giuridica stabilita dal diritto dello Stato membro che investe di una pubblica funzione – compito di interesse pubblico o esercizio di pubblici poteri – l'amministrazione) dovesse essere costituita esclusivamente da una norma di legge o, nei casi previsti dalla legge, di regolamento[115]. L'interpretazione del Garante della *privacy*, valorizzando l'avverbio '*esclusivamente'*, sosteneva che la base giuridica non potesse limitarsi ad investire il titolare del trattamento di un compito di interesse pubblico o dell'esercizio di poteri pubblici, ma doveva specificare espressamente il tipo di trattamento (di quali dati e per quali finalità) che poteva essere attuato[116]. Quindi, diversamente dal Regolamento, ci doveva essere una legge attributiva dello specifico

[114] E ciò, come viene sempre sottolineato in B.Ponti, *attività amministrativa e trattamento dei dati personali*, cit., p.76 specialmente in ragione del termine 'adeguare' contenuto nei summenzionati paragrafi dell'art. 6 ed in particolare per come è formulato terzo paragrafo.

[115] In tal senso prevedeva l'art. 2-*ter* d.lgs. n. 196/2003.

[116] In tal senso B. Ponti, *attività amministrativa*, p.95. Ed effettivamente si può prendere a titolo di esempio il provvedimento n. 289 del 1° settembre 2022, con il quale il Garante ha emesso un parere favorevole sullo schema di protocollo di intesa tra il Ministero del turismo, le Regioni e Province autonome avente finalità di realizzazione e gestione della Banca dati delle strutture ricettive e degli immobili destinati alle locazioni brevi.

potere di effettuare il trattamento, comprensiva anche delle finalità e delle modalità di questo, in linea quindi con il principio di legalità (a meno tale legge non avesse delegato un regolamento di specificare e circostanziare il trattamento).

Il decreto legge n. 139/2021 (decreto capienze), convertito nella legge n. 205/2021, ha apportato diverse modifiche, alleggerendo di fatto i presupposti per l'utilizzo dei dati personali da parte delle pubbliche amministrazioni. Queste modifiche sono intervenute per diverse esigenze e in particolari circostanze, tra le quali assumono rilievo la difficoltà che si sono riscontrate nel gestire l'emergenza pandemica causata dal Covid – 19[117], l'esigenza di assicurare una maggiore interoperabilità tra banche dati (molto incentivato dal PNRR)[118] ed infine la consapevolezza relativa alle opportunità offerte dall'analisi e dall'elaborazione dei dati.[119]

Per quanto riguarda i presupposti che integrano la liceità del trattamento assume particolare rilievo l'art. 2- *ter*, il quale al primo comma prevede che "*La base giuridica prevista dall'articolo 6, paragrafo 3, lettera b), del regolamento è costituita da una norma di legge o di regolamento o da atti amministrativi generali".* L'innovazione più rilevante[120] consiste nell'introduzione, quale base giuridica che legittima il trattamento, anche gli atti amministrativi generali.[121] In questo modo, una volta investita di una

[117] F. Vari, F. Piergentili, *"To no other end, but the... Safety, and publick good of the People": le limitazioni alla protezione dei dati personali per contenere la pandemia di Covid-19*, in *Rivista AIC*, 2021, pp. 328-342.

[118] G. Buttarelli, *L'interoperabilità dei dati nella Pubblica Amministrazione*, in V. Bontempi (a cura di), *Lo Stato digitale nel Piano nazionale di ripresa e resilienza*, Roma, Roma TrE-Press, 2022, pp. 141-147.

[119] M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data,* cit*.*, pp. 259-262.

[120] Oltre alla modifica dell'avverbio '*esclusivamente'*, che era stato utilizzato dal Garante della privacy per sostenere l'interpretazione secondo la quale la base giuridica dovesse includere anche la tipologia di dati da trattare e la finalità.

[121] La dottrina ha sottolineato il pericolo insito nella difficoltà di classificazione degli atti amministrativi generali rispetto ai regolamenti. I primi dovrebbero essere infatti atti formalmente amministrativi e non rappresentare una fonte di diritto. Tuttavia la giurisprudenza spesso ha dovuto indagare la natura sostanziale dell'atto, andando al di là del *nomen iuris* o della forma, in modo da farlo rientrare o meno nella categoria di regolamento e, dunque, sottoporlo a controllo giurisdizionale. In tal senso L. Califano, *La protezione dei dati personali: natura, garanzie e bilanciamento di un diritto fondamentale,* Torino, Giappichelli, 2023, pp.193-222.

pubblica funzione, l'amministrazione non sarà più vincolata da una fonte esterna in merito alla tipologia del trattamento ed alle finalità (specifiche) di questo, perché potrà emanare un atto amministrativo generale, che indichi quando l'ente possa trattare i dati personali (al di fuori, quindi, di una norma di legge).

Si permette quindi all'amministrazione di avere un margine di discrezionalità molto maggiore rispetto all'assetto previgente, in quanto potrà individuare con l'atto amministrativo generale quali trattamenti siano strumentali al compito di interesse pubblico che le sono stati affidati ed, una volta individuato il trattamento, specificarne i dati e le caratteristiche.[122] Il tutto risulta coerente con una delle colonne portanti del

[122] Per evidenziare la differenza rispetto al regime precedente si può fare l'esempio dell'attività di monitoraggio del Ministro della Salute per affrontare e combattere la crisi sanitaria causata dal Covid-19 utilizzando l'*app* di *contact-tracing* "Immuni". In tal caso, infatti, la base giuridica del trattamento e la finalità dovevano essere previste da un atto legislativo. A tal proposito il Garante della *Privacy* nel "*Provvedimento di autorizzazione al trattamento dei dati personali effettuato attraverso il Sistema di allerta Covid 19- App Immuni*" del primo giugno 2020, afferma (nella sezione denominata '*Base giuridica del trattamento, volontarietà e finalità perseguite*) che "*la realizzazione dell'app Immuni si colloca nel contesto normativo stabilito dall'art. 6 del d.l. n. 28/2020 con il quale è stata istituita la piattaforma unica nazionale per la gestione Sistema di allerta Covid-19 …[omissis] La predetta disposizione, fra l'altro, prevede alcuni requisiti fra loro strettamente connessi quali: 1.1) la volontarietà dell'istallazione dell'app; 1.2) il perseguimento di alcune specifiche finalità; 1.3) l'utilizzo di dati pseudonimizzati.*". L'art. 6 del d.l. n. 28/2020 dispone: "*Al solo fine di allertare le persone che siano entrate in contatto stretto con soggetti risultati positivi e tutelarne la salute attraverso le previste misure di prevenzione nell'ambito delle misure di sanità pubblica legate all'emergenza COVID-19, è istituita una piattaforma unica nazionale per la gestione del sistema di allerta [omissis]. Il Ministero della salute, in qualità di titolare del trattamento, si coordina [omissis] per gli ulteriori adempimenti necessari alla gestione del sistema di allerta e per l'adozione di correlate misure di sanità pubblica e di cura".* Ebbene in tale circostanza vi era l'esigenza di effettuare determinati trattamenti, ma il Ministero della salute non avrebbe potuto effettuare tali trattamenti, in quanto non previsti da un atto legislativo, cioè una base giuridica idonea a legittimare il trattamento ai sensi dell'art. 6 lett. e) del GDPR. Ed infatti l'art. 6 del decreto legge n. 28/2020 costituiva la base giuridica del trattamento, base giuridica che esplicitava anche la finalità del trattamento. Alla luce delle modifiche intervenute con il decreto capienze invece sarebbe bastato un atto amministrativo generale adottato dall'amministrazione competente per poter avere una base giuridica idonea, che prevedesse il potere di trattare i dati, disciplinasse le finalità specifiche (in quanto quelle generali sono sempre dedotte dalla norma attributiva del potere) e le modalità del trattamento. È da sottolineare che nel trattamento dei dati, effettuato su base volontaria attraverso il *software* di *contact-tracing* (Immuni), è emersa la "*funzione sociale*" del trattamento e la dimensione

GDPR, ossia il principio di responsabilizzazione ("*accountability*"), previsto al secondo paragrafo dell'art. 5, secondo cui il titolare del trattamento è competente per il rispetto dei principi fondamentali della materia (previsti dal primo paragrafo dell'art. 5) ed è in grado di comprovarlo[123].

Quindi, a differenza della disciplina nazionale precedente, sarà compito dell'amministrazione determinare, attraverso atti amministrativi generali in linea con i compiti e poteri ad essa affidati, il fine pubblico (specifico) che giustifica l'uso dei dati personali, senza la necessità di una legge di rango primario. E questo regime sarà applicabile non solo all'acquisizione del dato, ma anche alla comunicazione con altri soggetti pubblici.

Un'altra novità importante è prevista al comma 1-*bis* dell'art. 2-*ter*, il quale prevede che per determinate categorie di soggetti pubblici (individuati nella norma) il trattamento "*è anche*[124] *consentito se necessario per l'adempimento di un compito svolto nel pubblico interesse o per l'esercizio di pubblici poteri ad esse attribuiti"*, configurando di fatto uno schema non dissimile dalla *necessary clause* del Regolamento.

Tuttavia, questo speciale titolo di legittimazione non è consentito per tutte le amministrazioni, ma solamente per quanto riguarda i trattamenti realizzati dalle amministrazioni pubbliche e dalle società a controllo pubblico statale (o, nel caso di

---

collettiva di quest'ultimo, seppur con tutte le criticità dovute al pieno rispetto del GDPR. Per un approfondimento sulla relazione tra protezione dei dati e *app* immuni si veda V. Cuffaro, R. D'Orazio, *La protezione dei dati personali ai tempi dell'epidemia*, in *Il Corriere Giuridico*, 6/2020, pp.729-739; D. Poletti, *Il trattamento dei dati inerenti alla salute nell'epoca della pandemia: cronaca dell'emergenza*, in *Persona e mercato*, 2/2020, pp.65-76; D. Poletti, *Contact tracing a App Immuni: atto secondo*, in *Persona e mercato*, 1/2021, pp.92-101.

L'alleggerimento della regolamentazione sull'uso dei dati personali è destinato ad avere conseguenze anche nelle attività condotte dall'Agenzia delle Entrate e dalla Guardia di Finanza legate al contrasto dell'evasione fiscale. In questa prospettiva, infatti, si apre la strada all'incrocio dei dati, in assenza di una norma di legge che lo consenta, come si vedrà nel quarto capitolo.

[123] Per un approfondimento sul principio di *accountability* si rinvia a G. Finocchiaro, *Il principio di accountability*, in *Giurisprudenza italiana*, vol.171, 12/2019, pp. 2778-2783; L. Califano, V. Fiorillo, F. Galli, *La protezione dei dati personali: natura, garanzie e bilanciamento di un diritto fondamentale*, Torino, Giappichelli, 2023, pp. 137-193.

[124] 'Anche' nel senso che si aggiunge al titolo di legittimazione previsto dal primo comma dell'art. 2- *ter*.

gestori di servizio pubblico, anche locale), ad esclusione in questi ultimi casi delle attività svolte in regime di libero mercato. Quindi si compie una distinzione tra, da un lato, amministrazioni pubbliche e società a controllo pubblico statale (o, nel caso di gestori di servizio pubblico, anche locale) e, dall'altro lato, gli altri soggetti che svolgono compiti di interesse pubblico[125].

Quindi l'effetto ottenuto è quello di rendere meno gravoso il presupposto di liceità ed ampliare lo spazio di iniziativa del titolare del trattamento, attraverso la previsione degli atti amministrativi generali (che possono specificare la base giuridica) e attraverso la possibilità (per alcune amministrazioni) di effettuare un trattamento che risulti necessario al perseguimento di un compito di interesse pubblico.

Questo effetto si ripercuote sia sull'acquisizione dei dati presso l'interessato, sia sulla comunicazione[126] e sulla diffusione[127] di tali dati.

La comunicazione tra titolari che trattano dati per compiti di interesse pubblico o connessi all'esercizio di poteri pubblici seguirà il regime poc'anzi descritto per quanto

---

[125] La distinzione tra i soggetti pubblici viene chiarita in M. Iaselli, *La normativa di riferimento*, in G. Cassano, V. Colarocco, G.B. Gallus, F.P. Micozzi (a cura di), *Il processo di adeguamento al GDPR* (a cura di), II ed., Milano, Giuffrè, 2022, pp.1-35, effettuata a partire dal testo dell'art. 1- *bis*, che elenca i soggetti legittimati ad avvalersi di questo speciale titolo di legittimazione per effettuare il trattamento. Ai sensi di tale disposizione il trattamento potrà essere effettuato da "*un'amministrazione pubblica di cui all'articolo 1, comma 2, del decreto legislativo 30 marzo 2001, n. 165, ivi comprese le autorità indipendenti e le amministrazioni inserite nell'elenco di cui all'articolo 1, comma 3, della legge 31 dicembre 2009, n. 196, nonché da parte di una società a controllo pubblico statale o, limitatamente ai gestori di servizi pubblici, locale, di cui all'articolo 16 del testo unico in materia di società a partecipazione pubblica, di cui al decreto legislativo 19 agosto 2016, n. 175, con esclusione, per le società a controllo pubblico, dei trattamenti correlati ad attività svolte in regime di libero mercato*".

[126] Ai sensi del quarto comma della lettera a) dell'art. 2-*ter* del d.lgs. n. 196/2003 si intende per comunicazione: *"il dare conoscenza dei dati personali a uno o più soggetti determinati diversi dall'interessato, dal rappresentante del titolare nel territorio dell'Unione europea, dal responsabile o dal suo rappresentante nel territorio dell'Unione europea, dalle persone autorizzate, ai sensi dell'articolo 2-quaterdecies, al trattamento dei dati personali sotto l'autorità diretta del titolare o del responsabile, in qualunque forma, anche mediante la loro messa a disposizione, consultazione o mediante interconnessione*".

[127] Ai sensi del quarto comma della lettera b) dell'art. 2-*ter* del d.lgs. n. 196/2003 si intende per diffusione: "*il dare conoscenza dei dati personali a soggetti indeterminati, in qualunque forma, anche mediante la loro messa a disposizione o consultazione*".

riguarda i dati "comuni": potrà legittimamente avvenire, sia basandosi su una base giuridica definita dal comma 1 (e quindi anche quando le finalità del trattamento ed i dati trattati sono specificati in atti amministrativi generali), sia quando necessaria ai sensi del comma 1-*bis* (adempimento di un compito svolto nel pubblico interesse o per l'esercizio di pubblici poteri)[128]. In questo modo viene resa più flessibile la circolazione dell'informazione all'interno del patrimonio informativo pubblico[129].

Invece per quanto riguarda la diffusione e la comunicazione a soggetti che intendono trattare i dati per altre finalità, esse sono sempre ammesse se previste ai sensi del comma 1 o se necessarie ai sensi del comma 1-bis, ma in quest'ultimo caso "*ne viene data notizia al Garante almeno dieci giorni prima dell'inizio della comunicazione o diffusione*" ai sensi del terzo comma dell'art. 2-*ter*.

La disciplina relativa ai dati particolari è invece prevista agli articoli 2-*sexies*[130] e 2-*septies*. L'art. 2- *sexies* specifica i presupposti per il trattamento dei dati particolari (esclusi i dati genetici, biometrici e relativi alla salute, per i quali va rispettato anche un regime addizionale previsto all'art. 2-*septies*) effettuato per motivi di interesse pubblico rilevante ai sensi del paragrafo 2, lettera g) dell'art. 9 del GDPR. L'articolo del codice della Privacy prevede determinate basi giuridiche e predetermina i motivi di interesse pubblico rilevante che legittimano il trattamento. Tra i motivi che consentono tale

[128] E ciò secondo quanto previsto dal secondo comma dell'art. 2- *ter*.

[129] Nella sua versione precedente, infatti, l'articolo 2-*ter* del Codice della Privacy, richiedeva la previsione di una legge (o nei casi previsti dalla legge, di regolamento) per effettuare la comunicazione dei dati tra le amministrazioni. La previsione legislativa era necessaria nel caso in cui la finalità del richiedente non fosse compatibile con la finalità per cui i dati erano stati raccolti. Nel caso in cui, invece, le amministrazioni avessero ritenuto che le finalità fossero compatibili, allora in tal caso era previsto solamente un obbligo di segnalazione preventiva al Garante della *privacy*, la cui mancata risposta nell'arco di 45 giorni, era parificata a silenzio – assenso. Con il decreto capienze la comunicazione che non rispetti il vincolo di compatibilità delle finalità può essere prevista anche da un atto amministrativo generale o, in assenza, se necessaria all'espletamento della pubblica funzione, previa notizia al Garante dieci giorni prima di effettuare la comunicazione. Diversamente la comunicazione dei dati comuni che rispetta tale vincolo è ammessa anche se prevista da atto amministrativo generale o se necessaria per l'espletamento della pubblica funzione.

[130] In merito si veda F. Cortese, *Art. 2 sexies*, in R. D'Orazio, G. Finocchiaro, O. Pollicino, G. Resta (a cura di), *Codice della privacy e data protection*, Milano, 2021, pp. 1043 ss.

trattamento (elencati al secondo comma) vi rientrano ad esempio: tenuta degli atti e dei registri dello stato civile, delle anagrafi della popolazione residente in Italia e dei cittadini italiani residenti all'estero, e delle liste elettorali, nonché rilascio di documenti di riconoscimento o di viaggio o cambiamento delle generalità; attività dei soggetti pubblici dirette all'applicazione, anche tramite i loro concessionari, delle disposizioni in materia tributaria e doganale, comprese quelle di prevenzione e contrasto all'evasione fiscale; attività di controllo e ispettive; trattamenti effettuati a fini di archiviazione nel pubblico interesse o di ricerca storica.

Quindi le amministrazioni investite del perseguimento di questi interessi pubblici rilevanti potranno eseguire il trattamento se autorizzate da una specifica base giuridica. E l'innovazione più importante[131], anche qui, consiste proprio nell'aver previsto quale base giuridica, oltre la legge e il regolamento, gli atti amministrativi generali. Quindi l'amministrazione potrà tramite un proprio atto specificare quali dati trattare, le operazioni eseguibili e il motivo di interesse pubblico rilevante.

Rimane immutato invece l'art. 2-*septies*, che ammette il trattamento di dati genetici, biometrici e relativi alla salute per il perseguimento delle finalità di cui al paragrafo 2 dell'articolo 9 del GDPR ed in osservanza di apposite misure di garanzie disposte dal Garante della privacy.

Al termine dell'analisi effettuata nel primo capitolo, è possibile effettuare delle valutazioni in merito al difficile bilanciamento tra, da un lato le esigenze di circolazione ed acquisizione dei dati e, dall'altro, le normative che ostano a tali esigenze (nel corso dell'analisi sono state analizzate principalmente la Direttiva 96/9/CE ed il GDPR).

Per quanto riguarda la Direttiva 96/9/CE, nel corso della trattazione sono emerse in particolare le criticità correlate al *diritto sui generis*, il cui effetto risulta essere quello di

[131] È anche da segnalare il comma 1-*bis* di tale articolo, poiché è volto a delineare una regolamentazione tale da consentire l'interconnessione e la circolazione dei dati sanitari, esigenza maturata nel corso della pandemia, così come affermato nel 2022 dal Consiglio superiore di sanità nella *Proposta per lo schema di riforma dei sistemi informativi sanitari*.

"*freno totale*"[132] per il mercato dei dati e per la *data economy*. Per smorzare tale effetto si è agito prima attraverso la ODD ed in seguito, senza soluzione di continuità, con il DGA, impedendo l'applicazione del diritto *sui generis* per le banche dati di proprietà di enti pubblici al fine di favorire il riutilizzo delle informazioni; ciononostante, allo stato, manca un reale coordinamento con la Direttiva 96/9/CE, che rimane inalterata nel suo nocciolo duro, e l'effetto collaterale delle recenti modifiche è quello di omettere di "*intervenire sulle più pressanti problematiche ed introducendo addirittura ulteriori elementi di dubbio"*[133].

La disciplina del GDPR, invece, risulta avere molti risvolti critici quando vengono avvertite esigenze di circolazione dei dati, come quelle che si sono avute nel corso della pandemia[134] (e che si stanno avvertendo anche in virtù delle modifiche strutturali ed organizzative delle banche dati al fine di uniformare il patrimonio informativo)[135]. Il modello del GDPR risulta incentrato sulla "protezione" del dato a scapito dell'accesso, del riutilizzo e della circolazione. Si potrebbe parlare di un modello "*unipolare*"[136] che, tuttavia, in seguito al cambio di rotta che sta si attraversando per via della Strategia europea per i dati, sembra che stia cedendo il passo verso un modello "*multipolare*"[137], in cui non si può far a meno di far circolare il dato nel suo rispettivo mercato.

---

[132] C. Sganga, *Ventisei anni di Direttiva database alla prova della nuova strategia europea per i dati: evoluzioni giurisprudenziali e percorsi di riforma*, in *Il diritto dell'informazione e dell'informatica*, cit., p.702.

[133] C. Sganga, *Ventisei anni di Direttiva database alla prova della nuova strategia europea peri dati: evoluzioni giurisprudenziali e percorsi di riforma*, in *Il diritto dell'informazione e dell'informatica*, cit., p.703.

[134] Gli ostacoli dell'assetto normativo del GDPR vengono avvertiti specialmente nell'ambito della ricerca; ad esempio, a sottolineare le criticità del GDPR per lo sviluppo della ricerca biomedica si veda M. Shabani, *The Data Governance Act and the EU's move towards facilitating data sharing*, in *Molecular Systems Biology*, cit., pp.1-3. L'autrice, oltre a rappresentare le criticità di tale assetto normativo, vede con favore il percorso intrapreso con il DGA, sebbene riconosca l'importanza di tutelare i dati sensibili e che non sarà semplice coordinare tale normativa con il GDPR, afferma che molto dipenderà dagli sviluppi futuri della disciplina.

[135] Si veda in merito a quest'ultimo punto il secondo capitolo del presente lavoro.

[136] G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, cit., p.974.

[137] G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, cit., p.975.

Nell'attuale contesto economico e tecnologico, le cui potenzialità connesse ai *Big Data* e all'intelligenza artificiale sembrano essere molto promettenti, un approccio basato principalmente sulla protezione rischierebbe di non essere foriero delle innovazioni che è possibile sviluppare; in particolare, la Strategia europea per i dati prende in considerazione il nuovo approccio che si sta affermando nell'economia: il dato da qualcosa di "personale" viene visto come "materia prima-petrolio"[138]. Si sta forse assistendo al passaggio dalla "*protezione*" alla "*governance*" dei dati[139]. Tuttavia, sembrerebbe complicato soddisfare le nuove – così pressanti – esigenze relative al mercato dei dati, non incidendo in modo deciso su alcune caratteristiche del GDPR, la cui disciplina è caratterizzata da una tensione intrinseca con tali esigenze[140]. Anche a livello nazionale, attraverso il c.d. decreto capienze, ci si è spostati più verso un modello orientato a garantire una maggiore acquisizione e circolazione dei dati personali. In ogni caso, sebbene l'assetto complessivo rimanga ancora incentrato sulla "protezione" del dato, gli interventi a livello europeo (e nazionale) rappresentano sicuramente uno sforzo apprezzabile e smorzano il modello "unipolare", intervenendo in una materia dove agli estremi opposti si ritrovano dei diritti e delle esigenze, il cui bilanciamento risulta di non facile soluzione. A tutto ciò si deve aggiungere che alcuni principi cardine del GDPR, come il principio di finalità, subiscono delle gravi minacce e contraddizioni con le attuali tecniche di conservazione dei dati e di analisi algoritmica di questi ultimi[141].

Vi è da considerare inoltre che gli sforzi compiuti per conseguire una più agevole acquisizione e circolazione dei dati all'interno del settore pubblico (quindi ponendo a parte gli assetti relativi al riutilizzo delle PSI e degli *open data*) vedrebbero i loro benefici

[138] Sull'importanza della circolazione dei dati, specialmente per le attuali tecnologie, si rimanda al primo paragrafo del terzo capitolo del presente lavoro.

[139] In tal senso si veda F. Bravo, *Il principio di solidarietà tra data protection e data governance*, in *Il diritto dell'informazione e dell'informatica*, cit., p.518. L'autore afferma che questa potrebbe essere la seconda transizione che interessa questa materia, essendo la prima rappresentata dal passaggio dal concetto di *privacy* a quello di protezione dei dati.

[140] G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, cit., p.990.

[141] Per le tensioni con i modelli di analisi algoritmica e di conservazione dei dati si veda, rispettivamente, il paragrafo due e tre del terzo capitolo del presente lavoro.

frustrati (nell'ottica di una pubblica amministrazione fruitrice dei dati, per rafforzarne il potere conoscitivo) se non fossero accompagnati da apposite modifiche a livello organizzativo e strutturale, in grado di trarre i massimi benefici da tale *framework* normativo.

Nel capitolo seguente, si cercheranno quindi di delineare alcuni profili relativi all'organizzazione delle banche dati pubbliche.

CAPITOLO II

BANCHE DATI PUBBLICHE: PROFILI ORGANIZZATIVI E FUNZIONE AMMINISTRATIVA

1. Profili organizzativi e funzione amministrativa

In questo paragrafo verranno affrontate le principali modifiche intervenute a livello organizzativo che coinvolgono le banche dati. Modifiche che, in base alla loro conformazione, configurano una determinata funzione amministrativa. Ed infatti l'organizzazione non rappresenta un fine in sé, ma è un mezzo, ed ha una funzione servente rispetto alle esigenze cui vuole rispondere[142]. Per questo motivo si rende necessario prendere in considerazione le principali soluzioni organizzative adoperate dalle pubbliche amministrazioni e, attraverso un ragionamento induttivo, ricavare la funzione amministrativa che il legislatore italiano ed Europeo ha delineato[143].

Per individuare quali siano le principali soluzioni tecnologiche adoperate, occorre analizzare il percorso politico e legislativo degli ultimi anni relativo a tale ambito, sottolineando in particolare gli investimenti attuati, che sono un'utile indicazione per individuare la direzione intrapresa a livello organizzativo e funzionale.

Gli interventi partono, a monte, dall'ampia strategia europea in materia di dati menzionata in precedenza, che, tra le altre cose, mira alla realizzazione del mercato

[142] Tale impostazione, che dovrebbe essere sicuramente seguita dai soggetti pubblici la cui azione è sempre improntata ad un fine, appare prettamente logica ed è evidenziata in dottrina, ad esempio in G. Carullo, *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa*, cit., p.38., dove si legge che "*è necessario che la funzione organizzatrice sia effettivamente esercitata dominando la tecnologia di cui l'amministrazione si avvale, così che sia la prima (la funzione) a determinare il modo di essere della seconda (la tecnologia). In altri termini è necessario che sia l'amministrazione a definire quali strumenti tecnologici adottare. Non il contrario*".

[143] Una volta delineata l'organizzazione per come risulta alla luce della legislazione vigente, nel terzo capitolo, si potrà analizzare se tale organizzazione sia adeguata anche per rispondere allo sfruttamento delle potenzialità dei *Big Data* e alla funzione di analisi algoritmica dei dati.

unico digitale. Con la recente crisi pandemica si è avuto uno dei maggiori impulsi ai cambiamenti organizzativi, che erano già in corso d'opera, per via di tale strategia. La pandemia dovuta al virus SARS-CoV-2 (dal quale deriva la malattia di COVID-19)[144] ha mostrato l'esigenza della digitalizzazione ed ha portato, al tempo stesso, ad un incremento di velocità di questa.

Inoltre, l'emergenza sanitaria ha fatto acquisire la consapevolezza della centralità della trasformazione digitale, anche in ambito politico[145], ed ha fatto emergere in modo drammatico e prepotente la necessità di dotarsi di adeguate tecnologie, soprattutto per garantire la comunicazione tra banche dati (interoperabilità) e l'acquisizione a grandi velocità, in tempo reale di dati e lo scambio di questi tra le banche dati pubbliche nazionali ed europee[146].

Alla crisi pandemica ha fatto seguito un programma a livello europeo – il *Next Generation EU* (NG-EU) – che, da una parte, ha fornito agli Stati membri gli strumenti essenziali per affrontare la crisi emergenziale e, dall'altra, ha sfruttato tale situazione per implementare ulteriori obiettivi e per rinforzare le politiche europee[147].

Tale programma ha avuto origine con la comunicazione della Commissione del 2020, denominata "*Il momento dell'Europa: riparare i danni e preparare il futuro per la prossima generazione*"[148] e, dopo che si è raggiunto un accordo politico con gli Stati

---

[144] Diversi studi sottolineano la correlazione tra pandemia e cambiamento tecnologico. Cambiamento che si è avuto in moltissimi settori, tra cui quello sanitario, che risulta il più inciso. Sottolineano questi cambiamenti D.Vargo, L.Zhu, B.Benwell, *Digital technology use during COVID-19 pandemic: A rapid review*, in *Human Behavior and Emerging Technologies*, 2020, vol. 3, n. 1, pp. 13– 24.

[145] Ne sottolinea l'importanza, anche in ambito politico A. Baldassarre, *Dalla gestione dell'emergenza a una visione condivisa di futuro: il ruolo del digitale per lo sviluppo sostenibile del paese*, in *FPA - Annual Report 2020*, Edizioni FORUM PA, 2021.

[146] M. Bassil, *Interoperabilità e data governance nella nuova normalità, cosa resterà di quest'anno*, in *FPA - Annual Report 2020*, Edizioni FORUM PA, 2021.

[147] Per approfondimenti sul programma NG-UE si veda C. Pagliarin, C. Perathoner, S. Laimer, *Il Next Generation EU e i piani nazionali di ripresa e resilienza*, Milano, Giuffrè, 2023; C. Alcidi, D. Gros, *Next Generation EU: A Large Common Response to the COVID-19 Crisis*, Vol.55, n. 4, 2020, pp. 202-203; M. Mazzarella, C. Ramotti, *Pandemia e governo digitale*, in *Giornale di diritto amministrativo*, n. 3/2022, pp. 415-423.

[148] COM(2020) 456 final, 27 maggio 2020.

membri in merito al contenuto del programma, ha trovato parte della sua realizzazione nel Regolamento del 12 Febbraio 2021, n. 241, che ha istituito "*il dispositivo per la ripresa e resilienza"* e che "*stabilisce gli obiettivi del dispositivo, il suo finanziamento, le forme di finanziamento dell'Unione erogabili nel suo ambito e le regole di erogazione di tale finanziamento*" (art. 1 del Regolamento).

Il "dispositivo per la ripresa e la resilienza" o il *Recovery and Resilience Facility* (RRF) rappresenta lo strumento principale (ed il maggiormente finanziato) delle misure previste all'interno del fondo NGUE[149], ed infatti lo stanziamento delle risorse, ai sensi del primo comma dell'art. 6 del Regolamento è pari a 672.5 miliardi di euro, ottenuti mediante prestiti contratti dalla Commissione sul mercato dei capitali, suddivisi in 390 miliardi sotto forma di sovvenzioni e 360 miliardi in prestiti.

Per accedere alle risorse del dispositivo ogni Stato membro, è tenuto a sottoporre alla Commissione Europea un "Piano nazionale"[150], ossia un documento strategico che

[149] Tale fondo è previsto dal programma *NextGenerationEU* e prevede investimenti per un imposto pari a 806,9 miliardi di euro (importo espresso a prezzi correnti -2023. Equivale a 750 miliardi di euro a prezzi del 2018). NGEU costituisce il più ingente pacchetto di misure di stimolo mai finanziato in Europa con il fine di un'Europa più ecologica, digitale e resiliente. I fondi previsti dal fondo NGUE, unitamente a quelli del bilancio del quadro finanziario pluriennale (QFP) 2021-2027 stanziano un totale di 2.018 miliardi di euro (espresso a prezzi correnti -2023, 1.800 miliardi nel 2018). Tali fondi sono utilizzati non solo per garantire la ripresa dell'Europa all'indomani della pandemia, ma anche per affrontare le numerose difficoltà in corso, come quella scaturita in seguito dell'aggressione della Russia nei confronti dell'Ucraina. Invece *NextGenerationEU* è un programma temporaneo, previsto appositamente per riparare i danni causati dalla pandemia e per implementare nuovi obiettivi ed il "dispositivo" costituisce il fulcro di tale programma. Cifre e dati sono stati reperiti dal sito ufficiale della Commissione europea https://commission.europa.eu/strategy-and-policy/recovery-plan-europe_it.

[150] Secondo le linee guida (rese pubbliche il 22 gennaio 2021 dalla Commissione) per la redazione dei piani nazionali di ripresa e resilienza, gli Stati membri devono specificare come contribuiranno alla transizione digitale ed ogni Piano deve prevedere un minimo del 20% dei fondi a disposizione sia destinato alla trasformazione digitale. Inoltre, gli Stati membri sono invitati a spiegare come il piano contribuirà a migliorare le loro prestazioni digitali, misurate in base alle all'Indice dell'economia e della società digitali (DESI) e agli obiettivi delineati nella comunicazione "*Shaping Europe's digital future*" (COM/2020/67 final). Il *Digital Economy and Society Index* (DESI) è un indicatore lanciato dalla Commissione Europea nel 2014 per valutare i progressi dei paesi europei verso la digitalizzazione economica e sociale, promuovendo l'obiettivo di un mercato unico digitale. Questo indice aggrega diversi indicatori in quattro aree chiave: capitale umano, che valuta le competenze richieste per sfruttare le opportunità della

delinea gli investimenti, le riforme e gli obiettivi che si intendono conseguire attraverso i fondi europei richiesti. Inoltre, gli importi del relativo fondo sono versati solo al momento del conseguimento degli obiettivi concordati nel piano nazionale[151].

Per beneficiare del sostegno previsto dal dispositivo (RRF) ogni Stato europeo ha adottato il proprio Piano di ripresa e resilienza (attraverso i quali vengono presentate le riforme e gli investimenti che intendono attuare entro la fine del 2026), compresa l'Italia, il cui PNRR è stato valutato come idoneo prima dalla Commissione e poi dal Consiglio europeo, che lo ha approvato con decisione di esecuzione il 13 luglio 2021, così come disposto dalla procedura prevista nel Regolamento n. 241/2021[152].

Il Piano nazionale è suddiviso in sei missioni[153] e la prima missione, quella rilevante ai nostri fini, è denominata "*digitalizzazione, innovazione, competitività, cultura e turismo*",

---

società digitale; connettività, che considera l'espansione e la qualità della banda larga e l'accesso da parte dei vari *stakeholder*; integrazione delle tecnologie digitali, che esamina la digitalizzazione delle aziende e l'utilizzo di canali online per le vendite; e servizi pubblici digitali, che valuta la digitalizzazione della pubblica amministrazione, concentrandosi sull'*e-Government*. Per l'edizione 2022 del DESI, il cui monitoraggio spetta alla Commissione, l'Italia si colloca al diciottesimo posto fra i ventisette Stati membri dell'UE (nel 2017 si collocava, invece, al venticinquesimo posto); per i dati relativi al 2022 si veda la pubblicazione della Commissione relativa all'Italia reperibile sul portale dell'istituzione europea https://digital-strategy.ec.europa.eu/it/policies/desi-italy; per approfondimenti sui miglioramenti ottenuti dall'Italia nella classifica DESI si veda P. Piras, *Il tortuoso cammino verso un'amministrazione nativa digitale*, in *Il diritto dell'informazione e dell'informatica*, 1/2020, pp.43-65 e per ragionamenti critici su tal indice A. J. Tarjáni, N. Kalló, I. Dobos, *Evaluation of Digital Development Based on the International Digital Economy and Society Index 2020 Data*, in *Statistika: statistics and economic journal*, vol.3, 2023, pp. 355–373.

[151] Ciò in base a quanto emerge dal sito ufficiale della Commissione. https://commission.europa.eu/business-economy-euro/economic-recovery/recovery-and-resilience-facility_it. A tal proposito attraverso la modifica dell'art. 18- *bis* del CAD ad opera del d.l. 31 maggio 2021, n. 77, (decreto semplificazioni *bis*), sono stati attribuiti all'AGID dei poteri di accertamento e sanzionatori in merito ad eventuali violazioni degli obblighi di transizione digitale previsti dal Piano di ripresa.

[152] Capo III del Regolamento (artt. 17 ss).

[153]Piano reperibile in molteplici fonti, tra cui la pagina ufficiale della Camera dei deputai https://www.camera.it/temiap/2021/06/25/OCD177-4986.pdf. Le altre missioni del PNRR, oltra alla prima, sono: rivoluzione verde e transizione ecologica; infrastrutture per una mobilità sostenibile; istruzione e ricerca; inclusione e coesione; salute. Attraverso il Piano il Governo ha richiesto 191, 5 miliardi di euro, ossia la totalità delle risorse previste dal dispositivo destinate

che a sua volta è suddivisa in tre componenti, la prima "*digitalizzazione, innovazione e sicurezza nella PA*" (per la quale è stato previsto uno stanziamento pari a 9,75 miliardi di euro), la seconda "*digitalizzazione, innovazione e competitività nel sistema produttivo*" (stanziamento pari a 23,89 miliardi di euro) e la terza "*turismo e cultura 4.0*" (stanziamento pari a 6, 68 miliardi di euro). Per conseguire al meglio i risultati relativi alla prima componente è stata concepita la strategia Italia Digitale 2026, che fa perno in particolare sul rafforzamento delle infrastrutture digitali e della connettività a banda ultra-larga sul territorio nazionale e sulla digitalizzazione della PA[154].

La digitalizzazione, così come delineata nel NG-EU e nel PNRR, non è un obiettivo univoco, ma rappresenta una necessità trasversale, attraverso la quale è possibile efficientare ogni ambito, che sia politico, economico, amministrativo o sociale[155].

Una volta descritti i propositi fissati a livello nazionale , vale la pena soffermarsi sui principali interventi ed investimenti previsti dal programma Italia Digitale 2026 e dalla prima componente della prima missione del PNRR, per analizzare in quali tecnologie sono stati attuati i maggiori investimenti.

Il programma previsto in Italia Digitale 2026 prevede sette investimenti principali[156]:

---

all'Italia. Per un approfondimento sul PNRR italiano, sui rapporti di esso con le istituzioni e l'ordinamento, si veda D. De Lungo, F. S. Marini, *Scritti costituzionali sul Piano Nazionale di Ripresa e Resilienza*, Torino, Giappichelli, 2023.

[154] Tutti gli obiettivi e le strategie sono elencati al sito https://padigitale2026.gov.it/. I due obiettivi (reti ultraveloci e digitalizzazione della PA) sono consequenziali: lo sviluppo del primo è funzionale al raggiungimento del secondo. In tal senso A. F. Spagnuolo, E. Sorrentino, *Alcune riflessioni in materia di trasformazione digitale come misura di semplificazione*, in *federalismi.it*, n. 8/2021, pp. 275-287. Il portale PAdigitale2026 è gestito dal Dipartimento per la trasformazione digitale (DTD) – dipartimento istituito presso la Presidenza del Consiglio dei Ministri con Dpcm del 19 giugno 2019 - ed ha l'obiettivo di guidare la transizione digitale della pubblica amministrazione, infatti questo rappresenta l'unico punto di accesso per le Amministrazioni ai fondi assegnati, per il controllo del raggiungimento degli obiettivi e per ottenere assistenza. In merito a quest'ultimo punto I. Macrì, *Il PNRR italiano per la digitalizzazione e l'innovazione della Pubblica Amministrazione*, in *Azienditalia*, 2022, n. 1, pp. 38 ss.

[155] Così come affermato in P. Clarizia, G. Sgueo, *Lo stato digitale nel PNRR: la digitalizzazione come necessità trasversale*, Irpa, Osservatorio sullo Stato digitale, 2021.

[156] Pubblicati sul sito del DTD https://innovazione.gov.it/italia-digitale-2026/il-piano/digitalizzazione-della-pa/.

1) "*Infrastrutture digitali*" (900 milioni di euro), per attuare un processo di razionalizzazione dei molti *data center* distribuiti sul territorio nazionale, a partire da quelli meno efficienti e sicuri, secondo un approccio "*cloud first*";
2) "*Abilitazione e facilitazione migrazione al cloud*" (1 miliardo di euro), un programma di incentivazione e per facilitare il trasferimento di *database* e applicazioni, supportando la migrazione dei dati delle pubbliche amministrazioni, sia centrali che locali, verso il *cloud*;
3) "*Dati ed interoperabilità*" (650 milioni di euro), tale investimento prevede due iniziative mirate a rinnovare la struttura e il metodo di collegamento tra i *database* delle amministrazioni, allo scopo di creare un sistema di banche dati pubbliche interconnesse. Questo favorirà un risparmio economico per le amministrazioni e un risparmio di tempo per i cittadini. La prima iniziativa riguarda lo sviluppo della Piattaforma Digitale Nazionale Dati (PDND) per assicurare l'interoperabilità dei dati pubblici, semplificando per gli enti pubblici la fornitura di servizi (in maniera più sicura, rapida ed efficace) e permettendo ai cittadini di evitare la riconsegna di informazioni già in possesso della PA. La piattaforma renderà infatti le informazioni sui cittadini disponibili ed accessibili "*una volta per tutte*". La seconda iniziativa prevede invece l'attuazione dello Sportello Digitale Unico (*Single Digital Gateway*)[157];
4) "*Servizi digitali e cittadinanza digitale*" (2,01 miliardi di euro), un investimento mirato a semplificare la vita digitale dei cittadini, attraverso servizi pubblici più efficienti.[158];

[157] Il "*Single Digital Gateway*" (sportello unico digitale) è previsto dal Regolamento (UE) 2018/1724 ed è una piattaforma online progettata per semplificare l'accesso ad informazioni, procedure amministrative ed a servizi di assistenza di uno Stato membro necessari sia ai cittadini che alle imprese di un altro Stato membro dell'Unione.

[158] Stando a quanto riportato sul sito del DTD "*La trasformazione dell'architettura digitale della PA, dal cloud all'interoperabilità dei dati, è accompagnata da investimenti mirati a semplificare la vita digitale dei cittadini, attraverso migliori servizi pubblici*", è chiara quindi la correlazione tra i cambiamenti organizzativi e le migliorie relative ai servizi pubblici, che, sebbene sono investimenti distinti, pare evidente che molte modifiche organizzative, come il *cloud* e l'interoperabilità, sono volte ad ottenere dei servizi pubblici migliori, un fine importantissimo per le pubbliche amministrazioni, come è dimostrato anche dall'ingente investimento.

5) "*Cybersecurity*" (620 milioni di euro), tale investimento è necessario e direttamente proporzionale all'aumento della digitalizzazione, è infatti evidente che un incremento di quest'ultima esponga la società a un maggior rischio di attacchi informatici, come frodi, ricatti informatici o attacchi terroristici. Bisogna quindi introdurre misure per potenziare le difese *cyber* del paese, tra cui l'implementazione completa delle normative sul "*Perimetro di Sicurezza Nazionale Cibernetica*";
6) "*Digitalizzazione delle grandi amministrazioni centrali*" (610 milioni di euro), ad esempio alcuni ministeri, l'INPS, la Guardia di Finanza, il Consiglio di Stato e così via;
7) "*Competenze digitali di base*" (200 milioni di euro), un investimento volto a potenziare l'alfabetizzazione digitale dei cittadini.

Per quanto riguarda invece la prima componente della prima missione del PNRR, essa ha l'obiettivo di "*rendere la Pubblica Amministrazione la migliore "alleata" di cittadini e imprese, con un'offerta di servizi sempre più efficienti e facilmente accessibili*"[159] e per conseguire tale obiettivo si interviene a livello organizzativo, puntando sul *cloud*, sull'interoperabilità e sulla *cybersecurity* e si interviene anche a livello di erogazione dei servizi (sebbene, abbiam detto, strettamente correlati alle modifiche organizzative) e di alfabetizzazione digitale.

Alla luce degli ultimi interventi normativi è evidente che tra i maggiori interventi sul piano dell'organizzazione e della disposizione del patrimonio informativo (quindi sull'organizzazione delle banche dati) vi sono il *cloud computing*, l'interoperabilità e la *cybersecurity*.

Prima di esaminare tali concetti pare opportuno fare un'ulteriore precisazione preliminare in merito alla competenza relativa ai cambiamenti organizzativi a livello nazionale. Per quanto riguarda tale ambito, la lettera r) del secondo comma dell'art. 117 della Costituzione afferma che lo Stato ha la competenza legislativa esclusiva del «*coordinamento informativo statistico e informatico dei dati dell'amministrazione*

[159] Già dall'obiettivo (p.84 del PNRR) si scorge il modello di amministrazione incentrato sull'intermediazione e sulla fornitura dei dati, più che su una loro fruizione ed elaborazione.

statale, regionale e locale». Tale materia, peraltro, in base agli orientamenti della Corte costituzionale[160], non riguarda solamente la definizione delle regole tecniche per la trasmissione dei dati (e quindi l'implementazione dei meccanismi di interoperabilità)[161], fondamentali per assicurare degli *standard* comuni, ma concerne anche ulteriori esigenze, quali la spesa pubblica ed il coordinamento della finanza pubblica relativi a tale ambito[162].

Una volta precisato che è essenzialmente il legislatore nazionale a stabilire regole in tale ambito, possiamo analizzare le soluzioni tecnologiche *supra* esposte: l'interoperabilità ed il *cloud computing*.

2. Interoperabilità

L'interoperabilità, *latu sensu* considerata, può essere definita come la capacità (in assenza di barriere tecniche, organizzative, semantiche e giuridiche) di comunicazione e trasferimento effettivo ed automatizzato[163] di dati tra diversi componenti *hardware* o *software* autonomi.

---

[160] Tra cui si segnalano le sentenze Corte Cost.,10 gennaio 2004 n. 17 e Corte Cost., 5 maggio 2008 n. 133. In quest'ultima sentenza si legge ad esempio "*la citata disposizione deve essere intesa nel senso che lo Stato disciplina il coordinamento informatico, oltre che per mezzo di regole tecniche, anche quando sussistano esigenze di omogeneità ovvero anche «profili di qualità dei servizi» e di «razionalizzazione della stessa», funzionali a realizzare l'intercomunicabilità tra i sistemi informatici delle amministrazioni*".

[161] In tal senso F. Cardarelli, *3-bis. Uso della telematica*, in M. A. Sandulli (a cura di), *Codice dell'azione amministrativa*, Milano, Giuffrè, 2010, pp. 427-428.

[162] B. Ponti, *Coordinamento e governo dei dati nel pluralismo amministrativo,* M. Pietrangelo (a cura di), *Scritti in memoria di Isabella D'Elia Ciampi*, in Informatica e diritto, vol. XVII, 2008, n. 1-2, p. 430.

[163] Scambiare i dati in modo automatizzato significa anche elidere l'interazione tra i funzionari pubblici, nel senso che all'ufficio richiedente basterà effettuare la richiesta di dati all'amministrazione erogatrice che, nel caso in cui sussista la legittimazione, trasferirà i dati all'amministrazione richiedente. Ciò è evidenziato in G. Carullo, *Big Data e Pubblica Amministrazione nell'era delle banche dati interconnesse*, cit., p.188, ove l'autore afferma che "*un ente potrebbe avere accesso alle informazioni detenute da un'altra amministrazione senza la necessità — quantomeno tecnica — di alcuna interazione tra i funzionari. All'ufficio procedente basterebbe richiedere, attraverso il proprio sistema informatico, i dati di cui necessita, ed*

Per chiarire perché l'interoperabilità è stata definita in tal senso occorre sviluppare il ragionamento lungo due direttici: la prima, che chiarisce le questioni tecniche della definizione e, la seconda, che chiarisce il perché per aversi l'interoperabilità non ci si può soffermare solo sulle questioni tecniche.

Per avere delucidazioni dal punto di vista tecnico si può far riferimento alla definizione di tale nozione contenuta nel vocabolario ISO: "*capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units*"[164]. Il soggetto della nozione è rappresentato dall' "unità funzionale", che il vocabolario ISO definisce come "*entity of hardware or software, or both, capable of accomplishing a specified purpose*"[165]. Inoltre si può anche far riferimento alla definizione data dall'*Institute of Electrical and Electronics Engineers (*IEEE*)* "*the ability of two or more systems or components to exchange information and to use the information exchanged*"[166]. Da queste definizioni si può ricavare di come il soggetto dell'interoperabilità tecnica sia una componente *hardware* o *software* autonoma (*capable of accomplishing a specified purpose*), che sia in grado non solo di "osservare" i dati di un'altra componente *hardware* o *software*, ma anche di acquisirli e di utilizzarli. In conseguenza del fatto che l'interoperabilità è un'esigenza che si pone con singole componenti *hardware* o *software*, ne deriva che anche in una stessa infrastruttura informatica (si pensi ad un CED, che solitamente ospita una o più banche dati e, dove solitamente vi sono più sistemi funzionalmente autonomi) si pone il problema dell'interoperabilità e quindi ne deriva che questa non rileva solo per quanto riguarda le

automaticamente potrebbe recuperare le informazioni richieste dal sistema messo a disposizione da un'altra pubblica amministrazione"

[164] Così gli *standards* del vocabolario ISO/IEC 2382:2015 definiscono l'interoperabilità al punto 2121317. Il vocabolario è reperibile al seguente link https://www.iso.org/obp/ui/en/#iso:std:iso-iec:2382:ed-1:v2:en

[165] Definizione n. 2121310 dello stesso vocabolario.

[166] Definizione che si può reperire da un documento relativo ad una conferenza dell'ottobre 2013 concernente, tra le altre cose, l'interoperabilità. https://ieeexplore.ieee.org/document/6690487.

relazioni intersoggettive tra pubbliche amministrazioni, ma anche per le relazioni intraorganiche, persino all'interno di un medesimo ufficio[167].

Queste definizioni sembrano trovare un loro corrispondente coerente nella definizione legislativa contenuta nella lettera dd) del primo comma dell'art. 1 del decreto legislativo 7 marzo 2005, n. 82 (CAD- Codice dell'amministrazione digitale), il quale dispone che "*ai fini del presente codice*" si intende per interoperabilità la "*caratteristica di un sistema informativo, le cui interfacce sono pubbliche e aperte, di interagire in maniera automatica con altri sistemi informativi per lo scambio di informazioni e l'erogazione di servizi*". Il CAD utilizza la nozione generica di "*sistema informativo*"[168] (e quindi tale può essere considerata la singola unità funzionale) che, se interoperabile, può scambiare informazioni ed erogare servizi nei confronti di un'altra unità funzionale. Anche la Commissione europea ha definito l'interoperabilità come "*garanzia di comunicazione effettiva tra componenti digitali quali dispositivi, reti o archivi di dati*"[169].

Per quanto riguarda invece la seconda linea direttrice che ci permette di sottolineare di come l'interoperabilità sia un concetto che vada inteso a tutto tondo e non solo da un punto di vista tecnico, possiamo far riferimento alla distinzione tra gli ambiti operativi dell'interoperabilità contenuta all'interno del Quadro Europeo di Interoperabilità (QEI)[170]. Il QEI, infatti, elenca quattro livelli operativi: tecnico, organizzativo, giuridico e

---

[167] Ragiona su tale conseguenza G. Carullo, *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa* , cit., p.134-138.

[168] In ambito tecnico la nozione di sistema informativo assume un determinato ed ampio significato, esso infatti prende in considerazione tutto ciò che concerne la gestione e l'organizzazione delle informazioni: regole, personale preposto alla gestione ed alla circolazione delle informazioni, modulistica ed il sistema informatico eventualmente utilizzato. Quest'ultimo si distingue dal sistema informativo perché prende in considerazione in particolar modo gli strumenti informatici che si utilizzano (*server*, computer, rete di collegamento, *software*, *database* e così via). Quindi il sistema informativo è un concetto più ampio e può includere al suo interno anche il sistema informatico. In tal senso M. De Ghetto, *Corso di basi di dati I*, Youcanprint, Lecce, 2020, p.11-12. Peraltro, il CAD utilizza più volte la locuzione di "*sistema informativo*" (ad esempio all'interno degli artt.50, 60, 62- *ter* etc.).

[169] Comunicazione del 6 maggio 2015 relativa alla strategia per il mercato unico in Europa (COM(2015)192).

[170] Il *New European Interoperability Framework* (cfr. *infra*) prevede la distinzione tra i livelli di interoperabilità (*interoperability layers*) al terzo capitolo. Peraltro, tale distinzione non è

semantico. L'aspetto giuridico concerne la possibilità di organizzazioni differenti, che operano in contesti normativi differenti (per via della loro attività, funzione o localizzazione geografica), di poter interagire tra di loro in assenza di ostacoli normativi. È quindi il primo aspetto che va preso in considerazione e, a livello europeo, considerando i diversi contesti normativi degli Stati membri, assume un peso di un certo rilievo[171].

L'interoperabilità organizzativa riguarda invece, in primo luogo, la necessità di un'organizzazione di definire il proprio modello strutturale e di erogazione di un servizio pubblico[172] e, in secondo luogo, di formalizzare come avvengono le relazioni e le comunicazioni con le altre organizzazioni[173].

Tale ambito è particolarmente importante ai fini di una erogazione di servizi pubblici digitali più efficienti[174].

Il QEI con l'interoperabilità semantica punta invece a garantire che il formato ed il contenuto delle informazioni scambiate rimangano inalterati nel corso delle interazioni, assicurando che "*what is sent is what is understood"*[175]. Si raccomanda inoltre di sviluppare una strategia di gestione delle informazioni a un livello alto per prevenire

---

contenuta solamente a livello europeo, ma riconosciuta anche dall'ISO (*standard* ISO 19650), dall'IEEE ed anche in dottrina, ad esempio in H. Kubicek, R. Cimander, H.J. Scholl*, Organizational Interoperability in E-Government - Lessons from 77 European Good-Practice Cases,* 2011.

[171] Nel QEI (p.27) è infatti presa in considerazione anche l'eventuale necessità di modificare la legislazione: "*This might require that legislation does not block the establishment of European public services within and between Member States and that there are clear agreements about how to deal with differences in legislation across borders, including the option of putting in place new legislation*".

[172] Raccomandazione 28 del QEI riguardante il *Business process alignment*: "*Document your business processes using commonly accepted modelling techniques and agree on how these processes should be aligned to deliver a European public service"*.

[173] Raccomandazione 29 del QEI riguardante l'*organisational relationship*: "*Clarify and formalise your organisational relationships for establishing and operating European public services".*

[174] Così come sottolineato nel QEI ed in dottrina da V. Margariti, D. Anagnostopoulos, A. Papastilianou, T. Stamati, S. Angeli, *Assessment of organizational interoperability in e-Government: a new model and tool for assessing organizational interoperability maturity of a public service in practice,* ICEGOV 2020: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, 2020, p.300.

[175] Così come riportato nell'*European Interoperability Framework* (o QEI), p.29

frammentazione e duplicazione, utilizzando strumenti come accordi sui dati di riferimento, tassonomie, vocabolari controllati e *tehsauri* per garantire la standardizzazione dei dati (evitare la duplicazione dei dati è proprio un aspetto cruciale che dovrebbe assicurare una banca dati, ed è un elemento che la dovrebbe contraddistinguere dal mero archivio elettronico)[176].

Infine, il QEI include anche le norme relative al modello di interoperabilità tecnica che riguardano "*the applications and infrastructures linking systems and services*"[177], ovvero interfacce, interconnessioni, servizi di integrazione tra dati e, in particolare, protocolli di comunicazione sicuri. Quest'ambito è quello più affine alla definizione del vocabolario ISO/IEC vista in precedenza, seppur più circoscritta.

È quindi alla luce di queste considerazioni che si può ricavare la definizione ampia di interoperabilità *supra* esposta e che verrà utilizzata nel proseguo.

I maggiori benefici derivanti dall'interoperabilità per le pubbliche amministrazioni risiedono nel fatto che queste non debbano chiedere dati, documenti o informazioni ad un cittadino che abbia già comunicato gli stessi dati ad un'altra amministrazione (che possiede unità funzionali e banche dati differenti), ma che possano agevolmente reperire tali dati dall'altra amministrazione, realizzando così un patrimonio informativo statale (ed, in una prospettiva a lungo termine, europeo) unico ed accessibile per tutte le amministrazioni (nei limiti in cui siano legittimate ad accedere ai dati), realizzando così il principio del *once only*[178].

---

[176] R. Borruso, S. Russo, C. Tiberi, *L'informatica per il giurista,* cit., p.291.

[177] Così come riportato nell'*European Interoperability Framework* (o QEI).

[178] M. Cardone, D. Foà, *La valorizzazione del patrimonio informativo nell'ambito delle strategie di digitalizzazione della Pubblica Amministrazione*, in *Munus: Rivista giuridica dei servizi pubblici*, 3/2020, p. 609.

In altre parole, così come affermato in V. Patruno, M. M. Ragone, *Dati fulcro delle strategie per la PA digitale: tra strategie europee, interoperabilità e ruolo dei data manager*, in *FPA - Annual Report 2023*, Edizioni FORUM PA, 2023, p. 125, l'interoperabilità dei dati "*consente di immaginare non più tante pubbliche amministrazioni diverse e indipendenti l'una dall'altra, ma un'unica grande PA integrata, con la quale cittadini e imprese possono interagire in base alle proprie necessità*". Un altro beneficio rilevante che può apportare è correlato alla portabilità dei dati, in quanto se un'infrastruttura è interoperabile si potranno effettivamente trasmettere i dati nei confronti di un'altra infrastruttura. Imporre degli standard generalizzati di interoperabilità sarebbe importante non solo ai fini di una comunicazione più efficace tra banche dati pubbliche,

Dopo aver analizzato in cosa consiste ed i benefici dell'interoperabilità, bisogna analizzare come effettivamente si fa ad ottenerla. A tal proposito la tecnologia che maggiormente è in grado di ottenere l'interoperabilità sono le API (*application programming interface*), definite quali "*protocolli utilizzati come interfaccia di comunicazione tra componenti software*"[179]. Specificamente, queste interfacce digitali potenziano l'interazione tra dispositivi *software* e *hardware* con l'ambiente esterno, garantendo, da una parte, la disponibilità di infrastrutture e linguaggi in grado di acquisire e trasferire vasti volumi di dati e, dall'altra, l'accessibilità aperta a *dataset* d'interesse[180].

Pertanto, le API agiscono come un'interfaccia di comunicazione tra il *server* (erogatore), ossia l'applicativo del mittente, e il *client* (fruitore), l'applicativo del destinatario, eliminando la necessità di un meccanismo di funzionamento condiviso tra questi ultimi. Questo è particolarmente significativo considerando che ogni programma è sviluppato con linguaggi e strutture spesso diversi, per i quali è normalmente necessario creare connettori specifici per garantire l'interoperabilità con gli altri applicativi. Tuttavia, in

---

quindi per gli scambi A2A, ma per reperire informazioni da privati in un'ottica B2A (e viceversa). Sull'importanza dell'interoperabilità ai fini della portabilità si veda O. Borgogno, *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I.*, cit., p.694 e D. C. Cravo, *How to Make Data Portability Right More Meaningful for Data Subjects?,* in *European Data Protection Law Review*, cit., pp.56-57. Correlata a questi temi è la problematica del *vendor lock-in*, che potrebbe verificarsi in contesti caratterizzati da scarsa interoperabilità. Sul punto si veda lo studio preparato da *PwC EU Services* in occasione del programma ISA2*, D05.02 Big Data Interoperability Analysis*, p.29. Il concetto di "*lock-in* tecnologico" si riferisce alla situazione in cui un investimento in una tecnologia, che successivamente si rivela meno efficiente rispetto ad altre opzioni sul mercato, diventa troppo oneroso da abbandonare.

[179] A. Rezzani, *Big data. Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati*, Santarcangelo di Romagna, Maggioli, 2013, p.43. L'autore chiarisce che esse consistono in "*insiemi di routine, strutture dati e/o variabili che permettono al programmatore di richiamare le funzionalità di un'applicazione di terze parti*". Inoltre, viene sottolineata l'importanza che queste interfacce hanno nell'acquisizione di dati nei *big data,* esse infatti vengono utilizzate e messe a disposizione dalle *big tech*. Un esempio sono le *Graph API* di *Facebook*, che consentono di interfacciarsi con la piattaforma di *social network* e di esaminare tutti i contenuti pubblici (o accessibili tramite amicizia) che rispondono ai criteri di ricerca desiderati. Altri esempi sono le *Twitter* API e le API dei motori di ricerca come *Google* e *Yahoo*.

[180] Cfr. O. Borgogno, *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I.,* cit.*,* p. 689

questo contesto, il *client* deve solo conoscere le regole specifiche per richiedere i dati al *server* tramite l'API, senza dover comprendere le regole che governano il funzionamento dell'applicativo del *server* che risponde alla richiesta.

Sono quindi evidenti i vantaggi che le API possono offrire, ed infatti esse sono state poste a fondamento dell'interoperabilità sia a livello europeo[181] che nazionale. A tal fine l'AgID, attraverso le Linee Guida adottate ai sensi dell'art. 71 del CAD con determinazione del 11 ottobre 2021 n. 547[182], ha fornito indicazioni sulle tecnologie da adottare per le pubbliche amministrazioni, introducendo il Modello di Interoperabilità delle PA (MoDI). Il MoDI facilita la collaborazione tra le amministrazioni e tra queste e soggetti terzi, attraverso l'uso di soluzioni tecnologiche che permettono l'interazione e la condivisione di informazioni.

Quindi l'interoperabilità, realizzata attraverso le API, è stata posta al centro dell'attenzione da parte dell'ordinamento nazionale ed europeo, in quanto permette effettivamente a più unità funzionali di comunicare e trasferire dati.

2.1. L'interoperabilità nel contesto giuridico europeo

A livello europeo l'interoperabilità è alla base del mercato unico digitale[183] ed assume particolare rilevanza il QEI (quadro europeo di interoperabilità, anche detto EIF, ossia *European interoperability framework*).

Tale quadro, adottato per la prima volta nel 2004 in base alle previsioni contenute nella comunicazione della Commissione *eEurope* 2005[184], è stato modificato più volte, da

[181] La Commissione europea ne ha evidenziato l'utilità per garantire una corretta interoperabilità tra soggetti pubblici e privati, così come affermato in O. Borgogno, G. Colangelo, *Data sharing and interoperability: Fostering innovation and competition through APIs*, in *Computer Law & Security Review*, 2019, vol. 35, n. 5.

[182] In particolare, nella determinazione sono state adottate le "*Linee guida Tecnologie e standard per la sicurezza dell'interoperabilità tramite API dei sistemi informatici*" e le ""*Linee guida sull'interoperabilità tecnica delle Pubbliche Amministrazioni*".

[183] Commissione Europea, *Strategia per il mercato unico digitale in Europa*, COM (2015) 192 final, 6 maggio 2015.

[184] Comunicazione della Commissione europea, eEurope 2005: una società dell'informazione per tutti, COM(2002) 263 definitivo. Nel testo si legge, tra le azioni proposte (p.12), che la

ultimo, nel 2017 attraverso l'adozione del *New European Interoperability Framework*. L'obiettivo di fondo di tale quadro è quello per cui "*Le pubbliche amministrazioni dovrebbero fornire servizi pubblici digitali chiave, interoperabili e incentrati sull'utente alle imprese e ai cittadini a livello nazionale e dell'Unione, favorendo la libera circolazione delle merci, delle persone, dei servizi e dei dati in tutta l'Unione"*[185]. Quindi gli obiettivi dell'interoperabilità secondo la Commissione sono essenzialmente due: l'efficientamento dei servizi pubblici ed il sostegno per la libera circolazione di merci, persone e servizi. Questi obiettivi sono perseguiti nell'EIF attraverso la fissazione di principi, modelli e raccomandazioni comuni (la suddivisione tra i livelli di interoperabilità, vista *supra*, rappresenta un esempio di *best practice* comune).

Tale quadro è rivolto a tutte le pubbliche amministrazioni (europee e nazionali, regionali e locali) che si occupano di progettare, sviluppare ed erogare servizi pubblici europei[186]. Inoltre prevede tre tipi di interazioni: tra pubbliche amministrazioni (A2A - sia dell'UE che di uno Stato membro); tra pubbliche amministrazioni e imprese (A2B); tra pubbliche amministrazioni e cittadini (A2C).

L'importanza dell'interoperabilità è sottolineata anche da altre misure e normative che, seppur in modo frammentario, ne prevedono la disciplina[187].

Commissione per la fine del 2003 definirà "*una disciplina per la interoperabilità intesa a promuovere la fornitura di servizi paneuropei di e-government ai cittadini e alle imprese. Il documento affronterà la questione dei contenuti dell'informazione e raccomanderà alcune iniziative e specificazioni di natura tecnica per aggregare i sistemi informativi della Pubblica Amministrazione in tutta l'UE; si baserà su standard aperti e incoraggerà l'impiego di software libero (open software)*".

[185] Comunicazione della Commissione, Quadro europeo di interoperabilità - Strategia di attuazione, COM(2017) 134 final, 23 marzo 2017, p.5

[186] EIF, p.6 "*This document is addressed to all those involved in defining, designing, developing and delivering European public services*".

[187] Si pensi, ad esempio, alla proposta del *Data Act*, ossia del *Regolamento del parlamento europeo e del consiglio riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo (normativa sui dati)*, COM/2022/68 final, che contiene un capo denominato "interoperabilità" (artt.28-30) e predispone regole per garantire l'interoperabilità tra diversi settori. Si pensi al Regolamento 1024/2012/UE – cd Regolamento IMI – istitutivo del sistema di informazione del mercato interno, che ha l'obiettivo di attuare un efficace cooperazione amministrativa tra Stati membri e tra gli Stati membri e la Commissione, e per farlo prevede (al Considerando n. 2): "*un'applicazione software accessibile tramite internet, sviluppata dalla Commissione in*

Tuttavia, sebbene il valore dell'interoperabilità fosse già chiaro e sottolineato in differenti misure, si riscontravano ancora problematiche relative ad una concreta attuazione dell'interoperabilità, soprattutto a causa della natura non vincolante dell'EIF[188].

Per questo motivo la Commissione ha avanzato nel 2022 una proposta di Regolamento che "*stabilisce misure per un livello elevato di interoperabilità del settore pubblico*

collaborazione con gli Stati membri, al fine di assistere gli Stati membri nell'attuazione concreta dei requisiti relativi allo scambio di informazioni stabiliti in atti dell'Unione fornendo un meccanismo di comunicazione centralizzato che faciliti lo scambio di informazioni transfrontaliero e la mutua assistenza". Tra le Decisioni del Parlamento e del Consiglio europeo figura invece, quale principale strumento attuativo dell'EIF, la decisione (UE) 2015/2240, cd ISA2, che sostituisce la precedente decisione adottata nel 2009, ISA. Inoltre, nel programma predisposto dalla decisione ISA2 viene chiaramente sottolineata la necessità di adottare *standard* comuni a livello europeo, per eliminare il rischio che vengano adottate soluzioni di interoperabilità differenti nei diversi Stati membri (Considerando 21). Si pensi anche al *National Interoperability Framework Observatory* (NIFO), la cui missione principale è quella di monitorare l'attuazione del quadro europeo di interoperabilità (QEI) e contribuire a promuovere la modernizzazione delle pubbliche amministrazioni. In tal modo, mira a diventare una comunità online di pratiche e la principale fonte di informazioni riguardanti la pubblica amministrazione digitale e le questioni di interoperabilità in Europa. Si pensi, ancora, al *Joinup*, una piattaforma online creata nel 2021 dalla Commissione, con l'obbiettivo di diventare uno sportello unico per soluzioni digitali interoperabili, aperte e gratuite nel settore delle ICT e uno spazio online per i professionisti e gli appassionati di *e-Government* per condividere e conoscere i servizi pubblici e le iniziative digitali. Tale piattaforma vuole far emergere potenzialità di soluzioni libere ed *open source*, la cui importanza è confermata anche "dal basso" nel Parere del Comitato europeo delle regioni sulla normativa su un'Europa interoperabile - 2023/C 257/06 – in cui, al punto 11, viene affermato che le soluzioni libere ed *open source* "*rappresenterebbero un indubbio vantaggio per gli enti locali e i cittadini europei, in quanto contribuirebbero in maniera decisiva al conseguimento dell'obiettivo di condivisione e riutilizzo delle soluzioni di interoperabilità*" ed inoltre i *software open source* rappresentano il metodo principale per evitare l'effetto *lock-in* di tali enti (punto 13).

[188] M. Niestadt, *Interoperable Europe act,* European Parliament DG EPRS /Members' Research Service, 2024. L'autrice, membro del servizio di ricerca del Parlamento europeo, dopo aver constatato che l'unione europea già dispone di strutture che promuovono l'interoperabilità tra le pubbliche amministrazioni e dopo aver rilevato l'importanza dell'EIF (che ha fornito orientamenti sulla fornitura di servizi pubblici europei da oltre 15 anni), ha riscontrato (p.2) che: "*Evaluations of the current interoperability policy show, however, that being a non-binding measure, the EIF is not sufficient to remove all barriers for the EU public sector and that it has led to limited cross-border public services in the EU*". Testo reperibile al seguente sito https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)745711.

nell'Unione (normativa su un'Europa interoperabile)"[189] - (c.d. *Interoperable Europe act*) – e che ha proprio l'obiettivo di rendere il quadro precedente maggiormente vincolante ed unitario, non più adottato su base volontaria[190].

Inoltre, mira a creare degli *standard* comuni in tutta l'Europa in modo da assicurare l'interoperabilità transfrontaliera[191]. Ed infatti "*regolamentare l'interoperabilità transfrontaliera costituisce un prerequisito fondamentale per lo sviluppo e il completamento ulteriori di tutti gli aspetti esistenti del mercato unico digitale*"[192].

Con tale iniziativa si potrà garantire un approccio UE coerente e incentrato sull'uomo dell'interoperabilità; creare una struttura di governance dell'interoperabilità che aiuti le pubbliche amministrazioni e il settore privato a lavorare insieme; stabilire un ecosistema di soluzioni di interoperabilità per il settore pubblico dell'UE; ridurre la burocrazia per i cittadini e le imprese ed impedire che la limitatezza dei servizi pubblici comporti che i cittadini debbano fornire sempre le stesse informazioni[193].

L'obiettivo di incrementare l'interoperabilità e la digitalizzazione dei servizi pubblici è individuato dall'Unione europea come uno dei "*quattro punti cardinali*" da raggiungere entro il 2030 dalla Comunicazione della Commissione "*Bussola per il digitale 2030: il modello europeo per il decennio digitale*"[194]. Entro tale data è previsto l'ambizioso

---

[189] COM(2022) 720 final.

[190] Obiettivo sottolineato nello stesso testo della proposta (p.2), dove viene affermato che: "La *necessità di un'azione più incisiva nel settore è stata riconosciuta e un'azione concreta è stata annunciata in diverse comunicazioni della Commissione, tra cui le comunicazioni "Plasmare il futuro digitale dell'Europa", "Una strategia europea per i dati", "Individuare e affrontare le barriere al mercato unico" e "Digitalizzazione della giustizia nell'Unione europea. Un pacchetto di opportunità*"".

[191] Secondo l'art. 2 della proposta per interoperabilità transfrontaliera s'intende: "la *capacità dei sistemi informatici e di rete di essere utilizzati dagli enti pubblici nei diversi Stati membri e nelle istituzioni, negli organismi e nelle agenzie dell'Unione per interagire gli uni con gli altri, condividendo dati mediante comunicazione elettronica*".

[192] Come si può leggere all'interno della proposta, p. 2.

[193] M. Niestadt, *Interoperable Europe act*, cit., p.1. Inoltre, l'autrice riprende l'importanza delle API per l'interoperabilità transfrontaliera, in quanto permettono interazioni *machine-to-machine* (p.4).

[194] COM(2021) 118 final, p.13.

obiettivo di garantire il 100% dei servizi pubblici principali disponibili online per le imprese e i cittadini europei.

Per garantire una buona riuscita dell'interoperabilità, la Commissione rammenta che bisogna mettere a disposizione del pubblico soluzioni conformi ai principi del QEI di apertura, neutralità tecnica e sicurezza[195]. Particolare rilevanza assume il principio dell'*open source* dei dati e dei *software*, che mira a permettere un'effettiva condivisione delle informazioni non coperte da restrizioni (come quelle relative alla protezione dei dati personali o alla proprietà intellettuale) tra amministrazioni differenti[196]. Oltre ad efficientare l'attività amministrativa, è logico comprendere che il principio di apertura ha forti connessioni con il principio di trasparenza amministrativa e permette un controllo più intenso sui processi decisionali delle pubbliche amministrazioni.

L'interoperabilità costituisce quindi un importante obiettivo stabilito a livello unionale. Tuttavia, non mancano aspetti di criticità relativi alla proposta della Commissione per un'Europa interoperabile, tra i quali segnaliamo i rilievi effettuati dall'*European data protection supervisor* (EDPS)[197] e quelli effettuati dal Comitato europeo delle regioni (CdR)[198]. L'EDPS, dopo aver visto con favore l'iniziativa della Commissione, sottolinea che "*interoperability of network and information systems across sectors of public administration affects one of the most fundamental principles of data protection, the principle of purpose limitation*"[199]. In altre parole, l'interoperabilità, data la maggiore possibilità di condivisione e circolazione dei dati, potrebbe creare attrito con il principio di limitazione della finalità previsto dal GDPR. Il CdR, invece, sottolinea più volte le possibili difficoltà degli enti locali e regionali ad adeguarsi a tale normativa, date le loro risorse limitate. Il punto di vista del CdR non va sottovalutato, in quanto gli enti locali e regionali sono i più vicini ai cittadini e possono ottenere le informazioni più precise sulle

[195] Considerando n. 21 della proposta sulla normativa per un'Europa interoperabile.

[196] S. Aliprandi, *Interoperability and Open Standards: The key to true openness and innovation*, in *The Journal of Open Law, Technology and Society* (Jolts), 2011, Vol. 3 n. 1, pp. 5-24.

[197] Opinion 1/2023 *on the proposal for an interoperable Europe Act*.

[198] Parere del Comitato europeo delle regioni sulla normativa su un'Europa interoperabile, (2023/C 257/06).

[199] EDPS, Opinion 1/2023, p.2.

priorità e sulle aspettative dei cittadini. Quindi per fare in modo che l'interoperabilità sia un processo democratico e dal basso, sarà fondamentale un coinvolgimento sufficiente e attivo degli enti locali e regionali nel monitoraggio delle priorità dei cittadini[200].

2.2. L'attuazione dei meccanismi di interoperabilità nell'ordinamento giuridico nazionale

Terminati di vedere gli sviluppi in tema di interoperabilità a livello europeo, si vedrà adesso come essa si declina a livello nazionale.

Il principale documento di indirizzo strategico a livello nazionale è il Piano Triennale per l'informatica nella Pubblica Amministrazione (di seguito "Piano")[201], fondamentale per tutte le attività informative delle amministrazioni[202]. I piani introdotti sinora sono sei: il primo, relativo al triennio 2017-2019, che introduce un modello strategico per l'informatica nella PA; il secondo, relativo al periodo 2019-2021, che va ad implementare il piano precedente; il terzo, relativo al periodo 2020-2022, era incentrato più sul controllo dei risultati e sulle azioni previste; il quarto, relativo al periodo 2021-2023, risente e si allinea agli obblighi previsti dal PNRR ed è incentrato sulla vigilanza degli obblighi di transizione digitale della PA; il sesto, relativo al triennio 2022-2024, è allineato con la strategia "Italia Digitale 2026" e con la prima componente della prima missione del PNRR; il nuovo piano 2024-2026 si inserisce nel contesto del programma strategico "*Decennio Digitale 2030*"[203].

---

[200] Motivazione in merito ad un emendamento di un Considerando della proposta della Commissione, contenuta nel parere *supra* menzionato del Comitato.

[201] Il Piano stabilisce un modello guida per l'evoluzione dell'informatica pubblica in Italia, delineando i principi architettonici essenziali, le norme per usabilità e interoperabilità, e chiarisce la metodologia di classificazione delle spese in tecnologie dell'informazione e della comunicazione (ICT).

[202] I piani sono reperibili sul sito ufficiale dell'AgID https://www.agid.gov.it/it/agenzia/piano-triennale.

[203] Programma istituito con la Decisione (UE) 2022/2481 del Parlamento Europeo e del Consiglio del 14 dicembre 2022, i cui obiettivi sono articolati in quattro dimensioni: competenze digitali, servizi pubblici digitali, digitalizzazione delle imprese e infrastrutture digitali sicure e sostenibili.

Già il primo Piano dedicava un capitolo (il quinto) sull'interoperabilità, con l'obiettivo di favorire una condivisione trasparente di dati, servizi e piattaforme, secondo un approccio API *first* per assicurare la massima interoperabilità. Inoltre, dedicava una parte al progetto *Data & Analytics Framework* (DAF), il cui obiettivo principale era quello di "*sviluppare e semplificare l'interoperabilità dei dati pubblici tra PA, standardizzare e promuovere la diffusione degli open data, ottimizzare i processi di analisi dati e generazione di conoscenza*", con l'idea di "*aprire il mondo della Pubblica amministrazione ai benefici offerti dalle moderne piattaforme per la gestione e l'analisi dei big data*"[204].

Il progetto del DAF ha trovato la sua evoluzione[205] con l'introduzione della Piattaforma Digitale Nazionale Dati (PDND), prevista all'art. 50-*ter* del CAD, articolo introdotto a seguito delle modifiche effettuate dal d.lgs. 13 dicembre 2017, n. 217, e modificato più volte, da ultimo, con il d.l. 24 febbraio 2023, n. 13, convertito con modificazioni dalla Legge 21 aprile 2023, n. 41.

L'art. 50-*ter* prevede obiettivi specifici, da realizzare mediante un'infrastruttura tecnologica, corredati da obblighi e sanzioni in caso di inadempimento. Gli obiettivi previsti dall'articolo sono essenzialmente due: "*favorire la conoscenza e l'utilizzo del patrimonio informativo detenuto, per finalità istituzionali, dai soggetti di cui all'articolo 2, comma 2*"[206] e "*garantire la condivisione dei dati tra i soggetti che hanno diritto ad*

[204] Piano triennale 2017-2019. https://docs.italia.it/italia/piano-triennale-ict/pianotriennale-ict-doc/it/2017-2019/doc/09_data-analytics-framework.html.

[205] Il passaggio dal DAF alla PDND è chiarito più diffusamente nel terzo paragrafo del terzo capitolo. Come viene chiarito nella deliberazione – *Dati e interoperabilità* – del 16 febbraio 2023 della Corte dei conti (n. 16/2023/G), volta ad indagare l'attuazione dell'investimento 1.3 del PNRR "*Dati e interoperabilità*", la PDND dopo una prima fase sperimentale (c.d. di adozione controllata) ha raggiunto la piena operatività ed ha consentito il raggiungimento della *Milestone* del 31 dicembre 2022 prevista dal sistema ReGIS (modalità unica attraverso cui le amministrazioni possono adempiere agli obblighi di monitoraggio, rendicontazione e controllo delle misure e dei progetti finanziati dal PNRR) (si veda la pagina 43 del rapporto); Per alcune funzionalità della PDND, lo stato attuale ed il cronoprogramma della PDND si rimanda alla deliberazione della Corte dei conti appena citata (pp.36-40).

[206] In virtù del richiamo al secondo comma dell'art. 2 del CAD, l'ambito di applicazione soggettivo si estende anche ai gestori dei servizi pubblici, alle società a controllo pubblico e agli organismi di diritto pubblico. Per la rilevanza della nozione di organismo di diritto pubblico e del connotato

accedervi ai fini dell'attuazione dell'articolo 50 e della semplificazione degli adempimenti amministrativi dei cittadini e delle imprese"[207].

La PDND, la cui progettazione, sviluppo, realizzazione e gestione è attribuita alla Presidenza del Consiglio dei ministri[208], è costituita da un'infrastruttura tecnologica attraverso la quale è possibile realizzare l'interoperabilità dei sistemi informativi e delle basi di dati delle pubbliche amministrazioni autorizzate per le finalità su menzionate. L'infrastruttura è dotata di meccanismi che consentano l'accreditamento, l'identificazione e la gestione dei livelli di autorizzazione dei soggetti abilitati ad operare sulla stessa[209]. L'interoperabilità avviene "*attraverso la messa a disposizione e l'utilizzo,*

"sostanziale" di pubblica amministrazione si veda F. Gosis, *Ente pubblico*, in *Enciclopedia del diritto*, Annali VII, 2014, pp. 416-420 e A. Maltoni, *Esercizio privato di pubbliche funzioni*, in *Enciclopedia del diritto*, Annali I, 2007, pp. 570 ss.

[207] A ben vedere la seconda finalità, la condivisione dei dati è improntata alla realizzazione di due obiettivi: l'attuazione dell'art. 50 CAD e l'implementazione del principio di *once only*. Oltre queste finalità l'AgID, nel Piano 2022-2024 (cap.5), prevede che la PDND "*in futuro, dovrà consentire anche l'accesso ai big data prodotti dalle amministrazioni e l'elaborazione di politiche data-driven*", come verrà chiarito nel terzo capitolo.

[208] Ed infatti, a tal fine, il comma 2- *bis* stabilisce che: "*Il Presidente del Consiglio dei ministri o il Ministro delegato per l'innovazione tecnologica e la transizione digitale, ultimati i test e le prove tecniche di corretto funzionamento della piattaforma, fissa il termine entro il quale i soggetti di cui all'articolo 2, comma 2, sono tenuti ad accreditarsi alla stessa, a sviluppare le interfacce di cui al comma 2 e a rendere disponibili le proprie basi dati*". Per l'inadempimento dell'obbligo di rendere disponibili e accessibili le proprie basi dati è prevista una specifica sanzione, poiché secondo il comma quinto dell'art. 50-*ter* tale inadempimento costituisce un "*mancato raggiungimento di uno specifico risultato e di un rilevante obiettivo da parte dei dirigenti responsabili delle strutture competenti e comporta la riduzione, non inferiore al 30 per cento, della retribuzione di risultato e del trattamento accessorio collegato alla performance individuale dei dirigenti competenti, oltre al divieto di attribuire premi o incentivi nell'ambito delle medesime strutture*".

[209] Per aderire alla PDND è sufficiente essere catalogati su IPA, lo possono fare quindi tutte le pubbliche amministrazioni, i gestori di servizi pubblici e le società controllate (in futuro l'accesso sarà esteso alle imprese private in qualità di fruitori). In tal senso chiarisce la guida *online* predisposta da PagoPA S.p.A, società pubblica con la *mission* di diffondere i servizi pubblici digitali ed individuata come ente gestore della PDND. La guida ha l'obiettivo di essere maggiormente "pragmatica" rispetto alle linee guida AgID, poiché "*Laddove le linee guida descrivono ciò che PDND Interoperabilità dovrà essere a regime, questa guida fornisce tutti i dettagli necessari all'utilizzo di ciò che è disponibile oggi*". https://docs.pagopa.it/interoperabilita-1/ . A tal fine occorre chiarire che il secondo comma

da parte dei soggetti accreditati, di interfacce di programmazione delle applicazioni (API)"[210], e non potrebbe essere diversamente, dato che le API sono state individuate come la soluzione "principe" per realizzare l'interoperabilità e, quindi, la comunicazione tra i *data center* delle pubbliche amministrazioni con la PDND. A tal fine è previsto che le interfacce siano sviluppate dai soggetti abilitati con il supporto della Presidenza del Consiglio dei ministri (in conformità alle Linee guida AgID in materia interoperabilità) e raccolte in un apposito catalogo, "il catalogo API", reso disponibile dalla Piattaforma ai soggetti accreditati. Le amministrazioni pubbliche sono tenute ad accreditarsi alla piattaforma, a sviluppare le interfacce e a rendere disponibili le proprie basi di dati[211]. Quindi la piattaforma interviene per autorizzare lo scambio dati tra i due soggetti,

---

dell'art. 50-*ter* attribuisce all'AgID il compito di implementare il corretto funzionamento della piattaforma attraverso lo strumento delle linee guida. Queste ultime sono state adottate (ai sensi degli artt.50-*ter* e 71 del CAD) il dieci dicembre 2021, sentito il Garante per la protezione dei dati personali e acquisito il parere della Conferenza unificata.

[210] I soggetti accreditati sono i soggetti pubblici previsti al secondo comma dell'art. 2 del CAD, che hanno aderito alla PDND. Così come viene chiarito nelle linee guida AgID, *Linee Guida sull'infrastruttura tecnologica della Piattaforma Digitale Nazionale Dati per l'interoperabilità dei sistemi informativi e delle basi di dati*, 10.12.2021, p.18.

[211] Le linee guida dell'AgID sulla PDND, che si rivolgono ai soggetti di cui all'articolo 2, comma 2, del CAD, distinguono le tipologie di soggetti che interagiscono con la piattaforma in aderenti, erogatori e fruitori. Il soggetto aderente è quel soggetto che aderisce, attraverso il processo di adesione, all'Infrastruttura interoperabilità PDND in modo tale da essere abilitato erogare e/o usufruire dei servizi mediante le funzionalità dell'infrastruttura.

I soggetti erogatori (che sono tenuti ad implementare le API e a registrarle nel catalogo API reso disponibile dall'infrastruttura) sono quelle pubbliche amministrazioni che, attraverso le API, mettono a disposizione determinati *e- service* ed il proprio patrimonio informativo detenuto nelle loro banche dati per finalità istituzionali, nei confronti di altri soggetti (quelli di cui all'art. 2 comma 2 del CAD e soggetti privati che agiscono in qualità di soggetti fruitori), che hanno il diritto di accedervi – soggetti fruitori – per le finalità dell'art. 50 CAD o per la semplificazione degli adempimenti amministrativi dei cittadini e delle imprese.

In altre parole, un soggetto aderente, a seconda dei casi, potrà rivestire il ruolo di erogatore, fruitore o entrambi i ruoli (poiché può essere erogatore di determinati *e- service* e fruitore di altri). In tal modo quando un fruitore trova sul catalogo un *e-service* di cui vuole usufruire, precedentemente caricato da un erogatore, inoltra una richiesta di fruizione e l'erogatore attiva la richiesta di fruizione. Il fruitore dichiara poi per quale finalità ha bisogno di accedere al dato dell'erogatore. L'accesso al dato avviene attraverso un *voucher*, ossia un *token* con una scadenza definita, che sarà firmato dall'infrastruttura PDND interoperabilità e permetterà di creare un canale sicuro per lo scambio di informazioni tra erogatore e fruitore.

attraverso il catalogo API, la registrazione delle finalità e autorizzando il fruitore ad accedere al servizio dell'erogatore attraverso un *voucher*, evitando che per ogni scambio dati tra i due soggetti ci debbano essere accordi o convenzioni esterne. La piattaforma connette i *server* dei due soggetti attraverso un'interfaccia comune[212].

Inoltre, il comma sesto dell'art. 50-*ter* prevede che l'accesso attraverso la Piattaforma Digitale Nazionale Dati non modifica la disciplina relativa alla titolarità del trattamento, ferme restando le responsabilità ex art. 28 GDPR[213] in capo al soggetto gestore della piattaforma e le responsabilità dei soggetti accreditati che trattano i dati in qualità di titolari autonomi del trattamento[214].

La PDND, oltre all'infrastruttura tecnologica dedicata all'interoperabilità, è costituita anche da un'infrastruttura separata nella quale sono messi a disposizione, su richiesta

---

[212] Interconnettendo e rendendo interoperabili i sistemi informativi pubblici si permette di partecipare "*alla creazione di un sistema informativo pubblico unitario, che diviene base conoscitiva cui l'amministrazione pubblica può riferirsi per l'esercizio delle funzioni istituzionali*", come affermato in I. Alberti, *La creazione di un sistema informativo unitario pubblico con la Piattaforma digitale nazionale dati*, in *Le istituzioni del federalismo*, n. 2, 2022, p.476.
Per una delucidazione in merito al funzionamento della PDND e del ruolo della piattaforma quale intermediario tra erogatore e fruitore, si può vedere l'immagine della presentazione di gennaio 2023 (slide n. 18) messa a disposizione dal DTD attraverso il portale PAdigitale2026 https://assets.innovazione.gov.it/1675333598-misura-131_pdnd_avvisi-comuni_regioni_casi-d-uso.pdf .

[213] Tale articolo disciplina il rapporto tra, da un lato, il responsabile della protezione dei dati e, dall'altro, il titolare e il responsabile del trattamento (prevedendo responsabilità specifiche di questi ultimi nel caso di determinati inadempimenti).

[214] In merito a questo punto, il Garante per la protezione dei dati personali, nell'esprimere il parare favorevole alle Linee guida AgID sull'infrastruttura tecnologica PDND, ha rimarcato con rilievi positivi la differenziazione dei ruoli predisposta nell'ambito delle Linee guida: il Gestore agisce come titolare del trattamento per le attività necessarie all'implementazione e alla gestione dell'Infrastruttura interoperabilità PDND; ogni Aderente resta autonomo titolare del trattamento dei dati personali che rende disponibili o di cui fruisce nell'interazione con altro Aderente per mezzo dell'Infrastruttura interoperabilità PDND; l'Aderente che agisca come Capofila per conto di altri Aderenti deve essere designato da questi responsabile del trattamento. Parere sullo schema di "*Linee Guida sull'infrastruttura tecnologica della Piattaforma Digitale Nazionale Dati per l'interoperabilità dei sistemi informativi e delle basi di dati*" - 16 dicembre 2021 [9732758].

Presidenza del Consiglio dei ministri, i dati aggregati e anonimizzati delle pubbliche amministrazioni al fine di supportare politiche pubbliche basate sui dati[215].

È da sottolineare inoltre che il settimo comma dell'art. 50- *ter*, permette alle pubbliche amministrazioni di continuare a utilizzare anche i sistemi di interoperabilità già previsti dalla legislazione vigente. Tale disposizione non è esente da profili di criticità[216], infatti il mantenimento di tale facoltà potrebbe avere impatti negativi sulla *standardizzazione* delle tecnologie destinate ad assicurare l'interoperabilità e potrebbe non rappresentare uno stimolo per le pubbliche amministrazioni ad aderire alla PDND, in virtù del fatto che le soluzioni tecnologiche esistenti sono già conosciute dal personale amministrativo.

Il processo di accreditamento e di messa a disposizione delle basi di dati delle singole amministrazioni, previsto dall'art. 50-*ter* CAD, è un "*meccanismo attuativo a formazione progressiva*"[217], in quanto è previsto in fase di prima applicazione ed in via prioritaria "*l'interoperabilità con le basi di dati di interesse nazionale di cui all'articolo 60, comma 3-bis e con le banche dati dell'Agenzie delle entrate individuate dal Direttore della stessa Agenzia*".

A questo punto occorre chiarire cosa si intende e quali sono le basi di dati di interesse nazionale, poiché l'art. 50-*ter* CAD richiama il comma 3-*bis* dell'art. 60 CAD che elenca alcune basi di interesse nazionale. Una base di interesse nazionale rappresenta "*l'insieme delle informazioni raccolte e gestite digitalmente dalle pubbliche amministrazioni, omogenee per tipologia e contenuto e la cui conoscenza è rilevante per lo svolgimento delle funzioni istituzionali delle altre pubbliche amministrazioni, anche*

[215] Le tipologie, i limiti, le finalità e le modalità di messa a disposizione di queste tipologie dei dati aggregati ed anonimizzati sono identificate nella strategia nazionale dati, stabilita con decreto adottato dal Presidente del Consiglio dei ministri, di concerto con il Ministero dell'economia e delle finanze e il Ministero dell'interno, sentito il Garante per la protezione dei dati personali e acquisito il parere della Conferenza Unificata di cui all'articolo 8 del decreto legislativo 28 agosto 1997, n. 281. Anche l'inadempimento delle pubbliche amministrazioni di mettere a disposizione tali tipologie di dati è sanzionato con la previsione di cui al quinto comma dell'art. 50-*ter*.

[216] A. Sandulli, *Pubblico e privato nelle infrastrutture digitali nazionali strategiche*, in *Rivista trimestrale di diritto pubblico*, n. 2, 2021, p.518.

[217] *Ibidem*.

solo per fini statistici, nel rispetto delle competenze e delle normative vigenti"[218]. Inoltre, ogni base di interesse nazionale, per ciascuna tipologia di dati, costituisce un "*un sistema informativo unitario che tiene conto dei diversi livelli istituzionali e territoriali e che garantisce l'allineamento delle informazioni e l'accesso alle medesime da parte delle pubbliche amministrazioni interessate*"[219]. Quindi ogni base di dati, i cui dati sono omogenei per tipologia e contenuto, costituisce un sistema informativo, che deve possedere le caratteristiche minime di sicurezza, accessibilità e interoperabilità. In concreto, le pubbliche amministrazioni responsabili di tali basi di dati consentono l'utilizzo delle informazioni ad altre pubbliche amministrazioni mediante l'infrastruttura interoperabilità PDND[220]. Le basi di interesse nazionale sono individuate, in fase di prima applicazione, dallo stesso articolo 60, il quale, a sua volta, rimette all'AgID il compito di rilevarne di ulteriori "*tenuto conto delle esigenze delle pubbliche amministrazioni e degli obblighi derivanti dai regolamenti comunitari*"[221]. L'art. 60 del CAD individua dodici basi di dati di interesse nazionale, tra le quali, si segnala: repertorio nazionale dei dati territoriali[222]; l'anagrafe nazionale della popolazione residente[223]; banca dati nazionale

[218] Art. 60, co. 1, CAD. Secondo quanto riportato dall'AgID, le basi di dati di interesse nazionale sono "*basi di dati affidabili, omogenee per tipologia e contenuto, rilevanti per lo svolgimento delle funzioni istituzionali delle Pubbliche amministrazioni e per fini di analisi. Esse costituiscono l'ossatura del patrimonio informativo pubblico, da rendere disponibile a tutte le PA, facilitando lo scambio di dati ed evitando di chiedere più volte la stessa informazione al cittadino o all'impresa*". https://www.agid.gov.it/it/dati/basi-dati-interesse-nazionale .

[219] Art. 60, co. 2, CAD.

[220] Questo è quanto previsto dal comma 2-*bis* dell'art. 60, introdotto dal d.lgs. 217/2017, così come modificato dal decreto legge 16 luglio 2020, n. 76 (Decreto Semplificazioni), convertito nella legge 11 settembre 2020, n. 120.

[221] Art. 60, co. 3-*ter*, CAD. Inoltre, ulteriori basi di dati sono disciplinate dal contesto normativo del CAD.

[222] Il Repertorio Nazionale dei Dati Territoriali (Rndt, di cui è titolare l'AgID) è il catalogo nazionale dei metadati concernenti sia i dati territoriali che i servizi ad essi collegati forniti dalle amministrazioni pubbliche. Gestito dall'Agid e basato su tecnologia *open source*, il portale offre due caratteristiche principali: la possibilità di consultare i metadati, aperta a tutti, e la capacità di gestire i metadati, esclusiva per le amministrazioni pubbliche autorizzate. in tal senso è riportato nel portale gestito dall'AgID https://geodati.gov.it/geoportale/che-cos-e-il-rndt .

[223] L'Anagrafe Nazionale della Popolazione Residente (ANPR- di cui è titolare il Ministero dell'interno) rappresenta una base dati centralizzata a livello nazionale, che integra informazioni demografiche dei cittadini italiani residenti sia in Italia che all'estero. Questo sistema unificato

dei contratti pubblici di cui all'articolo 62-*bis*; il sistema informativo dell'indicatore della situazione economica equivalente (ISEE).

Tra quelle individuate dall'AgID e disciplinate dal contesto normativo del CAD segnaliamo: la Base dati catastale[224]; l'Indice dei domicili digitali delle pubbliche

non solo previene la duplicazione dei dati, ma offre anche ai cittadini la possibilità di controllare, correggere le proprie informazioni demografiche e accedere a servizi anagrafici in un'unica piattaforma, indipendentemente dal loro comune di residenza. Attraverso il portale dell'Anagrafe nazionale, utilizzando la propria identità digitale (come il Sistema pubblico di identità digitale, la Carta d'identità elettronica o la Carta nazionale dei servizi), i cittadini possono ottenere autonomamente e senza costi diverse tipologie di certificati digitali.

Il percorso per portare tutti i comuni italiani dentro l'Anagrafe Nazionale della Popolazione Residente si è completato il 17 gennaio 2022 con il subentro del comune di San Teodoro. Tale banca dati rappresenta un "*laboratorio*" di interoperabilità, reso possibile dalle ultime linee guida AgID. Similmente al Fascicolo Sanitario Elettronico (FSE) rappresenta un esempio virtuoso del nuovo paradigma di unificazione del patrimonio informativo reso possibile dall'interoperabilità. A tal proposito si veda la deliberazione della Corte dei conti n. 16/2023/G (pp.28-29)

[224] La banca dati catastale, di cui è titolare l'Agenzia delle Entrate, rappresenta un archivio che raccoglie e descrive le proprietà immobiliari registrate presso il Catasto. Questo sistema consente di accedere e consultare i dati in modo telematico, permettendo di ottenere documenti di visure catastali online, gli stessi che si potrebbero acquisire fisicamente presso gli uffici del catasto. Secondo quanto disposto dal provvedimento dell'Agenzia delle Entrate (Prot. n. 24383/2021), a partire dal primo febbraio 2021, è stato progressivamente implementato su tutto il territorio nazionale, ad eccezione delle aree gestite dalle Province Autonome di Trento e Bolzano, il Sistema Integrato del Territorio (SIT). Questo sistema è utilizzato dall'Agenzia delle Entrate per eseguire i compiti relativi al catasto e ai servizi geotopocartografici, nonché per la gestione dell'anagrafe immobiliare integrata. Il provvedimento stabilisce anche le modalità di consultazione e accesso agli atti e ai documenti catastali, sia presso gli uffici dell'Agenzia, sia online, oltre alle modalità di accesso telematico alla banca dati catastale da parte dei sistemi informativi delle pubbliche amministrazioni. Si vedano in tal senso il portale di catasto in rete (https://www.catastoinrete.it/banca_dati_catastale.asp) e di autonomie locali italiane (https://aliautonomie.it/2021/02/01/banche-dati-catastali-lagenzia-delle-entrate-attiva-il-sit-per-la-consultazione/).

amministrazioni e dei gestori di pubblici servizi (IPA)[225]; l'Anagrafe tributaria[226]; il Catalogo dei dati delle Pubbliche amministrazioni[227].

È da sottolineare che le basi di interesse nazionale, come tutte le banche dati pubbliche, possono contenere dati accessibili da chiunque (come il repertorio nazionale dei dati territoriali) e dati riservati (vista la natura personale di questi), con la conseguenza che l'accesso a tale banca dati sarà subordinato a specifici requisiti ed autorizzazioni (come l'ANPR o l'anagrafe tributaria). In merito alle banche dati che contengono dati personali è stato prospettato il pericolo che il principio di limitazione delle finalità possa costituire

[225] L'Indice delle Pubbliche Amministrazioni (IPA) è una banca dati pubblica, alimentata, in conformità con la normativa (CAD), dalle informazioni fornite dalle Pubbliche Amministrazioni e dai Gestori di Pubblici Servizi. Questo indice ha lo scopo di divulgare l'indirizzo di PEC (Posta Elettronica Certificata) dei Gestori di Pubblici Servizi e delle Aree Organizzative Omogenee (uffici di protocollo) delle Pubbliche Amministrazioni, facilitando lo scambio di documenti informatici attraverso vie ufficiali. Inoltre, fornisce il codice univoco per l'identificazione degli uffici delle Pubbliche Amministrazioni, necessario per la corretta trasmissione delle fatture elettroniche. L'IPA costituisce quindi un registro aggiornato delle Pubbliche Amministrazioni e dei Gestori di Pubblici Servizi, liberamente accessibile, che include, oltre ai dati obbligatori per legge, informazioni riassuntive sull'ente come l'indirizzo, il codice fiscale, il rappresentante legale, il sito web e i canali *social*. Include anche dettagli sulla struttura organizzativa (quali, uffici, responsabili, organigramma, email, indirizzi e telefono) e sui servizi offerti (includendo descrizioni, canali web di erogazione e contatti email). In tal senso, viene riportato nel sito istituzionale dell'AgID.

[226] L'Anagrafe tributaria, di cui è titolare l'Agenzia delle Entrate, è un *database* nazionale che raccoglie, ordina ed elabora le informazioni ed i dati ottenuti dalle dichiarazioni e segnalazioni fatte agli uffici finanziari. Questo include anche dati e informazioni che possono essere pertinenti per questioni fiscali. L'Anagrafe tributaria è funzionale all'attività di accertamento tributario. In tal senso nella rivista *online* dell'Agenzia delle Entrate, FiscoOggi. Per l'importanza di tale banca dati si rimanda al quarto capitolo.

[227] Il catalogo nazionale dei dati delle amministrazioni pubbliche italiane è costituito dai metadati relativi ai dati rilasciati in formato aperto dalle amministrazioni. Questo catalogo funge da punto di accesso per gli utenti ai dati aperti delle amministrazioni. È stato creato dall'Agenzia per l'Italia digitale ed è attivo dal novembre 2011. La pubblicazione e l'aggiornamento dei dati avvengono tramite un processo collaborativo coordinato dall'Agenzia in collaborazione con gli enti pubblici che forniscono dati aperti. Il catalogo, oltre a essere uno strumento di promozione delle politiche di valorizzazione del patrimonio informativo nazionale, offre risorse utili per amministrazioni e sviluppatori riguardanti i dati aperti. Sul sito è possibile cercare metadati di *dataset* pubblicati dalle amministrazioni e utilizzare questi metadati attraverso le API. Il catalogo contribuisce allo sviluppo del Portale europeo dei dati, allo stesso modo in cui lo fanno i cataloghi degli altri Paesi europei. In tal senso è riportato sul portale del catalogo https://www.dati.gov.it/ .

un vincolo alla concreta realizzazione delle trasformazioni organizzative in atto (per quanto riguarda la PDND e le basi di dati di interesse nazionale), in quanto viene imposto il vincolo di compatibilità tra la finalità del trattamento successivo e quello che ne ha giustificato la raccolta[228].

3. *Cloud*

La poca sicurezza e la frammentarietà dei *data center* delle pubbliche amministrazioni poco si concilia con un contesto tecnologico, nel quale sono necessarie infrastrutture che siano in grado di gestire grandi dimensioni di dati senza interruzioni e in sicurezza. Nell'attuale scenario una risposta a tali esigenze può consistere nel *cloud computing*, su cui l'Europa e, di conseguenza, l'Italia, attribuiscono rilevante importanza (come anche rilevato in base agli investimenti effettuati). È bene precisare che il *cloud* può essere visto come una soluzione alternativa all'interoperabilità, in quanto i dati raccolti, confluendo in un'infrastruttura comune, superano le problematiche e le difficoltà relative all'interconnessione e all'interoperabilità tra banche dati[229].

Il *cloud* può essere definito come uno spazio sull'infrastruttura di rete in cui le risorse informatiche (come *hardware*, *storage*, *database*, *networks*, sistemi operativi ed anche intere applicazioni *software*) sono disponibili istantaneamente, su richiesta[230]. Il

---

[228] B. Ponti, *Attività amministrativa e trattamento dei dati personali*, cit., p.59.

[229] M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data*, cit., p.105.

[230] S. Manvi, G. Shyam, *Cloud Computing Concepts and Technologies*, Taylor & Francis Group, LLC, 2021, pp.10-24. I vantaggi del *cloud* sono evidenti, specialmente rispetto alla molteplicità dei servizi che può offrire. Alla nozione *supra* menzionata vanno aggiunti ulteriori elementi lato utenti, così come affermato dal *National institute of standards and technology*, che attengono alle modalità di accesso al *cloud* (*self-service* - gli utenti attingono alle risorse del *cloud* in modo automatizzato, senza la necessità di un intervento umano) e la posizione geografica di essi (che non rileva ai fini dell'utilizzo delle risorse del *cloud*). Peraltro, vi sono tre modelli comuni di servizi offerti sul cloud: l'*Infrastructure-as-a-Service* (IaaS), il *Software as-a-Service* (SaaS) e la *Platform-as-a-Service* (PaaS).
Nel primo modello le funzionalità fornite all'utente sono le risorse dell'infrastruttura: capacità di elaborazione e di calcolo, archiviazione, reti, server e altre risorse informatiche fondamentali (si pensi all'esecuzione di un'applicazione ad uso intensivo di CPU o di memoria effettuata utilizzando *Amazon* IaaS *Cloud*). Nel secondo modello la funzionalità fornita all'utente è quella

presupposto della tecnologia *cloud* è la connessione ad internet, per cui le risorse informatiche sono offerte all'utente come servizi su internet, senza la necessità di installare il *software* o di essere dotati della piattaforma *hardware*. Il grande, conseguente vantaggio è quello che l'amministrazione e la manutenzione del *software* e dell'*hardware* sono centralizzate e gestite dal *cloud service provider* (CSP), perciò l'utente non dovrà farsi carico di tali attività (ma, solamente ed eventualmente, del canone da pagare). Un altro vantaggio di tali infrastrutture è che solitamente l'architettura è scalabile, per cui sarà in grado di gestire e di rispondere ad un incremento del traffico sulla rete (quindi ad un maggior utilizzo da parte degli utenti delle sue risorse), senza subire *crash* o *downtime*[231]. Sono quindi evidenti i vantaggi che tale tecnologia può offrire. Infatti, oltre alla scalabilità e costi accessibili, può offrire capacità di memoria e potenza di calcolo. La capacità di memoria può essere sfruttata in particolare per agevolare la migrazione al *cloud* dei dati contenuti nei *data center* di piccole e medie amministrazioni (con conseguenti vantaggi anche in termini di sicurezza)[232], mentre la potenza di calcolo può essere sfruttata per analizzare ed elaborare i dati, ai fini di una migliore erogazione dei servizi pubblici.

Si vedrà adesso come l'Europa e l'Italia intendono sfruttare tali potenzialità.

---

di accedere alle applicazioni ed ai *software* eseguiti su un'infrastruttura *cloud*, eliminando la necessità dell'installazione (si pensi *Microsoft Office* 365, *Gmail*, *Adobe Reader* ed a *File* PDF che utilizzano *Google Apps* senza l'installazione del *software MS Office*). Nel terzo modello la funzionalità fornita all'utente è una piattaforma su cui è possibile sviluppare applicazioni, utilizzando linguaggi di programmazione, librerie, servizi e strumenti supportati dal fornitore (ne è un esempio la creazione e la distribuzione di un'applicazione utilizzando *Google Cloud Platform* o *Microsoft Azure*). Per un approfondimento su tali servizi si veda S. Manvi, G.Shyam, *Cloud Computing Concepts and Technologies*, cit., pp.10-24; Huawei Technologies Co., Ltd. Author, *Cloud Computing Technology*, Springer nature Singapore, 2023, pp. 4-59; T. Erl, R. Puttini, Z. Mahmood, *Cloud computing: concepts, technology, & architecture*, Upper Saddle River, Prentice Hall, 2013.

[231] A. Rezzani, *Big data. Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati*, cit., pp.9-10. Per la definizione di scalabilità si veda A. B. Bondi, *Characteristics of scalability and their impact on performance*, 2000, https://dl.acm.org/doi/10.1145/350391.350432.

[232] A livello nazionale tale tecnologia può svolgere un ruolo importante per la sicurezza delle infrastrutture pubbliche e dei dati. in tal senso I. Macrì, *Dalle infrastrutture digitali delle Amministrazioni al cloud*, *il nuovo regolamento per la sicurezza dei dati e dei servizi pubblici*, in *Azienditalia*, 2022, n. 3, pp.488 ss.

3.1. Il *cloud* nel contesto giuridico europeo

A livello europeo il *Cloud*, come l'interoperabilità, riveste un'importanza centrale. La Commissione ha previsto, nella Comunicazione *Bussola Digitale 2030*, l'ambizioso obiettivo che il "*75 % delle imprese europee utilizzerà servizi di cloud computing, big data e intelligenza artificiale*"[233].

Tuttavia, sempre nella stessa Comunicazione, si prevede che una percentuale sempre crescente di dati sarà trattata ai margini della rete (*edge of the network*), più vicino agli utenti e a dove vengono generati i dati; perciò, questo cambiamento "*richiederà lo sviluppo e la diffusione di tecnologie di elaborazione dei dati fondamentalmente nuove che comprendano i margini della rete, abbandonando i modelli infrastrutturali centralizzati basati su cloud*. *L'Europa deve rafforzare le proprie infrastrutture e capacità cloud per tenere conto di queste tendenze verso una maggiore distribuzione e decentralizzazione delle capacità di elaborazione dei dati e per colmare le lacune in termini di offerta adeguata di servizi cloud in modo da soddisfare le esigenze delle imprese e della pubblica amministrazione europee"*[234].

[233] Comunicazione della Commissione, *Bussola per il digitale 2030: il modello europeo per il decennio digitale*, COM(2021) 118 final, p.11. Nella Comunicazione viene anche riconosciuta (nota n. 8) la vulnerabilità tecnologica dell'Europa, la sua crescente dipendenza da tecnologie critiche, spesso non appartenenti all'UE, e le altre dipendenze da alcune grandi aziende tecnologiche. A tal fine le statistiche riportate nella Comunicazione sono che il 90 % dei dati dell'UE è gestito da imprese statunitensi, meno del 4 % delle principali piattaforme online sono europee e i *microchip* fabbricati in Europa rappresentano meno del 10% del mercato europeo. A questa dipendenza tecnologica si aggiunge la problematica relativa allo scarso utilizzo del *cloud* da parte del settore economico, infatti nel 2019 in Europa solo una azienda su quattro ed una PMI su cinque (e si tenga presente che he le piccole, medie e microimprese costituiscono il 99% delle imprese dell'UE) utilizzano il *cloud computing* per le loro operazioni quotidiane (queste ultime statistiche sono riportate nella *brochure* della Commissione Europea del 23 Settembre 2019, *Cloud and Edge Computing: a different way of using IT — Brochure,* https://digital-strategy.ec.europa.eu/en/library/cloud-and-edge-computing-different-way-using-it-brochure).

[234] COM(2021) 118 final, p.8. Si prevede che nel 2025 ci sarà un'inversione di tendenza: l'ottanta per cento di tutti i dati saranno elaborati in dispositivi intelligenti, noti come *edge computing*, ai margini della rete https://digital-strategy.ec.europa.eu/it/policies/cloud-computing .

In altre parole, dato che una buona parte dei dati sarà trattata ai margini della rete, non si potrà puntare solamente sul *cloud*, basato su remoti e centralizzati[235] (e forse poco sicuri)[236] *data center*, ma bisognerà implementare dei nodi distribuiti che elaborino i dati ai margini della rete. A tal fine la Commissione nella Comunicazione *Bussola Digitale 2030* ha previsto, insieme all'obiettivo precedentemente menzionato sul *cloud* (e complementare a questo), l'obiettivo di installare e distribuire nell'UE "*10 000 nodi periferici ("edge nodes") a impatto climatico zero e altamente sicuri*"[237] in modo da garantire l'accesso a servizi di dati a bassa latenza (pochi millisecondi) ovunque si trovino le imprese. Attraverso l'utilizzo del *cloud* e dell'*edge* si vuole sfruttare tecnologie emergenti, quali l'intelligenza artificiale, l'analisi con i *Big Data* e *l'internet of things*[238].

3.2. L'attuazione del *cloud* nell'ordinamento giuridico nazionale

In Italia, invece, il tema del *cloud* si è iniziato a sviluppare sostanzialmente con il Piano Triennale per l'informatica nella Pubblica Amministrazione 2017-2019, che ha introdotto il Modello *Cloud* della PA, descrivendo le infrastrutture IT e i servizi *cloud* qualificati da AgID a disposizione delle amministrazioni pubbliche, delineando un programma di abilitazione al *cloud*. Quest'ultimo, sviluppato dal Team Digitale, includeva un *kit* con metodologie, strumenti e buone pratiche per guidare le amministrazioni nella creazione

[235] I servizi *cloud*, infatti, possono offrire un modello di archiviazione ed elaborazione dei dati, sia in *data center* centralizzati che in dispositivi connessi distribuiti vicino all'utente (al "bordo" della rete - *at the "edge"* of the network). In tal senso *Cloud ed Edge Computing: un modo diverso di usare l'IT — Brochure*.

[236] La Commissione ha infatti il compito di sviluppare politiche e regole che proteggano gli utenti del *cloud*, rendano i servizi *cloud* più sicuri, garantiscano una concorrenza leale e creino le condizioni quadro ottimali per un fiorente settore europeo del *cloud*. In tal senso *Cloud ed Edge Computing: un modo diverso di usare l'IT — Brochure.* Quindi la Commissione, tra le altre cose, sta provvedendo a delineare un quadro coerente di norme per regolamentare e certificare i servizi di *cloud* esistenti in merito alla portabilità e protezione dei dati, sicurezza ed efficienza energetica.

[237] COM(2021) 118 final, p.9.

[238] In tal senso viene riconosciuto nella brochure – *cloud and edge computing.* In virtù della connessione del tema dell'*edge computing* con i *Big data*, di quest'argomento se ne parlerà più diffusamente nel terzo capitolo.

di strategie di migrazione al *cloud*, e un *framework* per l'organizzazione delle unità operative responsabili dell'attuazione del programma. L'AgID ha successivamente stabilito i criteri per la qualificazione dei fornitori di servizi *cloud* e dei servizi SaaS per il *cloud* della PA con le circolari n. 2 e 3 del 9 aprile 2018, istituendo in tal modo un *marketplace* di servizi forniti da privati o società in *house* per la PA[239].

La necessità di una migrazione al *cloud* è sottolineata anche dalla circolare AgID n. 1/2019, in cui è emerso che vi sono undicimila *data center* al servizio di ventiduemila pubbliche amministrazioni e spesso in pessime condizioni (inoltre, secondo la mappatura del Governo presentata all'interno del PNRR, il novantacinque percento dei CED non soddisfa i requisiti minimi di sicurezza)[240].

In tale contesto i vantaggi relativi alla migrazione al *cloud* consisterebbero principalmente nella razionalizzazione delle infrastrutture (e quindi nell'ottimizzazione dei costi) e nella *performance* dell'infrastruttura[241].

---

[239] Si veda nel merito il portale Cloud Italia, reperibile al seguente indirizzo https://cloud.italia.it/strategia-cloud-pa/ ed E. Carloni, *Algoritmi su carta. Politiche di digitalizzazione e trasformazione digitale delle amministrazioni*, in *Diritto Pubblico*, 2/2019, pp. 363-391.

[240] Inoltre, secondo l'ultimo rapporto Clusit (Associazione Italiana per la Sicurezza Informatica), il ventitré per cento degli attacchi *cyber* in Italia sono subiti dal settore pubblico e viene inoltre affermato che i *target* maggiormente colpiti sono rappresentati dalle piccole amministrazioni locali, le aziende di servizi, gli studi professionali ed il settore industriale-manifatturiero (in merito a queste statistiche si veda *Clusit - Rapporto 2023 sulla Sicurezza ICT in Italia – edizione di metà anno, ottobre 2023*).

[241] In tal senso I. Macrì, *Cloud della Pubblica Amministrazione: una casa moderna per i dati degli Italiani*, in *Azienditalia*, 2021, n. 11, pp. 1847 ss. L'autrice evidenzia che le tecnologie *cloud* offrono risparmi anche nei costi di gestione, come quelli legati al consumo energetico. Utilizzando un'infrastruttura *cloud*, le singole amministrazioni pubbliche si alleggeriscono dal dovere di procurarsi i propri *server*, collegarli in rete e predisporre spazi adatti per ospitarli (sistemi di alimentazione, raffreddamento, antincendio, ecc.). Inoltre, il *cloud* assicura la fornitura di servizi sicuri anche in termini di *privacy*, affidabili, di alta qualità, efficienti ed efficaci. Poiché gran parte dell'innovazione in Italia è legata alla capacità di fornire servizi digitali a cittadini e imprese, più ampio sarà l'uso delle tecnologie *cloud* da parte delle amministrazioni, maggiori saranno i vantaggi diretti e indiretti per l'economia nazionale. Su tale tematica anche G. Sgueo, *I servizi pubblici digitali*, in V. Bontempi (a cura di), *Lo stato digitale nel piano nazionale di ripresa e resilienza*, Roma, RomaTrE-Press, 2022, il quale evidenzia che nonostante l'investimento iniziale richiesto per la migrazione, il *cloud* permette alle amministrazioni pubbliche di eliminare i costi associati all'acquisto e alla manutenzione dell'*hardware*.

Per questi motivi il sette settembre 2021 è stata presentata la “Strategia Cloud Italia” da parte del ministero per l’innovazione tecnologica e la transizione digitale[242]. Secondo tale strategia per i dati e per i servizi gestiti dalle amministrazioni pubbliche sarà effettuato: 1) un processo di classificazione 2) ed un piano di migrazione[243].

La classificazione dei dati e dei servizi delle amministrazioni pubbliche viene effettuata attraverso questionari sviluppati dall'Agenzia per la *cybersicurezza* nazionale (ACN), che le pubbliche amministrazioni compilano con l'assistenza del Dipartimento per la trasformazione digitale. Questo processo serve a determinare se ciascun servizio o insieme di dati può essere classificato come "*ordinario*"[244], "*critico*"[245] o "*strategico*"[246]. Data l’importanza di questa attività, visto che la classificazione dei dati e dei servizi influenzerà la scelta del tipo di *cloud* su cui destinare i dati dell'amministrazione, è prevista una fase di convalida e approvazione da parte del Dipartimento per la trasformazione digitale e dell'Agenzia per la *cybersicurezza* nazionale[247].

---

[242] “*Cloud Italia: presentati gli indirizzi strategici per la Pubblica Amministrazione*”, reperibile in https://innovazione.gov.it/notizie/articoli/cloud-italia-presentati-gli-indirizzi-strategici-per-la-pubblica-amministrazione/. Per la Strategia cloud Italia si veda https://docs.italia.it/italia/cloud-italia/strategia-cloud-italia-docs/it/stabile/index.html.

[243] La strategia governativa stabilisce scadenze abbastanza rigide: il processo di migrazione delle amministrazioni, con priorità per quelle centrali che possiedono *data center* con difetti strutturali e/o organizzativi o che non assicurano la continuità dei servizi, dovrebbe essere completato entro il 30 giugno 2026.

[244] I dati e i servizi definiti "*ordinari*" sono quelli la cui compromissione non causa l'interruzione di servizi statali né pregiudica il benessere economico e sociale della nazione. Ad esempio, i dati disponibili sui siti web pubblici delle Amministrazioni possono essere considerati ordinari.

[245] I dati e i servizi "*critici*" sono quelli la cui compromissione potrebbe influire negativamente sulle funzioni importanti per la società, compresa la salute, la sicurezza, e il benessere economico e sociale del Paese.

[246] I dati e i servizi "*strategici*" sono quelli la cui compromissione potrebbe avere conseguenze sulla sicurezza nazionale, come ad esempio i dati sanitari dei cittadini. Per le ultime tre definizioni indicate in nota si può consultare la Strategia Cloud Italia, reperibile sul portale del DTD https://innovazione.gov.it/dipartimento/focus/strategia-cloud-italia/.

[247] Il processo di classificazione effettuato dalle amministrazioni pubbliche e la successiva convalida da parte dell'ACN sono regolamentati in modo dettagliato dal nuovo regolamento sul *cloud*, che è stato adottato il 15 dicembre 2021 con la determinazione n. 628/2021 dell'Agenzia per l'Italia digitale (AGID) e successivamente integrato da atti dell'Agenzia per la *cybersicurezza* nazionale (determine n. 306/2022 e n. 307/2022).

Il piano di migrazione, che deve essere validato e confermato dal Dipartimento per la trasformazione digitale, prevede la destinazione dei dati e dei servizi delle pubbliche amministrazioni (ormai già classificati) in una determinata infrastruttura *cloud*. A tal fine la Strategia Cloud Italia prevede quattro tipologie di *cloud*: il *cloud* pubblico; il *cloud* pubblico criptato; il *cloud* privato/ibrido su licenza ed il *cloud* privato[248]. Questa categorizzazione, unita alla classificazione dei dati, è fondamentale per la gestione delle informazioni nelle Pubbliche Amministrazioni. Secondo la Strategia Cloud Italia, i dati e i servizi ordinari possono essere allocati in cloud pubblici qualificati o cloud pubblici criptati, i dati e i servizi critici in cloud pubblici criptati, cloud privati/ibridi su licenza o cloud privati qualificati, mentre i dati e i servizi strategici possono essere ospitati esclusivamente in cloud privati/ibridi su licenza o cloud privati qualificati.

Un importante tassello della Strategia *Cloud* Italia è rappresentato dal Polo Strategico Nazionale (PSN)[249], definito come una nuova infrastruttura informatica per la Pubblica

[248] Oltre ai servizi di Cloud Pubblico non qualificato (extra UE/UE), che non soddisfano i criteri tecnico-organizzativi e normativi definiti dalla Strategia, i servizi *cloud* possono essere suddivisi in diverse categorie: servizi di Cloud Pubblico qualificato (UE), conformi a legislazioni pertinenti come il GDPR e la direttiva NIS (*Network and Information Systems*), che garantiscono la localizzazione dei dati nell'UE e rispettano requisiti di sicurezza tecnico-organizzativi, basandosi tipicamente su sistemi di cifratura granulare gestiti dal fornitore CSP; servizi di Cloud pubblico con controllo *on-premise* dei meccanismi di sicurezza, noti come Cloud Criptato (IT), che aumentano significativamente il controllo sui dati e servizi, offrendo maggiore autonomia dai CSP extra-UE nella gestione e controllo delle infrastrutture tecnologiche; servizi cloud che permettono la localizzazione dei dati in Italia e un maggiore isolamento dalle "*region*" pubbliche dei CSP, che possono essere basate su tecnologia *hyperscaler* fornita da uno o più CSP, definiti Cloud privato/ibrido "su licenza" (IT), o su soluzioni basate su tecnologie commerciali qualificate tramite procedure di valutazione e certificazione tecnologica, noti come Cloud Privato Qualificato (IT). Per tali distinzioni si veda il portale *Docs Italia* nella sezione relativa alla strategia *cloud* per la pubblica amministrazione, https://docs.italia.it/italia/cloud-italia/strategia-cloud-italia-docs/it/stabile/4_la_strategia_cloud_per_la_pubblica_amministrazione.html.

[249] L'infrastruttura del PSN è stata promossa dalla Presidenza del Consiglio dei ministri attraverso il Dipartimento per la trasformazione digitale per razionalizzare i numerosi CED della PA, secondo quanto disposto dall'articolo 35 del decreto legge n. 76 del 16 luglio 2020. La Strategia *Cloud* Italia presentata il 7 settembre 2021 prevedeva la realizzazione del PSN, come una delle tre direttrici fondamentali della strategia.

Conformemente al Piano Triennale 2019-2021 ed alla circolare AgID n. 1/2019, l'Agenzia ha realizzato un censimento dei *data center* della pubblica amministrazione, rivelando che, dei 1252 *data center* censiti, soltanto 35 erano idonei all'uso da un polo strategico nazionale (PSN); 27

Amministrazione, composta da almeno quattro *data center* diversi, situati in due regioni italiane[250], per assicurare affidabilità, operatività continua in caso di guasti e autonomia.

sono stati inseriti nel gruppo A (considerati idonei a rimanere); mentre i rimanenti 1190 sono stati assegnati al gruppo B (da smantellare). Per tali riferimenti si veda la circolare AgID n. 1/2019 e la "*sintesi rapporto censimento Ict*", il cui file è reperibile al seguente indirizzo https://www.agid.gov.it/sites/default/files/repository_files/sintesi_rapporto_censimento_patrimonio_ict.pdf .

[250] Per la realizzazione di tale infrastruttura il PNRR ha individuato nel partenariato pubblico-privato lo strumento idoneo per la realizzazione del PSN. Di conseguenza il DTD ha promosso la procedura di "*project financing*", optando per la variante di cui all'art. 183, comma 15, del decreto legislativo 18 aprile 2016, n. 50, che prevede l'iniziativa degli operatori privati, invitati a predisporre un progetto di fattibilità dell'opera, da porre successivamente a base di una procedura di gara aperta anche alla partecipazione di altri concorrenti. Il DTD, dopo aver vagliato le proposte, ha selezionato, con decreto emanato il 27 dicembre 2021, il progetto di fattibilità dell'opera avanzato da TIM S.p.A., Enterprise Market, in qualità di mandataria del costituendo raggruppamento temporaneo (RTI) formato anche da CDP Equity S.p.A., Leonardo S.p.A., Sogei S.p.A. Quindi, con lo stesso decreto il raggruppamento da ultimo citato è stato nominato "soggetto promotore" dell'opera da realizzare.
Dopodiché il DTD (quale stazione appaltante), ha devoluto a Difesa Servizi S.p.A., società interamente partecipata dal Ministero della difesa (c.d. in *house*), il ruolo di centrale di committenza per l'aggiudicazione della concessione. Il bando di gara è stato pubblicato il 28 gennaio 2022 ed il 22 giugno 2022 la gara è stata aggiudicata al RTI facente capo a Fastweb S.p.A. (mandataria) e Aruba S.p.A. (mandate).
Vista la discrepanza tra soggetto aggiudicatario e promotore è stato possibile per quest'ultimo esercitare il diritto di prelazione in data 7 luglio 2022 (in base all'art. 183, comma 15, del d.Lgs. n. 50/2016), impegnandosi ad adempiere alle obbligazioni contrattuali alle "*medesime condizioni*" tecniche ed economiche offerte dal raggruppamento risultato aggiudicatario.
Il 26 agosto 2022 è stato firmato il contratto per l'avvio dei lavori di realizzazione e gestione del Polo Strategico Nazionale tra il capo del Dipartimento per la trasformazione digitale e il rappresentante legale della società di nuova costituzione (Polo Strategico Nazionale S.p.A), partecipata da TIM, Leonardo, CDP Equity e Sogei. La neo-società è attiva dal 21 dicembre 2022 nelle sedi di Acilia e Pomezia nel Lazio, Rozzano e Santo Stefano Ticino in Lombardia. Grazie alla realizzazione del PSN, la prima *milestone* della Missione 1, componente 1, investimento 1.1 Infrastrutture digitali del PNRR è stata completata. Il Polo inoltre contribuirà alla realizzazione dell'obiettivo previsto dal

L'intento del PSN è di ospitare dati e servizi critici e strategici delle Amministrazioni centrali (circa 200), delle aziende sanitarie locali e delle principali Amministrazioni locali (regioni, città metropolitane, comuni con più di 250mila abitanti). In relazione alla qualificazione dei servizi *cloud* menzionata precedentemente, il PSN fornisce servizi di tipo *cloud* pubblico criptato, *cloud* privato/ibrido "su licenza" o *cloud* privato qualificato.

---

programma Italia digitale 2026 di portare il 75% delle amministrazioni italiane ad utilizzare servizi in cloud entro il 2026.

Vi è da segnalare, tuttavia, che le società del RTI che non è risultato aggiudicatario della gara hanno impugnato diversi provvedimenti delle amministrazioni procedenti dinanzi al TAR Lazio (le istanze cautelari sono state invece respinte, consentendo di conseguenza la stipula della convenzione di concessione del 24 agosto 2022 tra il DTD e PSN S.p.A.), il quale ha accolto il ricorso, dichiarando inammissibile l'offerta presentata dal raggruppamento capitanato da TIM; la Presidenza del Consiglio dei ministri, il Ministero della difesa, il MEF e Difesa Servizi S.P.A. hanno chiesto la riforma della sentenza, ma il Consiglio di Stato (Cons. St., sez. V, 24 ottobre 2023, n. 9210) ha confermato la sentenza del TAR Lazio, nei cui confronti procede ora il giudizio risarcitorio. Il punto centrale della controversia verteva intorno alla legittimità dell'esercizio del dritto di prelazione, in quanto secondo l'art. 183, comma 15, del d.Lgs. n. 50/2016, il promotore si deve impegnare alle "*medesime condizioni*" offerte dall'aggiudicatario, sia dal punto di vista economico che tecnico, senza la possibilità di alterazioni o scostamenti; tale ultimo presupposto non sarebbe stato soddisfatto, con la conseguenza dell'illegittimità dell'esercizio del diritto della prelazione e della successiva concessione. Non è possibile affrontare le numerose problematiche, relative ai contratti pubblici, della sentenza, perciò per un approfondimento si rimanda alla sentenza stessa del Consiglio di Stato (sent. n. 9210/2023); con riguardo alla procedura di partenariato pubblico-privato ed al diritto di prelazione si rimanda a M. Ricchi, *L'Architettura dei Contratti di Concessione e di Partenariato Pubblico Privato nel Nuovo Codice dei Contratti Pubblici (d.lgs. 50/2016)*, in *Rivista giuridica del Mezzogiorno*, 3/2016, pp. 811-828 ed a E. Parisi, *Selezione del promotore e tutela giurisdizionale nel project financing a iniziativa privata*, in *Il diritto processuale amministrativo*, 4/2018, pp. 1483-1526.

Da ultimo, si segnala che la procedura di partenariato pubblico-privato è stata interessata dalla recente riforma al codice dei contratti pubblici ad opera decreto legislativo 31 marzo 2023, n. 36, che è entrato in vigore a partire dal primo aprile 2023. In merito a tale tema si può visionare A. Giovannini, *Il partenariato pubblico-privato nel nuovo codice dei contratti pubblici*, 2023, reperibile sul portale https://www.giustizia-amministrativa.it/-/158189-47.

Per permettere la migrazione dei dati e dei servizi critici delle pubbliche amministrazioni al *cloud*, è prevista una procedura che prevede tre passaggi al fine di stipulare il contratto con PSN S.p.A., le cui caratteristiche sono delineate nella convenzione stipulata tra il capo del dipartimento per la trasformazione digitale e PSN S.p.A il 24 agosto 2022, che ha ad oggetto «*l'affidamento in concessione dei servizi infrastrutturali e applicativi in cloud per la gestione di dati sensibili - "Polo Strategico Nazionale"*»[251], concessione della durata massima di tredici anni. Ebbene, secondo il procedimento delineato dalla convenzione, il primo passaggio consiste nell'invio nei confronti della società PSN da parte della pubblica amministrazione di un "*piano dei fabbisogni*", che specifica l'esigenza dell'amministrazione e le tipologie di servizi richiesti. La società PSN fornisce supporto alla PA per la redazione di tale piano[252]. Il secondo *step* prevede l'invio da parte del PSN alla pubblica amministrazione, entro 60 giorni dalla ricezione del piano dei fabbisogni, del progetto del piano dei fabbisogni con la descrizione dei servizi e la relativa quantificazione economica coerente con il listino di gara, allegando inoltre anche il piano di migrazione di massima. Infine, come ultimo passaggio, la pubblica amministrazione, se non desidera apporre modifiche al progetto del piano dei fabbisogni, viene stipulato il contratto tra le parti.

Generalmente, sebbene sia prevista per le Amministrazioni locali l'opzione di utilizzare *cloud* ibridi o privati, si prevede che i loro dati e servizi siano prevalentemente di tipo ordinario o critico, necessitando quindi di un *cloud* pubblico qualificato o criptato. Per le amministrazioni centrali, invece, si propende verso l'uso di *cloud* pubblici criptati o cloud privati, dato che si presume che i dati in loro possesso e i servizi forniti siano maggiormente di tipo critico e strategico. Di conseguenza, diventa chiaro che il valore

[251] È possibile visionare la convezione di concessione al seguente indirizzo https://www.polostrategiconazionale.it/app/uploads/2023/03/PSN-Concessione-Convenzione.pdf.

[252] Definito dalla convenzione (art. 2) come il "*documento formale predisposto dall'Amministrazione Utente, con l'ausilio del Concessionario, contenente per ciascuna categoria di servizi, indicazioni di tipo quantitativo di ciascun servizio che la stessa intende acquistare in cambio del pagamento di un prezzo*". Per agevolare la predisposizione di tale piano da parte dell'amministrazione è stato pubblicato sul portale del PSN un template, che chiarisce gli elementi che vanno specificati nel piano dei fabbisogni.

dei dati e dei servizi delle Pubbliche Amministrazioni sarà determinante nelle strategie di scelta del tipo di *cloud* più adatto per la migrazione.

In base alle modifiche organizzative analizzate all'interno del capitolo, si potrà adesso ragionare sul modello e sulla funzione amministrativa che si vuole perseguire. Per avere una visione più chiara in merito ai fini che si possono conseguire si può considerare che "*Gli obiettivi dell'informatizzazione della pubblica amministrazione erano già stati individuati almeno 26 anni or sono: a) nel miglioramento dei servizi; b) nella trasparenza dell'azione amministrativa; c) nel potenziamento dei supporti conoscitivi per le decisioni pubbliche; d) nel contenimento dei costi dell'azione amministrativa (d.lgs. 12 febbraio 1993, n. 39, art. 1*)"[253]. Attraverso l'interoperabilità si vuole garantire, in primo luogo, il miglioramento dei servizi pubblici ed, in secondo luogo, il potenziamento della decisione amministrativa, grazie al maggior numero di informazioni a supporto della decisione. Il *cloud* dovrebbe contenere i costi dell'azione amministrativa, anche se in ambito europeo attraverso l'*edge* ed il *cloud*, si vogliono potenziare anche le nuove tecnologie. Il paradigma degli *open data*, sposato dal legislatore europeo e nazionale, rafforza in particolare la trasparenza amministrativa. Questi sono gli obiettivi definibili "interni" della pubblica amministrazione, ma bisogna notare, che alcune modifiche, come quella degli *open data* (oltre a contribuire alla trasparenza amministrativa), sono capaci di avere un effetto notevole a livello sociale ed economico, in quanto viene messa a disposizione del pubblico un'ingente quantità di dati (ed anche si qualità). In ogni caso sembrerebbe che il modello predisposto dal legislatore nazionale ed europeo sia quello di un'amministrazione intermediaria di dati[254].

Nel capitolo che segue, dopo aver introdotto il cambio di paradigma all'approccio al dato (dovuto alle potenzialità di sfruttamento dei *Big Data*), si analizzerà se i profili di

[253] R.C. Perin, *Pubblica amministrazione e data analysis*, in R. C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p.11.

[254] A sostegno di tale ipotesi, anche M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data* , cit., p.287.

organizzazione amministrativa descritti in questo capitolo siano idonei anche a sfruttare le potenzialità correlate ai *Big Data*[255].

[255] Analisi che verrà condotta nel terzo paragrafo del terzo capitolo, mentre nel secondo paragrafo si analizzerà di come il paradigma dei *Big Data* influisce sulla normativa relativa alla circolazione dei dati, così come descritta nel primo capitolo.

CAPITOLO III

BIG DATA: POTENZIALITA' E CRITICITA'

1. Il paradigma dei *Big Data* e la correlazione con le tecniche di analisi dei dati

In questo paragrafo si esaminerà il cambio di paradigma all'approccio ai dati dovuto all'avvento dei *Big Data*. Inizialmente si inquadrerà il fenomeno dei *Big Data*, si sottolineerà di come i *Big Data* costituiscano un "ecosistema" favorevole per lo sviluppo dell'Intelligenza artificiale e parleremo delle metodologie di analisi ed elaborazione dei dati, in quanto inscindibilmente collegate ai *Big Data*. Introdotto in questo modo il tema, i paragrafi seguenti saranno dedicati all'analisi delle problematiche e delle relazioni che i *Big Data* hanno con il quadro normativo ed organizzativo descritto nel primo capitolo.
In assenza di definizioni vincolanti a livello normativo, per *Big Data* dobbiamo intendere un vero e proprio cambio di paradigma che concerne la raccolta, la circolazione e l'analisi di ingenti quantità di dati. I *Big Data* cambiano culturalmente l'approccio ai dati[256] e, solitamente, vengono identificati mediante alcune caratteristiche, esplicitate e compendiate nelle "3 V" (volume, varietà e velocità). Secondo tali caratteristiche per *Big Data* si intendono: "*high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization*"[257]. Per volume s'intende l'enorme quantità di dati

[256] M. Falcone*, Big Data e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica*, in *Rivista trimestrale di diritto pubblico*, n. 3, 2017, p. 608. L'autore sottolinea il passaggio da una logica *small data* ad una *Big Data*, che comporta il cambiamento di approccio, dal voler valorizzare solamente determinati dati rappresentativi della realtà a voler governare tutti i dati possibili. Questo cambiamento è osservabile in ogni momento che caratterizza la gestione dei dati: la raccolta, l'archiviazione, l'elaborazione e le finalità di utilizzo.
[257] M.A. Beyer e D. Laney, *The importance of Big Data: A definition*, Stamford, Gartner Retrieved, 21 June 2012. Dalla definizione emerge inoltre la correlazione tra le 3V e la necessità di nuove forme di analisi dei dati, perciò parlare di *Big Data* significa necessariamente parlare anche delle tecniche di analisi dei dati. E questa correlazione è evidenziata in diverse sedi. Ad esempio il Comitato della Convenzione del Consiglio d'Europa per la protezione dei dati (c.d. "Convenzione

generati e raccolti; per varietà s'intende che la tipologia di dati raccolti è di ogni tipo: strutturata, semi- strutturata e non strutturata (vista anche la diversità delle fonti da cui vengono raccolti tali dati, si pensi ad esempio alla *fingerprint*, all'*Internet of things*)[258]; per velocità s'intendono le tempistiche con cui il dato circola dal punto di origine alla banca dati e le tempistiche con cui vengono processati tali dati (c.d. *real-time action e real-time processing*)[259]. Ma la "V "più importante, espressione del lavoro sinergico delle tre precedenti è quella del "valore", che permette di ottenere quel *quid pluris* che i dati singolarmente considerati non hanno[260]. Tali caratteristiche rappresentano il *core*[261] dei *Big Data* e, a ben vedere, le modifiche riguardano molteplici aspetti, tra cui la raccolta e la circolazione dei dati, la costituzione di specifiche banche dati e l'utilizzo di determinati algoritmi di analisi.

Per segmentare le molteplici modifiche che tale paradigma comporta, può essere utile esaminare gli *step* della "*filiera dei Big Data*"[262] che porta all'estrazione della conoscenza. Tale filiera si compone di tre ordini principali di attività:

108") nelle "*Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data*" ha definito i *Big Data* come "*extremely large data sets that may be analysed computationally to extract inferences about data patterns, trends, and correlations*" ed ha sottolineato di come la maggior parte delle definizioni esistenti si focalizzano "*on the growing technological ability to collect process and extract new and predictive knowledge from great volume, velocity, and variety of data*". Nello stesso senso anche il WP29 nell'*opinion* 03/2013 *on purpose limitation* adottata nel 2013, in cui si precisa (p.35) che "*Big Data refers to the exponential growth both in the availability and in the automated use of information*".

[258] Per IoT si deve intendere la connessione tra diversi oggetti – sensori, telefoni cellulare, *wearable device*, RFID (*Radio- Frequency Identification*) – in modo da farli cooperare l'uno con l'altro al fine di completare un compito comune, attraverso l'uso di microprocessori presenti negli oggetti. In tal senso M. Chen, S. Mao, Y. Zhang, V. C. Leung. *Big data: related technologies, challenges and future prospects*, Springer, New York, 2014.

[259] M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., p.26.

[260] M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., p.10 ss.

[261] Nel tempo sono state individuate molte altre "V": ci limitiamo a segnalare in aggiunta solamente la "*veridicità*", ossia la qualità e significatività dei dati raccolti o elaborati.

[262] La filiera è stata così identificata nell'indagine conoscitiva svolta dalle tre autorità indipendenti; si veda in merito Autorità garante della concorrenza e del mercato (AGCM), Autorità per le garanzie nelle comunicazioni (AGCOM), Garante per la protezione dei dati personali (GPDP), *Indagine conoscitiva sui Big Data*, 10 febbraio 2020, pp. 8-22.

1. Raccolta, che comprende le fasi di generazione, acquisizione e memorizzazione dei dati;
2. Elaborazione, che comprende l'estrazione, l'integrazione e l'analisi dei dati;
3. Interpretazione dei dati analizzati e decisione.

La nostra attenzione si focalizzerà sulle prime due fasi, in quanto relative alle problematiche di raccolta e circolazione dei dati ed integrazione delle banche dati. Ci occuperemo dell'analisi dei dati e dell'interpretazione (dove un ruolo di primaria importanza è svolto dall'algoritmo) solo in quanto strettamente correlate ai dati.

Per quanto riguarda la raccolta, i primi due *step*[263] riguardano la generazione del dato e l'acquisizione, e le pubbliche amministrazioni hanno il continuo bisogno di acquisire dati: ciò ai fini dell'erogazione di servizi pubblici; per controllare i requisiti per concessioni ed autorizzazioni; efficientare servizi statali; efficientare servizi pubblici in un'ottica *smart city*[264]. La memorizzazione è il passaggio del dato dal dispositivo di acquisizione alla memoria di un sistema afferente all'elaborazione del dato.

---

[263] Generazione ed acquisizione si differenziano perché il primo è un processo "spontaneo", il dato viene generato infatti per il sol fatto (ad esempio) che l'utente sta navigando nel *web* o che un sensore rileva determinate caratteristiche nell'aria. La fase di generazione coinvolge il dispositivo che, grazie alle sue caratteristiche tecniche, ha generato un determinato dato. La fase di acquisizione è successiva e riguarda il passaggio del dato generato dal dispositivo verso una determinata banca dati, che entra quindi nella disponibilità di un determinato titolare.
La differenza può rilevare in particolare quando gli utenti, navigando in rete, generano un'ingente quantità di dati grezzi (c.d. *data exhaust*) – *cookies*, file temporanei, *logfiles*, parole digitate – che in seguito viene acquisita dai titolari del trattamento. I *cookies* sono file di testo che raccolgono le preferenze e le informazioni del consumatore attivo in un sito *web*, consentendone la profilazione e rivelandosi la più rilevante modalità di acquisizione di dati relativi all'utente. In tal senso M. Delmastro, A. Nicita, *Big Data. Come stanno cambiando il nostro mondo*, cit., p.10 ed AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.13. Tali dati sono infatti tutti molto rilevanti per i privati, perché consentono di effettuare al meglio uno degli obiettivi principali: la profilazione dell'utente. Tale obiettivo viene in rilievo anche in ambito pubblico nel caso in cui si volgiano erogare servizi pubblici "personalizzati" e nell'ambito dell'analisi del rischio fiscale (si veda il quarto capitolo della presente analisi).

[264] In quest'ambito in particolare la raccolta e l'elaborazione di una quantità massiva di dati si può rilevare particolarmente utile, specie attraverso l'*Internet of things*; si pensi, ad esempio, al monitoraggio del traffico statale e pedonale, al controllo della qualità dell'aria etc. La *Smart city* costituisce il terreno d'elezione per l'utilizzo dello IoT, infatti come viene precisato in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, pp.11-12, si potrebbe ipotizzare uno scenario

Durante la fase di elaborazione i dati vengono *in prims* organizzati (mediante l'estrazione e l'integrazione) e, grazie all'organizzazione, possono essere analizzati più agevolmente. Le tecniche di analisi, svolte mediante algoritmi, sono molteplici e costituiscono la c.d. *Big Data Analytics*. Si può effettuare una prima distinzione tra due tipologie di algoritmi[265]: quelli di interrogazione e quelli di apprendimento. Gli algoritmi di interrogazione[266] cercano di fornire risposte specifiche alle domande degli utenti, mentre quelli di apprendimento si concentrano sull'estrarre nuove informazioni, utilizzando tecniche di intelligenza artificiale, come il *machine learning*[267].

Ma una delle distinzioni più rilevanti tra le varie tecniche di analisi, come rilevato anche dalla *Federal Trade Commission* – l'agenzia governativa statunitense a tutela dei consumatori – è quella tra l'analisi descrittiva e l'analisi predittiva[268]. La prima consente di scoprire e riassumere i modelli o le caratteristiche che esistono negli insiemi di dati[269],

---

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*".

[265] La distinzione tra queste tecniche di analisi è individuata AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.15.

[266] Tali algoritmi permettono una ricerca più agevole nella banca dati in base alle richieste dell'utilizzatore.

[267] "*Per machine learning si intende una branca dell'Intelligenza Artificiale fondata sull'impiego della statistica bayesiana per la definizione di modelli predittivi derivati dall'analisi di ipotesi, conoscenze pregresse e altri dati relativi a specifici eventi (intensità, frequenza, ecc.)*". In tal senso O. Borgogno, *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I*, in *Il diritto dell'informazione e dell'informatica*, cit., p.689.

[268] Federal Trade Commission, *Big Data. A Tool for Inclusion or Exclusion? Understanding the Issues*, ftc Report, January 2016, p.4-5.

[269] In particolare, nell'analisi descrittiva vi rientra anche il c.d. *data mining* definito come: *" the automated process of extracting useful patterns from large data sets, and in particular, patterns that can serve as a basis for subsequent decision making*". Il *data mining* in altri contesti ricomprende anche l'analisi predittiva; infatti viene affermato che tale strumento serve a "*scoprire pattern nascosti e relazioni sottili tra i dati e di dedurre regole che permettono la predizione di futuri risultati*". In tal senso M.F. De Tullio, *La privacy e i Big Data verso una dimensione costituzionale Collettiva*, in *Politica del diritto*, 4/2016, p.639. Il perimetro della definizione risulta non chiaramente circoscritto e "*data mining remains an ambiguous term*", come affermato *in data mining: Committee on Governmental Affairs – us- Senate, Data Mining.*

la seconda si riferisce al metodo di modelli statistici per generare nuovi dati. Ebbene, entrambe queste tipologie di analisi richiedono un'ingente quantità di dati perché "*Developing and testing the models that find patterns and make predictions can require the collection and use of copious amounts of data*"[270]. E non potrebbe essere diversamente se si considera che, in particolar modo per l'analisi predittiva, il ragionamento dell'algoritmo è di tipo statistico-induttivo: più la mole di dati sarà ampia e più la base dell'inferenza sarà solida. È un'analisi probabilistica, che non può restituire certezze[271]. L'analisi effettuata dagli algoritmi può utilizzare un numero molto elevato di variabili, difficilmente valutabili a posteriori. È per questo che l'ultima fase della filiera consiste nell'interpretazione: l'analista dovrebbe comprendere il significato logico dei modelli o dei *pattern* dell'algoritmo, comprendendo le variabili che ha utilizzato e le connessioni tra i dati. Molto spesso però, data la difficoltà (se non l'impossibilità), di interpretare l'analisi dell'algoritmo, l'analista utilizza e "si fida" del risultato

---

Federal Efforts to Cover a Wide Range of Uses, Report to the Ranking Minority Member, Subcommittee on Financial Management, the Budget, and International Security, May 2004, reperibile su https://www.gao.gov/assets/gao-04-548.pdf . E' da sottolineare che l'*European Platform Undeclared Work* nel *toolkit data mining for more efficient enforcement* ha definito il *data mining* (p.1) come: "*a set of automated techniques used to extract buried or previously unknown pieces of information from large databases. By the use of data mining, correlations or patterns among dozens of fields in large relational databases will be identified*". Testo reperibile in file:///C:/Users/Utente/Downloads/Toolkit%20-%20Data%20Mining.pdf . Anche nel *Briefing* del Parlamento Europeo richiesto dal comitato JURI si distingue tra *data mining* e analisi predittiva
https://www.europarl.europa.eu/RegData/etudes/BRIE/2018/604942/IPOL_BRI(2018)604942_EN.pdf. Dobbiamo quindi concludere che gli ambiti della "*predictive analysis*" e del "*data mining*" sono distinti, seppur molto spesso vi siano confusioni tra tali termini. Vi è da considerare che con la Direttiva (UE) 2019/790 sul diritto d'autore il legislatore ha dato una definizione di *text and data mining*, su cui si tornerà *infra*.

[270] Federal Trade Commission, *Big Data. A Tool for Inclusion or Exclusion? Understanding the Issues*, ftc Report, January 2016, p. 5.

[271] Si veda a tal proposito V. Zeno-Zencovich, *Big Data e epistemologia giuridica*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. II, Roma, Roma TrE-Press, 2023, pp. 439-448; l'autore afferma che tradizionalmente il ragionamento è stato caratterizzato dal binomio "se/quindi", mentre i *Big Data*, per via delle loro caratteristiche, "*sono, solitamente, analizzati secondo una logica inferenziale, espressa dal binomio se/allora forse. Essa, naturalmente, apre la strada a molteplici soluzioni, ma solitamente verrà considerata solo quella che, sulla base dei dati disponibili, è considerata la più probabile"*.

dell'algoritmo "*non perché abbia compreso dal punto di vista logico le ragioni di queste connessioni, ma perché sa che correlazioni più ricorrenti hanno buone probabilità di ripetersi anche nei casi a venire*"[272]. In altre parole, l'algoritmo, sulla base di un'ingente quantità di dati e numerose variabili, restituisce un risultato probabilistico. La quantità di dati assume un ruolo ancora più rilevante in caso di utilizzo della tecnologia di *deep learning*[273]. Gli algoritmi di *deep learning* richiedono e sono capaci di sfruttare enormi quantità di dati (a differenza delle classiche applicazioni di *machine learning*, che performano meglio con minori quantità di dati).

Tutte queste motivazioni hanno portato la Commissione europea ad affermare che: "*L'aumento della potenza di calcolo e della disponibilità dei dati e il progresso negli algoritmi hanno reso l'IA una delle tecnologie più importanti del 21° secolo*"[274].

---

[272] M. F. De Tullio, *La privacy e i Big Data verso una dimensione costituzionale Collettiva*, cit., p.639.

[273] "*Per deep learning si intende una sottocategoria del machine learning finalizzata alla creazione di modelli di apprendimento automatico, basati sull'assimilazione di dati derivati da algoritmi complessi di calcolo (i.e. reti neurali)*". O. Borgogno, *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I.,* cit., p.690. L'importanza maggiore della quantità di dati è sottolineata più volte dalla letteratura scientifica. Ad esempio, in C.C. Aggarwal*, Neural Networks and Deep Learning, a textbook*, II ed., Springer, 2023, p.3 e 74, viene affermato che *". Recent years have seen an increase in data availability and computational power, which has led to a "Cambrian explosion" in deep learning technology*". Inoltre, l'autore sottolinea che la *performance* di un tradizionale algoritmo di *machine learning* è migliore per quantità di dati minori, invece con l'aumentare dei dati performa meglio il *deep learning*. Visto che negli anni recenti la quantità di dati è aumentata sempre di più il *deep learning* sembrerebbe una tecnologia che può rispondere alle esigenze di elaborazione dei *Big Data*. Quest'assunto viene confermato anche in C. Janiesch, P.Zschech, K.Heinrich, *Machine learning and deep learning*, *Electronic Markets*, 09/2021, volume 31, fascicolo 3, https://doi.org/10.1007/s12525-021-00475-2, dove viene affermato che: "*DL is particularly useful in domains with large and high-dimensional data, which is why deep neural networks outperform shallow ML algorithms for most applications in which text, image, video, speech, and audio data needs to be processed. However, for low-dimensional data input, especially in cases of limited training data availability, shallow ML can still produce superior results, which even tend to be better interpretable than those generated by deep neural networks*".

[274] Comunicazione della Commissione europea del 4 marzo 2019, COM(2018) 795. La Commissione è ben consapevole dell'importanza della quantità di dati per lo sviluppo di tali tecnologie, infatti in un'altra occasione ha dichiarato che "*Sono necessari ingenti volumi di dati per sviluppare l'IA. L'apprendimento automatico, un tipo di IA, opera mediante l'individuazione*

Tra quantità di dati ed IA vi è un rapporto di dipendenza bilaterale: l'IA ha bisogno di ingenti quantità di dati e questi ultimi hanno bisogno dell'IA per essere elaborati ed analizzati in modo efficace.

Tuttavia, affrontare tale discorso solamente dal punto di vista della quantità dei dati appare riduttivo. Ciò che ha permesso lo sviluppo di un ecosistema ideale per l'IA non è rappresentato infatti solamente dalla quantità dei dati, ma anche da un aumento della connettività, della capacità computazionale, della qualità dei dati e dell'eterogeneità delle fonti dei dati[275].

La qualità dei dati riveste un ruolo particolarmente importante in ambito pubblico e non solo per garantire un'analisi ottimale dei dati da parte dell'IA. La qualità dei dati pubblici è infatti imprescindibile sia per usi interni (decisionali, istruttori, di controllo, nell'ottica di un'amministrazione fruitrice dei dati), sia quando tali dati vengano messi a disposizione della collettività, secondo il paradigma degli *open data*. A ciò si deve aggiungere che, in virtù dell'attuale momento storico in cui si sta cercando di creare un patrimonio informativo comune attraverso l'integrazione e l'interoperabilità tra banche dati (dove lo scambio dei dati è automatico, senza l'intermediazione di un funzionario),

---

di modelli a partire dai dati disponibili e la successiva applicazione di questa conoscenza ai dati nuovi. Quanto più è grande il set di dati, tanto più accurata sarà l'individuazione delle relazioni anche impercettibili tra i dati", Comunicazione della Commissione europea del 26 giugno 2018 - COM(2018) 237 (avente per oggetto "*L'intelligenza artificiale per l'Europa*"), p.10.

[275] La relazione tra connettività e tecnologia computazione è affermata, ad esempio, in D. Iacovelli, M. Fontana, *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale: set di training, algoritmi e qualificazione dei dati. Profili critici*, in *Il diritto dell'economia*, cit., p.127, dove viene affermato che "*lo sviluppo degli ultimi anni di tecnologie di intelligenza artificiale è stato permesso innanzitutto da una disponibilità di grandi quantità di dati (big data), che possono essere archiviati, analizzati ed elaborati grazie all'aumento della capacità computazionale*". Nell'indagine conoscitiva delle tre autorità indipendenti – AGCOM, AGCM, GPDP, *Indagine conoscitiva sui Big Data*, p.17 – viene evidenziato che "*i miglioramenti registrati negli ultimi anni sono riconducibili non tanto agli algoritmi, che sostanzialmente non sono mutati rispetto al passato, quanto alla disponibilità di immensi quantitativi di dati, oltre che alla capacità computazionale alquanto più potente*". La relazione con l'eterogeneità delle fonti da cui provengono i dati, è stata rilevata, sempre nell'indagine conoscitiva appena menzionata (p.17), durante l'audizione di *Microsft*. Quest'ultima ha rilevato che: "*la precisione degli algoritmi aumenta con la diversità delle fonti di dati*". La qualità dei dati invece, specie nel settore pubblico, merita un discorso a parte.

un dato errato propagherà i suoi effetti negativi circolando all'interno della pubblica amministrazione[276] (a differenza del precedente modello a *silos*). Per questi motivi, nel sostenere che le banche dati pubbliche devono essere capaci di conservare massivamente le varie tipologie di dati, non si può anche sostenere che non debbano essere anche strumenti di garanzia di qualità dei dati *ivi* contenuti.

Tuttavia, delineare una nozione univoca e chiara di qualità non è semplice. Questo è dovuto principalmente al fatto che la qualità si lega con l'uso che si vuole fare di quel dato; infatti, a seconda della finalità, il dato dovrà essere "ripulito" o "elaborato" secondo determinate modalità, perciò è importante stabilire alcuni *standard* comuni, validi a prescindere dalle finalità[277].

A tal fine l'AgID, nella determinazione n. 183/2023[278], per definire la qualità ha fatto ricorso alla norma ISO/IEC 25012, secondo cui "*la qualità dei dati è il grado in cui le caratteristiche dei dati soddisfano esigenze espresse e implicite quando utilizzati in*

---

[276] E. Carloni, *Qualità dei dati, big data e amministrazione pubblica*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p.117.

[277] Ciò viene chiarito in E. Carloni, *Qualità dei dati, big data e amministrazione pubblica*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale* cit., p.122, *"Allora posto che la qualità si lega alle aspettative che abbiamo sul dato e sull'uso che ne faremo, quanto più ammettiamo usi secondari dei dati, tanto più diventa difficile assicurare la qualità del dato stesso. In altri termini, il dato, seppure di qualità adeguata per il suo uso primario, può rivelarsi incompleto, non aggiornato, inadeguato e persino erroneo se colto nella prospettiva di un uso secondario non previsto. Si pone sempre più un problema di qualità che è legato ad usi secondari dei dati, non previsti e non prevedibili nel momento in cui raccoglievamo i dati stessi: quindi dobbiamo porci l'obiettivo di garantire caratteristiche oggettive dei dati, standardizzando i requisiti di qualità di modo che i nuovi potenziali utilizzi possano ragionevolmente fondarsi su dati di qualità*".

[278] Con tale determinazione l'AgID ha adottato le "*Linee Guida recanti regole tecniche per l'apertura dei dati e il riutilizzo dell'informazione del settore pubblico*" ai sensi dell'articolo 12 del decreto legislativo 24 gennaio 2006, n. 36, disposizione introdotta a seguito delle modifiche del decreto legislativo 8 novembre 2021, n. 200, recepimento italiano della Direttiva 2019/1024 (c.d. Direttiva *Open Data*). Il tema degli *open data* è un tema che ha molti punti in comune con quello della qualità dei dati. Infatti, in tale sede l'AgID ha dedicato un paragrafo a parte per parlare della qualità dei dati, individuando alcune caratteristiche che si applicano a tutti dati e ricordando che "*La fase di valutazione della qualità dei dati è importante in tutti i sistemi informativi indipendentemente dalla scelta/necessità di procedere alla loro apertura*" (p.81).

specifiche condizioni". Dalla definizione emerge che la qualità dipende da determinate caratteristiche e dalle specifiche condizioni in cui sono utilizzati[279]. L'AgID raccomanda che per tutti i dati (e per quelli resi disponibili per il riutilizzo in particolare) vengano soddisfatte almeno quattro caratteristiche di qualità, fra le 15 previste dallo *standard* ISO/IEC 25012[280]. Tali caratteristiche sono: accuratezza; coerenza; completezza; attualità[281]. Inoltre, un ruolo di rilievo per aumentare la qualità dei dati è rivestito anche dai metadati, ossia dati che descrivono una qualche proprietà di un altro dato[282]. Lo standard ISO/IEC 25012 e le quattro caratteristiche individuate dall'AgID sono applicabili per tutti i tipi di dati, tuttavia, a conferma che in determinati casi non si può prescindere dalla finalità e dalla tipologia, per i dati territoriali è previsto uno standard specifico, ovvero l'ISO 19157 "*Geographic information - Data quality*".

Le caratteristiche viste sinora, che contrassegnano i dati pubblici, specie in un'ottica di garanzia delle informazioni che devono essere pubblicate ed aperte, rivestono un ruolo

---

[279] Le caratteristiche per valutare la qualità del dato, infatti, secondo lo *standard* ISO/IEC 25012:2008, divenuto norma italiana UNI CEI ISO/IEC 25012:2014, non sono solo "*inerenti*", cioè dipendenti dalle caratteristiche intrinseche del dato, ma anche "*inerenti e dipendenti dal sistema*" e "*dipendenti dal sistema*".

[280] Tale standard prevedeva quindici caratteristiche. Per misurare, invece, le caratteristiche, l'AgID assume come parametro lo standard ISO/IEC 25024, che definisce 63 misure di qualità applicabili alle 15 caratteristiche di qualità dei dati, con le relative funzioni di calcolo.

[281] Per accuratezza s'intende che "*il dato e i suoi attributi rappresentano correttamente il valore reale del concetto o dell'evento a cui ci si riferiscono*"; per coerenza che "*il dato e i suoi attributi non presentano contraddittorietà rispetto ad altri dati del contesto d'uso dell'amministrazione che detiene il dato*"; per completezza che "*il dato risulta esaustivo, sia per tutti i suoi valori attesi e sia rispetto alle entità relative (fonti) che concorrono alla definizione del procedimento a cui si riferisce*"; per attualità che " *il dato e i suoi attributi sono del "giusto tempo" (sono aggiornati) rispetto al procedimento a cui si riferiscono*". In tal senso, Determinazione AgID n. 183/2023, p.80. È da sottolineare che tali caratteristiche erano già state individuate dall'AgID con Determinazione Commissariale n. 68/2013, relativa alle regole tecniche per l'identificazione delle basi di dati critiche tra quelle di interesse nazionale specificate sulla base dell'art. 60 del CAD.

[282] Ad esempio, nell'Allegato 1 al documento "*Linee Guida sulla formazione, gestione e conservazione dei documenti informatici*" l'AgID definisce i metadati quali: "*Dati associati a un o documento informatico, a un fascicolo informatico o a un'aggregazione documentale per identificarli, descrivendone il contesto, il contenuto e la struttura*", https://www.agid.gov.it/sites/default/files/repository_files/allegato_1_glossario_dei_termini_e_degli_acronimi.pdf.

importante, come si è detto in precedenza, anche per lo sviluppo dell'IA[283] e delle tecnologie che si basano sui dati, così che i dati aperti e pubblici della pubblica amministrazione potranno avere un ruolo centrale per lo sviluppo di tali tecnologie e, più in generale, per lo sviluppo economico. Questo concetto è sottolineato dalle istituzioni europee nel primo considerando del *Data Act*[284], dove viene chiaramente affermato che le tecnologie - c.d. *data driven* (quindi anche l'IA) - stanno avendo un forte impatto in tutti i settori dell'economia e la qualità dei dati aumenta l'innovazione e la competitività della società, facendo anche presente che lo stesso *set* di dati può essere utilizzato per molteplici scopi, senza intaccarne la qualità[285]. La connessione tra qualità dei dati ed IA è prevista più dettagliatamente nella proposta del c.d. *AI act*[286], dove le istituzioni europee, dopo aver ribadito la connessione di tale normativa con quelle relative alla Strategia europea per i dati[287], affermano che: "*il sistema di IA, se non è addestrato con dati di elevata qualità, se non soddisfa requisiti adeguati in termini di*

[283] E questo sia per i soggetti pubblici che, in particolare, per quelli privati. Sotto questo punto di vista c'è da chiedersi se la filosofia degli *open data*, che si è ormai affermata, possa contribuire alla subalternità tecnologica dei soggetti pubblici rispetto a quelli privati, dato che sono soprattutto questi ultimi che si avvantaggiano di tali dati.

[284] Regolamento (UE) 2023/2854; per tale normativa, entrata in vigore l'11 gennaio 2024 e che diventerà applicabile a decorrere dal 12 settembre 2025 (art. 50 del *Data Act*), si rimanda al primo capitolo.

[285] Il primo considerando del *Data Act* rileva che: "*Negli ultimi anni le tecnologie basate sui dati hanno avuto effetti trasformativi su tutti i settori dell'economia. In particolare, la proliferazione di prodotti connessi all'internet delle cose ha aumentato il volume e il valore potenziale dei dati per i consumatori, le imprese e la società. Dati interoperabili e di elevata qualità provenienti da diversi settori aumentano la competitività e l'innovazione e garantiscono una crescita economica sostenibile. Lo stesso set di dati può essere potenzialmente utilizzato e riutilizzato per una varietà di scopi e in misura illimitata, senza alcuna perdita in termini di qualità o quantità*".

[286] *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*, COM(2021) 206 final. L'*AI act* è stato approvato il 14 giugno 2023 dal Parlamento Europeo.

[287] Nella proposta (p.6) si legge infatti che "*la promozione dell'innovazione basata sull'IA è strettamente legata all'Atto sulla governance dei dati, alla direttiva sull'apertura dei dati e ad altre iniziative nell'ambito della strategia dell'UE per i dati, che stabiliranno meccanismi e servizi affidabili per il riutilizzo, la condivisione e la messa in comune dei dati, essenziali per lo sviluppo di modelli di IA di alta qualità basati sui dati*".

accuratezza o robustezza, o se non è adeguatamente progettato e sottoposto a prova prima di essere immesso sul mercato o altrimenti messo in servizio, può individuare le persone in modo discriminatorio o altrimenti errato o ingiusto"[288].

Come è possibile notare da quest'affermazione i maggiori pericoli derivanti da un cattivo addestramento dell'IA (dove la qualità dei dati gioca un ruolo importante) vengono ricondotti a risultati discriminatori o ai cc.dd. *bias*. Questo perché vi è una stretta correlazione tra *input* (dati) ed *output* (decisione algoritmica), infatti "*Il processo decisionale automatizzato utilizza un set di dati, che rappresentano la base del processo di addestramento della macchina, e quindi la qualità degli input di formazione influisce sulla qualità degli output generati dal sistema"*[289] secondo un meccanismo di "*garbage in – garbage out*"[290].

Oltre alle caratteristiche qualitative individuate dall'AgID è possibile rinvenire in diverse previsioni legislative l'attenzione sulla qualità dei dati pubblici, ma queste normative sono frammentarie ed affrontano il problema da specifiche prospettive, tuttavia, contribuiscono ad irrobustire i requisiti di qualità[291]. Tra queste previsioni menzioniamo il principio di esattezza, valevole per i dati personali ed enunciato alla lettera d) dell'art. 5 del GDPR, il quale dispone che i dati personali debbano essere: "*esatti e, se necessario, aggiornati; devono essere adottate tutte le misure ragionevoli per cancellare o rettificare tempestivamente i dati inesatti rispetto alle finalità per le quali sono trattati*". Tale nozione incorpora dei requisiti di qualità, che sono essenziali per i dati personali, tra cui l'aggiornamento e la rettifica. Inoltre, emerge la natura relazionale di tale concetto, secondo cui l'esattezza va sempre parametrata alle finalità del trattamento[292]. Il principio

---

[288] Considerando n. 38 della proposta di Regolamento *AI act*.

[289] D. Iacovelli, M. Fontana, *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale*, cit., p.121.

[290] Espressione rinvenuta in G. F. Italiano, *Le sfide interdisciplinari dell'intelligenza artificiale*, in *Analisi giuridica dell'economia*, 1/2019, p.14.

[291] E. Carloni, *Qualità dei dati, big data e amministrazione pubblica*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* cit., p.124.

[292] Qui viene ripresa la centralità del principio di finalità nel GDPR e possiamo notare la relazione tra il principio di esattezza e la nozione di qualità *supra* descritta: entrambe dipendono dall'uso cui sono destinati i dati.

di esattezza "*mostra tutta la sua centralità proprio nei trattamenti articolati su filiere complesse, quali quelli operati mediante metodologie di AI. Come visto, dette tecniche sono influenzate dalla qualità dei dati in ingresso, infatti queste permettono la "creazione" di dati e di informazioni, spesso partendo raw data, mediante la combinazione e la ricerca di ricorrenze statistiche*"[293].

Per questi motivi possiamo concludere che le caratteristiche quantitative e qualitative dei dati sono un presupposto fondamentale per lo sviluppo efficace di tecnologie basate sull'intelligenza artificiale, anche quando l'algoritmo rimane immutato[294]. La scelta di prediligere le caratteristiche degli *open data*, sembra prendere in considerazione questi due aspetti; perciò, anche se tale scelta appare per lo più orientata in un'ottica di amministrazione intermediaria dei dati (più che fruitrice), pare essere correttamente ponderata alla luce della trasversalità e complessità che la disciplina in materia di dati abbraccia.

Al termine di questo paragrafo è stato rilevato che per rafforzare il potere conoscitivo sfruttando i *Big Data* sono necessari grandi quantità di dati, possibilmente di buona qualità. Nel paragrafo successivo l'analisi si concentrerà nell'esaminare se la disciplina che regola la circolazione dei dati è idonea a soddisfare le esigenze per sfruttare le potenzialità dei dati; quindi, si analizzerà se viene garantita un'ampia circolazione ed acquisizione dei dati e se sia possibile analizzare i dati personali con le tecnologie moderne, rispettando il GDPR.

---

[293] M. G. Peluso*, Intelligenza Artificiale e dati di qualità: la tecnologia come valido alleato*, in *Rivista di diritto dei media*, n. 2/2022, p.330. L'autrice si sofferma sull'impatto di tale principio sulla qualità dei dati e sottolinea la relazione del principio con la finalità di trattamento.

[294] In D. Iacovelli, M. Fontana, *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale*, cit., p.128, viene ripreso l'esempio del progetto "Norman" del *Massachusetts Institute of Technology*, in cui viene dimostrato che lo stesso algoritmo può portare a risultati molto diversi se "istruito" su *dataset* diversi. Nello stesso senso anche in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.17, viene sottolineato che: "*il dato, in quanto sorgente di informazione sul fenomeno che si intende studiare, rappresenta l'origine stessa dell'evoluzione degli algoritmi, cosicché è la disponibilità di nuove fonti di dati che consente il miglioramento degli algoritmi impiegati e/o lo sviluppo di nuovi algoritmi*".

2. Circolazione dei dati e *Big Data*: profili critici del GDPR

Si è visto che una normativa che incide trasversalmente sull'acquisizione e circolazione dei dati è costituita dalla Direttiva 96/9/CE sulla tutela delle banche dati. Il diritto *sui generis* ed il diritto d'autore potrebbero infatti impedire la circolazione ed il riutilizzo dei dati all'interno della banca dati protetta. Considerato che le moderne tecniche di analisi, come l'intelligenza artificiale, hanno bisogno di ingenti quantità di dati e che i *Big Data* costituiscono un ambiente favorevole all'utilizzo di tali tecniche, vi è da chiedersi in che misura una banca dati costituita da *Big Data* possa dar luogo a diritto d'autore o *sui generis*.

Il diritto d'autore, accordato alle banche dati in virtù dell'originalità della selezione o dell'organizzazione dei dati, non sembra conciliarsi con le "3V" che caratterizzano i *Big Data*; invero, queste caratteristiche restituiscono una configurazione dei *Big Data* come "*raccolte meramente casuali*"[295], in quanto vengono raccolti tutti i dati possibili, senza seguire uno specifico criterio. Ne deriva che il diritto d'autore non costituisce un impedimento per lo sfruttamento delle potenzialità dei *Big Data* attraverso le connesse tecniche di analisi algoritmica.

Il diritto *sui generis*, invece, potrebbe essere idoneo a tutelare le banche dati costituite da *Big Data*. Si è visto che per integrare i requisiti di tale diritto è necessario compiere un investimento volto alla raccolta dei dati (e non alla loro generazione)[296]. Tuttavia, molto spesso le raccolte di *Big Data* derivano dalla generazione di dati, effettuata specialmente da dispositivi o macchine (c.d. *machine generated data*), che, nel corso

[295] A. Amidei, M. Maggiolino, *Intelligenza artificiale, dati digitali e proprietà intellettuale*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., p. 57. Ciò non toglie che in un secondo momento i *Big Data* possano essere riordinati ed elaborati, potendo in seguito configurarsi un diritto d'autore o *sui generis* (cfr. Ibidem). In tal senso si veda anche G. Malgieri, *'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?*, in *Journal of Internet Law*, vol.20, n. 5, 2016, pp. 4 ss.

[296] E ciò deriva anche da un'interpretazione restrittiva della Corte di Giustizia dell'Unione europea. Si veda nel merito V. Falce, *Big Data, dataset e diritti esclusivi. Liaisons dangereuses tra innovazione e mercato*, in V. Falce, G. Ghidini, G. Olivieri (a cura di), *Informazione e Big Data tra innovazione e concorrenza*, Milano, 2018, pp. 123-128.

della loro attività generano ed organizzano molteplici dati[297]. In tali casi non è semplice distinguere se l'investimento sia volto alla creazione o all'organizzazione dei dati. Secondo l'opinione prevalente l'investimento deve essere finalizzato ad ottenere un *quid pluris* che determinati dati, già esistenti e generati, non hanno di per sé; in tal modo tutti quegli investimenti strumentali ad un servizio o ad un *software*, che vanno a creare, incidentalmente, banche dati "accessorie", solamente in funzione dell'erogazione di un servizio, che costituisce l'attività principale di un'organizzazione, non sarebbero tutelati dal diritto *sui generis*[298]. Tuttavia, ogni qual volta sia posto in essere un investimento autonomo e rilevante per l'organizzazione dei dati nel *database* il diritto *sui generis* dovrebbe essere applicabile, seppur con le varie e relative eccezioni previste a livello normativo[299].

---

[297] Si veda nel merito e per un approfondimento S. Scalzini, *Banche di dati, sfruttamento dei dati digitali e concorrenza*, Giappichelli, Torino, 2023, pp. 78-85. Si pensi ai dati generati da *wearable devices* (cc.dd. dispositivi *smart*, appartenenti all'IoT) o ai dati relativi alle caratteristiche del suolo, rilevati durante l'utilizzo di una macchina agricola.

[298] Si fa riferimento a tal proposito alla c.d. teoria dello *spin-off*. Tale teoria, secondo alcuni autori, è stata accolta anche dalla Corte di Giustizia europea; nel merito si veda E. Derclaye, *Databases sui generis right: should we adopt the spin-off theory?*, in *EIPR*, 2004, p. 402 ss e A. Ottolia, *Big Data e innovazione computazionale*, Torino, Giappichelli, 2017, p. 81.

[299] Vi è da considerare che, in ogni caso, ogni qual volta l'investimento sia diretto all'organizzazione dei dati, anche investendo in algoritmi diretti alla gestione dei dati, allora potrebbe aversi diritto *sui generis*; tale diritto quindi conserva un raggio d'azione ancora molto ampio; si veda in tal senso S. Scalzini, *Banche di dati, sfruttamento dei dati digitali e concorrenza*, Giappichelli, Torino, 2023, pp. 78-85. Considerato che la maggior parte delle banche dati potrebbe essere tutelata dal diritto *sui generis*, si pone il problema di limitare tale diritto nel caso in cui si debbano effettuare analisi che richiedono grandi quantità di dati. A tal fine, come affermato in S. Scalzini, M. Maggiolino, *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma?*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?*, cit., p. 186, sono state introdotte dalla Direttiva (UE) 2019/790 sul diritto d'autore (Direttiva c.d. *copyright*) due eccezioni al diritto *sui generis* riguardanti il *text and data mining* (TDM); quest'ultimo consiste, ai sensi dell'art. 2, paragrafo 1, Direttiva 2019/790 in "*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*". Tali eccezioni, previste dagli articoli 3 e 4 della citata Direttiva, sebbene volte ad una maggiore circolazione e riutilizzo dei dati, secondo la maggior parte della dottrina non sono in grado di rendersi promotrici delle innovazioni tecnologiche, in quanto sono subordinate a dei

L''assetto normativo che indirettamente costituisce il *framework* per regolamentare i *Big Data* (sebbene, occorre dire, nell'ordinamento giuridico europeo non vi sia una regolamentazione esplicita per i *Big Data*) non è orientato tuttavia all'inquadramento del fenomeno sotto il profilo della tutela del diritto d'autore e del diritto *sui generis*, ma risulta costituito dal Regolamento europeo n. 2016/679 sulla protezione dei dati personali e dal Regolamento europeo n. 2018/1807 sulla libera circolazione dei dati non personali[300].

I *Big Data* infatti necessitano di una regolamentazione che faciliti il più possibile l'acquisizione e la raccolta di dati, in modo che ne risultino potenziate tutte e tre le "V"

---

requisiti che ne limitano l'applicazione. Si veda nel merito per un approfondimento S. Orlando, *Il diritto di Text and Data Mining (TDM) non esiste*, in *Rivista italiana di informatica e diritto*, 5, 1/2023, pp. 67-81; L. Mansani, *Le eccezioni per estrazione di testo e dati, didattica e conservazione del patrimonio culturale*, in *AIDA. Annali italiani del diritto d'autore, della cultura e dello spettacolo*, 2019, pp. 3-21; S. Scalzini, *L'estrazione di dati e di testo per finalità commerciali dai contenuti degli utenti. Algoritmi, proprietà intellettuale e autonomia negoziale*, in *Analisi Giuridica dell'Economia*, 1/2019, pp. 395-423; C. Geiger, G. Frosio, O. Bulayenko, *Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU*, Centre for International Intellectual Property Studies (CEIPI), Research Paper No. 2019-08, reperibile a https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3470653.

Da ultimo, si segnala che anche il Regolamento (UE) 2023/2854 (*Data Act*) ha previsto un'eccezione al diritto *sui generis* all'art. 43; secondo quest'articolo tale diritto non si applica alle banche dati contenenti dati ottenuti o generati dall'utilizzo di un prodotto connesso o di un servizio correlato. In tal modo i *database* formati da *machine-generated data*, non dovrebbero essere tutelati dal diritto *sui generis*, favorendo l'accesso ai dati da parte degli utenti ed il loro riutilizzo da parte delle imprese. Tale previsione costituisce l'ennesima limitazione al diritto *sui generis*, come quelle presenti nel DGA, nella ODD o nella Direttiva *copyright*, tanto che "*every time the sui generis database right comes into conflict with the European Union's attempt to create a modern regulatory framework for the data economy, the SGDR* [*sui generis database right*] *loses out*"; in tal senso P. Keller, *A vanishing right? The Sui Generis Database Right and the proposed Data Act*, Wolters Kluwer blog, 2022, reperibile all'indirizzo https://copyrightblog.kluweriplaw.com/2022/03/04/a-vanishing-right-the-sui-generis-database-right-and-the-proposed-data-act/; si veda anche J. Kazeeva, *Sui Generis Intellectual Property Protection- Comparison of EU and U.S. Regulatory Approaches*, Springer, 2024, pp. 11-21.

[300] In tal senso F. Faini, *Big data, algoritmi e diritto*, in *Saggi – DPCE online*, 2019/3, p.1871.

(*volume*, *variety*, *velocity*)[301]. In particolare, in virtù della varietà dei dati, per i *Big Data* rilevano i c.d. "dati misti", ossia gli insiemi di dati composti sia da dati personali che da dati non personali, su cui la Commissione europea si è espressa con la comunicazione "*Guidance on the Regulation on a framework for the free flow of non-personal data in the European Union*"[302] del 29 maggio 2019. Alla luce delle disposizioni del RDNP e della Comunicazione della Commissione, i dati misti si possono presentare in due modalità differenti: quella in cui i dati personali e non personali siano indissolubilmente legati e quella in cui non lo siano. Nella prima situazione il RDNP lascia impregiudicata l'applicazione del GDPR[303] e secondo le linee guida il GDPR troverà applicazione anche nel caso in cui i dati personali costituiscono soltanto una piccola parte dell'insieme dei dati. Nella seconda situazione il RDNP si applica al *set* di dati non personali, ed invece, il GDPR si applica al *set* di dati personali.

Per quanto riguarda, invece, la disciplina riguardante la tutela dei dati personali, tale assetto normativo entra in crisi per via delle caratteristiche del trattamento effettuate mediante particolari algoritmi di analisi che agiscono sui *Big Data.* Infatti la normativa esposta nel primo capitolo risulta sicuramente utile per tutelare quelle categorie di dati personali (che necessitano di una maggiore tutela dal punto di vista della circolazione dei dati, delle caratteristiche dell'infrastruttura digitale e della banca dati) contenute all'interno delle banche dati tradizionali (relazionali e strutturate, che possono far eventualmente uso di algoritmi deterministici), ma entra in crisi quando si vogliano utilizzare *Big Data* e *data lake,* funzionali agli algoritmi di *machine learning* e di intelligenza artificiale e alle tecniche di *Big Data Analytics*, che analizzano l'intero "universo" dei dato e non solo un *cluster*.

[301] Sebbene bisognerebbe sempre garantire anche la "V" della veridicità, cioè quella relativa alla qualità e all'accuratezza dei dati, la quale, come si è visto, risulta essere un fattore cruciale anche per lo sviluppo di algoritmi ed IA più appropriate.

[302] Queste linee guida sono state emanate in conformità con le disposizioni dell'articolo 8, comma 3, del regolamento europeo 2018/1807. Tale articolo impone alla Commissione europea di fornire indicazioni sull'interazione tra i due regolamenti europei (GDPR e RDNP), prestando particolare attenzione sugli insiemi di dati che includono sia dati personali che non personali.

[303] Art. 2, paragrafo 2, reg. (UE) 2018/1807.

A collidere con la costruzione di un'infrastruttura *Big Data*, che raccolga massivamente i dati, è in particolare la disciplina che limita l'acquisizione e la circolazione dei dati all'interno del settore pubblico, che quindi limita la predisposizione di un ambiente favorevole a far lavorare l'algoritmo.

A collidere, invece, con la possibilità dell'utilizzo dell'algoritmo sulla massa di dati appare il principio di finalità, principio cardine del GDPR, come si è visto in precedenza.

Una doppia barriera quindi, che sembra costituire un limite allo sviluppo di tali tecnologie da parte delle pubbliche amministrazioni.

Nel proseguo faremo una prima distinzione con i soggetti privati, osservando come una diversa regolamentazione abbia effettivamente consentito un maggior sviluppo di tali tecnologie (paragone obbligatorio, visto che questo rappresenta un tassello della problematica relativa all'egemonia tecnologica dei privati rispetto alle pubbliche amministrazioni), in seguito cercheremo di esplicare meglio le caratteristiche delle barrire su esposte ed infine proveremo ad individuare qualche misura che possa comunque disciplinare tali trattamenti.

Nel primo capitolo si è vista la disciplina dell'acquisizione e circolazione dei dati personali in ambito pubblico. Tuttavia, una regolamentazione differente incide in maniera significativa sulla capacità di creare ed utilizzare infrastrutture capaci di raccogliere ed elaborare massivamente i dati[304].

[304] La regolamentazione si fa sempre più fitta attraverso un *climax* crescente a partire dai soggetti a cui non si applica il Regolamento, i soggetti privati a cui si applica il GDPR ed i soggetti pubblici sottoposti al GDPR. Probabilmente non è un caso infatti che le *Big tech* non siano europee, le stesse tre autorità garanti – AGCOM, AGCM, GPDP, *Indagine conoscitiva sui Big Data*, p.52 – infatti sottolineano che quello nord-americano è un "*ordinamento nel quale il fenomeno si è fatto strada veicolato dalle Big tech in assenza, è bene tenerlo presente approcciando la materia in esame, di un quadro normativo in materia di protezione dei dati personali comparabile con le garanzie offerte da quello europeo"*. Tuttavia, non si possono non considerare i rischi derivanti da una Regolamentazione laconica o assente, di cui lo scandalo di *Cambridge Analytica* risulta solo la punta dell'*iceberg*. Con la considerazione precedente si vuole solamente sottolineare l'importanza della normativa in tale ambito che riesce ad avere un grande impatto sulla costruzione di un'infrastruttura in grado di raccogliere massivamente i dati ed analizzarli, di fatto penalizzando in alcuni casi determinate tecnologie. Per una prospettiva a livello comparato della diversa regolamentazione delle piattaforme digitali si veda P. Guarda, G. Bincoletto, *Diritto comparato della privacy e della protezione dei dati personali*, Milano, Ledizioni, 2023; M. S. Erie,

Infatti, nel settore pubblico i presupposti per la raccolta e l'utilizzo dei dati sono nettamente distinti. In ambito privato i presupposti sono tipicamente individuati nella lett. a) (consenso) ed f) (legittimo interesse) dell'art. 6 del GDPR. Le finalità dei privati in questi casi sono autodeterminate, a differenza dei soggetti pubblici, in cui il fine è etero-determinato, individuato dal legislatore attraverso la norma attributiva del potere (sebbene il c.d. decreto capienze[305] abbia smorzato la connessione tra singolo trattamento e base giuridica legislativa, ma la norma attributiva del potere, la *mission*, rimane il cardine della liceità del trattamento, quale corollario del principio di legalità-indirizzo). Tali differenti basi giuridiche hanno consentito al settore privato di raccogliere massivamente i dati e di concentrarli in determinate infrastrutture, specialmente in virtù

T. Streinz, *The Beijing Effect: China's 'Digital Silk Road' as Transnational Data Governance*, in *N.Y.U. journal of international law & politics*, 1, 54, 2021.

E tutto ciò anche in considerazione di un difficile bilanciamento tra tutela dei dati personali e il principio di neutralità tecnologica, secondo il quale la normativa non dovrebbe imporre discriminazioni a favore dell'impiego di un tipo particolare di tecnologia, ma ci dovrebbe essere una libertà nello scegliere la tecnologia più adeguata ai propri bisogni (principio così definito sulla pagina ufficiale di EUR-lex https://eur-lex.europa.eu/IT/legal-content/summary/supporting-telecommunications-networks-and-digital-service-infrastructures-across-europe.html). Ed infatti, sebbene la normativa in ambito digitale cerchi sempre di informarsi su tale principio, alcuni autori sottolineano la difficoltà di conservare tale principio alla luce di nuove sfide e necessitò; si veda nel merito M. A. Scopelliti, *È ancora possibile la neutralità della tecnologica della normativa?*, in V. Falce (a cura di), *Strategia dei dati e intelligenza artificiale*, Torino, Giappichelli, 2023, pp. 213-223.

Disciplinando tale materia, bisogna tenere a mente anche l'importanza che da molti viene attribuita alla tecnologia dei *Big Data*, volta non solo ad una analisi dell'universo dei dati, ma che avvantaggia colui che la possiede per il sol fatto di possedere una tale quantità di dati, capace di costituire una barriera all'ingresso da parte dell'*incumbent* nei confronti delle imprese *newcomers*, fino ad essere paragonata ad una *essential facilities*. Non è possibile in questa sede analizzare le implicazioni e le disquisizioni che sollevano i *Big Data* anche nel settore concorrenziale, ma ci limitiamo a sottolineare che l'importanza di tale tecnologia, e la differenza tra chi la possiede e chi no, ha indotto in dottrina a parlare di "*big data divide*", M. Adrejevic, *The Big Data Divide*, in *International Journal of Communication*, 2014, p. 1673 ss. Si veda sul tema M. T. McCarthy, *The Big Data Divide and Its Consequences*, in S. Carta (a cura di), *Machine Learning and the City: Applications in Architecture and Urban Design*, Oxford, Wiley Blackwell, 2022, pp. 547-559; R. Dacar, *The Essential Facilities Doctrine, Intellectual Property Rights, and Access to Big Data*, in *IIC - International Review of Intellectual Property and Competition Law*, vol. 54, 2023, pp. 1487-1507.

[305] Decreto legge n. 139/2021, di cui si è visto nel primo capitolo.

della prestazione del consenso[306], in modo da rendere possibili le elaborazioni basate sulla *Big Data Analytics* e sul *machine learning*[307]. Per quanto riguarda invece la circolazione dei dati si è visto che l'ostacolo principale è rappresentato dal vincolo di compatibilità tra le finalità dei trattamenti. Vincolo superabile o in virtù del consenso o in virtù di un atto legislativo finalizzato agli interessi di cui all'art. 23, comma 1, GDPR. Anche in questo caso per i privati sarà più semplice effettuare un secondo trattamento con una finalità diversa rispetto al primo (acquisizione dei dati presso l'interessato), poiché basterà ottenere il consenso. Per i soggetti pubblici invece sarà necessario un atto legislativo che lo consenta[308].

---

[306] Tra l'altro la possibilità di accumulare massivamente i dati è anche dovuta alla superficialità con cui l'interessato spesso presta il consenso che, se anche accompagnato da un non completo rispetto da parte del titolare dell'obbligo di un'informativa sul trattamento comprensibile e trasparente (artt. 13 e 14 del GDPR), conduce ad un'inconsapevole accettazione del trattamento, il cui consenso è visto come una semplice lungaggine burocratica. La maggior parte della dottrina sottolinea le criticità e le inefficienze del principio del "consenso informato", si veda in merito M. F. De Tullio*, La privacy e i Big Data verso una dimensione costituzionale collettiva*, in *politica e diritto*, cit., p.637-639; B. Carotti, *La politica europea sul digitale: ancora molto rumore*, in *Rivista trimestrale di diritto pubblico*, 4/2022, pp.999-1000; G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, 4/2022, 977-978.

[307] Tale assunto viene evidenziato da gran parte della dottrina. Ad esempio in B. Ponti, *L'amministrazione come fornitore e come fruitore di dati personali pubblici: sono praticabili soluzioni basate sulla Big Data Analytics/Machine Learning?*, in R. C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* cit., p. 246, dove viene sottolineato: "*Mentre le banche dati di dati personali sono strutturalmente distribuite all'interno del settore pubblico, gli attori privati [omissis] possono procedere alla loro illimitata concentrazione, e proprio sulla base del principio del consenso (e delle sue ormai evidenti debolezze). Cosa comporti questo in termini di differente praticabilità (nei due contesti) di soluzioni basate su BDA/ML non è difficile immaginarlo. In termini concreti, poiché i dati personali sono "distribuiti" può accadere (anzi, è l'occorrenza fisiologicamente largamente prevalente) che lo sviluppo di una soluzione di BDA/ML necessiti della integrazione tra banche dati differenti e diversamente dislocate, ossia di un trattamento preliminare: la comunicazione/duplicazione della banca dati".* Si noti che l'autore prospetta come possibile soluzione al problema delle banche dati dislocate e distribuite, la comunicazione tra di esse (quindi l'interoperabilità) o l'integrazione (che potrebbe essere raggiunta attraverso il *cloud*).

[308] Si veda, ad esempio, G. Resta, *Pubblico, privato, collettivo nel sistema europeo di governo dei dati*, in *Rivista trimestrale di diritto pubblico*, cit., p.978.

Oltre al principio di limitazione delle finalità, un ostacolo alla raccolta massiva di dati è rappresentato anche dai principi di minimizzazione e di conservazione, valevoli per entrambi i settori[309].

Dall'analisi della disciplina concernente la liceità del trattamento, le basi giuridiche e il principio di limitazione delle finalità possiamo dedurre la diversa entità della prima barriera, che rappresenta un primo ostacolo allo sviluppo dei *Big Data* contenenti dati personali in ambito pubblico.

In ogni caso, ammesso che una tale infrastruttura *Big Data* sia stata costituita, sarà necessario analizzare la compatibilità di un'infrastruttura del genere e dell'analisi massiva dei dati con il GDPR. Per chiarire quali sono i maggiori pericoli si può far riferimento alle considerazioni svolte dal Garante della *privacy* nell'indagine conoscitiva sui *Big Data*[310], ove afferma che: "*L'indeterminatezza delle finalità perseguite in concreto con i Big Data, i rischi di reidentificazione degli interessati e l'opacità delle logiche applicate al trattamento [omissis] entrano in conflitto con i requisiti richiesti dalla normativa internazionale ed europea in materia di protezione dei dati personali a tutela dei diritti e libertà degli individui".* Vengono sottolineate quindi le tre problematiche principali: l'indeterminatezza delle finalità; i rischi di reidentificazione e l'opacità delle logiche applicate al trattamento.

---

[309] Come risulta in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p. 53, "*Anche i principi di minimizzazione, di limitazione della finalità e di conservazione dei dati per il solo tempo indispensabile alla realizzazione del trattamento mal si attagliano a raccolte massive di dati, in ipotesi acquisiti, magari non per esigenze attuali ma in vista di future e solo ipotetiche necessità, per essere quindi riutilizzati per fini ulteriori non sempre compatibili con quelli originari*". Per una crisi anche di tali principi alla luce del mutato contesto tecnologico si veda, per il principio di conservazione, V. Pagnanelli, *Conservazione dei dati e sovranità digitale. Una rilettura della (big) data governance pubblica alla luce delle nuove sfide globali*, in *Rivista italiana di informatica e diritto*, 1/2021, pp. 13-28 e per il principio di minimizzazione T. Taini, *L'incidenza dei Big data e del machine learning sui principi alla base del Regolamento Europeo per la tutela dei dati personali (2016/679/UE) e proposte per una nuova normativa in tema di privacy*, in S. Bonavita (a cura di), *Società delle tecnologie esponenziali e General Data Protection Regulation: Profili critici nella protezione dei dati*, Milano, Ledizioni, 2018, pp. 35-65; S. Franca, *I dati personali nell'amministrazione pubblica. Attività di trattamento e tutela del privato*, Università degli Studi di Trento, Collana della facoltà di giurisprudenza, 44, 2023, p. 182.

[310] AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.68.

Nel proseguo analizzeremo la prima e la seconda problematica, tralasciando in questa sede la terza, in quanto legata più ad aspetti di funzionamento dell'algoritmo, che esulano dalla presente trattazione[311], che invece riguarda principalmente la creazione dell'infrastruttura capace di raccogliere massivamente i dati e l'effetto delle elaborazioni algoritmiche nei confronti dei principi sopra esposti.

Il primo problema riguarda la finalità del trattamento che, come si è visto nel primo capitolo, rappresenta una "pietra angolare" del GDPR. Quando si utilizzano alcune tecniche di analisi algoritmica, gran parte della dottrina sottolinea una "tendenziale" (definita da alcuni persino "ontologica") incompatibilità[312] tra finalità del trattamento e

[311] Per la stessa ragione nell'analisi non è stata presa in considerazione la disciplina che riguarda principalmente l'algoritmo e le tutele predisposte in caso di decisione algoritmica. Per tali motivi non sono stati analizzati, ad esempio, gli articoli 13, 14 e 15 del GDPR che, tra le altre cose, prevedono il diritto di conoscere l'esistenza di un processo decisionale automatizzato, compresa la profilazione o l'art. 22, che prevede la garanzia di "*non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato dei dati che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla persona*" con una logica in piena aderenza alla teoria *human in the loop*, considerata un importante garanzia nell'ambito della decisione automatizzata o algoritmica, in quanto si concreta nell'obbligo di un intervento umano nel processo decisionale. Il procedimento *human in the loop,* il diritto a conoscere la logica dell'algoritmo che ha portata alla decisione, mirano a configurare la "*democratizzazione*" della decisione algoritmica, espressione di M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data*, cit., p.210; per il principio *human in the loop* si veda B. Marchetti, *La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica*, in *BioLaw Journal –Rivista di BioDiritto*, 2/2021, pp. 367-385 e S. Foa, *Intelligenza artificiale e cultura della trasparenza amministrativa. Dalle "scatole nere" alla "casa di vetro"?,* in *Diritto amministrativo*, 3/2023, pp. 526-536. Per una riflessione di come tale principio venga adottato nelle pubbliche amministrazioni statali si veda E. Chiti, B. Marchetti, N. Rangone, *L'impiego di sistemi di intelligenza artificiale nelle pubbliche amministrazioni italiane: prove generali*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. I, Roma, Roma TrE-Press, 2023, pp. 150-152.

[312] Ad esempio, afferma una "*tendenziale incompatibilità*" B. Ponti, *L'amministrazione come fornitore e come fruitore di dati personali pubblici: sono praticabili soluzioni basate sulla Big Data Analytics/Machine Learning?*, in R. C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale, cit.,* , p.236, "*Le logiche del BDA/ML entrano in conflitto con questo modello di tutela, dal momento che loro caratteristica precipua è quella di consentire l'emersione di conoscenza nuova ed inattesa, a partire dalla rilevazione di correlazioni/interrelazioni non immediatamente evidenti.*" Sottolinea invece un'"*ontologica*" incompatibilità F. Faini, *Big Data e internet of things: data protection e data governance alla luce*

tali analisi. Ciò è dovuto al fatto che le tecniche di *machine learning* e di *Big Data analytics* consentono l'emersione di una conoscenza nuova ed inattesa. Quindi sarà difficile per il titolare esplicare le finalità all'interessato, in virtù del fatto che l'algoritmo può trovare correlazioni tra i dati che il titolare non poteva prefigurarsi nel momento dell'acquisizione del dato, in tal modo utilizzando i dati con una finalità diversa e difficilmente conoscibile per il titolare. Questa problematica investe sia il settore privato che pubblico, minando a ben vedere un principio cardine della disciplina ed ostando ad un'importante funzionalità di un'infrastruttura che raccolga in massa i dati[313].

Il secondo problema concerne invece i rischi di "*reidentificazione*" dell'interessato. Questa problematica riguarda le procedure di de-anonimizzione e di "*ricostruzione del dato personale*"[314] attraverso la combinazione dei dati accumulati massivamente e le tecniche di *Big Data Analytics*.

---

del regolamento europeo, in G. Cassano, V. Colarocco, G.B. Gallus, F.P. Micozzi (a cura di), *Il processo di adeguamento al GDPR*, cit., pp. 311-312. L'autrice afferma: "*Le caratteristiche dei big data e dell'IoT con le relative modalità di utilizzo dei dati mostrano ontologicamente elementi di criticità nel rispetto della disciplina in materia di protezione dei dati personali. Come esaminato nei precedenti paragrafi, nelle strategie che riguardano l'utilizzo di big data e dei dati prodotti dell'IoT, non è necessariamente predefinito a priori l'oggetto di indagine e non sono prevedibili al momento della raccolta le molteplici finalità raggiungibili, dal momento che le analisi e le elaborazioni sono capaci di condurre a interessanti risultati inattesi; le domande da porre per lo più non sono note ex ante, ma emergono proprio al momento della raccolta e dell'analisi dei dati. Questo aspetto è evidentemente problematico in modo particolare quando i dati sono personali. In specifico, l'utilizzo dei big data, laddove siano presenti dati personali, rende particolarmente difficile il rispetto del principio di limitazione della finalità, ossia la raccolta per finalità determinate, esplicite e legittime, e il successivo trattamento in modo che non sia incompatibile con tali finalità*". Continua l'autrice sottolineando l'incompatibilità: "*Su cosa saranno fornite le informazioni da parte del titolare del trattamento e su cosa esprimerà il consenso l'interessato, se non si conoscono preventivamente le finalità di utilizzo dei big data?*". A tal fine si veda anche M. Falcone, *Big Data e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica*, cit., p.615.

[313] Si rimanda alle considerazioni svolte nella nota precedente.

[314] In tal seno M. Delmastro, A. Nicita, *Big data. Come stanno cambiando il nostro mondo*, cit., p.34-39, ove gli autori chiariscono in prima battuta che i dati personali assumono valore in quanto grazie ad essi le tecniche di *Big data analytics* forniscono indicazioni sugli schemi tipizzati di comportamento individuale, rivelando come si comportano certe categorie di individui, assumendo in tal modo anche la veste di informazioni personali di carattere generale (non solo

Infatti, attraverso queste ultime tecniche è possibile ottenere dati personali, partendo anche da dati non personali. Inoltre, è da considerare il fatto che a partire da dati non personali, è possibile anche ottenere dei dati personali di un gruppo o di una comunità. Si faccia l'esempio di un titolare che abbia svariati dati relativi ad un territorio circoscritto, ma dei quali nessuno sia direttamente o indirettamente attribuibile ad un determinato individuo: in tal caso, comunque, si sarà profilato quel determinato territorio circoscritto, potendo il titolare erogare annunci di *marketing* o servizi ritagliati su misura del territorio, emergendo in tal modo problematiche afferenti più ad una tutela collettiva dei dati personali e non solo individuale.

Per quanto riguarda invece la de-anonimizzazione, tale pericolo "*si concretizza nelle inferenze che possono essere tratte su gruppi o individui da dati anonimi, grazie anche alla disponibilità di dati ausiliari riferibili alle persone: di fatto, nessun dato è totalmente anonimo e può finire per essere identificativo, quindi personale, e, come tale, esigere l'applicazione della relativa disciplina*"[315]. In altre parole, a causa della possibile riconversione del dato anonimo in personale[316], il principio del Regolamento per il quale

---

individuale). Ancora una volta queste predizioni sono possibili grazie alle analisi degli algoritmi sull'intero universo dei dati (e non solo su campioni).

In altre parole, per aversi la ricostruzione è necessario che l'utente generi un qualsiasi tipo di dati (anche non personali). Questi ultimi dati vengono aggregati in banche dati (in genere semi-strutturate) e le tecniche algoritmiche di *Big Data Analytics* individuano a quale "ideal- tipo" è ricondotto quell'individuo (ottenendo quindi un profilo caratterizzato da dati personali), i cui dati originari potevano essere anche non personali. È questa la "ricostruzione" del dato personale che, come sottolineano gli autori ci fa riflettere su alcune criticità: "*Peraltro, paradossalmente, le predizioni sono molto più esaustive delle informazioni consapevolmente lasciate dagli utenti. In altre parole, i modelli di Big Data Analytics permettono di 'ricostruire' dati personali, indipendentemente dal loro originario rilascio, rendendo del tutto superata la tradizionale classificazione tra dati personali e dati non personali. Ma anche quella tra dati strutturati e non strutturati ai fini dell'efficacia della profilazione. Una circostanza che deve farci riflettere circa l'efficacia di approcci segmentati (di tutela della privacy o di governance pubblica e privata nella gestione del dato) se ancorati a distinzioni nominalistiche*". In tal senso, op. ult. cit., p.36.

[315] F. Faini, *Big Data e internet of things: data protection e data governance alla luce del regolamento europeo*, in G. Cassano, V. Colarocco, G.B. Gallus, F.P. Micozzi (a cura di), *Il processo di adeguamento al GDPR*, cit., p.317.

[316] Peraltro, tale pericolo sembra essere stato preso in considerazione dal legislatore del Regolamento. Ad esempio, nel Considerando n. 26 viene sottolineata l'importanza di aver riguardo alle tecnologie e agli altri dati a disposizione del titolare del trattamento, per distinguere

i dati anonimi sono esentati dalla disciplina del GDPR si rivela di non facile applicazione[317].

---

tra dato anonimo e non. Infatti, il Considerando chiarisce "*Per accertare la ragionevole probabilità di utilizzo dei mezzi per identificare la persona fisica, si dovrebbe prendere in considerazione l'insieme dei fattori obiettivi, tra cui i costi e il tempo necessario per l'identificazione, tenendo conto sia delle tecnologie disponibili al momento del trattamento, sia degli sviluppi tecnologici. I principi di protezione dei dati non dovrebbero pertanto applicarsi a informazioni anonime, vale a dire informazioni che non si riferiscono a una persona fisica identificata o identificabile o a dati personali resi sufficientemente anonimi da impedire o da non consentire più l'identificazione dell'interessato*". Si noti inoltre che la pseudonimizzazione non comporta l'anonimizzazione del dato, infatti con la prima è sempre possibile attribuire i dati pseudonimizzati ad un determinato interessato (con l'ausilio di informazioni aggiuntive a disposizione del titolare e tenute separatamente), mentre i dati anonimi dovrebbero essere non riferibili in alcun modo all'interessato. Qui non si vuole parlare della pseudonimizzazione, accorgimento tecnico di notevole rilievo (specie in caso di *data breach* e di *data leak*), ma di una vera e proprio de- anonimizzazione del dato. Anche nelle "*Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data*", adottate il 23 gennaio 2017 dal Comitato della Convenzione sulla protezione delle persone rispetto al trattamento automatizzato di dati a carattere personale (Convenzione 108) viene preso in considerazione l'aspetto problematico dell'anonimizzazione, sottolineando che la prova sull' adeguatezza delle misure di anonimizzazione spetta ai titolari, in ottemperanza al principio di *accountability*. L'importanza della questione è presa in considerazione anche dal Considerando n. 9 del REGOLAMENTO (UE) 2018/1807 (RDNP), dispone che: "*Se i progressi tecnologici consentono di trasformare dati anonimizzati in dati personali, tali dati sono trattati come dati personali e si applica di conseguenza il regolamento (UE) 2016/679"*.

Si noti che se il titolare rendesse anonimo il dato, potrebbe utilizzare tecniche di *Big data analytics* e *machine learning* senza preoccuparsi della normativa europea. Se lo stesso titolare riuscisse a risalire ai dati in precedenza anonimizzati si avrebbe un aggiramento della normativa.

[317] Peraltro, lo stesso rischio sin qui esaminato, che comporta una sfumatura della differenza tra, da un lato, dati anonimi e non personali e, dall'altro lato, dati personali, si pone anche con riferimento alla distinzione tra dati sensibili e comuni, poiché i primi possono essere ricavati combinando tra loro dati comuni. Come viene sottolineato in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.53, tale rischio è stato evidenziato anche dalla risoluzione del 14 marzo 2017 adottata dal Parlamento europeo.

Per tutti questi motivi in dottrina è stata avanzata l'idea di "*A risk based approach* [omissis] *in which different degrees of re-identifiability are recognized*", a differenza di quello individuato dal GDPR, che può essere definito come "*binary or 'black and white' approach"*, in quanto focalizzato su una distinzione netta tra dato personale e non personale; in tal senso e per un approfondimento si veda J. Bholasing, *How Technological Advances in the Big Data Era Make it Impossible to Define the 'Personal' in GDPR's 'Personal Data'*, in *European Data Protection Law Review*, 3/2022, pp. 346-361.

Alla luce delle considerazioni sin ora svolte, ed al di là delle garanzie tecniche ed obbligatorie che devono essere prese in considerazione quando si utilizzano *Big Data* ed algoritmi[318], occorre quindi valutare come i principi analizzati siano di ostacolo per la creazione di banche dati capaci di conservare massivamente dati personali da parte delle pubbliche amministrazioni. Invero se la progettazione di una tale banca dati sembra apparire più difficoltosa per i soggetti pubblici (seppur con le agevolazioni nel caso di trattamenti di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici o per le motivazioni di interesse pubblico rilevante) rispetto ai soggetti privati, rimane pur sempre percorribile senza troppe difficoltà, soprattutto alla luce delle modifiche intervenute con il c.d. decreto capienze (ponendo a parte il discorso sull'alfabetizzazione digitale e sulle competenze digitali della PA).

Per quanto riguarda invece l'utilizzo su dati personali di tecniche di *Big Data Analytics* e di *machine learning*, esse presentano in radice delle problematiche che si scontrano con alcuni dei principi fondanti del GDPR, primo fra tutti, il principio di limitazione delle finalità. Se aggiungiamo anche il rischio di "ricostruzione" del dato personale, si comprende perché si parla di contrapposizione "ontologica" o di "paralizzazione"[319] della *privacy*.

---

[318] Ci limitiamo a segnalare che, ai sensi dell'articolo 37 del GDPR, per la natura di tali trattamenti, che utilizzano *Big Data* ed algoritmi (che quindi spesso consistono nel "*monitoraggio regolare e sistematico degli interessati su larga scala*") e per i soggetti pubblici (autorità pubblica o organismo pubblico) la nomina del "*responsabile della protezione dei dati*" (RPD o DPO) sarà obbligatoria. Un'altra garanzia rilevante per queste tipologie di trattamenti, la cui importanza viene sottolineata dal Garante della privacy in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.66, è la valutazione d'impatto sulla protezione dei dati (o DPIA – *Data Protection Impact Assessment*) di cui all'art. 35 del GDPR, che consiste in un documento che valuta, anche sulla base del registro dei trattamenti, il rischio nei confronti degli interessati causato dai trattamenti che svolge. I trattamenti automatizzati, compresa la profilazione, comportano l'obbligo di tale documento. La dottrina sottolinea che anche i principi di *privacy by design* e di *privacy by default* (art. 25 GDPR, primo e secondo paragrafo) aiutano a delineare una gestione giuridica dei *Big Data*, tuttavia, non vengono meno aspetti problematici "ontologici" poc'anzi esposti. F. Faini, *Big Data e internet of things: data protection e data governance alla luce del regolamento europeo*, in G. Cassano, V. Colarocco, G.B. Gallus, F.P. Micozzi (a cura di), *Il processo di adeguamento al GDPR*, cit., p.313.

[319] Termine presente in V.M. Schönberger, K. Cukier, *Big Data*, cit., p.206-212.

Tali questioni risultano di ostacolo sia per il settore pubblico che per quello privato, delineando, sotto alcuni aspetti, una vera e propria crisi della normativa e rendendo difficile l'attuazione di trattamenti che siano effettivamente GDPR *compliant*. Tutte queste motivazioni rendono ostica una soluzione da parte di un'amministrazione basata su *Big Data* ed analisi algoritmica dei dati personali.

3. Le soluzioni tecnologiche per gestire i *Big Data* e le loro relazioni con l'interoperabilità ed il *cloud*

Nel primo capitolo si è visto di come l'organizzazione delle banche dati e dei *data center* delle pubbliche amministrazioni stia attraversando un processo di razionalizzazione e di uniformazione del patrimonio informativo, specialmente grazie all'interoperabilità ed al *cloud*.

In questo paragrafo si tenterà di analizzare invece quali siano le tecnologie adatte per sfruttare i *Big Data* (oltre alle criticità degli strumenti informatici adoperati dalle pubbliche amministrazioni) e si concentrerà l'attenzione sulla relazione tra, da un lato, interoperabilità e *cloud* e, dall'altro, i *Big Data*.

L'insufficienza degli attuali strumenti informatici per gestire le potenzialità dei *Big Data* è legata a tutte le fasi della gestione dei dati: raccolta, conservazione ed elaborazione[320]. Questa incapacità appare correlata soprattutto alla difficoltà per i *database* relazionali di gestire dati di tipo non strutturato[321], che rappresentano l'ottanta per cento dei dati[322] e di cui probabilmente vi sarà ancora un aumento con il proliferare dell'Iot[323].

---

[320] Tale aspetto è sottolineato in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.7, dove si afferma che "*per Big Data si intende una collezione di dati che non può essere acquisita, gestita ed elaborata da strumenti informatici, da software e da hardware "tradizionali" in un tempo tollerabile*". È sottolineato altresì in M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data*, cit., p.242.

[321] I dati di tipo strutturato possono essere ordinati e facilmente elaborati in *database* relazionali. In tal senso D. Iacovelli, M. Fontana, *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale,* cit., p.128.

[322] D.E. Holmes, *Big Data. A very short introduction*, Oxford, 2017, p.5 ss.

[323] P. Azad, N.J. Navimipour, A.M. Rahmani et al., *The role of structured and unstructured data managing mechanisms in the Internet of things. Cluster Comput* 23, 1185–1198 (2020), ove si

I *database* relazionali sono strumenti tradizionali di immagazzinamento dati, che archiviano dati pre-elaborati in colonne e tabelle[324]. Tale banca dati non è idonea ad immagazzinare e ad elaborare i *Big Data*. Per unificare, valutare e gestire un'enorme quantità di dati eterogenei serve un *repository*[325] idoneo affinché l'analisi e la *query*[326]

---

legge che "*The rise of the IoT will change traditional database systems to adapt to huge volumes of unstructured and semi-structured data from different sources*".

[324] I *database* relazionali presentano una struttura a righe, colonne e tabelle. Stabiliscono inoltre le relazioni tra tabelle ed tra i campi di una singola tabella. Quest'architettura delinea lo schema del *database*, di cui si è accennato nel primo capitolo. Invece i *database* non relazionali si strutturano sovente in base ad una relazione chiave/valore. In tal senso M. Aldinucci, *La pubblica amministrazione con i Big Data*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* cit., e G. Carullo, *gestione, fruizione e diffusione dei dati*, cit., p.70.

[325] Termine utilizzato per indicare il contenitore dei dati e che potrebbe essere tradotto con archivio, deposito o magazzino.

[326] La *query* è una modalità di interrogazione del *database*, attraverso la quale è possibile visualizzare, recuperare ed anche modificare i dati (eliminazione, inserimento ed aggiornamento).

Tra i linguaggi di query più utilizzati vi è SQL (*Structured Query Language*), che permette di gestire i dati contenuti nei database relazionali, che archiviano dati strutturati.

Tuttavia, sempre a causa dell'aumento esponenziale della quantità e dell'eterogeneità dei dati che si è avuta nel corso degli anni, le tecnologie applicate ai *database* hanno subito variazioni considerevoli ed un'altra tipologia di database è cresciuta esponenzialmente, i c.d. *databse* NoSQL (ossia "*not only SQL*"). In tal senso viene chiarito in W. Khan, T. Kumar, C. Zhang, K. Raj, A.M. Roy, B. Luo, *SQL and NoSQL Database Software Architecture Performance Analysis and Assessments—A Systematic Literature Review.,* in *Big Data and Cognitive Computing 2*, n. 7.97, 2023. Gli autori sottolineano che la scelta di adottare un database SQL o NoSQL dipende dalle esigenze che un'organizzazione vuole soddisfare, ma che n linea generale "*an organization can utilize an SQL database if it places a priority on data standardization and consistency. NoSQL should be utilized when a business has a large quantity of unstructured data and data availability is a high requirement. A relational database may be preferred over a NoSQL database for the aggregation of small datasets, and vice versa for big data analytics*". Quindi sottolineano che la letteratura scientifica solitamente evidenzia che "*NoSQL databases, with their specifically tailored structures, may be the best option for big data analytics*". Gli autori sottolineano inoltre la difficoltà di rendere tra loro interoperabili i database NoSQL e di permettere la portabilità dei dati, poiché attualmente i *cloud service provider* (che possiedono e mettono a disposizione degli utenti i database NoSQL) utilizzano protocolli ed interfacce incompatibili tra di loro. Inoltre, altri autori, efficientemente chiariscono che "*Relation database can only work with structured data, so there is a need of NoSql Database management system, which can work with no-structured data*". In tal senso R. Čerešňák, M. Kvet, *Comparison of query performance in relational a non-*

siano efficienti. Per affrontare tale problematica solitamente si fa riferimento ai concetti di *data lake* o di *data warehouse*[327].

Per *data warehouse* s'intende un *repository* che memorizza dati strutturati, filtrati e pre-elaborati, che sono stati trattati per uno scopo specifico ai fini dell'analisi. Invece il *data lake* è un *repository* centralizzato per i dati grezzi (c.d. *raw data*, ossia dati non strutturati, semi-strutturati o strutturati), non necessariamente pre-elaborati in precedenza ed il cui scopo di utilizzo non è definito e che, quindi, potenzialmente, potrebbero essere utilizzati per qualsiasi finalità[328].

---

relation databases, Transportation Research Procedia, vol. 40, Elsevier BV, 2019, pp. 170–177. Crossref.

In altre parole, un *databse* NoSQL sarà preferito per l'analisi dei *Big Data*, dato il volume e la varietà dei dati, mentre sarà preferito un database SQL in presenza di dati strutturati. Si ricordi che la *query* è sempre un modo di interrogare i dati e, potenzialmente, potrebbe essere utilizzata anche in un *data lake*.

[327] "*In broad terms, storage solutions in a big data context can be classified as either data warehouses or data lakes*". In tal senso PwC EU Services, D05.02 *Big Data Interoperability Analysis*, p.30.

[328] Si è voluto delineare solamente i tratti salienti di queste due metodologie di archiviazione dei dati, alternative al database tradizionale. Ma vi sono altri aspetti caratterizzanti da sottolineare. Il data *warehouse* è infatti capace di archiviare grandi quantità di dati, ma in genere utilizza schemi predefiniti come i *database* relazionali. Per poter inserire i dati nell'interfaccia del *database*, è necessaria una fase di pre-elaborazione dei dati, in cui i *set* di dati vengono puliti, filtrati e strutturati preventivamente. Queste tecniche di pre-elaborazione vengono chiamate strumenti di ETL (*Extract/Transform/Load*), e consistono proprio nel processo di raccolta dei dati da un numero illimitato di sorgenti e nella loro successiva organizzazione e centralizzazione in un unico *repository*. Effettuando la pre-elaborazione la qualità dei dati tende ad essere affidabile. Il fine di tale *repository* è specialmente quello di essere d'ausilio nei processi decisionali attraverso l'analisi e il *data mining*.

Invece il *data lake* differsice sotto molteplici aspetti. L'aspetto principale è che esso può archiviare *raw data* senza organizzarli in uno schema fisso (vi è un'assenza di uno schema definito *ex- ante*, come per i *database*) e senza che essi siano stati precedentemente pre-elaborati ("lago" di dati). Quindi sarà possibile prima memorizzare i dati e successivamente elaborarli. Data l'eterogeneità dei dati è essenziale che vi sia un catalogo di metadati associati ai *raw data*, per abilitare le funzionalità del *data lake*. I *raw data* comportano l'ulteriore problematica di essere di difficile utilizzo, perciò, in genere, sarà necessaria una figura professionale specifica, il *data scientist*, per utilizzare tali dati (quest'ultimo dovrà recuperare i dati dal *data lake* – pescare i dati – ed elaborarli per il caso d'uso). Il *data lake* fornisce un'interfaccia di accesso comune a tutti i dati grezzi (essi potrebbero anche confluire in un *data warehouse* per poi essere analizzati meglio). Come viene chiarito in L.Annet, D.Laurent,

Il *datawarehouse* è una soluzione efficiente per immagazzinare ed analizzare grandi quantità di dati, ma si riscontrano difficoltà con dati non strutturati, non pre-elaborati.
Si è fatta sentire quindi la necessità di disporre di soluzioni migliori per archiviare e analizzare grandi quantità di dati semi-strutturati e non strutturati, al fine di ottenere informazioni rilevanti ed analisi efficienti. Gli approcci tradizionali di ETL sono troppo inefficienti per gestire tali tipologie di dati e ciò ha consentito lo sviluppo del *data lake*[329].
Ed è proprio il *data lake* il paradigma tecnologico attuale, quale strumento per sfruttare le potenzialità dei *Big Data* e delle analisi algoritmiche, che semplifica le fasi di raccolta e (soprattutto) di conservazione dei dati.
Per questi motivi, occorre interrogarsi se il settore pubblico stia predisponendo modifiche a livello organizzativo che vadano in tale direzione. Nel proseguo analizzeremo

---

C.Madera, *Data lakes*, Vol.2, ISTE Ltd, London, 2020, il concetto di *data lake* è stato introdotto per la prima volta nel 2010 da James Dixon, CTO di Penthao, in un blog , in cui si aspettava che i *data lake* sarebbero stati enormi insiemi di dati in fila, strutturati o meno, a cui gli utenti avrebbero potuto accedere per scopi di campionamento, estrazione o analisi. Inoltre, come viene chiarito in A.Nambiar, D. Mundra, *An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management*, in *Big Data and Cognitive Computing*, 2022,n. 6, pp. 132, sempre secondo Dixon *"whilst a data warehouse seems to be a bottle of water cleaned and ready for consumption, then "Data Lake" is considered as a whole lake of data in a more natural state*".
I dati presenti nel *data lake*, oltre ad essere vari, non hanno una struttura organizzativa specifica e sono selezionati ed organizzati come e quando necessario. I *data lake* sono inoltre ideali per l'utilizzo di tecniche di *machine learning* e per l'analisi dei dati in tempo reale.
Per un analisi più approfondita sui *data lakes* si rinvia a L.Annet, D.Laurent, C.Madera, *Data lakes*, op.cit. Invece, per una disamina più approfondita in merito alla differenza tra *data lake* e *data warehouse* si rinvia a A.Nambiar, D. Mundra, *An Overview of Data Warehouse and Data Lake, op.cit.* Infine, per un approfondimento sul processo di ETL, si rinvia a A.A. Yulianto, *Extract Transform Load (ETL) Process in Distributed Database Academic Data Warehouse*, in *APTIKOM Journal on Computer Science and Information Technologies*, vol. 4, no. 2, Institute of Advanced Engineering and Science, July 2019, pp. 64-71.
Ai nostri fini interessa in particolare sottolineare che la tecnologia del *data lake* sta emergendo grazie alla capacità di gestire *Big Data*, che i dati in esso contenuti sono grezzi e non sempre archiviati per una finalità specifica e che esso funge da interfaccia comune per tutti i dati archiviati.

[329] In A.Nambiar, D. Mundra, *An Overview of Data Warehouse and Data Lake, op.cit.,* viene affermato che "*Traditional schema-on-write approaches such as the extract, transform, and load (ETL) process are too inefficient for such data management requirements. This gave rise to another popular modern enterprise data management scheme known as data lakes*".

i rapporti dell'interoperabilità e del *cloud*, per esaminare se siano capaci di sfruttare le potenzialità connesse ai *Big Data*, specialmente con riferimento alle fasi di raccolta e conservazione del dato.

Se l'interoperabilità rappresenta una sfida complessa già di per sé, in un contesto *Big Data* le sfide si accentuano ancor di più, e sono legate alle difficoltà di rendere interoperabili fonti eterogenee di dati, che spesso hanno bisogno di essere processati anche in *real-time*[330] (e che quindi difficilmente possono essere elaborati)[331]. Vengono intaccati tutti e quattro i livelli di interoperabilità previsti dall'*European Interoperability Framework*[332]: legale, organizzativa, semantica e tecnica. Ad esempio, la diversità della tipologia e delle strutture dei dati (varietà) incrementa la difficoltà di integrazione tra i dati, danneggiando l'interoperabilità semantica e tecnica. Allo stesso modo la necessità di effettuare analisi in *real-time* (velocità), sottrae tempo per elaborare i dati dopo che sono stati acquisiti, intaccando l'interoperabilità tecnica. O, ancora, vi possono essere conflitti a livello di schema (varietà - ad esempio vengono usati nomi diversi per lo stesso concetto)[333], intaccando quindi l'interoperabilità semantica. In realtà, se venisse

[330] Si veda su tale tema D05.02 *Big Data Interoperability Analysis* redatto dai servizi europei di PwC. In alcuni casi si ha bisogno di processare i dati in tempo reale, ad esempio nel caso in cui si voglia monitorare l'andamento del traffico, monitorare le reti elettriche o identificare le frodi.

[331] I dati elaborati (quindi ripuliti o che sono stati immagazzinati in seguito ad ETL) possono essere riconosciuti più facilmente da un'altra unità funzionale che adotta categorie simili per ordinare i dati. Se i dati non sono stati elaborati e categorizzati, sarà difficile per un'altra unità funzionale leggere quel dato. Le esigenze di *real- time* accentuano tale problema, perché i dati, una volta acquisiti e non pre-elaborati, dovranno essere analizzati immediatamente, senza la possibilità di un'eventuale elaborazione successiva all'acquisizione.

[332] PwC EU Services, D05.02 *Big Data Interoperability Analysis*.

[333] Quando si integrano dati da diverse fonti, ogni fonte potrebbe offrire una descrizione diversa dello stesso concetto. Ad esempio, un *set* di dati potrebbe riferirsi ad un'opera di "scrittore A" come "romanzo", mentre un altro potrebbe definire tale opera come dell'"autore B" e come "libro". Tuttavia, "scrittore A" e "autore B" potrebbero in realtà essere la stessa persona e le opere, sebbene classificate come "romanzo" e "libro", potrebbero coincidere. In tal senso PwC EU Services, D05.02 Big Data Interoperability Analysis, p.19. A tal proposito distinguiamo i concetti di "formato interoperabile" o "interoperabilità dei dati" con quello di "interoperabilità" in senso stretto. Mentre il primo rappresenta il modello minimo dei dati in modo che siano scambiabili (come lo stesso formato strutturato), il secondo attiene alla capacità del Sistema di comunicare, di eseguire programmi o di trasferire dati ad un'altra unità funzionale, senza la necessità di conoscere le caratteristiche di ogni unità funzionale; si veda a tal proposito *European*

attribuito un maggior peso alla veridicità, anche a scapito del volume e della velocità (quando possibile), i problemi relativi all'interoperabilità diminuirebbero in modo considerevole (ed infatti la veridicità sta assumendo sempre più importanza nei contesti *Big Data*). Si potrebbe ritenere anche che una maggiore interoperabilità vada anche a favore di una raccolta voluminosa di dati di qualità (poiché i dati potranno essere trasferiti da un'unità funzionale ad un'altra), quindi gli sforzi fatti in tal senso non possono che giovare ad un ampio patrimonio informativo, che sia anche di qualità[334].

Il concetto di *data lake* è invece connesso alle fasi di raccolta e di conservazione dei dati. Esso potrebbe essere lo strumento ideale per sfruttare i *Big Data*, i sistemi di IA (*machine learning*) ed incrociare basi di dati (come l'anagrafe tributaria con le altre banche dati dell'agenzia delle entrate e con i dati provenienti dai *social network*)[335]. Quindi all'interno delle infrastrutture pubbliche bisognerebbe implementare anche un *data lake*, in cui confluiscano i dati delle differenti banche dati delle altre amministrazioni pubbliche (che potrebbero essere medie o piccole amministrazioni), che sono caratterizzate più da sistema di raccolta e conservazione tradizionale[336].

Commission, *GDPR Data Portability and Core Vocabularies*, 2018 e D. C. Cravo, *How to Make Data Portability Right More Meaningful for Data Subjects?,* in *European Data Protection Law Review*, cit., p.56.

[334] Ad esempio, in M. Aldinucci, *La pubblica amministrazione con i Big Data,* in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p.225, viene affermato che "*La struttura a silos su cui si basa il sistema informativo della pubblica amministrazione italiana risulta inadeguato all'approccio della big data analysis che, a contrario, richiede la piena comunicabilità tra sistemi informativi affinché il reperimento dei dati su cui condurre sperimentazioni sia quanto più facile e mirato*". Invece, per un approfondimento circa la difficoltà di mantenere i livelli di interoperabilità in un contesto *Big Data* si veda PwC EU Services, D05.02 *Big Data Interoperability Analysis*.

[335] M. Aldinucci, *La pubblica amministrazione con i Big Data,* in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,* cit., sottolinea l'importanza del *data lake* e della sua importanza per incrociare basi di dati (il cui incrocio, altrimenti, sarebbe molto più difficoltoso e consisterebbe sostanzialmente nella creazione di un altro database che funga da "ponte" tra i *database* che si vuole incrociare).

[336] Sull'importanza e la necessità di dotare l'amministrazione di un'organizzazione basata su sistemi OLAP (*On Line Analytics Processing* - quale il *data lake*) e sistemi OLTP (*On Line Transaction Processing* – quali i tradizionali *database*) connessi e cooperanti si veda M. Aldinucci, *La pubblica amministrazione con i Big Data,* in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p.227. L'autore afferma

Per verificare se la pubblica amministrazione italiana abbia effettivamente implementato un sistema basato su un *data lake* è possibile prendere ad esame il progetto del *Data & Analytics Framework* (DAF) e la sua evoluzione nell'infrastruttura interoperabilità PDND. Il DAF era stato previsto dal piano triennale per l'informatica 2017-2019 e l'idea era quella di aprire il mondo della pubblica amministrazione ai benefici offerti dalle moderne piattaforme per la gestione e l'analisi dei *Big Data*.

Per gestire tale progetto era stato nominato nel 2016 un Commissario straordinario per l'attuazione dell'Agenda digitale[337]. La struttura commissariale prevedeva un Team per la Trasformazione Digitale[338], al cui interno era stato costituito il *Big Data Team* della PA (BDT-PA), composto da specialisti (*data scientist* e *big data architect*)[339].

---

infatti "*Parrebbe quindi che il paradigma data lake sia la soluzione adatta per l'organizzazione dei dati della Pubblica Amministrazione. Purtroppo la risposta non è così semplice. Un data lake è tecnicamente un sistema OLAP (On Line Analytics Processing) che è adatto a cercare nuove relazioni fra i dati ma è totalmente inadeguato a mantenere la consistenza e la correttezza dei dati nel tempo per cui i tradizionali data silos (tecnicamente detti OLTP - On Line Transaction Processing) basati su schemi sono ancora l'unica soluzione. Di fatto oggi è necessario utilizzare insieme data silos e data lake*". Bisogna precisare inoltre che con i termini OLAP e OLTP ci si riferisce a differenti ai metodi di elaborazione dei dati. Il sistema OLTP si occupa di gestire principalmente le transizioni ordinarie ("*transaction*") e le operazioni giornaliere, quali le classiche transizioni finanziarie come il pagamento con bancomat, attraverso l'inserimento, la modifica e la cancellazione dei dati. È progettato per elaborare le transazioni recenti in modo rapido ed accurato, il suo scopo principale, quindi, consiste proprio nell'elaborare le transazioni di *routine*. Il sistema OLAP permette di eseguire un'analisi su un archivio di dati centralizzato e voluminoso, come un *data lake* o una *data warehouse*. L'analisi aggregata dei dati contenuti all'interno di questi *repository*, permette di ottenere *report* e tendenze, che possono essere di supporto per le decisioni (a differenza dei sistemi OLTP, che sono sistemi operativi). Dato che i due sistemi hanno funzioni diverse, un'organizzazione potrebbe usare entrambi questi sistemi. Infatti si potrebbero inviare i dati relativi alle transazioni ordinarie (gestiti da sistemi OLTP con database relazionali) verso i sistemi OLAP per effettuare un'analisi più complessa.
Tali sistemi sono riportati in questi termini in G. Satyanarayana Reddy et. al., in *IJCSE - International Journal on Computer Science and Engineering*, Vol. 02, n. 9, 2010, pp. 2865-2873 e sul sito di IBM https://www.ibm.com/it-it/topics/oltp.

[337] Istituito con decreto del Presidente del Consiglio dei ministri (Dpcm) del 16 settembre 2016.

[338] Così come risulta dalla pagina della Presidenza del Consiglio dei ministri https://teamdigitale.governo.it/it/chi-siamo.

[339] Come si evince dalla piattaforma NoiPA, realizzata dal Dipartimento dell'Amministrazione Generale del personale e dei servizi (DAG) del Ministero dell'Economia e delle Finanze (MEF), https://noipa.mef.gov.it/cl/web/guest/-/news-daf-l-infrastruttura-su-cui-corrono-i-dati-della-

Il sistema predisposto dal DAF, oltre a prevedere un gruppo di specialisti, si basava anche su una componente di architettura *Big Data*[340]. Il gruppo di specialisti aveva il compito, tra le altre cose, di analizzare i dati e sviluppare modelli di *machine learning*, mentre la componente di architettura *Big Data* aveva il compito di centralizzare e conservare i dati all'interno di un *data lake* e di processarli attraverso i *data engine*[341]. In tal modo i dati eterogeni (*raw data*) prodotti dalle pubbliche amministrazioni[342], venivano memorizzati nel *data lake* per essere poi armonizzati ed elaborati, sia in *batch*[343] che in *real- time*, dai *Big Data engine*, che implementavano anche modelli di *machine learning*.

Tale modello rimase invariato anche nel successivo piano triennale AgID 2019- 2021, sebbene il progetto avesse cambiato nome da DAF a PDND[344] ed il Team per la

---

pa#:~:text=Il%20Big%20Data%20Team%20della,analisi%20e%20sviluppo%20di%20applicazioni
.

[340] In tal senso viene chiarito dall'ex *Chief Data Officer* del *team* per la trasformazione digitale https://medium.com/team-per-la-trasformazione-digitale/dati-interoperabili-open-pubblica-amministrazione-data-analytics-framework-4eb53dafd618.

[341] Come riportato dall'ex *Chief Data Officer* la componente di architettura *Big Data* ha il compito di "*centralizzare e conservare (data lake), elaborare e armonizzare (data engine), e distribuire (layer di comunicazione) i dati, sia in modalità batch che stream real-time. Il tutto privilegiando, ove possibile, software open source*".

[342] Quali le basi di dati che le PA generano per svolgere il proprio mandato istituzionale, i dati generati dai sistemi informatici delle Pubbliche amministrazioni come *log* ed altri ed i dati autorizzati provenienti dal web e dai *social network* di potenziale interesse della Pubblica amministrazione. In tal senso quanto previsto dal piano triennale AgID 2017- 2019.

[343] "*Particolari procedure attivate da un operatore su masse di dati di grandi dimensioni*". In tal senso M. Iaselli, *Il contratto di outsourcing in ambito informatico*, *MediaLaws*, comments, 2011.

[344] Così come risulta chiaramente dal Piano 2019- 2021, infatti "*La Piattaforma Digitale Nazionale Dati (PDND), precedentemente nota con il nome di Data & Analytics Framework (DAF), fa parte delle attività atte a valorizzare il patrimonio informativo pubblico nazionale*" e l'idea era sempre quella di aprire il mondo della PA ai benefici offerti dalle piattaforme "*per la gestione e l'analisi dei big data*". Viene fatto sempre riferimento all'architettura *Big Data* presente anche nel DAF. Tali funzioni della PDND emergono anche dal "*Piano triennale- aggiornamento 2020- 2022", dove viene affermato che la PDND "e permette di valorizzare il patrimonio informativo pubblico attraverso l'introduzione di tecniche moderne di analisi di grandi quantità di dati (BigData)".* Inoltre il collegamento con i *Big Data* e la PDND persiste anche dal piano triennale 2020- 2022. Si può consultare a tal fine il portale di Docs Italia e quello dell'Agid per gli aggiornamenti https://docs.italia.it/italia/piano-triennale-ict/pianotriennale-ict-doc/it/2020 -2022/06_capitolo_3_piattaforme.html , https://www.agid.gov.it/it/agenzia/piano-triennale . Tra l'altro la configurazione della PDND quale *data lake* in cui conservare dati in massa poneva i

trasformazione digitale avesse concluso il suo mandato a fine 2019[345]. In altre parole, vi era un interesse concreto nello sviluppo di una piattaforma unica in cui confluissero i dati delle pubbliche amministrazioni, in un'ottica *Big Data*[346]. Tale progetto, oltre a rafforzare il potere conoscitivo della PA attraverso lo sfruttamento dei *Big Data*, aveva anche gli obiettivi (parimenti importanti) di sviluppare e semplificare l'interoperabilità dei dati pubblici tra PA e di standardizzare e promuovere la diffusione degli *open data*. E sembrerebbe che proprio verso queste ultime direzioni abbia virato il progetto. Ed infatti nel "*Piano triennale per l'informatica nella PA - Aggiornamento 2021 – 2023*" ci si riferisce alla PDND, quale piattaforma che "*permette di aprire canali tra le PA e, così, farle dialogare, realizzando l'interoperabilità, attraverso l'esposizione di API. La Piattaforma concretizza il principio "once-only" e in futuro, dovrà consentire anche l'analisi dei big data prodotti dalle amministrazioni, resi disponibili nel data lake, per l'elaborazione di politiche data-driven*", mentre nel "*Piano Triennale per l'informatica nella PA - Aggiornamento 2022 – 2024*" viene definita come una piattaforma che "*permette di autorizzare e autenticare le PA alla comunicazione tra i loro sistemi informativi e alla condivisione dei dati a loro disposizione, realizzando l'interoperabilità attraverso*

classici problemi legati alla tutela dei dati personali, che sono stati opportunamente sottolineati dal Garante della privacy in AGCM, AGCOM, GPDP, *Indagine conoscitiva sui Big Data*, p.68, ove si legge che "*La creazione della suddetta Piattaforma comporta un accentramento e una duplicazione di tutti i dati detenuti dalle pubbliche amministrazioni per finalità del tutto generiche, realizzando di fatto una concentrazione presso un unico soggetto di informazioni, anche sensibili e sensibilissime, con evidenti rischi di vulnerabilità dei dati stessi ovvero di possibili usi distorti*".

[345] Così come risulta dal portale del Governo. La struttura commissariale (e quindi anche il Team) ha infatti concluso il suo mandato il 31 dicembre 2019 e tali progetti sono confluiti nel Dipartimento per la trasformazione digitale (DTD) della Presidenza del Consiglio dei ministri (istituito con DPCM del 19 giugno 2019) e nella società PagoPA S.p.a. Dipartimento che attualmente non è più affidato ad un Ministro, ma ad un Segretario di Stato.

[346] La dottrina ha sottolineato l'importanza del DAF e dei primi progetti della PDND, quali sistemi per sfruttare le potenzialità dei *Big Data*. Sul punto si veda F. Costantino, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Diritto pubblico*, fasc.n. 1, 2019, pp.63-65 e F. Sarpi, *La regolazione di domani. Come adeguare il processo normativo alle sfide dell'innovazione*, in *Rivista Italiana di Politiche Pubbliche*, fasc. n. 3, 2018, p.464 ss. Inoltre, secondo l'autore Costantino, come sopramenzionato, la PDND, intesa quale struttura per lo sfruttamento dei *Big Data*, è prodromica all'utilizzo dell'IA.

l'esposizione di servizi digitali implementati dalle necessarie API. La Piattaforma contribuisce alla realizzazione del principio once only e in futuro, dovrà consentire anche l'accesso ai big data prodotti dalle amministrazioni l'elaborazione di politiche data-driven".

L'attenzione attuale nei confronti della PDND viene rivolta soprattutto al versante dell'interoperabilità[347] ai fini della realizzazione del principio del *once only*, mentre si afferma che solamente "in futuro"[348] consentirà di sfruttare le potenzialità connesse ai *Big Data*. Essa, secondo l'art. 50-*ter* CAD risulta infatti incentrata sull'autorizzazione dello scambio dati tra soggetto erogatore e fruitore, attraverso un sistema basato sul catalogo API e sui *voucher*, non rappresentando, quindi, un *repository* per i flussi eterogenei di dati delle pubbliche amministrazioni, quale potrebbe essere un *data lake*[349].

Per quanto riguarda il *cloud* invece, esso può rivestire un'importanza centrale per lo sfruttamento dei *Big Data*[350]. Si vedranno perciò le relazioni della Strategia *cloud* Italia e degli interventi europei in materia con il paradigma dei *Big Data*.

La strategia italiana è incentrata sul principio *cloud first* e sulla migrazione dei dati delle pubbliche amministrazioni in infrastrutture *cloud*. I CSP (*cloud services provider*) possono offrire sia servizi per immagazzinare *Big Data*, sia servizi per effettuare l'analisi e l'elaborazione dei dati. Non bisognerebbe però correre il rischio, pur di sfruttare le potenzialità dei *Big Data*, di delegare il controllo di queste infrastrutture completamente ai privati[351]. Ed il PSN forse potrebbe rispondere in parte anche a quest'esigenza[352]. Il

---

[347] Ed è indicativo il fatto che la PDND costituisca una parte importante del capitolo cinque sull' interoperabilità del Piano-aggiornamento 2021- 2023, mentre non viene menzionata nel capitolo relativo all'interoperabilità nel Piano- aggiornamento 2020- 2022.

[348] Nel piano triennale per l'informatica nella PA - Edizione 2024-2026 (versione accessibile), oltre a scomparire ogni riferimento in merito al data lake, non vi è menzione della possibilità di sfruttare i Big Data attraverso la PDND.

[349] Sull'art. 50-*ter* CAD e sul funzionamento della PDND si rimanda al secondo capitolo della presente analisi.

[350] A. Rezzani, *Big data. Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati*, cit., p.10

[351] M. Falcone, *Ripensare il potere conoscitivo pubblico tra algoritmi e big data* , cit., p.106

[352] Ed infatti il capitale sociale di PSN S.p.A allo stato attuale risulta così ripartito: TIM per il 45 %, Leonardo per il 25%, CDP Equity per il 20% e Sogei per il 10%. In virtù delle partecipazioni dirette ed indirette del MEF in queste società, risulta che il MEF ha un controllo o un'influenza

template del piano dei fabbisogni scaricabile sul portale del PSN, tra l'altro, prevede uno specifico servizio relativo ai *Big Data* che "*consente la costruzione di Data Lake as a service, servizi di analisi dati batch, stream e real-time con scalabilità orizzontale*". Il PSN quindi, attraverso una soluzione *Platform as a service* (Paas), fornisce il *data lake* come una piattaforma "*pronta all'uso che dispone di tutte le funzionalità necessarie a sviluppatori, Data Scientist e analisti per archiviare facilmente dati di tutte le dimensioni, forme e velocità*"[353]. Anche il *batch* ed il *real- time processing* costituiscono una soluzione PaaS che fornisce una piattaforma pronta all'uso ed una soluzione che consente "*l'analisi di grandi moli di dati*"[354]. Sembrerebbe quindi che, attraverso una soluzione nazionale, si voglia offrire alle pubbliche amministrazioni i vantaggi di archiviare e di analizzare i *Big Data*. Una soluzione che ben si concilia con il rafforzamento di un potere conoscitivo secondo il paradigma della *data- driven administration*[355].

---

dominante su Polo Strategico Nazionale. Per approfondimenti si veda il portale del PSN https://www.polostrategiconazionale.it/chi-siamo/governance/societa-trasparente/#:~:text=In%20ottica%20di%20trasparenza%2C%20vogliamo,e%20Sogei%20per%20il%2010%25.

[353] All. A della Convenzione del 24 agosto 2022 avente ad oggetto la concessione del PSN.

[354] *Ibidem*.

[355] La pubblica amministrazione basata sui dati (*data driven public administration*) "*is envisaged as an all-encompassing public administration reform, fundamentally changing the way democratic systems engage with and learn about citizens*", come affermato in H. Broomfield, L. Reutter, *Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation*, in *Scandinavian Journal of Public Administration*, 25 (2), 2021, p. 75. Gli autori, analizzando le sfide poste per l'implementazione di una PA *data driven* (con un *focus* nel contesto norvegese), sottolineano le problematiche tradizionali a livello tecnico per creare tale modello "*Data-driven public administration is predominantly perceived by policymakers as a technical issue, requiring technical infrastructure to provide access to high volumes of good quality data. This manifests itself in the prioritisation of the material aspect of the task, such as the purchasing of cloud solutions and building national data and API catalogues and data lakes*" (op. ult. cit., p. 83). Vi è tuttavia da sottolineare che gli autori, attraverso un approccio critico, evidenziano la necessità di intervenire in molteplici settori ed a livello culturale per ottenere un tale cambiamento. Per ulteriori approfondimenti si rimanda a C. Van Ooijen, B. Ubaldi, B. Welby, *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*, OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris, 2019.

A livello europeo vi è da segnalare l'obiettivo, predisposto nella Comunicazione della Commissione *Bussola per il digitale 2030* e ribadito nella Decisione che istituisce il programma strategico per il decennio digitale 2030[356], che consiste nell'installare e distribuire all'interno dell'UE diecimila nodi periferici (*edge nodes*). Tale obiettivo, complementare alla strategia sul *cloud*, permette di elaborare in parte e di gestire i *Big Data* prodotti al margine della rete[357].

[356] Decisione (UE) 2022/2481 del Parlamento e del Consiglio del 14 dicembre 2022 che istituisce il programma strategico per il decennio digitale 2030. Tale decisione si pone in continuità con la comunicazione della Commissione e ne rappresenta la concretizzazione, basandosi su meccanismi di cooperazione e di vigilanza finalizzati al conseguimento dei quattro punti cardinali individuati nella comunicazione: competenze digitali, infrastrutture digitali, digitalizzazione delle imprese e digitalizzazione dei servizi pubblici (cfr. art. 3, primo e decimo Considerando della Decisione); in tal senso e per un approfondimento si veda L. Válková, *Bussola per il digitale 2030: il modello europeo per il decennio digitale*, in *Rivista di diritto internazionale privato e processuale*, 2/2021, pp.469-473 ed S. Bernasconi, *Il Programma strategico dell'Unione europea per il decennio digitale 2030*, in *Rivista di diritto internazionale privato e processuale*, 1/2023, pp. 253-255.

[357] Si fa riferimento alla tecnologia emergente dell'*edge computing*, che si sta sviluppando particolarmente in seguito all'effetto dirompente dello IoT e che ha trovato la sua definizione formale nel 2015 nel *white paper* "*Edge computing: Vision and challenges*", nel quale è definito come una tecnologia abilitante che consente di eseguire i calcoli ai margini della rete, sui dati a valle per conto dei servizi *cloud* e sui dati a monte per conto dei servizi IoT. In altre parole, l'*edge layer* (al cui interno si trovano gli *edge server*) funge da ponte tra i dispositivi IoT (che generano i dati e che vengono invati all'*edge layer*, dove potranno essere in parte elaborati) e l'infrastruttura *cloud* (dove viene gestito il calcolo e l'archiviazione globale/condivisa). L'*edge layer* elabora i dati inviati dai dispositivi IoT (*upstream*) ed invierà solamente i dati che non riesce totalmente ad analizzare al cloud, il quale restituirà i dati analizzati all'*edge layer* (*downstream*). I vantaggi della costituzione dell'*edge layer* derivano dalla riduzione della latenza nel calcolo e nell'utilizzo della rete (perché non viene usata la rete internet per arrivare al *cloud*) e da una maggiore sicurezza dei dati. Infatti, le criticità connesse all'utilizzo esclusivo del *cloud* derivano dal fatto che, sebbene essa sia in grando di gestire i *Big Data*, tale tecnologia fallisce quando l'infrastruttura è distante (quindi aumentano i costi di trasmissione), le fonti che generano i dati sono distribuite ed è necessario effettuare elaborazioni quasi in tempo reale. Per questo è necessario effettuare i calcoli ai margini della rete, in luoghi più vicini alle fonti di generazione dei dati. Per tutti questi motivi è possibile considerare le tecnologie dell'*edge* e del *cloud* quali tecnologie complementari, capaci di lavorare in modo sinergico. Per una spiegazione ulteriore dell'*edge computing* si veda K. A. Kumari, G.S.Sadasivam, D. Dharani, M. Niranjanamurthy, *Edge Computing Fundamentals, Advances and Applications*, 1st Edition, Boca Raton, 2021, pp.12-20. A tal fine gli autori propongono un chiaro esempio per comprendere le criticità del *cloud* e le connessioni con l'*edge*: "*Consider a smart building with video cameras to detect the movement*

Ed infatti si vuole dotare l'Europa di sistemi interconnessi, affidabili, interoperabili e sostenibili basati sul *cloud-to-edge* (infrastrutture, piattaforme e servizi). Tali infrastrutture e servizi serviranno anche come spazi comuni di dati e permetteranno una rapida crescita di tecnologie emergenti, quali IA e *Big data*[358]. Per monitorare l'implementazione degli *edge nodes* (ossia i nodi periferici) sono previsti quattro rapporti che controlleranno l'andamento degli obiettivi fissati nel decennio digitale con la Comunicazione *Bussola Digitale 2030*[359].

Attraverso questa strategia si vuole quindi potenziare sia il *cloud* che *l'edge computing*, riconoscendo tuttavia che "*se non saranno accompagnati da capacità di calcolo all'avanguardia, gli ecosistemi cloud e edge non potranno apportare pieni benefici alle imprese e alle amministrazioni pubbliche europee*"[360]. Le tecnologie dell'*edge* e del *cloud*, unitamente alla capacità di calcolo, saranno quindi essenziali per sfruttare in modo efficiente innovazioni quali, l'intelligenza artificiale, l'analisi con i *Big Data* e *l'internet of things*[361]. Inoltre sarà necessario sviluppare un mercato del *cloud* equo

---

of people. Video signal is continuously streamed to the cloud server. The motion detection application in the cloud server can detect features and store to the database. As large volume of video data is streamed continuously to the cloud, significant network bandwidth is consumed. The cloud server should also analyze the video footage from all the cameras simultaneously. These issues can be overcome if the motion detection application could be executed close to the cameras. Only the required clips can then be transferred to the cloud for further analysis and storage".

È proprio la consapevolezza della complementarità del *cloud* e dell'*edge*, che ha spinto la Commissione a sviluppare la distribuzione degli *edge nodes* (nodi periferici). Questi ultimi sono definiti dall'art. 2, paragrafo 1, n. 6, Decisione (UE) 2022/2481, come "*capacità distribuita in rete di elaborazione dei dati e situata in prossimità o in corrispondenza del terminale (endpoint) fisico dal quale sono generati, che offre capacità di calcolo e conservazione distribuiti per l'elaborazione di dati a bassa latenza*".

[358] Tali sforzi sono chiariti Work programme 2023-2024, annex I digital Europe, C(2023) 8620 final, p.29.

[359] A tal fine si segnala che il primo rapporto, che ha il compito di monitorare lo sviluppo dei diecimila nodi periferici, è già stato pubblicato. *Edge Deployment Data Report, Edge Observatory for Digital Decade - Edge Computing Nodes: Characterisation, Deployment Monitoring and Trajectories, September 2023*.

[360] COM(2021) 118 final, p.9.

[361] Nella brochure – *cloud and edge computing* viene riconosciuto che: "*Cloud and edge computing unlocks access to future and emerging technologies, such as artificial intelligence, the*

(permettendo il passaggio tra diversi fornitori di servizi cloud veloce, gratuito e tecnologicamente fluido, evitando l'effetto *lock-in*) e regole chiare sulla circolazione dei dati per garantire che circolino liberamente (a tal fine il GDPR ed il regolamento sulla libera circolazione dei dati non personali giocano un ruolo importante)[362], in modo da garantire flussi continui e sicuri tra i dispositivi, le infrastrutture *edge* e quelle *cloud*.

All'esito di dell'analisi svolta nel presente paragrafo si può dedurre che il quadro organizzativo delineato nel primo capitolo, incentrato sul *cloud* e sull'interoperabilità in un'ottica di una funzione amministrativa di intermediazione dei dati, è funzionale anche alla creazione di un'organizzazione capace di sfruttare le potenzialità dei *Big Data* ed il potere conoscitivo della pubblica amministrazione. In altre parole, l'organizzazione modellata in base alla funzione intermediaria del dato sembra rispondere anche alle esigenze di una funzione conoscitiva, basata sulle analisi algoritmiche dei *Big Data*.

All'esito di questo capitolo sono emerse principalmente le criticità del GDPR rispetto allo sfruttamento dei *Big Data*. Tuttavia, si è vista l'importanza relativa alla capacità di gestire ed ottenere queste tipologie di dati. Dopo aver delineato nel presente capitolo, in via astratta i profili critici, che tale cambiamento di paradigma comporta con l'assetto delineato nel primo e nel secondo capitolo, nel quarto capitolo si analizzerà un caso studio, in cui l'amministrazione si ritrova a dover gestire i *Big Data*[363] ed in cui il

---

Internet of Things and blockchain. It plays a key role in fostering a competitive and innovative European economy in the digital age." In diversi punti viene ripresa anche la connessione con i *Big Data* e con la *data analysis*. Anche in altri contesi viene ripetuto più volte che *l'edge computing*, in combinazione con IoT e *cloud*, offre all'Europa una nuova opportunità per soddisfare la domanda di infrastrutture di elaborazione dati di nuova generazione (si veda la *Declaration of the European Alliance for Industrial Data, Edge and Cloud* e quanto riportato sul portale della Commissione https://digital-strategy.ec.europa.eu/it/policies/iot-investing#ecl-inpage-lmrs16r4).

[362] Così come si riporta sul portale della Commissione europea https://digital-strategy.ec.europa.eu/it/policies/cloud-computing.

[363] La circostanza per cui l'amministrazione finanziaria deve gestire i *Big Data* emerge dalle audizioni del direttore dell'Agenzia delle Entrate svolte nel 2021 dal Senato e dalla commissione parlamentare di vigilanza sull'anagrafe tributaria, in cui si legge che "*L'Agenzia delle entrate è titolare di un ampio patrimonio informativo derivante da numerose banche dati di grandi dimensioni, eterogenee per struttura e contenuto, e soggette a forte dinamismo (big data)*"; in tal senso è riportato in Senato della Repubblica-VI Commissione Finanze e tesoro, *Audizione del*

potenziamento del potere conoscitivo della Pubblica Amministrazione potrebbe risultare un fattore cruciale per migliorare l'azione amministrativa, dovendo rispettare la problematica particolarmente rilevante della tutela dei dati personali.

Direttore dell'Agenzia delle entrate - Progetti di digitalizzazione e innovazione tecnologica del settore fiscale, Roma, 2021.

CAPITOLO IV

BIG DATA E FISCO

1. *Big Data* e Fisco

In un sistema ideale tutti i cittadini dovrebbero pagare i tributi per garantire il c.d. "*interesse fiscale*"[364] dello Stato, in modo tale da garantire il corretto funzionamento della macchina amministrativa, in un'ottica del *welfare state*[365]. In questo Stato ideale ogni cittadino dovrebbe adempiere spontaneamente al pagamento del tributo, senza troppi interventi da parte dell'amministrazione finanziaria. Tuttavia, è quasi impossibile (o forse impossibile) raggiungere una simile condizione idilliaca, per questo c'è bisogno di un complesso sistema ben congegnato che permetta ai contribuenti di pagare le tasse ed all'amministrazione finanziaria di riscuoterle. Per ottenere un simile obiettivo bisogna implementare molteplici ambiti del settore fiscale: la legislazione (in modo che sia chiara e corretta); la scelta dei contribuenti da sottoporre a controllo; la fase istruttoria e di controllo; le fasi dell'accertamento e della riscossione; potenziare i rapporti tra gli enti dell'amministrazione finanziaria e così via[366].

In particolare, la scelta dei contribuenti da controllare rappresenta una fase particolarmente delicata, in quanto non è possibile, per l'amministrazione finanziaria,

[364] La giurisprudenza costituzionale ha elaborato il concetto di "*interesse fiscale*", quale interesse collettivo all'acquisizione delle risorse tributarie e quale principio e valore primario dell'ordinamento (Corte Cost. 21 maggio 2001, n. 155; Corte Cost. 17 giugno 1975, n. 164).

[365] Tale concetto ricomprende tutti quei diritti sociali che lo Stato dovrebbe garantire ai cittadini, ricomprendendo gli ambiti della sanità, dell'istruzione, dei servizi sociali ed assistenziali, delle pensioni e della previdenza sociale. Ma il tributo è connesso in generale ai diritti della collettività, senza di essi, infatti, la macchina amministrativa non potrebbe erogare e garantire correttamente i servizi. È stato infatti affermato che i tributi raffigurano il "*costo dei diritti*". In tal senso S. Holmes, Cass R. Sustain, *Il costo dei diritti: perché la libertà dipende dalle tasse*, Bologna, Il Mulino, 2000.

[366] Si veda ad esempio S. Vaccari, *Funzione tributaria e diritto amministrativo*, in *Diritto pubblico*, 2/2022, pp. 493-545; M. Beghin, *Diritto tributario-Principi, istituti e strumenti per la tassazione della ricchezza*, Giappichelli, Torino, 2011.

sottoporre a controllo, anno dopo anno, tutti i contribuenti[367]. E questo è dovuto al principio di economicità: l'amministrazione finanziaria non possiede in concreto le risorse per controllare nella sostanza tutte le dichiarazioni, i controlli generalizzati possono avvenire solo con metodologie di controllo completamente automatizzate[368]. Si rende necessario allora selezionare i contribuenti da sottoporre a controllo e, a tal fine, si compie una complessa analisi in base alle macro-tipologie di contribuenti (per esempio grandi contribuenti, piccole e medie imprese, professionisti)[369] ed in base alle linee guida predisposte annualmente dall'Agenzia delle Entrate (AdE) e dalla Guardia di Finanza (GdF), che stilano delle apposite "liste selettive"[370], sulla base delle risultanze acquisite

[367] A. Carinci, T. Tassani, *Manuale di diritto tributario*, Giappichelli editore, Torino, VI ed., 2023, p.233. Tale attività di selezione, propedeutica ad un eventuale attività istruttoria e di accertamento, è particolarmente importante per evitare che le risorse dell'amministrazione si disperdano per effettuare controlli che risultano essere infruttuosi.

[368] Ibidem. Un esempio di controllo automatizzato è la liquidazione automatica dell'imposta (art. 36- *bis* D.P.R. 29 settembre 1973, n. 600 e art. 54- *bis* D.P.R. 26 ottobre 1972, n. 633), attraverso la quale si effettua una riliquidazione della dichiarazione, attraverso la correzione di eventuali errori formali o di calcolo che possono averla inficiata. Non è un controllo sostanziale, che richiede un dispendio di maggiori risorse dell'amministrazione finanziaria.

[369] Ad esempio, la circolare AdE n. 21/E del 20 giugno 2022 (avente ad oggetto gli *Indirizzi operativi e linee guida per il 2022 sulla prevenzione e contrasto all'evasione fiscale, nonché sulle attività relative al contenzioso tributario, alla consulenza e ai servizi ai contribuenti)* elenca strategie differenziate per eseguire la selezione dei contribuenti e per indirizzare i controlli in base alle categorie dei soggetti. Ad esempio, per la categoria "*imprese di medie e piccole dimensioni*", viene stabilito che "*l'attività di analisi del rischio e di selezione dei soggetti - ritenuti a maggiore pericolosità fiscale - dovrà essere condotta dalla Direzioni regionali concentrando le risorse e l'impegno nell'utilizzo delle nuove piattaforme di analisi avanzata dei dati, che consentono l'interoperabilità delle informazioni disponili e la personalizzazione dei dataset in base alle specificità dei singoli percorsi di analisi del rischio*", oppure che "*l'attività di analisi e selezione sarà indirizzata prioritariamente nei confronti di coloro che presentano volumi di ricavi o di fatturato non in linea con quanto dichiarato da soggetti che presentano la medesima struttura operativa*". Viene affermato di attribuire particolare rilevanza agli ISA (indici sintetici di affidabilità fiscale) per questa categoria di soggetti o di indirizzare la selezione nei confronti di quei soggetti che portano in detrazione costi non inerenti all'attività esercitata o che hanno un elevato importo di costi c.d. "residuali". Nella stessa maniera, altre circolari, relative ad altre annualità, prevedono differenti linee guida in base alle categorie di contribuenti (ad es. si veda circolare AdE dell'8 agosto 2019, n. 19).

[370] le liste selettive sono gli "*elenchi di contribuenti scelti, in base a particolari criteri e indici di pericolosità fiscale, dall'Amministrazione finanziaria per essere sottoposti al controllo*". In tal

grazie alle attività di analisi effettuate per mezzo di alcune banche dati, come l'anagrafe tributaria e l'anagrafe dei conti correnti[371]. Inoltre, in tale fase, particolare rilevo assume l'attività di "analisi del rischio fiscale"[372], che risente positivamente dell'impatto dell'interoperabilità tra le banche dati dell'anagrafe tributaria e dell'intelligenza artificiale[373].

---

senso il portale della rivista online dell'AdE https://www.fiscooggi.it/glossario/termine/liste-selettive.

[371] A. Carinci, T. Tassani, *Manuale di diritto tributario*, cit., p. 233. Tale fase potrebbe costituire un terreno fertile per lo sfruttamento delle potenzialità dei *Big Data*, in quanto rappresenta una fase preliminare all'istruttoria ed alla decisione amministrativa, è infatti puramente orientativa e di iniziativa: l'amministrazione finanziaria deve redigere delle liste selettive e scegliere quali contribuenti controllare e quest'attività, tradizionalmente, è considerata libera, nel senso di essere insuscettibile di incidere sulle posizioni soggettive dei contribuenti (a differenza, ad esempio, della fase decisoria o istruttoria, maggiormente vincolante per l'amministrazione). In quest'ultimo senso si veda F. Gallo, *Discrezionalità (voce), diritto tributario*, in *Enciclopedia del diritto*, agg. III, 1999, 9, p.539. Tuttavia, non vi è un'unica prospettiva, in merito si veda G. Vanz, *I principi della proporzionalità e ragionevolezza nelle attività conoscitive e di controllo dell'Amministrazione finanziaria*, in *Diritto e pratica tributaria*, 5/2017, pp.1912 ss.

[372] Secondo la definizione contenuta nell' "*Informativa sulla logica sottostante i modelli di analisi del rischio basati sui dati dell'archivio dei rapporti finanziari*", pubblicata sul portale dell'Agenzia delle Entrate (AdE), l'analisi del rischio ricomprende "*le tecniche, le procedure e gli strumenti informatici utilizzati per individuare i contribuenti che presentano un elevato rischio fiscale, inteso come il rischio di operare, o aver operato, in violazione di norme di natura tributaria ovvero in contrasto con i principi o con le finalità dell'ordinamento tributario; una volta individuate le posizioni fiscalmente rischiose, le stesse sono trasmesse alle articolazioni organizzative che si occupano dei controlli, che effettuano ulteriori approfondimenti e valutazioni al fine di individuare i soggetti nei cui confronti avviare un'attività istruttoria*". In un secondo momento, dopo un'analisi preliminare, le risorse dell'amministrazione finanziaria potranno quindi essere indirizzate per svolgere un'istruttoria mirata, attraverso i tradizionali poteri, quali gli accessi, ispezioni e verifiche.

[373] A. Bongi, *Ebook Intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, Wolters Kluwer Italia, Milano, 2023, p.7. L'autore fa inoltre presente che all'interno dell'organizzazione dell'Agenzia delle Entrate è presente il "*Settore Analisi del Rischio e Ricerche per la Tax Compliance*", in cui operano due dirigenti con funzioni di *data scientist*. Peraltro, gli strumenti della *data economy*, come sottolineato in A. Contrino, *Digitalizzazione dell'amministrazione finanziaria e attuazione del rapporto tributario: questioni aperte e ipotesi di lavoro nella prospettiva dei princìpi generali*, in *rivista di diritto tributario*, fasc.2, 2023, pp. 105 ss., possono avere un ruolo importante nel corso di tutta la fase attuativa

L'attività di selezione dei contribuenti e di analisi del rischio, come ogni altro obiettivo di politica fiscale, viene redatta in base agli atti di indirizzo del Ministero dell'Economia e delle Finanze (MEF) e segue quanto stabilito nelle convenzioni stipulate tra le Agenzie ed il MEF. Gli atti di indirizzo vengono emanati ogni triennio e assegnano le linee generali da seguire a tutte le Agenzie fiscali, le quali emanano circolari contenenti indirizzi operativi cui gli uffici periferici devono attenersi per i controlli da effettuare nell'anno[374]. Per analizzare le problematiche relative all'analisi dei dati ed allo sfruttamento delle banche dati può essere utile esaminare questi atti.

L'ultimo atto di indirizzo del MEF è relativo al triennio 2024-2026 ed è rivolto a ciascuna articolazione dell'amministrazione finanziaria[375], mentre l'ultima convenzione del MEF con l'AdE e l'Agenzia delle Entrate-Riscossione (AdER) è relativa al triennio 2023- 2025[376] ed è specifica per quanto riguarda i rapporti con queste due ultime agenzie. Vista la

---

del tributo (che ricomprende, ad esempio, anche i metodi di accertamento, la riscossione, il processo). Tuttavia, il naturale campo d'applicazione dello sfruttamento dei *Big Data* è rappresentato dalla fase istruttoria, condizionata maggiormente dalle banche dati, così come emerge dal decreto del 15 luglio 2021 del Ministro dell'Economia e delle finanze. In tale ultimo senso si veda S. A. Parente, *Le funzioni giurisdizionali e di accertamento tributario nell'epoca della data economy e dei sistemi di cloud computing: l'ausilio di intelligenze artificiali, big data e algoritmi informatici*, in *Euro-Balkan Law and Economics Review*, università degli studi di Bari Aldo Moro, n. 1, 2023. Ed infatti, in ambito internazionale, è proprio il campo dell'analisi del rischio fiscale ai fini della selezione e dell'adempimento spontaneo dei contribuenti ad essere maggiormente interessato dai sistemi moderni di analisi dei dati. Si veda A. Ribes Ribes, *la inteligencia artificial al servicio del «compliance tributario»*, in *Revista española de derecho financiero*, 2020, 125 ss. Peraltro, la tipologia di IA che si rivela più promettente per effettuare l'analisi del rischio fiscale da parte dell'Agenzia delle Entrate è quella "*cognitiva*", ossia quella utilizzata per l'analisi di grandi moli di dati per supportare processi decisionali (come le tecniche di *machine learning*); si rilevano utili anche le tecniche di *text mining* che rientrano nella "*IA di facilitazione*". In tal senso e per tale categorizzazione dell'IA si veda quanto affermato dal direttore centrale tecnologie e innovazione dell'Agenzia delle Entrate, G. Buono, *Rapporto 4/2022 – Intelligenza artificiale e amministrazioni centrali*, in *BioLaw Journal – Rivista di BioDiritto*, n. 1/2022, pp. 261-264.

[374] A. Carinci, T. Tassani, *manuale di diritto tributario*, cit., p.233.

[375] Il provvedimento, ossia l'"*Atto di indirizzo per il conseguimento degli obiettivi di Politica fiscale 2024-2026*" è reperibile sul portale del MEF.

[376] "*Convenzione con l'Agenzia delle Entrate - triennio 2023 – 2025*", disponibile sul portale del MEF.

coincidenza delle linee direttive impartite all'AdE nei due provvedimenti, si può prendere in esame solo la convenzione[377].

Nella convenzione si afferma che "*la strategia digitale dell'Agenzia si basa su sette direttive strategiche, che ne rappresentano di fatto i pilastri, vale a dire i riferimenti nella definizione delle iniziative progettuali e nella valutazione delle priorità di azione*"[378] e, per quanto interessa ai nostri fini, la seconda direttiva strategica ha ad oggetto la "*valorizzazione del patrimonio informativo*" e consiste in una valorizzazione dell'enorme quantità di dati a disposizione dell'AdE (presenti in particolare nell'anagrafe tributaria), attraverso la "*sfida*" di "*applicare paradigmi nuovi e tecnologie innovative a grandi moli di dati, spesso distribuiti su numerosi sistemi, con livello di strutturazione spesso non elevato per loro natura, con livelli di qualità da controllare con processi specifici, tramite un approccio multidisciplinare (fiscale, statistico, informatico, matematico) e l'acquisizione e lo sviluppo di competenze specializzate nell'analisi avanzata dei dati*". In particolare, nell'ambito strategico d'intervento relativo ai controlli[379], l'Agenzia s'impegna a potenziare i controlli fiscali puntando sulla "*qualità dei controlli attraverso selezioni più mirate dei contribuenti a maggiore rischio di evasione, rese possibili dall'applicazione di strumenti di data analysis più avanzati – quali lo sfruttamento dei big data – e dall'interoperabilità delle banche dati, favorita dalla pseudonimizzazione delle informazioni*". Inoltre, sempre secondo quanto dispone la convenzione, la valutazione del rischio si baserà anche sullo sfruttamento delle informazioni provenienti da enti

[377] Ad esempio, l'atto di indirizzo del MEF, nelle disposizioni relative all'AdE (allegato A), nell' area di intervento denominata "*Area contrasto – controlli e risoluzione delle controversie fiscali*", prevede orientamenti sostanzialmente analoghi all'omologa area di intervento prevista nella convenzione. In linea generale, in entrambi i provvedimenti, si ritiene che: "*il potenziamento delle infrastrutture tecnologiche, l'interoperabilità delle banche dati e gli strumenti di data analysis potranno contribuire al potenziamento dei controlli effettuati dall'Amministrazione finanziaria, consentendo selezioni sempre più mirate di contribuenti a maggiore rischio di evasione e riducendo così l'invasività dei controlli*"; In tal senso dispone l'*Atto di Indirizzo per il conseguimento degli obiettivi di Politica fiscale 2024-2026*.

[378] All. 2, Piano dell'agenzia delle entrate, p. 50.

[379] Art. 5 della convenzione triennale; il documento è reperibile al seguente link https://www.finanze.gov.it/export/sites/finanze/.galleries/Documenti/Varie/Convenzione-con-lAgenzia-delle-Entrate-triennio-2023-2025.pdf.

esterni, di quelle di natura finanziaria contenute nell'Archivio dei rapporti finanziari, nonché dei dati derivanti dalla fatturazione elettronica.

Su tali dati ed informazioni a disposizione dell'AdE sarà possibile migliorare la qualità dei controlli attraverso "*nuovi strumenti e progetti di analisi avanzata dei dati, l'applicazione di tecniche come l'intelligenza artificiale, il machine learning e il text mining*"[380].

Da tali indicazioni si può dedurre che l'amministrazione finanziaria, per potenziare la selezione dei contribuenti a rischio, necessita di analizzare una vasta mole di dati, che possono essere anche non strutturati[381]. Per affrontare tale problematica la convenzione suggerisce di potenziare l'interoperabilità (soprattutto tra le banche dati dell'anagrafe tributaria) e di sfruttare in tal modo tutte le banche dati cui ha accesso l'amministrazione finanziaria (che costituiscono dei veri e propri *Big Data*)[382], anche attraverso le moderne tecniche di analisi dei dati (per esempio di *machine learning*).

Realizzare tali obiettivi non è una sfida di poco conto, soprattutto per via della complessità delle banche dati a disposizione dell'amministrazione finanziaria, prima fra

[380] Occorre sottolineare che un incentivo all'innovazione tecnologica in quest'ambito proviene anche dal PNRR, dal piano di riforma concordato con il Consiglio dell'Unione europea e dai finanziamenti di cui ha beneficiato l'Agenzia delle Entrate in relazione al progetto "*A data driven approach to tax evasion risk analysis in Italy*" (comunicato stampa dell'Agenzia delle Entrate del 4 marzo 2021); attraverso quest'ultimo progetto la Commissione europea, dal 2021, supporta l'Agenzia delle Entrate per adottare un approccio *data driven* e l'IA per contrastare l'evasione. In merito a quest'ultimo punto si veda M. Pierro, G. Ragucci, *le analisi del rischio di evasione tra selezione dei contribuenti da sottoporre a controllo e accertamento "algoritmico"*, in G. Ragucci (a cura di), *Fisco digitale cripto-attività, protezione dei dati, controlli algoritmici*, Torino, Giappichelli, 2023, pp. 79-112;

[381] Tale indicazione emerge anche dall'audizione del cinque maggio 2021 del direttore dell'AdE, Ernesto Maria Ruffini, il quale ha spiegato che il patrimonio informativo a disposizione dell'Agenzia delle Entrate si compone di numerose banche dati di grandi dimensioni, eterogenee per struttura e contenuto e soggette a rapidi cambiamenti. L'audizione è stata svolta dalla Commissione parlamentare di vigilanza sull'anagrafe tributaria, avente per oggetto la tematica relativa alla "*Digitalizzazione e interoperabilità delle banche dati fiscali*".

[382] I Big Data, intesi in questo senso come un'ingente mole di informazioni di carattere personale, rafforzano il potere conoscitivo dell'amministrazione finanziaria aumentando l'eterogeneità e l'ampiezza del patrimonio informativo fiscale. In tal senso A. Purpura, *La frontiera dei Big data*, in G. Palumbo, *Fisco e privacy. Il difficile equilibrio tra lotta all'evasione e tutela dei dati personali*, Pacini Giuridica, Pisa, 2021, p.79-80.

tutte, l'anagrafe tributaria[383], che assume un importante ruolo per la lotta all'evasione e per la selezione dei contribuenti da sottoporre a controlli fiscali. Vi è da considerare che l'amministrazione finanziaria in determinati casi potrà sfruttare anche il potenziale dei dati derivanti dalle fonti aperte[384]. L'anagrafe tributaria raccoglie numerosissime

[383] Di cui si è già accennato nel secondo capitolo, in quanto annoverata tra le basi di dati di interesse nazionale.

[384] Specialmente le fonti aperte e i dati accessibili da chiunque possono rivelarsi utili per le indagini fiscali, sebbene sarà imprescindibile effettuare dei più rigidi controlli relativi alla qualità dei dati. le fonti aperte andranno ad integrare i processi di valutazione tradizionale dei rischi di non compliance come evidenziato nella circolare AdE n. 16/E del 2016 "*Dal punto di vista operativo, alle notizie ritraibili dalle banche dati si aggiungono quelle che non pervengono da altre fonti, ivi incluse fonti aperte*"; tale concetto è ribadito nella legge di bilancio 2020. La capacità di analizzare i dati, spesso destrutturati, provenienti da fonti aperte necessiterà dell'IA per evidenziare le correlazioni non immediatamente evidenti tra questi dati e quelli già in possesso dall'amministrazione finanziaria, i quali, in ogni caso, costituiscono le evidenze principali su cui basarsi. In merito alle possibilità di ricorrere alle fonti aperte nel senso appena descritto si veda O. Signorile, *La ricerca di dati su fonti aperte come nuovo strumento delle indagini fiscali*, in G. Ragucci (a cura di), *Fisco digitale cripto-attività, protezione dei dati, controlli algoritmici*, Torino, Giappichelli, 2023, pp. 113-127.
Un esempio di utilizzo di fonti aperte è quello relativo al reperimento di informazioni provenienti da articoli di giornale, *social network* e siti *web*, nel caso del rilevamento algoritmico di un significativo scostamento tra le spese sostenute e quelle dichiarate. In quest'ultimo senso si veda I. Alberti, *La partecipazione procedimentale per legittimare gli algoritmi nel procedimento amministrativo*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., 293; L'autrice prosegue facendo presente che l'Agenzia delle Entrate può utilizzare la conoscenza disponibile online per acquisire informazioni sulle specifiche delle transazioni immobiliari e sulla posizione geografica degli immobili; queste informazioni vengono poi confrontate con le quotazioni riportate dall'Osservatorio del mercato immobiliare per assicurarsi dichiarazioni sottostimate riguardo al valore di acquisto di un immobile.

informazioni[385] e il suo funzionamento si basa sulla connessione di dette informazioni con il codice fiscale del contribuente[386].

[385] Come precisato in A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.77, le informazioni contenute nell'anagrafe ricomprendono i dati che pervengono all'Agenzia delle Entrate attraverso comunicazioni effettuate "*sia direttamente dai contribuenti o da loro intermediari abilitati (dichiarazioni dei redditi, atti soggetti a registrazione, pagamenti con Modelli F24 e F23, dichiarazioni di inizio attività`, fatturazione elettronica, corrispettivi telematici, ecc.), sia da enti esterni (istituti finanziari, società fornitrici di utilities, assicurazioni, CCIAA, INPS, INAIL, Comuni, ecc.)*" e tra queste ultime vi sono anche le amministrazioni fiscali estere. Un'efficace panoramica delle numerose banche dati di cui si compone l'anagrafe tributaria, utilizzate per comporre il *dataset* di analisi, è contenuta nella valutazione di impatto sulla protezione dei dati, pubblicata il 30 luglio 2022 dal Garante della *privacy* [doc. web n. 9808839], relativa al trattamento "*Analizzare rischi e fenomeni evasivi/elusivi tramite l'utilizzo dei dati contenuti nell'Archivio dei rapporti finanziari e l'incrocio degli stessi con le altre banche dati di cui dispone l'Agenzia delle entrate*", *trattamento di cui all'art. 1, comma 684, della legge 27 dicembre 2019, n. 160*. Il Garante rileva infatti che per creare il *dataset* di analisi verranno utilizzate numerose banche dati, tra le quali, ad esempio, oltre all'archivio dei rapporti finanziari, dati anagrafici, dichiarativi, accertamenti e controlli, fatture elettroniche (cc.dd. "dati fattura" e "dati fattura integrati"), catasto, osservatorio mercato immobiliare, versamenti F24 e F23, Anagrafe nazionale popolazione residente, depositi IVA (dichiarazione sostitutiva dei requisiti di affidabilità, prestazione della garanzia), deroga alla limitazione all'uso del contante, canone TV, dati INPS (aziende e contributi artigiani e commercianti), contratti di locazione breve, utenze (gas, elettriche, idriche e telefoniche), etc. Tra i dati più recentemente aggiunti all'anagrafe, è importante sottolineare sia le informazioni estratte dalle fatture elettroniche, che, secondo il decreto legislativo n. 124/2019, sono integralmente acquisite e conservate per otto anni a beneficio delle autorità fiscali, sia quelle ottenute dagli indicatori di affidabilità fiscale (ISA), che saranno impiegati, come stabilito dalla normativa, per le attività di selezione e analisi del rischio di evasione fiscale.

Tra i dati elencati particolarmente rilevanti sono quelli relativi alla fatturazione elettronica; infatti tra i principali fattori del recupero di *compliance* nell'IVA, conseguito nel 2019, vi è stata l'introduzione della fatturazione elettronica, in detto periodo infatti è stato stimato un maggior introito, attribuibile a tale adempimento digitale, tra 1,7 e 2 miliardi di euro; in tal senso si veda G. Liberatore, *Lotta all'evasione con consenso 2.0: meglio coordinarsi con l'UE*, in *Il fisco*, 12/2022, p. 1131.

Vista la sensibilità dei dati contenuti nell'anagrafe tributaria, nel 1976 la gestione del suo sistema informatico è stata affidata al Ministero dell'Economia e delle Finanze ed è sottoposta al controllo della commissione bicamerale di vigilanza, composta da deputati e senatori. La SOGEI S.p.a. (*partner* istituzionale di cui si è già visto in merito al PSN), è incaricata dell'implementazione e della gestione informatica dell'anagrafe dal 1976.

[386] A. Bongi, *Intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte,* cit., p.81.

L'anagrafe tributaria è disciplinata dal D.P.R. 29 settembre 1973, n. 605 ("*Disposizioni relative all'anagrafe tributaria e al codice fiscale dei contribuenti*"), che al all'art. 1 chiarisce che essa "*raccoglie e ordina su scala nazionale i dati e le notizie risultanti dalle dichiarazioni e dalle denunce presentate agli uffici dell'amministrazione finanziaria e dai relativi accertamenti, nonché i dati e le notizie che possono comunque assumere rilevanza ai fini tributari*" ed inoltre sulla base di questi dati provvede "*alle elaborazioni utili per lo studio dei fenomeni fiscali*". L'art. 7 del D.P.R. n. 605/1973, invece, disciplina gli obblighi comunicativi che determinati soggetti hanno nei confronti dell'anagrafe.
Ad ogni obbligo di comunicazione corrisponde una sezione dell'anagrafe tributaria e tra i vari obblighi comunicativi, particolare rilevanza assumono quelli di cui al sesto comma dell'art. 7 del D.P.R. n. 605/1973, che pone le basi per la disciplina dell'archivio dei rapporti finanziari, anche detto archivio dei conti correnti, costituito all'interno dell'anagrafe tributaria. Secondo il sesto comma gli intermediari e gli operatori finanziari sono tenuti a rilevare i dati identificativi di ogni soggetto che intrattenga con loro un qualsiasi rapporto o effettui un'operazione di natura finanziaria[387]. In altre parole, all'interno dell'archivio dei rapporti finanziari sono contenute informazioni importanti,

[387] Secondo il sesto comma dell'art. 7, D.P.R. n. 605/1973, "*Le banche, la società Poste italiane Spa, gli intermediari finanziari, le imprese di investimento, gli organismi di investimento collettivo del risparmio, le società di gestione del risparmio, nonché ogni altro operatore finanziario, fatto salvo quanto disposto dal secondo comma dell'articolo 6 per i soggetti non residenti, sono tenuti a rilevare e a tenere in evidenza i dati identificativi, compreso il codice fiscale, di ogni soggetto che intrattenga con loro qualsiasi rapporto o effettui, per conto proprio ovvero per conto o a nome di terzi, qualsiasi operazione di natura finanziaria ad esclusione di quelle effettuate tramite bollettino di conto corrente postale per un importo unitario inferiore a 1.500 euro; l'esistenza dei rapporti e l'esistenza di qualsiasi operazione di cui al precedente periodo, compiuta al di fuori di un rapporto continuativo, nonché la natura degli stessi sono comunicate all'anagrafe tributaria, ed archiviate in apposita sezione, con l'indicazione dei dati anagrafici dei titolari e dei soggetti che intrattengono con gli operatori finanziari qualsiasi rapporto o effettuano operazioni al di fuori di un rapporto continuativo per conto proprio ovvero per conto o a nome di terzi, compreso il codice fiscale*". Inoltre, diversi provvedimenti (es. provvedimento dell'Agenzia delle Entrate del 23 maggio 2022) hanno ampliato nel tempo i dati da comunicare ed il numero di soggetti tenuti a comunicare periodicamente tali dati. È da sottolineare che con il decreto semplificazioni (D.L. 16 luglio 2020, n. 76 - *Misure urgenti per la semplificazione e l'innovazione digitale*) potranno accedere alle informazioni di tale archivio, oltre all'Agenzia delle Entrate, alla Guardia di Finanza e dei concessionari della riscossione, anche i comuni e gli enti locali.

quali quelle relative ai titoli, conti correnti e deposito, con relativi saldi iniziali, finali ed intermedi. Tale archivio è definito il "*cuore*" [388] dell'anagrafe tributaria.

Anche gli attuali punteggi ISA entrano a far parte delle informazioni contenute in detto archivio ai fini della selezione e dell'analisi del rischio dei contribuenti. In tal modo l'anagrafe risulta composta dai dati derivanti dalle comunicazioni periodiche delle banche e degli altri intermediari, da punteggi ISA e da altre e numerose informazioni.

È dunque chiaro che l'anagrafe tributaria costituisca una risorsa preziosa per l'amministrazione finanziaria e la necessità di sfruttare a pieno le banche dati che fanno parte di essa (specialmente l'archivio dei rapporti finanziari), per ottimizzare l'analisi del

---

[388] A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., *passim*. Le interazioni tra il settore finanziario e fiscale assumono una rilevanza centrale, specialmente nell'ottica del riutilizzo dei dati finanziari ai fini fiscali, considerando anche il fatto il mondo finanziario e bancario utilizzano "*le informazioni quali parte dei propri processi decisionali quotidiani (dalla concessione di prestiti, alla gestione dei portafogli di investimento, ecc.). Il settore finanziario, inoltre, genera grandi quantità di dati, che vengono accumulati sia all'interno delle istituzioni finanziarie sia al loro esterno (ad esempio dai social media)*". In tal senso e per le relazioni tra importanza dei dati e settore finanziario si rimanda a S. Alvaro, L. Marzialetti, *Neutrality of financial regulation. Il (nuovo) ruolo del regolatore*, in A. Genovese e V. Falce (a cura di), *La portabilità dei dati in ambito finanziario*, Quaderni FinTech, CONSOB, 2021, p. 57. Per un approfondimento sull'archivio dei rapporti finanziari si può vedere Corte dei conti, *l'utilizzo dell'anagrafe dei rapporti finanziari ai fini dell'attività di controllo fiscale*, Deliberazione 26 luglio 2017, n. 11/2017/G. Tra i dati assumono rilievo in particolare quelli derivanti dai pagamenti elettronici effettuati tramite POS. Tale processo è gestito dai Prestatori di Servizi di Pagamento (PSP) in collaborazione con PagoPA S.p.A., che successivamente mette a disposizione dell'Agenzia delle Entrate tali informazioni. Le modalità operative sono dettagliate nel provvedimento direttoriale dell'Agenzia delle Entrate n. 253155, pubblicato il 30 giugno 2022. In tal senso si veda A. Mastromatteo, *Arriva l'algoritmo antievasione, ecco a cosa serve e come funziona*, in *Agenza Digitale*, 5 luglio 2022, disponibile su https://www.agendadigitale.eu/documenti/arriva-lalgoritmo-antievasione-ecco-a-cosa-serve-e-come-funziona/. In tale ambito rileva anche l'art. 94 della Direttiva (UE) 2015/2366 (c.d. PSD2), il quale stabilisce che "*gli Stati membri autorizzano il trattamento dei dati personali da parte di sistemi di pagamento e di prestatori di servizi di pagamento se necessario per garantire la prevenzione, l'indagine e l'individuazione dei casi di frode nei pagamenti*".

rischio fiscale, oltre ad essere sottolineato nella convenzione tra AdE e MEF e nell'atto di indirizzo del MEF, era già stata avvertita da diverso tempo.

Il primo passo in avanti decisivo è stato fatto dalla legge 27 dicembre 2019, n. 160 (legge di bilancio 2020 - *Bilancio di previsione dello Stato per l'anno finanziario 2020 e bilancio pluriennale per il triennio 2020-2022*), art. 1, commi da 681 a 686[389], che ha posto al centro del contrasto all'evasione la selezione dei contribuenti e l'analisi del rischio e le cui disposizioni costituiscono il *framework* normativo per il contrasto moderno all'evasione, basato sull'interoperabilità delle banche dati e sulla digitalizzazione[390].

Il comma 682, art. 1, legge n. 160/2019, dispone infatti che per le attività di analisi del rischio, con riferimento ai dati contenuti nell'archivio dei rapporti finanziari, l'Agenzia delle Entrate[391] si avvale di tecnologie, delle elaborazioni e delle interconnessioni con le

[389] Come rilevato in I. Alberti, *La partecipazione procedimentale per legittimare gli algoritmi nel procedimento amministrativo*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, cit., p.292, il valore conoscitivo delle banche dati era già stato riconosciuto nella circolare AdE n. 16/E del 28 aprile 2016. Quest'ultima inoltre rileva l'importanza della qualità dei dati al fine di supportare le attività di analisi del rischio e di selezione dei contribuenti da controllare. Il passo in avanti è dovuto al fatto che i commi da 681 a 686 dell'art. 1 della legge appena menzionata hanno introdotto per la prima volta le basi normative per l'utilizzo dell'intelligenza artificiale nelle attività di analisi e di selezione del rischio fiscale. Prima l'Agenzia applicava a fatica tali strumenti, vista l'assenza di un'esplicita normativa.

[390] A. Bongi, *Intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.49 e 14. L'autore fa presente che tali disposizioni normative incoraggiano una lotta all'evasione basata sull' "*utilizzo di strumenti informatici ed elaborazioni statistico-matematiche in grado di leggere e valutare la miriade di informazioni contenute nell'anagrafe tributaria*".

[391] Anche la Guardia di finanza può utilizzare i dati contenuti nell'archivio dei rapporti finanziari per le stesse finalità di cui al comma 682, art. 1 ed avvalendosi delle moderne tecnologie ed interconnessioni tra banche dati. Infatti, il comma 686, art. 1 così dispone: "*Per le stesse finalità di cui al comma 682, la Guardia di finanza utilizza i dati contenuti nell'Archivio dei rapporti finanziari con le medesime modalità disciplinate dai commi da 681 a 685, avvalendosi delle tecnologie, delle elaborazioni e delle interconnessioni con le altre banche dati di cui è titolare*". Sarà quindi fondamentale realizzare un coordinamento efficace tra Agenzia delle Entrate e Guardia di finanza. A tal fine l'art. 6 del decreto del Ministero dell'economia e delle finanze del 28 giugno 2022 sancisce che per le finalità di cui al comma 686 dell'art. 1 della l.160/2019 i dati contenuti nell'archivio dei rapporti finanziari sono resi disponibili dall'Agenzia alla Guardia di finanza, che li utilizza avvalendosi delle tecnologie e delle interconnessioni con le altre

altre banche dati di cui dispone allo scopo di far emergere posizioni da sottoporre a controllo e incentivare l'adempimento spontaneo[392].

La legge n. 160/2019 attribuisce particolare rilevanza, in *primis*, all'analisi dei dati contenuti nell'archivio dei rapporti finanziari e, in *secundis*, lascia aperta la possibilità di interconnessione con le altre banche dati di cui dispone l'Agenzia.

In tale contesto si inserisce l'applicativo *software* Ve.R.A. (acronimo di "verifica dei rapporti finanziari"), il quale permette all'amministrazione finanziaria, secondo quanto previsto dai commi da 681 a 686 dell'art. 1 della legge 27 dicembre 2019, n. 160, di svolgere le analisi del rischio di evasione partendo dai dati dell'archivio dei rapporti finanziari ed integrandoli con gli altri elementi presenti in anagrafe tributaria[393]. In tal

---

banche dati di cui è titolare, nell'ambito di un "*accordo convenzionale*", che disciplina le modalità di accesso all'archivio, i termini e le misure di sicurezza.

[392] Secondo quanto dispone il comma 682: "*Per le attività di analisi del rischio di cui all'articolo 11, comma 4, del decreto-legge 6 dicembre 2011, n. 201, convertito, con modificazioni, dalla legge 22 dicembre 2011, n. 214, con riferimento all'utilizzo dei dati contenuti nell'archivio dei rapporti finanziari, di cui all'articolo 7, sesto comma, del decreto del Presidente della Repubblica 29 settembre 1973, n. 605, e all'articolo 11, comma 2, del decreto-legge 6 dicembre 2011, n. 201, convertito, con modificazioni, dalla legge 22 dicembre 2011, n. 214, l'Agenzia delle entrate, anche previa pseudonimizzazione dei dati personali, si avvale delle tecnologie, delle elaborazioni e delle interconnessioni con le altre banche dati di cui dispone, allo scopo di individuare criteri di rischio utili per far emergere posizioni da sottoporre a controllo e incentivare l'adempimento spontaneo".* Il comma 4 dell'art. 11 del decreto legge 6 dicembre 2011, n. 201, convertito, con modificazioni, dalla legge 22 dicembre 2011, n. 214, disponeva l'acquisizione di ulteriori informazioni all'anagrafe tributaria, in modo da agevolare la predisposizione delle liste selettive dei contribuenti con maggior rischio di evasione, attraverso procedure centralizzate. In tale ultimo senso si veda B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, cit., p.151. Occorre sottolineare che il tenore letterale della norma non pone limiti alle interconnessioni che è possibile effettuare tra le banche dati, infatti, come rileva l'Agenzia delle Entrate nel *Documento di Valutazione di Impatto sulla Protezione dei Dati (Stralcio)*, l'interconnessione tra le banche dati, potrà "*essere effettuata anche con eventuali banche dati che dovessero rendersi disponibili in futuro, considerato che il più volte citato comma 682 richiama «le altre banche dati di cui dispone [l'Agenzia delle entrate]», senza porre alcun vincolo temporale*".

[393] Ve.R.A. è un algoritmo che ha la capacità di incrociare contemporaneamente milioni di dati per effettuare l'analisi del rischio fiscale. In tal senso C. Francioso, *Intelligenza artificiale nell'istruttoria tributaria e nuove esigenze di tutela*, in *Rassegna tributaria*, 2023, pp. 47 ss. Occorre precisare che solo nel 2022 Ve.R.A. esce dalla fase di sperimentazione e diventa, a regime, un mezzo di supporto istruttorio per l'analisi del rischio fiscale. Secondo alcuni autori

modo potrebbe inaugurarsi una "*nuova stagione di contrasto all'evasione e di compliance fiscale, anche in chiave PNRR*"[394], che si basa proprio sull'incrocio fra i dati dell'archivio dei rapporti finanziari con gli altri di cui dispone l'amministrazione finanziaria. Sulla base di queste informazioni l'applicativo Ve.R.A. permetterà di predisporre le liste selettive centralizzate per l'attività di controllo[395]. Quindi, per formare il *dataset* di analisi[396], su cui si baserà l'algoritmo, vengono sfruttate tutte le banche dati dell'anagrafe tributaria ed il primo problema è rappresentato proprio dall'esigenza di dover interconnettere tutte le banche dati che fanno parte dell'anagrafe[397]. In ogni caso la selezione dei contribuenti sarà "*tanto più efficace quante*

non è casuale che tale passaggio sia avvenuto in concomitanza con le modifiche apportate dal decreto capienze (d.l. n. 139/2021) al codice *privacy*, in quanto solamente grazie a queste ultime modifiche l'amministrazione titolare del trattamento possiede spazi di manovra più ampi. In merito a quest'ultimo punto si veda B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, cit., pp.158-159; M. Conigliaro, *Lotta all'evasione con l'intelligenza artificiale*, in *Il Fisco*, 2022, n. 32/33, pp. 3107 ss.; C. Francioso, *Intelligenza artificiale nell'istruttoria tributaria e nuove esigenze di tutela*, cit., pp. 47-94.

[394] A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, p.97.

[395] Infatti, secondo la circolare AdE n. 21/E del 20 giugno 2022, l'analisi del rischio di evasione basata sui dati dell'archivio dei rapporti, sarà potenziata mediante l'elaborazione di "*nuove liste selettive per l'attività di controllo, che saranno rese disponibili mediante l'applicativo Ve.R.A*". in tal modo, a livello centrale, verranno formati elenchi realizzati grazie all'utilizzo integrato delle informazioni comunicate dagli operatori finanziari all'archivio dei rapporti finanziari e degli altri elementi presenti in anagrafe tributaria. Tali elenchi permetteranno a ciascuna Direzione regionale e provinciale di indirizzare l'ordinaria attività di controllo nei confronti delle posizioni a più elevato rischio di evasione.

[396] Definito quale l'"*l'insieme dei dati selezionati ai fini di cui all'articolo 1, commi da 682 a 686, della legge 27 dicembre 2019, n. 160, per verificare la presenza dei rischi fiscali identificati in esito alle tecniche ed ai modelli di analisi utilizzati*". In altre parole, il *dataset* di analisi rappresenta quei dati utilizzati dalle tecniche di analisi per identificare il rischio fiscale. La definizione appena riportata è contenuta *Documento di Valutazione di Impatto sulla Protezione dei Dati (Stralcio),* emanato dall'AdE in ottemperanza dell'articolo 1, comma 684, della legge 27 dicembre 2019, n. 160, in relazione al trattamento consistente nell' "*Analizzare rischi e fenomeni evasivi/elusivi tramite l'utilizzo dei dati contenuti nell'Archivio dei rapporti finanziari e l'incrocio degli stessi con le altre banche dati di cui dispone l'Agenzia delle entrate*".

[397] È possibile dedurre che il potenziale delle banche dati fiscali non sia ancora totalmente sfruttato dalla relazione sul rendiconto annuale dello Stato (relativa al 2022 e pubblicata il 28 giugno 2023) effettuata dalla Corte dei conti. Quest'ultima infatti afferma che "*una maggiore*

maggiori sono le informazioni disponibili e quanto maggiore è il livello qualitativo delle stesse"[398].

Il trattamento previsto dalla legge n. 160/2019 pone numerose problematiche in relazione alla disciplina sulla tutela dei dati personali. Nel paragrafo seguente si vedrà come si cerca di risolvere le questioni inerenti a tale disciplina.

2. Analisi del rischio fiscale e tutela dei dati personali

L'amministrazione finanziaria potrà quindi incrociare basi di dati ed effettuare trattamenti algoritmici. E questo soprattutto grazie alla legge n. 160/2019. Numerosi sono, tuttavia, i profili problematici che si pongono in relazione alla tutela dei dati personali, alcuni dei quali sono affrontati dalla stessa legge n. 160/2019[399]. Si ritiene opportuno, quindi, analizzare in via preliminare le disposizioni di quest'ultima legge.

Il comma 681, art. 1 della legge di bilancio 2020, apporta due modifiche significative al codice della *privacy* (d.lgs. n. 196/2003): secondo una prima modifica, relativa all'art. 2-*sexies*, le attività di "*prevenzione e contrasto all'evasione fiscale*" diventano motivi di interesse pubblico rilevante che legittimano ad effettuare il trattamento delle categorie

---

frequenza dei controlli fiscali potrebbe e dovrebbe integrare l'utilizzazione in chiave (prima di tutto) preventiva della ingente mole di dati a disposizione dei sistemi informativi (tra i quali, i dati descrittivi delle fatture elettroniche emesse e ricevute, i corrispettivi comunicati telematicamente e i movimenti risultanti dall'Anagrafe dei rapporti finanziari e dai pagamenti elettronici), già normativamente prevista, in buona parte, ma ancora non compiutamente realizzata. Ciò nell'ottica di una maggiore efficienza dell'attività accertativa e della connessa funzione di deterrenza rispetto a comportamenti non conformi alla normativa tributaria".

[398] A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.77.

[399] Il tema della protezione dei dati personali in relazione alle banche dati fiscali sta assumendo sempre maggiore rilevanza, poiché in queste ultime sono conservati una moltitudine di dati personali e vi è una necessità continua anche di scambiarli con altri enti ed autorità, anche straniere, ai fini della repressione di illeciti internazionali. Ed infatti, attraverso tali dati, si procede ad una vera e propria "profilazione" del contribuente. Per un approfondimento si veda A. Contrino, *Banche dati tributarie, scambio di informazioni fra autorità fiscali e "protezione dei dati personali": quali diritti e tutele per i contribuenti?,* in R*ivista di diritto tributario supplemento online*, Pacini giuridica, 2019.

particolari di dati personali previste dall'art. 9 del GDPR; la seconda modifica, invece, relativa all'art. 2-*undecies*, permette di limitare i diritti degli interessati nel caso in cui l'esercizio di tali diritti possa arrecare pregiudizio effettivo e concreto agli "*interessi tutelati in materia tributaria e allo svolgimento delle attività di prevenzione e contrasto all'evasione fiscale*"[400]. Per effetto di tali modifiche, la lotta all'evasione fiscale diventa

[400] Dispone infatti il comma 681, art. 1, della legge di bilancio che:
"*In considerazione dei rilevanti obiettivi di interesse pubblico di prevenzione e contrasto all'evasione, al codice di cui al decreto legislativo 30 giugno 2003, n. 196, sono apportate le seguenti modificazioni:*

a) all'articolo 2-sexies, comma 2, lettera i), dopo la parola: «doganale» sono aggiunte le seguenti: «, comprese quelle di prevenzione e contrasto all'evasione fiscale»;

b) all'articolo 2-undecies, comma 1, dopo la lettera f) è aggiunta la seguente: «f-bis) agli interessi tutelati in materia tributaria e allo svolgimento delle attività di prevenzione e contrasto all'evasione fiscale»;

c) all'articolo 2-undecies, comma 3, le parole: «e) ed f)», ovunque ricorrono, sono sostituite dalle seguenti: «e), f) e f-bis)»".

La prima modifica, relativa all'art. 2-*sexies*, introducendo tra i motivi di interesse pubblico rilevante quelli relativi al contrasto dell'evasione fiscale, si avvale di quanto disposto dall'art. 9 GDPR, che permette di effettuare il trattamento di categorie particolari di dati qualora tale trattamento sia necessario per motivi di interesse pubblico rilevante sulla base del diritto dell'Unione o dello Stato membro (lett.g), paragrafo 2, art. 9 GDPR). Giova ricordare che la modifica apportata dalla legge n. 160/2019 è intervenuta in un momento in cui solo una base legislativa poteva legittimare un trattamento avente ad oggetto categorie particolari di dati personali, in ossequio alla *strict legality rule*. Tale base legislativa non sarebbe stata necessaria invece, se le modifiche apportate con il decreto capienze fossero già entrate in vigore. Come sottolineato nel primo capitolo, infatti, attualmente il trattamento di dati particolari per motivi di interesse pubblico rilevante può essere effettuato anche se previsto da un atto amministrativo generale.

La seconda modifica, relativa all'art. 2-*undecies*, rubricato '*limitazioni ai diritti dell'interessato*', si avvale del margine di manovra previsto dall'art. 23 GDPR, che permette di limitare i diritti degli interessati attraverso "*misure legislative"* previste dal diritto dello Stato membro e finalizzate al perseguimento di interessi previsti dallo stesso articolo 23, tra i quali, alla lettera e), figura la salvaguardia di "*altri importanti obiettivi di interesse pubblico generale dell'Unione o di uno Stato membro, in particolare un rilevante interesse economico o finanziario dell'Unione o di uno Stato membro, anche in materia monetaria, di bilancio e tributaria, di sanità pubblica e sicurezza sociale*". E le modifiche apportate all'art. 2-*undecies* permettono proprio di limitare i diritti degli interessati previsti dal GDPR per contrastare l'evasione fiscale (lett. f-*bis*), art. 2-*undecies*). Tuttavia, la misura legislativa, ai sensi dell'art. 23, paragrafi 1 e 2, dovrà rispettare pur sempre una serie di requisiti minimi. E tale ultima previsione risulta coordinata con il terzo comma dell'art. 2- *undecies*, il quale stabilisce che i diritti degli interessati sono esercitati

una priorità nazionale ed i cittadini non saranno più in grado di usufruire di diversi diritti stabiliti dalla normativa sulla *privacy*, come la richiesta di accesso ai dati o la presentazione di reclami in situazioni di errori o violazioni delle norme relative al trattamento dei dati personali[401].

Le limitazioni dei diritti dell'interessato e degli obblighi del titolare e del responsabile del trattamento, per effetto degli articoli 23 GDPR e 2-*undecies* del codice *privacy*, devono essere previste da disposizioni di legge o di regolamento e, per questo, il comma 683, art. 1, della legge n. 160/2019, demanda la normativa di dettaglio ad un decreto del Ministro dell'economia e delle finanze[402].

---

conformemente a "*disposizioni di legge o di regolamento*" (e, si noti, non da atti amministrativi generali), che devono recare misure dirette a disciplinare gli ambiti di cui all'art. 23, paragrafo 2, GDPR. Quest'ultima condizione risulta soddisfatta, in quanto, la legge n. 160/2019, al comma 683 demanda ad un decreto del Ministro dell'economia e delle finanze di specificare tali requisiti, che devono essere rispettati dal trattamento.

[401] Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.51. Nel corso dell'analisi non ci si è soffermati in particolare sui diritti degli interessati, quanto sui principi che regolano l'acquisizione e la circolazione dei dati (quale il principio di limitazione delle finalità), ci si limita quindi a sottolineare che la ricomprensione di tale trattamento quale trattamento atto a realizzare un obiettivo di "*rilevante interesse pubblico*", permetterà di comprimere molti diritti degli interessati.

[402] Il comma 683, art. 1, legge n. 160/2019, assegna a tale decreto il compito di definire le limitazioni e le modalità di esercizio dei diritti degli interessati, di adottare misure adeguate a tutela dei diritti e delle libertà degli interessati e di definire il contenuto minimo di alcune garanzie previste dall'art. 23, paragrafo 2, GDPR (es. l'art. 23 GDPR prevede che la misura legislativa, che limita i diritti degli interessati, debba contenere disposizioni specifiche riguardanti: finalità del trattamento; periodi di conservazione dei dati; categorie di dati personali; portata delle limitazioni introdotte; indicazione precisa del titolare del trattamento etc.). Infatti, secondo quanto dispone il comma 683, art. 1: "*Nel rispetto delle disposizioni di cui all'articolo 2-undecies, comma 3, del codice di cui al decreto legislativo 30 giugno 2003, n. 196, nonché dell'articolo 23, paragrafo 1, del regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, considerati i principi di necessità e di proporzionalità, limitatamente al trattamento dei dati contenuti nell'archivio dei rapporti finanziari di cui al comma 682, con decreto del Ministro dell'economia e delle finanze, da emanare entro novanta giorni dalla data di entrata in vigore della presente legge, sentiti il Garante per la protezione dei dati personali e l'Agenzia delle entrate, sono definite: a) le specifiche limitazioni e le modalità di esercizio dei diritti di cui agli articoli 14, 15, 17, 18 e 21 del regolamento (UE) 2016/679, in modo da assicurare che tale esercizio non possa arrecare un pregiudizio effettivo e concreto all'obiettivo di interesse pubblico; b) le disposizioni specifiche relative al contenuto minimo essenziale di cui all'articolo*

Il comma 682, art. 1, legge n. 160/2019, già esaminato, specifica le finalità del trattamento, la tipologia e le modalità del trattamento (compresa l'indicazione delle banche dati utilizzabili).

I commi 684 e 685 dell'art. 1, invece, impongono all'Agenzia delle Entrate di effettuare una "*valutazione unitaria di impatto sulla protezione dei dati*" (DPIA), sentito il Garante della *privacy*, prima di iniziare il trattamento, in piena aderenza con il principio di

23, paragrafo 2, del regolamento (UE) 2016/679; c) le misure adeguate a tutela dei diritti e delle libertà degli interessati.".

Il decreto attuativo è stato emanato dal MEF il 28 giugno 2022 e si apre, all'art. 1 (rubricato '*definizioni*'), con una serie di nozioni caratterizzanti il GDPR ed "*in certo senso nuove per il diritto tributario e che fanno il loro ingresso in questa delicata materia per la prima volta*", come affermato in A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.64. Ed infatti, nell'art. 1, sono richiamate le definizioni contenute nel GDPR per quanto riguarda le nozioni di dato personale, interessato, trattamento, e pseudonimizzazione. Invece, definendo con la nozione di "titolari del trattamento" l'Agenzia delle Entrate e la Guardia di finanza e con la nozione di "banche dati" gli archivi detenuti da questi ultimi enti, si prevede il significato della nozione e, al tempo stesso, alcune caratteristiche della disciplina.

responsabilizzazione di cui al GDPR[403]. Non è consentito iniziare il trattamento prima della valutazione d'impatto, a meno di espresso consenso del Garante delle *privacy*[404].

Alla luce della normativa primaria, vediamo come risulta caratterizzato il trattamento in seguito all'emanazione del decreto attuativo del MEF e della valutazione d'impatto *privacy* effettuata dall'AdE.

---

[403] Secondo quanto dispone il comma 684, art. 1 "*Nel rispetto del principio di responsabilizzazione, ai sensi dell'articolo 35 del regolamento (UE) 2016/679, il trattamento di cui al comma 682 è oggetto di una valutazione unitaria di impatto sulla protezione dei dati, effettuata dall'Agenzia delle entrate prima di iniziare il trattamento stesso, sentito il Garante per la protezione dei dati personali. Nella valutazione d'impatto sono indicate anche le misure necessarie e ragionevoli per assicurare la qualità dei dati*". Inoltre, come specificato dal comma 10, art. 5, d.m. del MEF 28 giugno 2022, la valutazione d'impatto va periodicamente aggiornata, sentito il Garante della privacy, quando insorgono variazioni del rischio rappresentato dalle attività relative al trattamento.

Ai sensi di queste ultime disposizioni, oltre che per via dell'art. 35 GDPR, la valutazione d'impatto risulta obbligatoria. Tuttavia, non è previsto in via generale, neanche dal GDPR, l'obbligo di pubblicare la valutazione d'impatto. Infatti, nelle "*Linee guida in materia di valutazione d'impatto sulla protezione dei dati e determinazione della possibilità che il trattamento "possa presentare un rischio elevato" ai fini del regolamento (UE) 2016/679*", adottate dal Gruppo di lavoro articolo 29 (WP29) per la protezione dei dati il 4 aprile 2017, come modificate e adottate da ultimo il 4 ottobre 2017, fatte proprie dal Comitato europeo per la protezione dei dati il 25 maggio 2018, "*la pubblicazione di una valutazione d'impatto sulla protezione dei dati non è un requisito giuridico sancito dal regolamento generale sulla protezione dei dati, è una decisione del titolare del trattamento procedere in tal senso. [….] Costituisce una prassi particolarmente buona pubblicare una valutazione d'impatto sulla protezione dei dati nel caso in cui individui della popolazione siano influenzati dal trattamento interessato. Nello specifico, ciò potrebbe essere il caso in cui un'autorità pubblica realizza una valutazione d'impatto sulla protezione dei dati*". In virtù di tali considerazioni in Garante della *privacy*, nella valutazione del 30 luglio 2022 della DPIA predisposta dall'Agenzia delle Entrate, ha formulato una precisa ingiunzione all'agenzia: "*pubblicare un estratto della valutazione di impatto sulla protezione dei dati, omettendo gli allegati e le parti che possono compromettere la sicurezza dei trattamenti*". L'Agenzia delle Entrate, a sua volta, ha provveduto a pubblicare uno stralcio del documento di valutazione di impatto sulla protezione dei dati, reperibile sul portale stesso dell'Agenzia https://www.agenziaentrate.gov.it/portale/documents/20143/5316839/Documento+di+Valutazione+di+Impatto+sulla+Protezione+dei+Dati+%28Stralcio%29.pdf/f8491e14-aaca-34f1-c157-65f964fbb0a3 .

[404] E ciò in virtù di quanto dispone il comma 685, art. 1, che sancisce "*Salvo che non sia stato espressamente autorizzato prima della data di entrata in vigore della presente legge dal Garante per la protezione dei dati personali, non è consentito il trattamento dei dati di cui al comma 682 prima della valutazione di impatto di cui al comma 684*".

Innanzitutto, occorre rilevare che, secondo il comma 2, art. 3, d.m. 28 giugno 2022, non sono trattati dati particolari di cui all'art. 9 GDPR, ma solamente dati comuni (relativi all'identità fisica ed economica, tra cui quelli contabili, finanziari e patrimoniali)[405]. Viene ulteriormente precisato che all'interno dei *dataset* di analisi non sono mai inseriti i dati particolari di cui all'art. 9 e quelli giudiziari di cui all'art. 10 del Regolamento (UE) 2016/679[406], poiché questi ultimi sono trattati esclusivamente nella fase di test[407].

Inoltre, i diritti degli interessati risultano compressi e, quindi, il Garante della *privacy* è, di fatto, "*l'unico soggetto che può pretendere il rispetto delle disposizioni di legge anche ai fini di una tutela dei contribuenti*"[408].

La valutazione d'impatto effettuata dall'Agenzia delle Entrate precisa che la finalità del trattamento ex comma 682, art. 1, legge n. 160/2019 è duplice e mira a: "*a) far emergere le posizioni da sottoporre a controllo con i tradizionali poteri istruttori di cui agli artt. 32 e 33 del d.P.R. 29 settembre 1973, n. 600, 51 e 52 del d.P.R. 26 ottobre 1972 n. 633 e 53-bis del d.P.R. 26 aprile 1986, n. 131; b) rilevare le anomalie da comunicare ai contribuenti, ai sensi dell'art. 1, commi 634-636, della legge 23 dicembre 2014, n. 190, per l'azione di*

---

[405] Secondo il comma 2, art. 3, d.m. 28 giugno 2022 per le finalità di analisi del rischio e per incentivare l'adempimento spontaneo "*sono trattati dati personali comuni, contenuti nelle banche dati, relativi all'identità anagrafica ed alla capacità economica, tra cui dati riguardanti le dichiarazioni fiscali, il patrimonio mobiliare e immobiliare, dati contabili e finanziari, dati dei pagamenti, dei versamenti e delle compensazioni, nonché i dati di profilazione relativi agli eventuali indicatori di rischio desunti o derivati attribuiti ai soggetti; non sono oggetto di trattamento nei dataset i dati di cui all'art. 9 del regolamento*".

[406] AdE, *Documento di Valutazione di Impatto sulla Protezione dei Dati (Stralcio)*, p.39. Oltre che nei *dataset* di analisi non sono mai inseriti neanche nei *dataset* di controllo, per questi intendendosi, ai sensi dell'art. 1, d.m. 28 giugno 2022 l' "*insieme delle posizioni fiscali dei contribuenti, caratterizzate dalla ricorrenza di uno o più rischi fiscali, nei confronti dei quali potranno essere avviate le attività di controllo ovvero le attività volte a stimolare l'adempimento spontaneo*".

[407] AdE, *Documento di Valutazione di Impatto sulla Protezione dei Dati (Stralcio)*, p.93. La caratteristica di non inserire tali dati nei dataset, oltre che a permettere all'AdE di non effettuare una DPIA su questi ultimi (p.94 della DPIA dell'AdE), comporta che non si possano avere discriminazioni effettuate dall'algoritmo basate su queste categorie di dati (p.66 della DPIA). Quindi si elimina in radice una sfaccettatura delle conseguenze pregiudizievoli che possono derivare da questa tipologia di trattamento.

[408] A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.125.

stimolo dell'adempimento spontaneo"[409]. Il trattamento, teleologicamente orientato a questi fini, si potrà basare sulle banche dati a disposizione dell'Agenzia delle Entrate e sulle analisi effettuate dalle moderne tecnologie informatiche, quindi sono impiegate "*metodologie basate su tecniche di analisi di carattere sia deterministico che stocastico, che vengono implementate sui soli dati contenuti nell'Archivio dei rapporti finanziari, ovvero su detti dati collegati, preventivamente o successivamente, con quelli contenuti nelle altre basi dati*"[410] ed, inoltre, possono essere utilizzare anche tecniche di *machine learning* o altre soluzioni di intelligenza artificiale.

Anche in questo specifico trattamento, come in tutti quelli basati su analisi probabilistiche ed inferenziali, si presenta la contraddizione, esaminata in precedenza, rispetto al principio di finalità, poiché queste analisi permettono di ottenere un *output* che non può essere preventivato, in tal modo risulta difficile determinare *ex ante* quali "*tipologie, quantità, varietà di dati risultano utili all'evidenziazione di una certa correlazione significativa*"[411] e per quale finalità "*quel dato trattamento risulterà utile,*

---

[409] Come viene precisato in B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, cit., p.154, la legge n. 160/2019 aggiunge una finalità (oltre a quella dell'analisi del rischio fiscale) e la possibilità di utilizzare le altre banche dati (oltre all'anagrafe tributaria) ad un trattamento che era già ammesso ai fini dell'analisi del rischio fiscale attraverso l'utilizzo dell'archivio dei rapporti finanziari.

[410] L'analisi stocastica, a differenza dell'analisi deterministica (che studia relazioni non probabilistiche), utilizza tecniche di *machine learning* (che, come noto, si basa su analisi probabilistiche) e questi metodi di analisi possono essere di tipo supervisionato o non supervisionato. Secondo il *Documento di Valutazione di Impatto sulla Protezione dei Dati (Stralcio)*, pubblicato sul portale dell'AdE l'obiettivo del *machine learning* supervisionato "*è quello di costruire, partendo da un'ipotesi inferenziale, un algoritmo in grado di elaborare delle previsioni sui valori di uscita di una variabile (output) rispetto ad una serie di altre variabili (input), identificando le relazioni esistenti fra esse. La variabile di output può essere, ad esempio, la classificazione di un soggetto come "rischioso" o "non rischioso", in base agli esiti di pregresse attività istruttorie o di accertamento condotte nei confronti di altri soggetti*". Invece, i metodi di *machine learning* non supervisionato "*non prevedono la presenza di un profilo noto a priori su cui addestrare l'algoritmo, ma sono orientati ad evidenziare in maniera autonoma eventuali schemi ricorrenti presenti nei dati. Il tipico esempio di utilizzo di metodi non supervisionati consiste nell'individuazione di gruppi di soggetti omogenei rispetto alle caratteristiche descritte dai dati*".

[411] B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, cit., p.155.

strumentale, necessario"[412]. Per questa ragione, le misure per mitigare e gestire i rischi per i diritti e le libertà degli interessati, vanno valutati in concreto, vale a dire "*tenendo in considerazione le caratteristiche delle banche dati di volta in volta utilizzate e i modelli di analisi impiegati (es. quelli di analisi deterministica o stocastica, oppure una combinazione di essi)*"[413]. In altre parole, non potendo definire e conoscere *ex ante* i criteri per selezionare le liste dei contribuenti con un maggior rischio[414], bisogna tenere traccia dell'analisi eseguita (basi di dati e modelli di algoritmi utilizzati) e verificare che la qualità e la quantità dei dati sia appropriata[415].

---

[412] *Ibidem*.

[413] Provvedimento n. 276 del 30 luglio 2022 del Garante della *privacy*, punto 4.

[414] B. Ponti, *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità,* p.157.

[415] Provvedimento n. 276 del 30 luglio 2022 del Garante della *privacy*, punto 4. Nonostante vi sia questa collisione con il principio di finalità, la normativa prevede sempre che i dati siano trattati per le finalità di cui al comma 682, art. 1, legge n. 160/2019 e conservati per un determinato periodo di tempo, in relazione alla finalità, anche se spesso ci si domanda della portata effettiva di tali disposizioni. Ad esempio, il d.m. 28 giugno 2022, oltre a riconoscere garanzie generiche e ridondanti (si veda il comma 6 dell'art. 4 ed il comma 5 dell'art. 5), sancisce, all'art. 3, commi 3 e 4, che i " *dataset sono conservati fino al secondo anno successivo a quello in cui matura la decadenza della potestà impositiva e, comunque, fino alla definizione di eventuali giudizi*" e che comunque in questo periodo venga escluso il trattamento dei *dataset* "*per finalità diverse dall'esercizio del diritto di accesso*". Decorsi tali periodi i *dataset* vengono cancellati, ferma restando la conservazione dei dati contenuti nelle banche dati dell'Agenzia secondo i criteri di conservazione stabiliti in relazione alle finalità per le quali ciascun dato è stato raccolto. In queste disposizioni è possibile osservare come il principio di finalità deve essere sempre correlato alla conservazione ed al trattamento del dato e del *dataset*. La possibilità di conservare i *dataset* dopo la scadenza della potestà impositiva dell'amministrazione finanziaria è stata considerata in dottrina *contra legem,* rispetto alla normativa del GDPR, poiché se la finalità del trattamento dovrebbe essere esclusivamente quella di cui al comma 682, art. 1, l. n. 160/2019 (analisi del rischio fiscale ed adempimento spontaneo), essa dovrebbe fisiologicamente esaurirsi con la cessazione della potestà impositiva. In quest'ultimo senso si veda C. Nocera, *Periodo di conservazione dei dati dei contribuenti non in linea con il GDPR*, in quotidianopiù, Giuffrè, 6 luglio 2022, reperibile al sito https://www.quotidianopiu.it/dettaglio/9951967/periodo-di-conservazione-dei-dati-dei-contribuenti-non-in-linea-con-il-gdpr . Sembrerebbe che con il decorrere del tempo e con il cessare della potestà impositiva muti il trattamento e la finalità, che diventa esclusivamente quella di accesso.

Oltre che con il principio di finalità, il Garante della *privacy* nel provvedimento n. 276 del 30 luglio 2022 sottolinea anche i pericoli, nonostante la pseudonimizzazione dei dati[416], in relazione alla re-identificazione degli interessati[417]. Per mitigare i rischi connessi alla sicurezza dei dati e alle libertà degli interessati è previsto che i dati non siano semplicemente pseudonimizzati con sostituzione o modifica delle informazioni, ma anche con perturbazioni delle variabili[418]. Tuttavia, nonostante tali tecniche adoperate dall'Agenzia delle entrate, il Garante, anche sotto questo profilo, ha formulato la specifica ingiunzione di "*adottare efficaci tecniche di pseudonimizzazione dei dati nell'ambito dei trattamenti in esame*"[419].

Anche la qualità dei dati, requisito di cui si è già discusso in precedenza per l'importanza dell'analisi algoritmica, assume un ruolo rilevante per l'analisi del rischio fiscale. Sarà quindi necessaria non solo una maggior quantità di dati, ma dati esatti, pertinenti e aggiornati (non obsolescenti)[420]. L'inesattezza del dato può condurre ad un'erronea

---

[416] Pseudonimizzazione prevista come facoltativa dall'articolo 1, comma 682, della legge 27 dicembre 2019, n. 160 e come obbligatoria dall'art. 5, comma 5, d.m. 28 giugno 2022.

[417] Tale rischio è stato già esaminato in astratto nel secondo capitolo e qui lo ritroviamo in un caso concreto.

[418] In tal senso art. 5, comma 5, d.m. 28 giugno 2022"*In particolare, l'Agenzia, anche per rafforzare le garanzie connesse al trattamento dei dati personali, effettua le elaborazioni finalizzate a far emergere le posizioni da sottoporre a controllo su dati preventivamente pseudonimizzati, attraverso metodi di sostituzione o modifica delle informazioni anagrafiche ovvero tramite perturbazioni delle variabili, al fine di impedire, in presenza di dati finanziari, l'identificazione diretta degli interessati*". Come sottolinea Bongi, *Intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.54, si tratta di tecniche nuove che si sono rese necessarie in seguito alle criticità rilevate dal Garante della *privacy* nel parere del 22 dicembre 2021.

[419] Il Garante, sempre nel provvedimento del 30 luglio, sottolinea molti aspetti critici relativi ad un'insufficiente pseudonimizzazione, evidenziando anche i *deficit* tecnici. È utile notare che, come affermato dal Garante, ridurre i rischi di re-identificazione significa anche rispettare i principi di "*minimizzazione dei dati, di integrità e riservatezza, e di privacy by design e by default e l'adempimento degli obblighi di sicurezza (artt. 5, par. 1, lett. c) e f), 25 e 32 del Regolamento)*". Limitare tali rischi è anche una questione di sicurezza informatica, un aspetto cruciale quando si mira a concentrare in un'unica piattaforma una grande mole di dati, che diventa un obiettivo preferenziale per gli attacchi informatici.

[420] In tal senso si è espresso il Garante *privacy* nel corso dell'audizione del 6 luglio 2021 presso la Commissione parlamentare di vigilanza sull'Anagrafe tributaria nell'Indagine conoscitiva "*Digitalizzazione e interoperabilità delle banche dati fiscali*". Lo stesso direttore dell'Agenzia delle

rappresentazione (cc.dd. *bias*) della capacità contributiva e rivelarsi come un vero e proprio "*boomerang*"[421] per l'amministrazione finanziaria, con un maggior dispendio di risorse, il pericolo di incorrere in sanzioni ed il pericolo per le garanzie e le libertà dei contribuenti.

Volendo provare a riassumere le condizioni principali, sottolineate dal Garante nel provvedimento del 30 luglio 2022, che deve rispettare l'AdE per effettuare il trattamento ai sensi della legge n. 160/2019, è possibile elencare: l'individuazione e la documentazione delle banche dati utilizzate; la pubblicazione di un estratto della valutazione d'impatto sulla protezione dei dati; garanzia di un intervento umano costante nel processo (*human in the loop*) ed adeguata formazione del personale coinvolto nel trattamento; adozione di tecniche efficaci di pseudonimizzazione; effettuare un costante ed adeguato controllo della qualità (e quantità) dei dati e dei modelli di analisi impiegati.

Al termine di questo capitolo si è visto che, per sfruttare in concreto le potenzialità dei dati ai fini dell'analisi del rischio fiscale, è necessario utilizzare le moderne tecnologie di analisi dei dati ed assicurare l'interoperabilità delle banche dati (*Big Data*) a disposizione dell'amministrazione finanziaria. Inoltre, è indispensabile, così come è anche emerso nel terzo capitolo, assicurare non solo la quantità dei dati, ma anche la qualità. Sarà anche indispensabile fugare quei rischi che si sono precedentemente esaminati in via astratta, in particolare il rischio di re-identificazione degli interessati.

Entrate, Ernesto Maria Ruffini, sempre in audizione presso la suddetta commissione parlamentare di vigilanza, ha sottolineato come spesso le banche dati cui attinge l'Agenzia abbiano dati "*eterogenei per struttura e dimensione, (..) soggetti a rapidi cambiamenti e, quindi, anche a obsolescenza*", con il rischio di disallineamento del patrimonio informativo tra un archivio e l'altro. Bisognerebbe quindi porre attenzione, prima ancora di incentivare l'interconnessione e l'integrazione tra più banche dati, su una qualità dei dati funzionali allo scopo. L'importanza della qualità dei dati è riconosciuta anche nel d.m. del MEF del 28 giugno 2022 all'art. 5, commi 2 e 4, in cui si dispone che vengano utilizzate "*tutte le misure necessarie*" per escludere dati personali inesatti o non aggiornati.

[421] A. Bongi, *intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, cit., p.125

In ogni caso, bilanciando gli interessi in gioco (tra tutela dei diritti personali ed esigenze fiscali)[422] e considerando che tali metodologie apportano risultati concreti[423], appare sicuramente condivisibile la strada percorsa da tale amministrazione, dovendo sempre, tuttavia, assicurare le garanzie del contribuente, compresa la tutela essenziale dei dati personali.

[422] Per il difficile bilanciamento tra questi due valori si veda L. Izzo, *Il difficile rapporto tra il diritto alla privacy e il dovere di contribuzione alla spesa pubblica*, in *European journal of privacy law e technologies*, 2/2022, pp. 188-214.

[423] Nella nota di aggiornamento del documento di economia e finanza (anno 2023) viene espressamente riconosciuto che i risultati relativi alle maggiori imposte accertate si devono anche all'introduzione di nuove metodologie operative, che includono l'applicazione di tecniche di intelligenza artificiale nel processo di valutazione del rischio fiscale (atto reperibile sul portale del MEF https://www.finanze.gov.it/export/sites/finanze/.galleries/Documenti/Varie/Allegato_NADEF2023_RAPPORTO_EVASIONE_28set_finale.pdf).

CONCLUSIONI

L'evoluzione tecnologica continua a correre a velocità elevate. L'effetto dirompente che l'IA assume nel contesto economico e sociale è sotto gli occhi di tutti; ciò che (forse) non emerge a prima vista è l'importanza dei dati e di una strategia digitale organica e sistematica per arrivare a sfruttare compiutamente l'IA e le moderne tecniche di analisi dei dati.

Queste ultime, infatti, si nutrono di dati e contribuiscono ad alimentare quello che si potrebbe definire un cambio di paradigma relativo al modo di intendere i dati. Essi non sono più visiti come qualcosa di personale, ma come una materia prima o (per usare una metafora più nota) come petrolio, e, di conseguenza, andranno lavorati, elaborati, analizzati, al fine di estrarre informazioni utili.

Questo cambio di paradigma sembra riflettersi a livello normativo, attraverso misure che mirano a facilitare l'acquisizione, l'accesso, la circolazione ed il riutilizzo dei dati. Ne sono un esempio il c.d. decreto capienze e la Strategia europea per i dati.

Attraverso tali misure normative sembra che si cerchi di scardinare (in parte) i tradizionali baluardi che limitano la circolazione dei dati, come alcuni diritti di esclusiva previsti dalla Direttiva 96/9/CE ed il modello incentrato sulla "protezione"[424] dei dati previsto dal GDPR, atto normativo che continua ad essere quello centrale all'interno della Strategia per i dati; si interviene, sul versante della c.d. Direttiva *database*, in particolare prevedendo numerose eccezioni al diritto *sui generis*, tanto da far configurare tali eccezioni come la regola stessa nei contesti tecnologici dove sono maggiormente sentite le esigenze relative alla circolazione ed allo sfruttamento dei dati[425].

Invece, sul versante del GDPR, si interviene cercando di garantire una più ampia circolazione ed utilizzo dei dati personali. Invero, il GDPR presenta profili di criticità non solo con riferimento alle nuove esigenze di circolazione dei dati, ma anche rispetto alle

[424] Tale modello è stato definito "*unipolare*" dalla dottrina.

[425] Si pensi all'eccezione prevista dal *Data Act* per i *database* contenenti *machine-generated data*, alle eccezioni di *text and data mining* previste dalla Direttiva *Copyright*, o a quelle previste dalla Direttiva *Open Data* e dal *Data Governance Act*.

capacità delle nuove tecnologie di destabilizzare alcuni assi portanti della normativa, sicché oggi si parla di paralizzazione della *privacy*. Quest'ultima circostanza è emersa specialmente con riferimento al principio di finalità ed alla distinzione tra dato personale e non personale[426].

In ogni caso, le recenti misure legislative appartenenti alla Strategia europea per i dati mirano a soddisfare le nuove esigenze circolatorie dei dati, cercando non di proteggere il dato dal mercato, ma disciplinando il mercato stesso. A tal fine vengono in rilievo le norme volte a regolare gli intermediari dei dati e la qualità degli stessi, i flussi di informazione G2B e B2G (c.d. *reverse PSI*) ed il riutilizzo dell'informazione pubblica.

Tutte queste circostanze hanno fatto emergere il possibile passaggio da una "tutela" ad una "*governance*"[427] dei dati, in cui si cerca di soddisfare le esigenze di circolazione dei dati, senza che vengano meno i presidi fondamentali per la tutela dei diritti degli interessati.

La tendenza della Strategia europea per i dati è quindi quella di garantire una maggiore produzione e riutilizzo dei dati. In virtù di questa circostanza, è possibile affermare che tali misure normative contribuiscono ad alimentare il fenomeno e lo sfruttamento dei *Big Data*, ponendo i presupposti necessari per un loro proficuo utilizzo/impiego.

Inoltre vi è da sottolineare che, in ogni caso, è previsto un aumento dei *Big Data* con il proliferare dello IoT, del 5G, delle necessità di analisi in *real time* e dei numerosi dati destrutturati in circolazione, spesso difficilmente pre-elaborabili.

I fattori normativi esposti in precedenza e queste circostanze fattuali portano a ritenere che sempre più in futuro dovrà essere affrontata la sfida dei *Big Data*, i quali possono rivelarsi come una difficoltà da gestire o come una potenzialità da sfruttare a seconda dell'implementazione di apposite soluzioni tecnologiche e sistemi *hardware* e *software* in grado di maneggiarli.

---

[426] Si è visto che il principio di finalità, non solo è caratterizzato da tensioni intrinseche con le tecniche di immagazzinamento (*data lake*) e di analisi inferenziali dei dati, ma costituisce un ostacolo anche all'accentramento del patrimonio informativo della pubblica amministrazione ed alle comunicazioni dei dati personali A2A.

[427] Oppure, per utilizzare parole differenti, da un modello "*unipolare*" ad uno "*multipolare*".

La pubblica amministrazione, in virtù del patrimonio informativo di cui dispone, dovrebbe essere in particolar modo interessata ai meccanismi capaci di sfruttare tali potenzialità. Sfruttare queste potenzialità significherebbe rafforzare il potere conoscitivo ed essere in grado di fruire dei dati attraverso un approccio *data driven*. I *Big Data* costituiscono infatti terreno fertile per le tecniche di analisi algoritmiche dei dati (come l'intelligenza artificiale, il *machine learning* o il *deep learning*); tra di essi vi è un rapporto di dipendenza bilaterale: esse hanno bisogno di ingenti quantità di dati (possibilmente anche di buona qualità) e questi ultimi hanno bisogno delle moderne tecniche algoritmiche per essere elaborati ed analizzati in modo efficace. Inoltre, costituirebbe un approccio laconico garantire un'ampia circolazione ed acquisizione dei dati in assenza di soluzioni organizzative in grado di gestire tali dati.

Se dall'analisi del *framework* normativo è emersa la circostanza per cui è possibile creare i presupposti per sfruttare i *Big Data*, si pone la questione di analizzare se le modifiche organizzative in atto siano effettivamente in grado di sfruttare le potenzialità dei *Big Data*, in modo da rinforzare il potere conoscitivo delle pubbliche amministrazioni e la fruizione dei dati stessi.

Dal presente studio è emersa la circostanza per cui, sebbene le modifiche organizzative *in fieri* (come l'interoperabilità, il *cloud*, le basi di dati di interesse nazionale) siano volte a delineare maggiormente un modello di amministrazione intermediaria dei dati[428], esse contribuiscono anche a migliorare un altro aspetto relativo alla funzione amministrativa dei dati: la fruizione degli stessi. Sicché viene riscontrata una parziale coincidenza dei mezzi per fini differenti: l'affermazione di una nuova pubblica amministrazione che si pone non solo come intermediaria dei dati, ma anche come loro fruitrice per le esigenze di carattere pubblico.

Quest'ultima circostanza emerge in particolare dalla strategia dell'amministrazione finanziaria per migliorare l'analisi del rischio fiscale. L'Agenzia delle Entrate, infatti, mira a gestire i propri *Big Data* attraverso l'interoperabilità[429] delle banche dati e le moderne

[428] Dettaglio che si rileva, ad esempio, del passaggio del progetto dell'allora DAF all'attuale PDND.

[429] Sebbene dalle analisi effettuate sia emerso che in un contesto *Big Data* siano più difficili da conservare i livelli di interoperabilità, gli interventi normativi in tale ambito non possono che

tecniche di analisi algoritmica, in modo da poter "profilare" i contribuenti (similmente al modo con cui le imprese profilano i propri clienti).

La circostanza appena menzionata emerge anche dal *template* del piano dei fabbisogni del Polo Strategico Nazionale e dalla strategia europea *cloud to edge* prevista nella Comunicazione *Bussola Digitale 2030*.

Il prevalente modello di amministrazione intermediaria dei dati emerge anche dall'analisi del *framework* normativo sulla circolazione dei dati esposto in precedenza. Si pensi, ad esempio alle norme volte a disciplinare il riutilizzo dell'informazione pubblica e al paradigma degli *open data* (fatto proprio dal legislatore europeo e nazionale)[430]; tali misure, previste dalla Direttiva *Open Data*, dal *Data Governance Act* beneficiano in particolar modo i privati e, più in generale, la collettività. Tuttavia, anche in tale *framework* normativo, si riscontrano interventi volti a rinforzare il potere conoscitivo, si pensi ai meccanismi di c.d. *reverse PSI* previsti dal *Data Act* o al c.d. decreto capienze.

In altre parole, dall'analisi combinata del *framework* normativo sulla circolazione dei dati e sulle modifiche organizzative in atto, si riscontra che, sebbene a livello normativo vengano creati i presupposti e si cerchi di alimentare maggiormente le potenzialità correlate ai *Big Data*, a livello organizzativo si sta costruendo un modello che restituisce principalmente alle pubbliche amministrazioni una funzione di mera intermediazione dei dati, che si concretizza attraverso i meccanismi dell'interoperabilità e degli *open data*, in modo da stimolare il progresso economico-sociale ed una fornitura più efficace dei servizi pubblici.

Tuttavia, tale modello consente in parte anche di rafforzare il modello di un'amministrazione fruitrice dei dati, così da potenziare il potere conoscitivo delle istituzioni pubbliche e permettere loro di sfruttare, forse in futuro, le potenzialità correlate ai *Big Data*.

---

apprezzarsi, in quanto l'interoperabilità è funzionale a molteplici scopi e, inoltre, anche per i *Big Data* dovrebbe essere sempre importante assicurare la qualità e l'interoperabilità del dato.

[430] Tale paradigma sembra principalmente volto all'intermediazione dei dati a vantaggio di tutta la collettività e l'economia. Tuttavia, bisogna considerare che la stessa amministrazione potrà usufruire dei dati aperti di un'altra amministrazione.

BIBLIOGRAFIA

Adrejevic M., *The Big Data Divide*, in *International Journal of Communication*, 8, 2014, pp. 1673 ss.

Aggarwal C.C*., Neural Networks and Deep Learning, a textbook*, II ed., Springer, 2023, pp. 3 e 74.

Alberti I., *La creazione di un sistema informativo unitario pubblico con la Piattaforma digitale nazionale dati*, in *Le istituzioni del federalismo*, n. 2/2022, pp. 473-494.

Alberti I., *La partecipazione procedimentale per legittimare gli algoritmi nel procedimento amministrativo*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 285-298.

Alcidi C., Gros D., *Next Generation EU: A Large Common Response to the COVID-19 Crisis*, Vol.55, n. 4, 2020, pp. 202-203.

Aldinucci M., *La pubblica amministrazione con i Big Data*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 225-231.

Aliprand S., *Interoperability and Open Standards: The key to true openness and innovation*, in *The Journal of Open Law, Technology and Society* (Jolts), Vol. 3, n. 1, 2011, pp. 5-24.

Alvaro S., L. Marzialetti, *Neutrality of financial regulation. Il (nuovo) ruolo del regolatore*, in A. Genovese e V. Falce (a cura di), *La portabilità dei dati in ambito finanziario*, Quaderni FinTech, CONSOB, 2021, pp. 39-74.

Amidei A., Maggiolino M., *Intelligenza artificiale, dati digitali e proprietà intellettuale*, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione?,* volume III, Bologna, Il Mulino, 2022, pp. 53-89.

Annet L., Laurent D., Madera C., *Data lakes*, Vol.2, ISTE Ltd, London, 2020.

Ascarelli T., *Teoria della concorrenza e dei beni immateriali: lezioni di diritto industriale,* II ed., Milano, Giuffrè, 1957.

Atzeni P., Ceri S., Paraboschi S., Torlone R., *Basi di dati. Modelli e linguaggi di interrogazione*, II ed., Milano, McGraw-Hill, 2006, p.5.

Azad P., Navimipour N.J., Rahmani A.M. et al., *The role of structured and unstructured data managing mechanisms in the Internet of things. Cluster Comput* 23, 2020.

Baldassarre A., *Dalla gestione dell'emergenza a una visione condivisa di futuro: il ruolo del digitale per lo sviluppo sostenibile del paese*, in *FPA - Annual Report 2020*, Edizioni FORUM PA, 2021.

BassiI M., *Interoperabilità e data governance nella nuova normalità, cosa resterà di quest'anno*, in *FPA - Annual Report 2020*, Edizioni FORUM PA, 2021.

Beghin M., *Diritto tributario-Principi, istituti e strumenti per la tassazione della ricchezza*, Giappichelli, Torino, 2011.

Bernasconi S., *Il Programma strategico dell'Unione europea per il decennio digitale 2030*, in *Rivista di diritto internazionale privato e processuale*, 1/2023, pp. 253-255.

Beye M.A, Laney D., *The importance of big data: A definition*, Stamford, Gartner Retrieved, 21 June 2012.

Bholasing J., *How Technological Advances in the Big Data Era Make it Impossible to Define the 'Personal' in GDPR's 'Personal Data'*, in *European Data Protection Law Review*, 3/2022, pp. 346-361.

Bombardelli M., *Dati personali (Tutela dei)*, in B.G. Mattarella e M. Ramajoli , *Funzioni amministrative – Enciclopedia del diritto - I tematici*, III volume, Milano, Giuffrè, 2020, pp. 360 ss.

Bondi B., *Characteristics of scalability and their impact on performance*, 2000, https://dl.acm.org/doi/10.1145/350391.350432.

Bongi A., *Ebook Intelligenza artificiale e fisco: come cambiano compliance, controlli e riscossione delle imposte*, Wolters Kluwer Italia, Milano, 2023.

Borgogno O., Colangelo G., *Data sharing and interoperability: Fostering innovation and competition through APIs*, in *Computer Law & Security Review*, vol. 35, n. 5, 2019.

Borgogno O., *Regimi di condivisione dei dati ed interoperabilità: il ruolo e la disciplina delle A.P.I.*, in *Il diritto dell'informazione e dell'informatica*, n. 3/2019, pp. 689-710.

Borruso R., Russo S., Tiberi C., *L'informatica per il giurista. Dal Bit a internet*, III ed., Milano, Giuffrè Editore, 2009.

Bravo F., *Il principio di solidarietà tra data protection e data governance,* in *Il diritto dell'informazione e dell'informatica*, n. 3/2023, pp. 481-518.

Bravo F., *Sul bilanciamento proporzionale dei diritti e delle libertà "fondamentali", tra mercato e persona: nuovi assetti nell'ordinamento europeo?,* in *Contratto e impresa*, n. 1/2018, pp.190-216.

Broomfield H., Reutter L., *Towards a Data-Driven Public Administration: An Empirical Analysis of Nascent Phase Implementation*, in *Scandinavian Journal of Public Administration*, 25 (2), 2021, p.75.

Buono G., *Rapporto 4/2022 – Intelligenza artificiale e amministrazioni Centrali*, in *BioLaw Journal – Rivista di BioDiritto*, n. 1/2022, pp. 261-264.

Buttarelli G., *L'interoperabilità dei dati nella Pubblica Amministrazione,* in V. Bontempi (a cura di), *Lo Stato digitale nel Piano nazionale di ripresa e resilienza*, Roma, Roma TrE-Press, 2022, pp. 141-147.

Buttarelli G., *La regolazione delle piattaforme digitali: il ruolo delle istituzioni pubbliche*, in *Giornale di diritto amministrativo*, 1/2023, pp. 116-127.

Califano L., Fiorillo V., Galli F., *La protezione dei dati personali: natura, garanzie e bilanciamento di un diritto fondamentale*, Torino, Giappichelli, 2023, pp. 137-193.

Califano L., *La protezione dei dati personali: natura, garanzie e bilanciamento di un diritto fondamentale,* Torino, Giappichelli, 2023, pp. 193-222.

Cardarelli F., *3-bis. Uso della telematica*, in M. A. Sandulli (a cura di) *Codice dell'azione amministrativa*, Milano, Giuffrè, 2010, pp. 427-428.

Cardarelli F., *Le banche dati pubbliche: una definizione*, in *Il diritto dell'informazione e dell'informatica*, n. 2/2002, pp. 321-341.

Cardarelli M.C., *banche dati e direttiva comunitaria: il diritto sui generis. La durata*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, pp. 64-85.

Cardone M., Foà D., *La valorizzazione del patrimonio informativo nell'ambito delle strategie di digitalizzazione della Pubblica Amministrazione*, in *Munus: Rivista giuridica dei servizi pubblici*, n. 3/2020, p. 609 ss.

Carinci A., Tassani T., *Manuale di diritto tributario*, Giappichelli editore, Torino, VI ed., 2023.

Carloni E., *Algoritmi su carta. Politiche di digitalizzazione e trasformazione digitale delle amministrazioni*, in *Diritto Pubblico*, 2/2019, pp. 363-391.

Carloni E., *Qualità dei dati, big data e amministrazione pubblica,* in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 117-131.

Carotti B., *La politica europea sul digitale: ancora molto rumore,* in *Rivista trimestrale di diritto pubblico*, n. 4/2022, pp. 997-1013.

Carullo G., *Big Data e Pubblica Amministrazione nell'era delle banche dati interconnesse, in Concorrenza e mercato*, n. 23/2016, pp. 181-204.

Carullo G., *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa*, Torino, Giappichelli, 2018.

Carullo G., *trattamento dei dati personali da parte delle pubbliche amministrazioni e natura del rapporto giuridico con l'interessato*, in *Rivista italiana di diritto pubblico comunitario*, n. 1-2/2020, pp. 134 ss.

Caso R., *Open data, ricerca scientifica e privatizzazione della conoscenza*, Trento law and technology research group, research paper n. 48, Università di Trento, 2022, pp. 1-33.

Cataleta A., *Innovazione della PA, un anno di fermento normativo tecnologico,* in *FPA-Annual Report 2023*, Edizioni Forum PA, 2023, pp. 140-142.

Cattari M., *Il nuovo Programma europeo "Europa digitale"(2021-2027) Proposta della Commissione Europea e documentazione*, in *DigItalia - Rivista del digitale nei beni culturali*, Roma, ICCU, Anno XV, Numero 1, 2020, pp. 125-130.

Čerešňák R., Kvet M., *Comparison of query performance in relational a non-relation databases, Transportation Research Procedia*, vol. 40, Elsevier BV, 2019, pp. 170-177.

Chen M., Mao S., Zhang Y., Leung V. C., *Big data: related technologies, challenges and future prospects*, New York, Springer, 2014.

Chimenti L., *Banche di dati e diritto d'autore*, Milano, Giuffrè, 1999, pp. 29 ss.

Chiti E., Marchetti B., Rangone N., *L'impiego di sistemi di intelligenza artificiale nelle pubbliche amministrazioni italiane: prove generali*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. I, Roma, Roma TrE-Press, 2023, pp. 132-160.

Clarich M., *Manuale di diritto amministrativo*, IV ed., Bologna, Il Mulino, 2019.

Clarizia P., Sgueo G., *Lo stato digitale nel PNRR: la digitalizzazione come necessità trasversale*, Irpa, Osservatorio sullo Stato digitale, 2021.

Cohen J.E., *Law for the Platform Econ*omy, 51 U.C. Davis L. Rev, 2017.

Conigliaro M., *Lotta all'evasione con l'intelligenza artificiale*, in *Il Fisco*, n. 32-33/2022, pp. 3107 ss.

Contrino A., *Banche dati tributarie, scambio di informazioni fra autorità fiscali e "protezione dei dati personali": quali diritti e tutele per i contribuenti?,* in *Rivista di diritto tributario supplemento online*, Pacini giuridica, 2019.

Contrino A., *Digitalizzazione dell'amministrazione finanziaria e attuazione del rapporto tributario: questioni aperte e ipotesi di lavoro nella prospettiva dei princìpi generali*, in *Rivista di diritto tributario*, n. 2/2023, pp. 105 ss.

Cortese B., *Commento all'art. 286 TCE*, in A. Tizzano (a cura di), *Trattati dell'Unione e della Comunità Europea*, Milano, Giuffrè, 2004, pp.1284-1287.

Cortese F., *Art. 2 sexies*, in R. D'Orazio, G. Finocchiaro, O. Pollicino, G. Resta (a cura di), *Codice della privacy e data protection*, Milano, 2021, pp. 1043 ss.

Costantino F., *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data*, in *Diritto pubblico*, n. 1/2019, pp. 43-70.

Cravo D.C., *How to Make Data Portability Right More Meaningful for Data Subjects?,* in *European Data Protection Law Review,* n. 1/2022, pp. 52-60.

Crovara M., *Informazioni non strutturate*, in *Atti del convegno-L'innovazione tecnologica e metodologica al servizio del mondo del lavoro-primo seminario-Roma, Centro Congressi Villa Eur*, Tipolitografia INAIL, Milano, 2009, pp. 10 ss.

Cuffaro V., D'Orazio R., *La protezione dei dati personali ai tempi dell'epidemia*, in *Il Corriere Giuridico*, n. 6/2020, pp. 729 ss.

D'Ancona S.M.A., *Trattamento e scambio di dati e documenti tra pubbliche amministrazioni, utilizzo delle nuove tecnologie e tutela della riservatezza tra diritto nazionale e diritto europeo*, in *Rivista italiana di diritto pubblico comunitario,* n. 3/2018, pp. 587-627.

Dacar R., *The Essential Facilities Doctrine, Intellectual Property Rights, and Access to Big Data*, in *IIC - International Review of Intellectual Property and Competition Law*, vol. 54, 2023, pp. 1487-1507.

De Ghetto M., *Corso di basi di dati I*, Lecce, Youcanprint, 2020.

De Lungo D., Marini F.S., *Scritti costituzionali sul Piano Nazionale di Ripresa e Resilienza*, Torino, Giappichelli, 2023.

De Tullio M.F., *La privacy e i big data verso una dimensione costituzionale Collettiva*, in *Politica del diritto*, 4/2016, pp. 637-696.

Delmastro M., Nicita A., *Big data. Come stanno cambiando il nostro mondo*, Bologna, Il Mulino, 2019.

Derclaye E., *Databases sui generis right: should we adopt the spin-off theory?*, in *EIPR*, 2004, pp. 81 ss.

Di Cataldo V., *Banche-dati e diritto sui generis: la fattispecie costitutiva*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, pp. 27 ss.

Di Cocco C., *Tutela delle banche di dati: patrimonio culturale e mercato unico digitale*, in *Aedon-Rivista di arti e di diritto on line*, n. 3/2020, https://aedon.mulino.it/archivio/2020/3/dicocco.htm.

Di Porto F., Grote T., Volpi G., Invernizzi R., *Talking at Cross Purposes? A computational analysis of the debate on informational duties in the digital services and the digital markets acts*, in G. Resta., V. Zeno -Zencovich (a cura di), *Governance of/through Big Data*, vol. I, Roma, Roma TrE-Press, 2023, pp. 299-352.

Durst L., *Oggetto e finalità: un nuovo statuto giuridico dei dati personali,* in R. Panetta (a cura di), *Circolazione e protezione dei dati personali, tra libertà e regole del mercato. Commentario al Regolamento UE n. 679/2016 e al d.lgs n. 101/2018,* Milano, Giuffrè, 2019, pp. 43 ss.

Engels B., *Data portability among online platforms*, in *Internet policy review-Journal on internet regulation*, vol.5, issue 2, 2016, pp. 1-17.

Erie M.S., Streinz T., *The Beijing Effect: China's 'Digital Silk Road' as Transnational Data Governance*, in *N.Y.U. journal of international law & politics*, 1, 54, 2021.

Erl T., Puttini R., Mahmood Z, *Cloud computing: concepts, technology, & architecture*, II ed., Upper Saddle River, Prentice Hall, 2023.

Faini F., *Big Data e internet of things: data protection e data governance alla luce del regolamento europeo*, in G. Cassano, V. Colarocco, G. B. Gallus, F. P. Micozzi (a cura di), *Il processo di adeguamento al GDPR*, 2022, pp. 311 ss.

Faini F., *Big data, algoritmi e diritto*, in *Saggi – DPCE online*, n. 3/2019, pp. 1871 ss.

Falce V., *Big Data, dataset e diritti esclusivi. Liaisons dangereuses tra innovazione e mercato*, in V. Falce, G. Ghidini, G. Olivieri (a cura di), *Informazione e Big Data tra innovazione e concorrenza*, Milano, 2018, pp. 123-128.

Falcone M.*, Big Data e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica*, in *Rivista trimestrale di diritto pubblico*, n. 3/2017, pp. 601-639.

Falcone M., *Ripensare il potere conoscitivo pubblico tra algoritmi e big data*, Napoli, Editoriale scientifica, 2023.

Finocchiaro G., *Il principio di accountability,* in *Giurisprudenza italiana*, vol.171, 12/2019, pp. 2778-2783.

Floridi L., *Pensare l'infosfera*, Raffaello Cortina, Milano, 2020.

Foa S., *Intelligenza artificiale e cultura della trasparenza amministrativa. Dalle "scatole nere" alla "casa di vetro"?,* in *Diritto amministrativo*, n. 3/2023, pp. 526-536.

Franca S., *I dati personali nell'amministrazione pubblica. Attività di trattamento e tutela del privato*, Università degli Studi di Trento, Collana della facoltà di giurisprudenza, 44, 2023.

Franca S., *La semplificazione nelle modalità di trattamento dei dati personali da parte della pubblica amministrazione*, in *Diritto pubblico*, 2/2021, pp. 635-665.

Francioso C., *Intelligenza artificiale nell'istruttoria tributaria e nuove esigenze di tutela*, in *Rassegna tributaria*, 2023, pp. 47 ss.

Frosini T.E., *L'ordine giuridico del digitale*, in *Ceridap*, 2/2023, pp.36-65.

Galetta D.U., *Accesso civico e trasparenza della Pubblica Amministrazione alla luce delle (previste) modifiche alle disposizioni del Decreto Legislativo n. 33/2013*, in *Federalismi.it*, n. 5/2016.

Galli C., Bogni M., *I requisiti per la tutela IP dei Big Data*, in V. Falce, G. Ghidini, G. Olivieri (a cura di) *Informazione e Big Data tra innovazione e concorrenza*, Milano, 2018, pp. 97 ss.

Gallo F., *discrezionalità (voce), diritto tributario*, in *Enciclopedia del diritto*, agg. III, 1999, 9, pp. 539 ss.

Geiger C., Frosio G., Bulayenko O, *Text and Data Mining: Articles 3 and 4 of the Directive 2019/790/EU*, Centre for International Intellectual Property Studies (CEIPI) Research Paper No. 2019-08, reperibile a https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3470653.

Ghidini G., *Rethinking intellectual property: balancing conflicts of interest in the constitutional paradigm*, Cheltenham, Northampton, E. Elgar, 2018.

Giannantonio E., *Banche dati (tutela delle)*, in *Enciclopedia del Diritto*, Aggiornamento V – 2001, Giuffrè, pp. 130 ss.

Giovannini A ., *Il partenariato pubblico-privato nel nuovo codice dei contratti pubblici*, 2023, reperibile sul portale https://www.giustizia-amministrativa.it/-/158189-47.

Gobbato S., *Verso l'attuazione della direttiva (UE) 2019/1024 sul riutilizzo degli open data della PA: nuove opportunità per le imprese,* in *MediaLaws*, n. 2/2020, pp.247-261.

Gosis F., *Ente pubblico*, in *Enciclopedia del diritto*, Annali VII, 2014, pp. 570 ss.

Guarda P., Bincoletto G., *Diritto comparato della privacy e della protezione dei dati personali*, Milano, Ledizioni, 2023.

Holmes D.E., *Big Data. A very short introduction*, Oxford, 2017.

Holmes S., Sustain Cass R., *Il costo dei diritti: perché la libertà dipende dalle tasse,* Bologna, Il Mulino, 2000.

Huawei Technologies Co., Ltd. Author, *Cloud Computing Technology*, Springer nature Singapore, 2023.

Iacovelli D., Fontana M., *Nuove sfide della tecnologia e gestione dei rischi nella proposta di regolamento europeo sull'intelligenza artificiale: set di training, algoritmi e qualificazione dei dati. Profili critici*, in *Il diritto dell'economia*, STEM Mucchi Editore, n. 3/2022, pp. 107-139.

Iaselli M., *Il contratto di outsourcing in ambito informatico*, in *MediaLaws*, 2011, https://www.medialaws.eu/il-contratto-di-outsourcing-in-ambito-informatico/.

Iaselli M., *La normativa di riferimento*, in G. Cassano, V. Colarocco, G. B. Gallus, F. P. Micozzi (a cura di), *Il processo di adeguamento al GDPR*, II ed., Milano, Giuffrè, 2022, pp. 1-35.

Italiano G. F., *Le sfide interdisciplinari dell'intelligenza artificiale*, in *Analisi giuridica dell'economia*, n. 1/2019, p. 14.

Izzo L., *Il difficile rapporto tra il diritto alla privacy e il dovere di contribuzione alla spesa pubblica*, in *European journal of privacy law e technologies*, n. 2/2022, pp. 188-214.

Janiesch C., Zschech P., Heinrich K., *Machine learning and deep learning*, *Electronic Markets*, 09/2021, volume 31, fascicolo 3.

Kazeeva J., *Sui Generis Intellectual Property Protection- Comparison of EU and U.S. Regulatory Approaches*, Springer, 2024, pp. 11-21.

Keller P., *A vanishing right? The Sui Generis Database Right and the proposed Data Act*, Wolters Kluwer blog, 2022, reperibile all'indirizzo https://copyrightblog.kluweriplaw.com/2022/03/04/a-vanishing-right-the-sui-generis-database-right-and-the-proposed-data-act/.

Kerber W., Govern*ance of IoT Data: Why the EU Data Act Will not Fulfill Its Objectives,* in *GRUR International,* n. 2/2023, pp. 120-135.

Khan W., Kumar T., Zhang C., Raj K., Roy A.M., Luo B., *SQL and NoSQL Database Software Architecture Performance Analysis and Assessments—A Systematic Literature Review.,* in *Big Data and Cognitive Computing 2*, n. 7.97, 2023.

Kubicek H., Cimander R., Scholl H.J*., Organizational Interoperability in E-Government - Lessons from 77 European Good-Practice Cases,* 2011.

Kumari K.A., Sadasivam G.S., Dharani D., Niranjanamurthy M., *Edge Computing Fundamentals, Advances and Applications*, 1st Edition, Boca Raton, 2021, pp. 12-20.

Liberatore G., *Lotta all'evasione con consenso 2.0: meglio coordinarsi con l'UE*, in *Il fisco*, n. 12/2022, pp. 1131 ss.

Libertini M., *Il Regolamento europeo sui mercati digitali e le norme generali in materia di concorrenza*, in *Rivista trimestrale di diritto pubblico*, 4/2022, pp. 1069 ss.

Lillo R. - *Chief Data Officer* del team per la trasformazione digitale -, *Data & Analytics Framework*, in *Medium*, 13 febbraio 2017, reperibile al link https://medium.com/team-per-la-trasformazione-digitale/dati-interoperabili-open-pubblica-amministrazione-data-analytics-framework-4eb53dafd618.

Macrì I., *Cloud della Pubblica Amministrazione: una casa moderna per i dati degli Italiani*, in *Azienditalia*, n. 11/2021, pp. 1847 ss.

Macrì I., *Dalle infrastrutture digitali delle Amministrazioni al cloud, il nuovo regolamento per la sicurezza dei dati e dei servizi pubblici*, in *Azienditalia*, n. 3/2022, pp.488 ss.

Macrì I., *Il PNRR italiano per la digitalizzazione e l'innovazione della Pubblica Amministrazione*, in *Azienditalia*, n. 1/2022, pp. 38 ss.

Malgieri G., *'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?*, in *Journal of Internet Law*, vol.20, n. 5, 2016, pp. 4 ss.

Maltoni A., *Esercizio privato di pubbliche funzioni*, in *Enciclopedia del diritto*, Annali I, 2007, pp. 570 ss.

Mansani L., *Le eccezioni per estrazione di testo e dati, didattica e conservazione del patrimonio culturale*, in *AIDA. Annali italiani del diritto d'autore, della cultura e dello spettacolo*, 2019, pp. 3-21.

Manvi S., Shyam G., *Cloud Computing Concepts and Technologies*, Taylor & Francis Group, LLC, 2021, pp. 10-24.

Marchetti B., *La garanzia dello human in the loop alla prova della decisione amministrativa algoritmica*, in *BioLaw Journal –Rivista di BioDiritto*, 2/2021, pp. 367-385.

Margariti V., Anagnostopoulos D., Papastilianou A., Stamati T., Angeli S., *Assessment of organizational interoperability in e-Government: a new model and tool for assessing organizational interoperability maturity of a public service in practice,* ICEGOV 2020: Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance, 2020, p.300.

Marrone P., *Democrazia oscura e algocrazia*, in *Endoxa – Prospettive sul presente*, 5/23, 2020.

Mastromatteo A., *Arriva l'algoritmo antievasione, ecco a cosa serve e come funziona*, in *Agenza Digitale*, 5 luglio 2022, disponibile su

https://www.agendadigitale.eu/documenti/arriva-lalgoritmo-antievasione-ecco-a-cosa-serve-e-come-funziona/.

Mazzarella M., Ramotti C., *Pandemia e governo digitale*, in *Giornale di diritto amministrativo*, n. 3/2022, pp. 415-423.

McCarthy M.T., *The Big Data Divide and Its Consequences*, in S. Carta (a cura di) *Machine Learning and the City: Applications in Architecture and Urban Design*, Oxford, Wiley Blackwell, 2022, pp. 547-559.

Merloni F., *Data analysis e capacità conoscitive delle pubbliche amministrazioni*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, 107-117.

Montagnani M.L., *Dati e proprietà intellettuale in Europa: dalla "proprietà" all'"accesso",* in *Il diritto dell'economia*, STEM Mucchi Editore, n. 1/2020, pp. 555 ss.

Nambiar A., Mundra D., *An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management*, in *Big Data and Cognitive Computing*, 6, 2022, pp. 132 ss.

Nicotra I.A., *pandemia costituzionale*, Napoli, 2021.

Niestad M., *Interoperable Europe act,* European Parliament DG EPRS /Members' Research Service, 2024.

Nocera C., *Periodo di conservazione dei dati dei contribuenti non in linea con il GDPR*, in quotidianopiù, Giuffrè, 6 luglio 2022, reperibile al sito https://www.quotidianopiu.it/dettaglio/9951967/periodo-di-conservazione-dei-dati-dei-contribuenti-non-in-linea-con-il-gdpr.

Orlando S., *Il diritto di Text and Data Mining (TDM) non esiste*, in *Rivista italiana di informatica e diritto*, 5, n. 1/2023, pp. 67-81.

Ottolia A., *Big Data e innovazione computazionale,* Torino, Giappichelli, 2017.

Pagallo U., *Big data, open data e black box society*, in R.C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 49-55.

Pagliarin C., Perathoner C., Laimer S., *Il Next Generation EU e i piani nazionali di ripresa e resilienza*, Milano, Giuffrè, 2023.

Pagnanelli V., *Conservazione dei dati e sovranità digitale. Una rilettura della (big) data governance pubblica alla luce delle nuove sfide globali*, in *Rivista italiana di informatica e diritto*, n. 1/2021, pp. 13-28.

Parente S.A., *Le funzioni giurisdizionali e di accertamento tributario nell'epoca della data economy e dei sistemi di cloud computing: l'ausilio di intelligenze artificiali, big data e algoritmi informatici*, in *Euro-Balkan Law and Economics Review*, università degli studi di Bari Aldo Moro, n. 1/2023.

Parisi E., *Selezione del promotore e tutela giurisdizionale nel project financing a iniziativa privata*, in *Il diritto processuale amministrativo*, 4/2018, pp. 1483-1526.

Parsier E., *The Filter Bubble: what the Internet is Hidining from you*, Londra, Viking, 2011.

Patruno V., Ragone M.M., *Dati fulcro delle strategie per la PA digitale: tra strategie europee, interoperabilità e ruolo dei data manager*, in *FPA - Annual Report 2023*, Edizioni FORUM PA, 2023, pp. 125-127.

Peluso M.G*., Intelligenza Artificiale e dati di qualità: la tecnologia come valido alleato*, in *Rivista di diritto dei media*, n. 2/2022, pp. 330 ss.

Perin R.C., *Pubblica amministrazione e data analysis,* in R. C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 11-19.

Pierro M., Ragucci G., *Le analisi del rischio di evasione tra selezione dei contribuenti da sottoporre a controllo e accertamento "algoritmico"*, in G. Ragucci (a cura di), *Fisco digitale cripto-attività, protezione dei dati, controlli algoritmici*, Torino, Giappichelli, 2023, pp. 79-112.

Piras P., *Il tortuoso cammino verso un'amministrazione nativa digitale*, in *Il diritto dell'informazione e dell'informatica*, n. 1/2020, pp.43-65.

Poletti D., *Contact tracing a App Immuni: atto secondo, in Persona e mercato,* n. 1/2021, pp.92-101.

Poletti D., *Il trattamento dei dati inerenti alla salute nell'epoca della pandemia: cronaca dell'emergenza*, in *Persona e mercato*, n. 2/2020, pp.65-76.

Ponti B., *attività amministrativa e trattamento dei dati personali – gli standard di legalità tra tutela e funzionalità*, I ed., Milano, Franco Angeli, 2023.

Ponti B., *Coordinamento e governo dei dati nel pluralismo amministrativo,* in M. Pietrangelo (a cura di), *Scritti in memoria di Isabella D'Elia Ciampi*, in *Informatica e diritto*, vol. XVII, n. 1-2/2008, pp. 430 ss.

Ponti B., *L'amministrazione come fornitore e come fruitore di dati personali pubblici: sono praticabili soluzioni basate sulla Big Data Analytics/Machine Learning?*, in R. C. Perin (a cura di), *L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale*, Torino, Università degli Studi di Torino, 2021, pp. 233-251.

Prainsack B., *Data Donation: How to resist the iLeviathan*, in J. Krutzinna, L. Floridi (a cura di), *The Ethics of Medical Data Donation*, Cham, Springer, 2019, pp. 9-22.

Purpura A., *La frontiera dei Big data*, in Palumbo G. (a cura di), *Fisco e privacy. Il difficile equilibrio tra lotta all'evasione e tutela dei dati personali*, Pacini Giuridica, Pisa, 2021, , pp. 79 ss.

PwC EU Services, D05.02 *Big Data Interoperability Analysis*.

Ramge T., Mayer V.-Schönberger , *Fuori i dati! Rompere i monopoli sulle informazioni per rilanciare il progresso*, Egea, Milano, 2021.

Reichel J., *The European Strategy for Data and Trust in EU Governance. The Case of Access to Publicly Held Data*, in *Ceridap*, 4/2023, pp. 129-158.

Reichman J.H*., la guerra delle banche dati. Riflessioni sulla situazione americana,* in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997.

Resta G., *Pubblico, privato, collettivo nel sistema europeo di governo dei dati,* in *Rivista trimestrale di diritto pubblico*, n. 4/2022.

Resta G., *Towards a unified regime of data-rights? Rapport de synthèse*, in T. Pertot (a cura di), *Rechte an Daten*, 2020, pp. 971-995.

Rezzani A., *Big data. Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati*, Santarcangelo di Romagna, Maggioli, 2013.

Ribes Ribes A., *La inteligencia artificial al servicio del «compliance tributario»*, in *Revista española de derecho financiero*, 188, 2020, pp. 125 ss.

Ricchi M., *L'Architettura dei Contratti di Concessione e di Partenariato Pubblico Privato nel Nuovo Codice dei Contratti Pubblici (d.lgs. 50/2016)*, in *Rivista giuridica del Mezzogiorno*, 3/2016, pp. 811-828.

Ricci A., *Sulla "funzione sociale" del diritto alla protezione dei dati personali,* in *Contratto e impr*esa, n. 2/2017, pp.586 ss.

Rodotà S., *Tecnologie e diritti,* Bologna, Il Mulino, 2021.

Ruschemeier H., *Data Brokers and European Digital Legislation*, in *European Data Protection Law Review*, n. 1/2023, pp.27-38.

Sandulli A., *Pubblico e privato nelle infrastrutture digitali nazionali strategiche*, in *Rivista trimestrale di diritto pubblico*, n. 2/2021, pp. 513-527.

Sarpi F., *La regolazione di domani. Come adeguare il processo normativo alle sfide dell'innovazione*, in *Rivista Italiana di Politiche Pubbliche*, n. 3/2018, pp. 464 ss.

Satyanarayana G., Reddy et. al., in *IJCSE - International Journal on Computer Science and Engineering*, Vol. 02, No. 09, 2010, pp. 2865-2873.

Scalzini S., *Banche di dati, sfruttamento dei dati digitali e concorrenza*, Giappichelli, Torino, 2023.

Scalzini S., *L'estrazione di dati e di testo per finalità commerciali dai contenuti degli utenti. Algoritmi, proprietà intellettuale e autonomia negoziale*, in *Analisi Giuridica dell'Economia*, 1/2019, pp. 395-423.

Scalzini S., Maggiolino M., *Disciplina delle banche di dati e questioni di accesso e riutilizzazione dei dati digitali per il funzionamento dei sistemi di intelligenza artificiale: verso la necessità di una riforma*?, in A. Pajno, F. Donati, A. Perrucci (a cura di), *Intelligenza artificiale e diritto: una rivoluzione*?, volume III, Bologna, Il Mulino, 2022, pp. 173-209.

Schönberger V.M., Cukier K., *Big Data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, V ed., Milano, Garzanti, 2013.

Scopelliti M. A., *È ancora possibile la neutralità della tecnologica della normativa?*, in V. Falce (a cura di), *Strategia dei dati e intelligenza artificiale*, Torino, Giappichelli, 2023, pp. 213-223.

Sganga C., *Ventisei anni di Direttiva database alla prova della nuova strategia europea peri dati: evoluzioni giurisprudenziali e percorsi di riforma,* in *Il diritto dell'informazione e dell'informatica*, n. 3/2022, pp.651-704.

Sgueo G., *I servizi pubblici digitali*, in V. Bontempi (a cura di), *Lo stato digitale nel piano nazionale di ripresa e resilienza*, Roma, RomaTrE-Press, 2022, pp. 119-127.

Shabani M., *The Data Governance Act and the EU's move towards facilitating data sharing,* in *Molecular Systems Biology,* 17, 2021.

Signorile O., *La ricerca di dati su fonti aperte come nuovo strumento delle indagini fiscali*, in G. Ragucci (a cura di), *Fisco digitale cripto-attività, protezione dei dati, controlli algoritmici*, Torino, Giappichelli, 2023, pp. 113-127.

Spada P., *Banche dati e diritto d'autore (il "genere" del diritto d'autore sulle banche di dati)*, in *AIDA-annali italiani del diritto d'autore, della cultura e dello spettacolo*, 1997, pp. 5 ss.

Spagnuolo F., Sorrentino E., *Alcune riflessioni in materia di trasformazione digitale come misura di semplificazione*, in *federalismi.it*, n. 8/2021, pp. 275-287.

Streinz T., *The Evolution of European Data Law*, in P. Craig e G. de Búrca (a cura di), *The Evolution of EU Law*, Oxford, Oxford University Press, 2021, pp. 902 ss.

Taddeo M., *Data Philantropy and Individual Rights,* in *27 Minds and Machines*, Vol.27, 2017, pp. 1-5.

Taini T., *L'incidenza dei Big data e del machine learning sui principi alla base del Regolamento Europeo per la tutela dei dati personali (2016/679/UE) e proposte per una nuova normativa in tema di privacy*, in S. BONAVITA (a cura di), *Società delle tecnologie esponenziali e General Data Protection Regulation: Profili critici nella protezione dei dati*, Milano, Ledizioni, 2018, pp. 35-65.

Tarjáni A.J., Kalló N., Dobos I., *Evaluation of Digital Development Based on the International Digital Economy and Society Index 2020 Data*, in *Statistika: statistics and economic journal*, vol.3, 2023, pp. 355–373.

Tigano F., *Protezione dei dati personali e pubblica amministrazione: alcuni spunti di riflessione,* in *Diritto e società*, n. 2/2022, pp. 418 ss.

Torchia L., *Lo Stato digitale. Una introduzione*, Bologna, Il Mulino, 2023.

Torregiani S., *Il dato non personale alla luce del Regolamento (UE) 2018/1807: tra anonimizzazione, ownership e Data by Design, in federalismi.it*, n. 18/2020, pp. 319 ss.

Troiano S*., Il diritto alla portabilità dei dati,* in N. Zorzi Galgano (a cura di) *Persona e mercato dei dati: riflessioni sul GDPR*, Milano, Wolters Kluwer, 2019, pp. 195 ss.

Vaccari S., *Funzione tributaria e diritto amministrativo*, in *Diritto pubblico*, n. 2/2022, pp. 493-545.

Válková L., *Bussola per il digitale 2030: il modello europeo per il decennio digitale*, in *Rivista di diritto internazionale privato e processuale*, 2/2021, pp.469-473.

Van Eechoud M., *A Serpent Eating Its Tail: The Database Directive Meets the Open Data Directive,* in *IIC - International Review of Intellectual Property and Competition Law*, Vol.52, 04/2021, pp. 375-378.

Van Ooijen C., Ubaldi B., Welby B., *A data-driven public sector: Enabling the strategic use of data for productive, inclusive and trustworthy governance*, OECD Working Papers on Public Governance, No. 33, OECD Publishing, Paris, 2019.

Vanz G., *I principi della proporzionalità e ragionevolezza nelle attività conoscitive e di controllo dell'Amministrazione finanziaria*, in *Diritto e pratica tributaria*, n. 5/2017, pp.1912 ss.

Vargo D., Zhu L., Benwell B., *Digital technology use during COVID-19 pandemic: A rapid review*, in *Human Behavior and Emerging Technologies*, vol. 3, n. 1, 2020, pp. 13– 24.

Vari F., Piergentili F., *"To no other end, but the... Safety, and publick good of the People": le limitazioni alla protezione dei dati personali per contenere la pandemia di Covid-19*, in *Rivista AIC*, 2021, pp. 328-342.

Vigorito A., *Government Access to Privately-Held Data: Business-to-Government Data Sharing. Voluntary and Mandatory Models*, in G. Resta, V. Zeno-Zencovich (a cura di), *Governance of/through Big Data*, vol. II, Roma, Roma TrE-Press, 2023, pp. 697-720.

Warren D., Brandeis L.D., *The right to privacy*, in *Harvard law review*, Vol.VI, n. 5, 1890.

Yeh C.L*., Pursuing consumer empowerment in the age of big data: A comprehensive regulatory framework for data brokers, in Telecommunications* policy, 05/2018, Volume 42, Fascicolo 4, pp. 282-292.

Yulianto A.A., *Extract Transform Load (ETL) Process in Distributed Database Academic Data Warehouse*, in *APTIKOM Journal on Computer Science and Information Technologies*, vol. 4, no. 2, Institute of Advanced Engineering and Science, July 2019, pp. 64-71.

Zeno-Zencovich V., *Big Data e epistemologia giuridica*, in G. Resta, V. Zeno -Zencovich (a cura di), *Governance of/through Big Data*, vol. II, Roma, Roma TrE-Press, 2023, pp. 439-449.

Zeno-Zencovich V., *Do "data markets" exist?,* in *MediaLaws*, n. 2/2019, pp. 22 ss.