La gestione degli archivi sanitari e la protezione dei dati personali, tra pseudonimizzazione e anonimizzazione

officinadellastoria.eu/it/2020/06/29/la-gestione-degli-archivi-sanitari-e-la-protezione-dei-dati-personali-tra-pseudonimizzazione-e-anonimizzazione/

Francesco Ciclosi

Giugno 29, 2020

Introduzione

In questo articolo si approfondirà l'utilizzo di alcune misure tecniche per la protezione dei dati, quali la pseudonimizzazione e l'anonimizzazione, con particolare riferimento alla loro applicazione nell'ambito degli archivi. Contestualmente si analizzeranno alcune delle criticità connesse alla gestione degli archivi sanitari, contenenti per loro natura dati personali appartenenti a particolari categorie di dati personali (come i dati genetici, i dati biometrici e i dati relativi alla salute), e a come poterle superare attraverso la scelta di misure tecniche che siano adeguate. Nell'ambito di questa analisi si considereranno, in aggiunta alle previsioni contenute nella normativa in materia di protezione dei dati personali (sia Unionista che della Repubblica Italiana) anche le linee guida dell'*European Archives Group*, nonché i risultati consolidati della letteratura scientifica in materia di protezione dei dati.

Il trattamento dei dati genetici, dei dati biometrici e dei dati relativi alla salute

La comprensione di quelle che potrebbero essere le criticità degli archivi sanitari con particolare riferimento ai dati personali, non può prescindere dalla preventiva definizione di quali categorie di dati potrebbero rientrare sotto questa denominazione. Facendo riferimento a quanto indicato nel Regolamento (UE) 2016/679[1] tali dati appartengono alle categorie particolari di dati personali ex articolo 9, con particolare riferimento alle sottocategorie dei dati genetici, dei dati biometrici e dei dati relativi alla salute.

Il *Regolamento generale sulla protezione dei dati* definisce dati genetici, "*i 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*"[2] [1]. Nello specifico il Regolamento (UE) 2016/679 fornisce anche ulteriori indicazioni in merito al

campione biologico oggetto di esame, specificando come tali informazioni univoche sulla persona fisica possano in particolar modo derivare "*dall'analisi dei cromosomi, dell'acido desossiribonucleico (DNA) o dell'acido ribonucleico (RNA), ovvero dall'analisi di un altro elemento che consenta di ottenere informazioni equivalenti*"[3] [1].

Vengono invece definiti dati biometrici, quei "*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* ***[4]*** *facciale o i dati dattiloscopici*"[5] [1].

Infine, il Regolamento (UE) 2016/679, definisce dati relativi alla salute, quei "*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*"[6] [1]. Il Considerando 35 fornisce ulteriori informazioni, specificando come si debba fare riferimento alle informazioni tali da rivelare lo stato di salute fisica o mentale di un interessato senza limiti temporali, essendo d'interesse sia quella passata, che quella presente, così come quella futura. Inoltre, sempre lo stesso Considerando precisa che i dati riguardanti lo stato di salute dell'interessato "*comprendono informazioni sulla persona fisica raccolte nel corso della sua registrazione al fine di ricevere servizi di assistenza sanitaria*"[7] [1], così come:

- "*un numero, un simbolo o un elemento specifico attribuito a una persona fisica per identificarla in modo univoco a fini sanitari* ***[8]****;*
- *le informazioni risultanti da esami e controlli effettuati su una parte del corpo o una sostanza organica, compresi i dati genetici e i campioni biologici*"[9] [1];
- "*qualsiasi informazione riguardante, ad esempio, una malattia, una disabilità, il rischio di malattie, l'anamnesi medica, i trattamenti clinici o lo stato fisiologico o biomedico dell'interessato, indipendentemente dalla fonte, quale, ad esempio, un medico o altro operatore sanitario, un ospedale, un dispositivo medico o un test diagnostico in vitro*"[10] [1].

Il principio generale sancito dall'articolo 9, paragrafo 1 del Regolamento (UE) 2016/679 è che è vietato trattare dati personali appartenenti a categorie particolari di dati personali, salvo che non si verifichino alcune specifiche circostanze[11]. Nella Tabella 1 sono sinteticamente descritte le casistiche per le quali, seppur a determinate condizioni, è ammissibile il trattamento dei dati appartenenti a categorie particolari di dati personali.

Tabella 1 – Casistiche in cui è ammissibile il trattamento dei dati appartenenti a particolari categorie di dati personali

Descrizione della casistica	Riferimento normativo
L'interessato ha prestato il proprio consenso esplicito al trattamento di tali dati personali per una o più finalità specifiche	GDPR, artt. 9(1) e 9(2)(a)
Il trattamento riguarda dati personali resi manifestamente pubblici dall'interessato	GDPR, art. 9(2)(e)
Il trattamento è necessario per assolvere gli obblighi ed esercitare i diritti specifici del titolare del trattamento o dell'interessato in materia di diritto del lavoro e della sicurezza sociale e protezione sociale	GDPR, art. 9(2)(b)
Il trattamento è necessario per tutelare un interesse vitale dell'interessato o di un'altra persona fisica	GDPR, art. 9(2)(c)
Il trattamento è effettuato, nell'ambito delle sue legittime attività e con adeguate garanzie, da una fondazione, associazione o altro organismo senza scopo di lucro che persegua finalità politiche, filosofiche, religiose o sindacali, a determinate e stringenti condizioni	GDPR, art. 9(2)(d)
Il trattamento è necessario per accertare, esercitare o difendere un diritto in sede giudiziaria	GDPR, art. 9(2)(f)
Il trattamento è necessario per motivi di interesse pubblico rilevante sulla base del diritto dell'Unione o degli Stati membri	GDPR, art. 9(2)(g)
Il trattamento è necessario per finalità di medicina preventiva o di medicina del lavoro, valutazione della capacità lavorativa del dipendente, diagnosi, assistenza o terapia sanitaria o sociale, ovvero gestione dei sistemi e servizi sanitari o sociali, sulla base del diritto dell'Unione o degli Stati membri	GDPR, artt. 9(2)(h), 9(3)
Il trattamento è necessario per motivi di interesse pubblico nel settore della sanità pubblica, quali la protezione da gravi minacce per la salute a carattere transfrontaliero o la garanzia di parametri elevati di qualità e sicurezza dell'assistenza sanitaria e dei medicinali e dei dispositivi medici, sulla base del diritto dell'Unione o degli Stati membri	GDPR, art. 9(2)(i)
Il trattamento è necessario a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici, sulla base del diritto dell'Unione o degli Stati membri	GDPR, artt. 9(2)(j), 89(1)

Tabella 1

In particolar modo, il Regolamento (UE) 2016/679 prevede che il trattamento a fini di archiviazione nel pubblico interesse, di ricerca scientifica o storica o a fini statistici sia soggetto a garanzie adeguate per i diritti e le libertà dell'interessato, specificando che "*tali garanzie assicurano che siano state predisposte misure tecniche e organizzative, in particolare al fine di garantire il rispetto del principio della minimizzazione dei dati*"[12] [1]. Più in dettaglio è previsto che tali misure "*possono includere la pseudonimizzazione, purché le finalità in questione possano essere conseguite in tal modo*"[13] [1], diversamente, nel caso in cui sia possibile conseguirle attraverso l'anonimizzazione, ovvero attraverso "*il trattamento ulteriore che non consenta o non consenta più di identificare l'interessato, tali finalità devono essere conseguite in tal modo*"[14] [1].

Inoltre, il *Regolamento generale sulla protezione dei dati* introduce anche la possibilità che il diritto dell'Unione o degli Stati membri preveda delle deroghe a specifici diritti degli interessati, nel caso in cui i dati personali siano trattati a fini di ricerca scientifica o storica, oppure a fini statistici, o anche per finalità di archiviazione nel pubblico interesse. In particolar modo, i diritti che possono essere soggetti a deroga sono quelli di accesso[15], di rettifica[16], di limitazione del trattamento[17], e di opposizione[18]. A queste limitazioni possono aggiungersi, ma nel solo caso di trattamenti effettuati per finalità di archiviazione nel pubblico interesse, le deroghe del diritto alla portabilità dei dati[19] e dell'obbligo di notifica[20] in caso di rettifica o cancellazione dei dati personali o limitazione del trattamento.

Il Regolamento (UE) 2016/679 prevede anche 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*"[21] [1]. Si tratta di un'opportunità colta dal legislatore italiano in occasione dell'adeguamento del d.lgs. 196/2003[22], a seguito del quale sono state stabilite delle ulteriori condizioni a cui un trattamento di dati genetici, biometrici e relativi alla salute deve essere conforme. Nello specifico, il novellato *Codice in materia di protezione dei dati personali* prevede che "*i dati genetici, biometrici e relativi alla salute, possono essere oggetto di trattamento in presenza di una delle*

*condizioni […] [*riassunte nella Tabella 1, nonché*] in conformità alle misure di garanzia disposte dal Garante*"[23] [2] e contenute in un apposito provvedimento adottato con cadenza almeno biennale. Il d.lgs. 196/2003 indica anche che il citato provvedimento del Garante dovrà essere adottato tenendo conto dei seguenti elementi specifici:

- "*delle linee guida, delle raccomandazioni e delle migliori prassi pubblicate dal Comitato europeo per la protezione dei dati e delle migliori prassi in materia di trattamento dei dati personali*"[24] [2];
- "*dell'evoluzione scientifica e tecnologica nel settore oggetto delle misure*"[25] [2];
- "*dell'interesse alla libera circolazione dei dati personali nel territorio dell'Unione Europea*"[26] [2].

Tali misure di garanzia vengono adottate relativamente a ciascuna delle citate categorie, "*avendo riguardo alle specifiche finalità del trattamento*"[27], nonché potendovi individuare "*ulteriori condizioni sulla base delle quali il trattamento di tali dati è consentito*"[28] [2].

Il novellato *Codice in materia di protezione dei dati personali* fornisce anche ulteriori dettagli relativi al contenuto delle misure di garanzia, chiarendo che queste "*individuano le misure di sicurezza, ivi comprese quelle tecniche di cifratura e di pseudonimizzazione, le misure di minimizzazione, le specifiche modalità per l'accesso selettivo ai dati e per rendere le informazioni agli interessati, nonché le eventuali altre misure necessarie a garantire i diritti degli interessati*"[29] [2]. D'interesse è anche la previsione che consente "*l'utilizzo dei dati biometrici con riguardo alle procedure di accesso fisico e logico ai dati da parte dei soggetti autorizzati*"[30] [2], purché ciò avvenga nel rispetto degli obblighi di sicurezza di cui all'articolo 32, del Regolamento (UE) 2016/679, nonché delle citate misure di garanzia.

Purtroppo, allo stato attuale dei fatti, non essendo ancora stato adottato il Provvedimento del *Garante per la protezione dei dati personali* contenente le misure di garanzia per il trattamento dei dati genetici, biometrici e relativi allo stato di salute, non sarà possibile fornire ulteriori indicazioni di dettaglio.

Gli archivi digitali e il GDPR, alla luce delle Regole deontologiche di settore

La finalità delle Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica è quella di "*garantire che l'utilizzazione di dati di carattere personale acquisiti nell'esercizio della libera ricerca storica e del diritto allo studio e all'informazione, nonché nell'accesso ad atti e documenti, si svolga nel rispetto dei diritti, delle libertà fondamentali e della dignità delle persone interessate*"[31] [3]. Il tutto con particolare riferimento a due diritti specifici, ovvero quello alla riservatezza e quello all'identità personale, che andranno tutelati con le soluzioni tecniche più adeguate, ivi incluse l'anonimizzazione e la pseudonimizzazione.

L'ambito di applicazione delle stesse è quello dei "*trattamenti di dati personali effettuati per scopi storici in relazione ai documenti conservati presso archivi delle pubbliche amministrazioni, enti pubblici ed archivi privati dichiarati di notevole interesse storico*"[32] [3].

Le Regole deontologiche contengono anche i "*principi-guida di comportamento dei soggetti che trattano per scopi storici dati personali conservati presso archivi pubblici e archivi privati dichiarati di notevole interesse storico*"[33] [3], con riferimento alle categorie degli archivisti[34] e degli utenti[35].

Per quanto concerne l'accesso agli archivi pubblici, dopo aver sancito che "*l'accesso agli archivi pubblici è libero*"[36] [3], le Regole tecniche vanno a definire quali sono i documenti che ai sensi delle leggi vigenti fanno eccezione a tale disposizione. Tra questi[37] si devono ricomprendere "*i documenti […] contenenti i dati di cui agli artt. 9, par. 1, e 10 RGPD, che divengono liberamente consultabili quaranta anni dopo la loro data*"[38] [3]. Si tratta di un termine che viene esteso a "*settanta anni se i dati sono relativi alla salute ovvero alla vita o all'orientamento sessuale oppure rapporti riservati di tipo familiare*"[39] [3].

Si consideri che è comunque possibile accedere a tali documenti prima della scadenza dei termini, richiedendo e ottenendo una speciale autorizzazione alla consultazione che sarà nel caso rilasciata dal Ministro dell'Interno[40]. Nel caso della presentazione della predetta richiesta di autorizzazione, l'utente dovrà trasmettere all'ente che conserva i documenti che si vuole consultare "*un progetto di ricerca che, in relazione alle fonti riservate per le quali chiede l'autorizzazione, illustri le finalità della ricerca e le modalità di diffusione dei dati*"[41] [3].

Nell'autorizzazione alla consultazione dei documenti, eventualmente rilasciata, possono essere contenute delle "*cautele volte a consentire la comunicazione dei dati senza ledere i diritti, le libertà e la dignità delle persone interessate*"[42] [3]. Le Regole deontologiche, dopo aver precisato che le cautele da adottare dipenderanno dai particolari obiettivi della ricerca, così come desumibili dal progetto, chiariscono anche in cosa possono consistere tali cautele, fornendo delle esemplificazioni a riguardo. Più in dettaglio tali cautele consistono "*nell'obbligo di non diffondere i nomi delle persone, nell'uso delle sole iniziali dei nominativi degli interessati, nell'oscuramento dei nomi in una banca*

dati, nella sottrazione temporanea di singoli documenti dai fascicoli o nel divieto di riproduzione dei documenti"[43] [3]; inoltre, particolare attenzione dovrà essere "*prestata al principio della pertinenza e all'indicazione di fatti o circostanze che possono rendere facilmente individuabili gli interessati*"[44] [3].

Particolari prescrizioni sono previste anche per la diffusione dei dati personali da parte degli utenti. Infatti, premesso che "*l'interpretazione dell'utente […] rientra nella sfera della libertà di parola e di manifestazione del pensiero costituzionalmente garantite*"[45] [3], questa dovrà svolgersi sempre "*nel rispetto del diritto alla riservatezza, del diritto all'identità personale e della dignità degli interessati*"[46] [3].

In particolar modo, nel caso in cui l'utente si trovi a effettuare dei riferimenti allo stato di salute delle persone fisiche, questo dovrà astenersi "*dal pubblicare dati analitici di interesse strettamente clinico e dal descrivere abitudini sessuali riferite ad una determinata persona identificata o identificabile*"[47] [3]. Inoltre, coerentemente con le previsioni di cui all'articolo 101, comma 2, del d.lgs. 196/2003, all'utente è consentito diffondere i dati personali solo "*se pertinenti e indispensabili alla ricerca e se gli stessi non ledono la dignità e la riservatezza delle persone*"[48] [3]. La valutazione di tale principio di pertinenza dovrà essere valutato al momento della diffusione dei dati, "*con particolare riguardo ai singoli dati personali contenuti nei documenti, anziché ai documenti nel loro complesso*"[49] [3].

Per l'effettiva implementazione delle cautele da adottare a protezione dei diritti e delle libertà delle persone fisiche, saranno di ausilio le misure tecniche della pseudonimizzazione e dell'anonimizzazione, di seguito descritte.

La pseudonimizzazione

Il paragrafo 5, dell'articolo 4, del Regolamento (UE) 2016/679 definisce la «pseudonimizzazione», come "*il trattamento dei dati personali in modo tale che i dati personali non possano più essere attribuiti a un interessato specifico senza l'utilizzo di informazioni aggiuntive*"[50] [1]. Dunque si richiede esplicitamente la presenza di una componente addizionale, rispetto ai dati personali originari, in grado di consentire l'attribuzione dei dati pseudonimizzati[51].

Lo stesso *Regolamento generale sulla protezione dei dati* pone poi delle condizioni specifiche che devono essere soddisfatte da queste informazioni aggiuntive, ovvero, da un lato, l'essere "*conservate separatamente*"[52] [1], e dall'altro, l'essere "*soggette a misure tecniche e organizzative intese a garantire che tali dati personali non siano attribuiti a una persona fisica identificata o identificabile*"[53] [1].

Quindi, la pseudonimizzazione si configura come un trattamento addizionale dei dati personali che richiede necessariamente la custodia separata dei dati pseudonimizzati e della parte addizionale. In generale il Regolamento (UE) 2016/679 non indica puntualmente il soggetto che deve custodire la parte addizionale, configurando pertanto come elemento di flessibilità la fattispecie che questa possa essere trattenuta direttamente dall'interessato e non dal titolare del trattamento. Si tratta dello scenario delineato dall'articolo 11 del GDPR, dove si precisa che "*se le finalità per cui un titolare del trattamento tratta i dati personali non richiedono o non richiedono più l'identificazione dell'interessato, il titolare del trattamento non è obbligato a conservare, acquisire o trattare ulteriori informazioni per identificare l'interessato al solo fine di rispettare il […] regolamento*"[54] [1].

A livello operativo, si può prendere un carattere dell'interessato ed effettuare un pre-trattamento, che dai dati personali originari, derivi un nuovo dato non immediatamente riferibile all'interessato. In particolar modo la conoscenza di questo nuovo dato non deve consentire di identificare da solo l'interessato, ovvero in sé deve essere un dato di tipo non identificativo. Si avrà dunque la presenza di due distinti elementi, uno non direttamente identificante, e un altro che, se utilizzato in aggiunta al primo, consente di risalire all'identità dell'interessato.

Da questo punto di vista il trattamento di pseudonimizzazione può essere considerato anche come una sorta di generatore di nuove identità, poiché consente di accompagnare all'identità di partenza quella che è ottenibile tramite l'utilizzo delle caratteristiche della funzione di pseudonimizzazione scelta.

Sebbene tale affermazione potrebbe essere non di immediata comprensione, va evidenziato come questa ricalchi quanto ognuno di noi ha modo di sperimentare direttamente nella propria esperienza quotidiana, laddove ogni persona fisica è interessata dall'avere associate molteplici identità che la caratterizzano, a seconda del contesto in cui si opera.

Del resto è possibile considerare l'uso che fin dall'antichità è stato fatto del patronimico, ovvero del "*nome o cognome, derivati dal nome del padre per mezzo di un suffisso*" [4], quale strumento di identificazione della persona fisica. Quindi, a titolo esemplificativo e senza bisogno di scomodare i classici della mitologia greca, è possibile pensare al patronimico *Ivanovič* o *Ivanovna*, a significare, rispettivamente, "*figlio di Ivan*" oppure "*figlia di Ivan*".

Ovviamente le identità associate a una persona fisica non hanno tutte lo stesso potere identificante, e soprattutto, anche considerando la medesima identità, la sua effettiva capacità di identificare un interessato dipende moltissimo dal contesto di riferimento. Un esempio può aiutarci a comprendere meglio questa affermazione. Consideriamo due persone fisiche Pietro e Paolo, dove il primo è il padre del secondo. Possiamo quindi costruire delle identità aggiuntive per *Pietro* e *Paolo*, denominandole rispettivamente, "*Padre di Paolo*" e "*Figlio di Pietro*". Nello specifico, il trattamento di pseudonimizzazione trasformerà il nome proprio *Pietro* (che costituisce la prima identità, quella principale) in "*Padre*", dato di per sé non identificante, ma che se utilizzato in combinazione con l'informazione aggiuntiva "*di Paolo*" consentirà di creare la nuova identità attribuibile all'interessato. Analogamente il nome proprio Paolo (identità principale) potrà essere trasformato in "*Figlio*", e quindi addizionato dell'informazione aggiuntiva "*di Pietro*", per creare la seconda identità dell'interessato.

Ciò ci consente di verificare in modo empirico e immediato come la scelta del contesto di riferimento non sia indifferente per la determinazione del potere identificante che ciascuna delle distinte identità ha per l'interessato. Infatti, se consideriamo il contesto del posto di lavoro di Pietro (il padre), il nome proprio è maggiormente identificante per Pietro, piuttosto che il suo essere il "*Padre di Paolo*", dato che tale circostanza magari risulta sconosciuta ai più. Così come, l'essere il "*Figlio di Pietro*" risulta maggiormente identificante per Paolo (il figlio), rispetto all'utilizzo del suo nome proprio, dato che nel contesto considerato, è presumibile che questo, pur non essendo conosciuto direttamente, sia sicuramente immediatamente conoscibile attraverso la relazione che lo lega al padre (questo sì conosciuto direttamente) ed espressa dall'identità aggiuntiva. Un discorso analogo e a parti inverse può essere effettuato se consideriamo il contesto della scuola frequentata da Paolo (il figlio), laddove questo verrà presumibilmente meglio identificato con il suo nome proprio (e non con l'essere il "*Figlio di Pietro*", dato che difficilmente questi sarà conosciuto nel contesto), mentre Pietro, sarà meglio identificato con la sua identità aggiuntiva che lo caratterizza come "*Padre di Paolo*", essendo quest'ultimo ben conosciuto nel contesto di riferimento.

Riassumendo, il trattamento della pseudonimizzazione può essere utilizzato per differenti scopi, ovvero sia per mascherare oppure offuscare l'identità di una persona fisica, sia come generatore di nuove identità associabili a un interessato. Quest'ultima circostanza trova la sua giustificazione nel fatto che in taluni casi può essere più tutelante per la persona fisica l'essere maggiormente identificata. Del resto si tratta di un concetto che trova un'analogia con quello più antico e noto di pseudonimo, definibile come un "*nome diverso da quello reale usato da uno scrittore, un poeta, un giornalista, un artista e sim. che non voglia o non possa firmare le proprie opere con il vero nome*" [5]. Tra l'altro lo pseudonimo può, a seconda dei casi, essere notorio o meno, e nel caso lo sia è equiparabile al nome anagrafico anche sotto un profilo giuridico, essendo "*tutelato […] con le stesse modalità che difendono il diritto al nome*" [5]. Un esempio paradossale in tal senso è dato dal nome Italo Svevo [6], pseudonimo di Aron Hector Schmitz, che risulta addirittura più notorio e identificante del nome reale dello scrittore.

Descritta per sommi capi cosa sia la pseudonimizzazione è importante comprendere per quale motivo il legislatore europeo l'abbia prevista come trattamento atto a contribuire alla protezione dei dati personali. Una risposta in tal senso è contenuta nell'articolo 25 del *Regolamento generale sulla protezione dei dati*, laddove viene descritto il concetto di protezione fin dalla progettazione[55]. Nello specifico, la normativa prescrive che "*il titolare del trattamento mette in atto misure tecniche e organizzative adeguate, quali la pseudonimizzazione, volte ad attuare in modo efficace i principi di protezione dei dati, quali la minimizzazione, e a integrare nel trattamento le necessarie garanzie al fine di soddisfare i requisiti del presente regolamento e tutelare i diritti degli interessati*"[56]. Quindi l'idea di fondo è quella di mettere in essere una serie di misure tecniche e organizzative (tra le quali rientra a pieno titolo la pseudonimizzazione) con lo scopo di determinare un effetto che sia misurabile, e quindi dimostrabile, relativamente all'applicazione dei principi[57] applicabili al trattamento dei dati personali.

Si tratta dunque di verificare se, nei particolari contesti analizzati, sia possibile garantire il miglior rispetto dei principi di protezione dei dati personali, nonché l'esercizio dei diritti degli interessati, in modo più proficuo attraverso una leva pseudonimica, piuttosto che attraverso l'identità primaria.

Prima di procedere oltre è opportuno evidenziare che la tecnica della pseudonimizzazione, seppur in molte circostanze risulti essere necessaria, non sia sempre il migliore strumento di tutela dei diritti e delle libertà delle persone fisiche. Infatti, come già evidenziato dal *Gruppo di lavoro articolo 29* nella sua Opinione 5/2014, "*la pseudonimizzazione non è un metodo di anonimizzazione. Si limita a ridurre la correlabilità di un insieme di dati all'identità originaria di una persona interessata, e rappresenta pertanto una misura di sicurezza utile*" [7].

Diventa quindi indispensabile operare realizzando "*un'accurata valutazione preliminare del tipo di tutela integrata nel trattamento che* [ci] *si prefigge*"[58] [8], a seguito della quale sarà possibile scoprire naturalmente quale sia la tipologia di trattamento successivo (se pseudonimizzazione o se anonimizzazione) più opportuno da implementare. A titolo esemplificativo, "*se […] dall'uso del dato trattato si generano comunque specifiche conseguenze sulla persona che*

richiedono il mantenimento dei diritti di accesso sul dato […], allora il titolare sta procedendo naturaliter ad un trattamento di pseudonimizzazione"[59] [8], dato che come già visto lo scopo di tale tipologia di trattamento è proprio quella di contrastare le possibili incertezze in sede di attribuzione del dato.

Il trattamento successivo realizzato tramite le tecniche di pseudonimizzazione può comportare effetti e livelli di tutela estremamente variegati, proprio derivanti dalle differenti tecniche[60] che si sceglieranno di utilizzare in tale processo. Infatti, "*i dati possono essere pseudonimizzati utilizzando sistemi tracciabili o non tracciabili, a seconda che ne sia possibile o meno la reidentificazione*"[61] [9]. Più nello specifico, come si vedrà successivamente da un punto di vista operativo "*nel caso dell'adozione di tecniche tracciabili si utilizzano delle liste di corrispondenza delle identità con i relativi pseudonimi, piuttosto che algoritmi crittografici bidirezionali; mentre, nel caso di adozione di tecniche non tracciabili, si è soliti far ricorso alla crittografia unidirezionale, tramite la quale è anche possibile creare dati anonimi*"[62] [9]. In altri termini, ciò vuol anche dire che "*il risultato della pseudonimizzazione può essere indipendente dal dato iniziale […] o può essere calcolato a partire dal valore originale di un attributo o insieme di attributi*"[63] [8].

Più in generale possiamo considerare il trattamento di pseudonimizzazione utile per garantire la salvaguardia di molti aspetti dei dati personali, tra cui vanno ricompresi l'integrità, la riservatezza, la correttezza e la limitazione sia della finalità che della conservazione.

Per comprendere come un trattamento di pseudonimizzazione possa essere di ausilio per garantire l'integrità dei dati possiamo considerare che la rilevazione degli errori risulta agevolata nel dominio del dato pseudonimizzato, rispetto a quanto lo sarebbe nel dominio del dato personale originario, in quanto una volta pseudonimizzato, il dato ha un potere identificante maggiore di quello originario. La spiegazione risiede nelle caratteristiche delle tecniche utilizzate dalla funzione di pseudonimizzazione, solitamente basata su strumenti crittografici[64]. Un esempio in tal senso è la funzione crittografica di hash[65], ovvero "*un algoritmo matematico che mappa dei dati di lunghezza arbitraria* [ricevuti in input] *(messaggio) in una stringa**[66]** binaria di dimensione fissa chiamata valore di hash*" [10], che viene restituita in output.

Tali funzioni si caratterizzano per il possesso delle seguenti proprietà [10]:

- identificano in modo univoco il messaggio codificato, ovvero a messaggi differenti, ancorché molto simili, corrisponderanno valori di hash molto differenti[67];
- sono deterministiche, ovvero dato lo stesso messaggio, gli attribuiscono sempre il medesimo valore di hash che non è intellegibile;
- consentono in modo semplice e veloce il calcolo di un valore di hash dato un messaggio sorgente;
- rendono impossibile, da un punto di vista computazionale, determinare qual è il messaggio corrispondente a un valore di hash noto;
- la distanza esistente tra i caratteri diversi in uscita è molto maggiore della distanza presente tra caratteri diversi in entrata, ovvero dei piccoli cambiamenti in ingresso inducono enormi cambiamenti in uscita[68].

Tra l'altro, in merito alla possibilità di realizzare la pseudonimizzazione di un dato tramite la codifica con chiave si è già espresso il *Gruppo di lavoro articolo 29* nella sua Opinione n. 4/2007, affermando che in tal caso "*le informazioni si riferiscono a persone contrassegnate da un codice, mentre la chiave che crea la corrispondenza tra il codice e i comuni identificatori (il nome, data di nascita, indirizzo) è tenuta separata*" [11].

Volendo riprendere l'esempio precedente dei due interessati *Pietro* (padre) e *Paolo* (figlio), potremmo dire che la chiave crittografica della funzione crittografica di hash corrisponde a quell'informazione aggiuntiva (rispettivamente "*di Paolo"* o "*di Pietro*") che, applicata al dato personale originario nell'ambito del trattamento di pseudonimizzazione, consente di ottenere in uscita un nuovo dato non direttamente attribuibile all'interessato (rispettivamente *Padre* o *Figlio*).

Per rendere la trattazione più concreta possiamo fare riferimento a due tragici fatti di cronaca relativi allo scambio di sacche di sangue, occorsi rispettivamente all'ospedale San Martino di Genova nel 2018[69] [12] e all'ospedale di Vimercate nel 2019[70] [13]. Il secondo caso, in particolar modo, è significativo nella sua tragicità, non solo perché purtroppo ha determinato il decesso della paziente, ma anche perché è stato originato da un caso di omonimia che sarebbe stato possibile evitare qualora si fossero utilizzati strumenti di pseudonimizzazione, implementati progettando l'apposita funzione in modo tale che generasse lo pseudonimo inserendovi all'interno anche degli opportuni metadati di contesto. Del resto, commentando il caso di specie, anche l'assessore Gallera rappresentava di fatto la necessità di operare nell'ottica dell'utilizzo di meccanismi di pseudonimizzazione, affermando inizialmente che "*l'identificazione del paziente e la tracciabilità di ogni prodotto somministrato rappresentano obblighi di legge ben regolamentati da Regione Lombardia*" [13], per poi proseguire specificando che "*ogni Azienda Socio Sanitaria Territoriale ha la possibilità di declinare queste prescrizioni avvalendosi degli strumenti che meglio ritiene: bracciali, codici a barre, microchip*" [13].

Va comunque considerato che l'utilizzo di misure tecniche adeguate a garantire e rendere verificabile l'integrità e l'accuratezza del dato, non esima il titolare del trattamento dall'adozione di misure organizzative adeguate, in mancanza delle quali si rischia di inficiare l'adeguatezza delle citate misure tecniche. Infatti, se a titolo esemplificativo si volesse gestire l'etichettatura delle sacche di plasma avvalendosi di un trattamento di pseudonimizzazione, basato sulla generazione di Qrcode o di Barcode da stampare e successivamente applicare alle sacche in modo da aumentarne l'identificabilità, non si potrebbe prescindere dall'adozione di opportune procedure organizzative, finalizzate alla strutturazione del processo in modo che non si verifichino errori proprio in sede di apposizione dei codici pseudonimici alla sacca di sangue corrispondente.

La pseudonimizzazione può essere utilizzata anche per garantire la riservatezza del dato personale, ma con opportune accortezze. Infatti, come già descritto precedentemente, se da un lato il trattamento di pseudonimizzazione è in grado di rendere il dato originario riservato, dall'altro ne aumenta il potere identificativo. Tale circostanza risulta essere particolarmente critica in quanto se non adeguatamente gestita potrebbe determinare la re-identificazione del dato. Si tratta di un aspetto ben noto e determinato dalle modalità operative del trattamento di pseudonimizzazione che viene realizzato in modo atomico, ignorando il contesto all'interno del quale il dato personale originario si trova.

Infatti, bisogna sempre considerare che "*la rimozione o la cifratura degli identificatori espliciti dei rispondenti (38) non sempre è una misura sufficiente, dato che la de-identificazione dei dati non fornisce alcuna garanzia di anonimato*" [9]. Ciò si verifica perché "*le informazioni rilasciate spesso contengono altri dati, come la razza, la data di nascita, il sesso e il codice postale, che possono essere collegati a informazioni pubblicamente disponibili per re-identificare (o limitarne l'incertezza in merito) i rispondenti dei dati, facendo così trapelare informazioni che non si pensava di rilasciare*"[71] [14] in origine.

Nello specifico, sebbene in alcuni casi sia possibile rilevare una corrispondenza esatta, tra le sorgenti dati de-identificate oggetto di rilascio da parte del titolare del trattamento e altre sorgenti dati altrove disponibili, individuando così in modo univoco il rispondente dei dati; in altri casi la sua individuazione univoca non è possibile e ci si deve necessariamente "*accontentare di restringere la corrispondenza a un insieme ristretto di tuple, individuando in tal modo un gruppo limitato di individui al cui interno si troverà il rispondente effettivo dei dati, pur non potendo determinare con certezza chi esso sia tra questi*" [9].

Ci vuol dire che la pseudonimizzazione da sola non è sufficiente, poiché per poter tutelare i diritti e le libertà delle persone fisiche si ha necessità anche di ridurre il potere identificativo del dato, potere che invece aumenterebbe se ci si limitasse semplicemente a pseudonimizzarlo. Emerge pertanto la necessità di operare sul contorno del dato, minimizzando la presenza degli elementi non strettamente necessari alla realizzazione del trattamento considerato, in modo da diminuire la superficie del cosiddetto quasi-identificatore. Dove con tale termine si è soliti indicare "*quell'insieme di attributi presente nella tabella privata che, essendo anche disponibili esternamente, sono quindi sfruttabili in combinazione per effettuare il collegamento tra sorgenti dati e, di conseguenza, re-identificare i rispondenti*"[72] [9]. Pertanto, si tratta di un insieme di attributi che, considerando la criticità delle proprie caratteristiche nel particolare contesto considerato, deve essere sottoposto a rilascio solo dopo attenta valutazione dei rischi connessi.

Una possibilità concreta è quella di agire sul dato originario per diluire la scala utilizzata per rappresentarlo, operando nell'ottica di minimizzarlo per generalizzazione. Uno degli effetti che si possono raggiungere operando in tal modo è quello di poter pervenire alla negatività del dato, ovvero di consentire a un interessato di negare che il dato gli appartenga, poiché è evidente che non appartiene solo a lui, ma a un cluster più ampio di interessati.

Ci sono più strade per realizzare la minimizzazione, la prima come già descritto è quella di ridurre il quantitativo di dati presenti, ma esiste anche una seconda via che consiste nell'aumentare in modo controllato il numero di dati, al fine di generare una maggiore complessità all'interno della quale sia possibile celare il dato particolare che si vuole proteggere, in quanto oggetto di tutela.

Nel prosieguo dell'articolo verranno descritte sinteticamente alcune di queste tecniche[73] che operano per generalizzazione e soppressione.

Il trattamento addizionale di pseudonimizzazione è utile anche per garantire la correttezza del trattamento, intesa come la capacità di aumentare il livello di fiducia[74] nell'attività svolta. Una delle possibilità per realizzare questa condizione è quella di fornire la chiave di cifratura direttamente all'interessato in modo che questi possa procedere a pseudonimizzare i dati personali in autonomia (dunque senza che il titolare del trattamento possa identificarlo) e solo successivamente trasmettere il dato già pseudonimizzato al titolare che provvederà a trattarlo.

Si tratta dello scenario descritto dall'articolo 11 del Regolamento (UE) 2016/679, in cui si prevede che "*se le finalità per cui un titolare del trattamento tratta i dati personali non richiedono o non richiedono più l'identificazione dell'interessato, il titolare del trattamento non è obbligato a conservare, acquisire o trattare ulteriori informazioni per identificare l'interessato*"[75] [1]. Il Considerando 57 chiarisce che il concetto di identificazione dell'interessato debba far riferimento anche all'utilizzo di un'identità digitale da parte di una persona fisica. Ciò potrebbe essere realizzato "*ad esempio mediante un meccanismo di autenticazione quali le stesse credenziali, utilizzate dall'interessato per l'accesso (log in) al servizio on line offerto dal titolare del trattamento*"[76] [1].

La conseguenza di tale circostanza sarà quella che il titolare del trattamento non sarà più in grado di identificare l'interessato e quindi, nel caso di specie non si applicheranno a questo i diritti[77] di accesso, rettifica, cancellazione, limitazione, notifica e portabilità. In ogni caso l'interessato ha sempre la possibilità di fornire delle informazioni addizionali al titolare del trattamento per rendersi identificabile e di conseguenza ripristinare la possibilità di esercitare i propri diritti. Del resto il Considerando 57 specifica che "*il titolare del trattamento non dovrebbe rifiutare le ulteriori informazioni fornite dall'interessato al fine di sostenere l'esercizio dei suoi diritti*"[78] [1], e l'articolo 11 prescrive che la non applicabilità dei diritti di cui agli articoli da 15 a 20 permane fintanto che "*l'interessato, al fine di esercitare i diritti di cui ai suddetti articoli, fornisce ulteriori informazioni che ne consentano l'identificazione*"[79] [1].

In merito a queste fattispecie è fondamentale sottolineare che la non applicabilità dei citati diritti dell'interessato non si configura come una sorta di esenzione riconosciuta al titolare del trattamento, bensì una maggiore forma di tutela dello stesso interessato che si realizza attraverso strumenti tecnologici. Infatti, all'interessato è sempre consentito agire nell'ottica di ripristinare la sua identità e retrocedere da tale ulteriore forma di tutela.

È possibile sfruttare il trattamento della pseudonimizzazione anche per garantire la limitazione della finalità e della conservazione. In tal caso sarà necessario costruire una funzione di pseudonimizzazione nella quale dovranno essere incluse, oltre all'identità della persona fisica, anche l'identità del soggetto con cui la persona fisica scambia dei dati, nonché il dato vero e proprio, ovvero la sua parte semantica. Sarà così possibile ottenere quella che in letteratura si chiama una *zero knowledge proof*, ovvero una sorta di impronta del dato, che rappresentando il minimo scambio informativo possibile.

Un esempio di utilizzo degli strumenti *zero knowledge proof* è dato dalla CNIL (l'autorità di controllo della Repubblica Francese), nell'ambito delle sue linee guida "*Blockchain. Solutions for a responsible use of the blockchain in the context of personal data*" [15] in cui si afferma che "*i dati personali dovrebbero essere preferibilmente registrati nella blockchain sotto forma di un commitment**[80]***"[81] [15]. Ma l'autorità di controllo francese va oltre fornendo delle indicazioni di dettaglio in merito a quale sia la tipologia di soluzione tecnica più adeguata a garantire la conformità con i principi di protezione dei dati personali, anche indicando un ordine di preferenza nella scelta. Più in dettaglio, la CNIL afferma che "*per quanto riguarda il dato personale aggiuntivo, al fine di garantire la conformità con la protezione dei dati fin dalla progettazione e per impostazione predefinita, nonché con gli obblighi di minimizzazione, […]* [si] *consigliano delle soluzioni nelle quali il dato è trattato esternamente alla blockchain oppure, nelle quali sono memorizzati nella blockchain i seguenti* [elementi] *in ordine di preferenza:*

a. *un commitment del dato;*
b. *lo hash del dato prodotto da una funzione di hash con chiave;*
c. *la cifratura del dato*"[82] [15].

L'elemento interessante è che qualora si dovesse verificare una violazione della protezione dei dati, a seguito della quale il dato pseudonimizzato dovesse essere rivelato, lo stesso non avrebbe alcuna utilità in quanto privo di valore semantico; infatti la tecnica utilizzata prevede che ci siano due distinte entità che congiuntamente operino sul dato per poterlo validare.

L'anonimizzazione

Nel Considerando 26, il Regolamento (UE) 2016/679 fornisce una definizione di informazioni anonime, asserendo che queste "*non si riferiscono a una persona fisica identificata o identificabile"*[83] [1], oppure si riferiscono ai *"dati personali resi sufficientemente anonimi da impedire o da non consentire più l'identificazione dell'interessato*"[84] [1], anche nel caso in cui si prendano in considerazione "*tutti i mezzi […] di cui il titolare del trattamento o un terzo può ragionevolmente avvalersi*"[85] [1].

In altri termini è possibile "*definire il dato anonimo come quel dato precedentemente corrispondente a una persona identificabile, che a seguito di idoneo trattamento non ne permette più l'identificazione*"[86] [9].

Il Gruppo di lavoro articolo 29, nella sua Opinione 5/2014 ha chiarito alcuni aspetti relativi alle tecniche di anonimizzazione, specificando che questa "*costituisce un trattamento successivo dei dati personali*" [7] ed "*è il risultato del trattamento di dati personali volto a impedire irreversibilmente l'identificazione*" [7]. Si consideri che il rispetto di quest'ultimo requisito è particolarmente oneroso e impegnativo, dato che richiede al titolare del trattamento di "*operare a ritroso, considerando in via preliminare tutti i mezzi che ragionevolmente potrebbero essere utilizzati per l'identificazione di una persona fisica*"[87] [9].

Si tratta di un'operazione lungi dall'essere banale, dato che tali mezzi sono di tipo eterogeneo, ricomprendendo tra di loro i mezzi economici, le risorse tecnologiche, le informazioni, le competenze specialistiche, nonché il tempo a disposizione dell'attaccante. Ma soprattutto, "*ogni valutazione […] non potrà non tener conto di elementi soggettivi, che possono variare in ragione del contesto*"[88] [8]. A titolo esemplificativo, va considerato che "*la disponibilità di certe risorse, che può essere «irragionevole» per taluni, potrà infatti non esserlo per altri soggetti*"[89] [8], così come differente sarà il grado motivazionale dei vari soggetti, eventualmente interessati a riuscire a stabilire una correlabilità tra il dato anonimizzato a disposizione e una specifica persona fisica. Parimenti, sarà necessario considerare che alcune tipologie di interessati sono più facilmente identificabili; a titolo esemplificativo, se dovessimo considerare il dato "posizione lavorativa", avremmo maggiore difficoltà a impedire la riconoscibilità di un interessato a partire dalla conoscenza del lavoro svolto, se questi ricopre il ruolo di Pontefice, rispetto al caso che avremmo se il ruolo dell'interessato fosse quello di un normale ministro del culto, così come di uno studente o di un impiegato.

In generale è ormai consolidato in letteratura che affinché un processo di anonimizzazione possa essere effettivamente considerato uno strumento di tutela integrato nel trattamento dei dati personali, e pertanto uno strumento di conformità con il criterio di protezione fin dalla progettazione ex articolo 25 del GDPR, è necessario che questi "*si basi su tecniche (di distorsione o di generalizzazione dei dati) riconosciute dalla comunità scientifica internazionale e che tenga conto degli aspetti contestuali idonei a valutare l'irragionevolezza dei mezzi*"[90] [8] necessari a de-anonimizzare i dati anonimizzati.

Pertanto, considerando che l'incremento dei dati ausiliari riferibili a una persona fisica alla quale si vuole cercare di correlare il dato anonimizzato determina un parallelo incremento della probabilità di riuscire a compiere tale attività di collegamento con successo, è necessario procedere periodicamente alla reiterazione della valutazione sull'irragionevolezza dei mezzi. Dunque, "*tale attività non andrà espletata una tantum, bensì ripetuta continuamente fintanto che perduri il ciclo di vita del trattamento nel cui ambito si vogliono proteggere i dati personali*"[91] [9].

Sempre il Gruppo di lavoro articolo 29 ha chiarito che "*i dati resi anonimi non rientrano nell'ambito di applicazione della legislazione in materia di protezione dei dati*" [7], sempre ovviamente che il trattamento di anonimizzazione sia stato effettuato in modo adeguato [7]. Parimenti il WP29[92] ha anche illustrato le principali tecniche di anonimizzazione, suddivise nelle due macro famiglie della randomizzazione[93] e della generalizzazione, con particolare riferimento all'aggiunta di rumore statistico (noise addition), all'aggregazione, alle permutazioni, alla differential privacy, a *k*-anonymity, a *l*-diversity e a *t*-closeness.

Il citato parere si è anche occupato di definire una metodica da utilizzare per la valutazione del livello di robustezza di ogni tecnica, che viene qui poi indagata nelle tre direttrici dell'individuazione, della correlabilità e della deduzione. Dunque, il WP29 ci invita a valutare una tecnica di anonimizzazione in base a cosa sia ancora possibile fare sul dato originario dopo la sua applicazione. In particolare si vuole valutare se:

- "*sia ancora possibile individuare una persona fisica (individuazione);*
- *sia ancora possibile collegare i dati relativi a una persona fisica (correlabilità);*
- *sia possibile dedurre informazioni riguardanti una persona fisica (deduzione)*"[94] [9].

Descrizione delle principali tecniche di anonimizzazione

Come precedentemente visto le tecniche di anonimizzazione si suddividono in due grandi famiglie, quella della randomizzazione e quella della generalizzazione.

La randomizzazione viene definita dal *Gruppo di lavoro articolo 29* come "*una famiglia di tecniche che modifica la veridicità dei dati al fine di eliminare la forte correlazione che esiste tra i dati e la persona*" [7]. Nella sua opinione 5/2014 il WP29 si concentra in particolar modo sull'analisi delle tecniche di randomizzazione, dell'aggiunta del rumore statistico, della permutazione e della privacy differenziale. Pur non potendo soffermarci in questa sede nella descrizione dei meccanismi alla base della privacy differenziale, si consideri che a livello pratico questa, "*comporta l'iniezione di «rumore» nelle risposte alle query effettuate sull'insieme di dati. Il rumore dovrebbe essere abbastanza grande da garantire l'anonimato a livello individuale, ma non abbastanza da influire sull'utilità della risposta*"[95] [16].

Dunque, è possibile affermare, seppur non in modo formale, che la privacy differenziale "*richiede che la distribuzione di probabilità sui risultati pubblicati di un'analisi, sia praticamente la stessa, indipendentemente dal fatto che un individuo sia presente o meno nell'insieme di dati oggetto di analisi*"[96] [9].

Invece, la generalizzazione viene definita dal Gruppo articolo 29 come una famiglia di tecniche che "*consiste nel generalizzare, o diluire, gli attributi delle persone interessate modificando la rispettiva scala o ordine di grandezza*" [7].

Nel prosieguo dell'articolo verranno illustrate in modo sintetico le caratteristiche e il funzionamento di alcune di queste tecniche, ovvero di k-anonymity, di l-diversity e di t-closeness. Si tratta di tecniche che "*possono essere efficaci per impedire l'individuazione di un rispondente, pur non garantendo un'anonimizzazione sempre efficace, dato che possono comunque esporre al rischio di correlabilità e deduzione, se non vengono implementate facendo attenzione a ponderare gli aspetti quantitativi*"[97] [9].

La prima delle tecniche analizzate si occupa della protezione dell'identità dei rispondenti in uno scenario in cui si opera al rilascio di microdati (e quindi dati veritieri e granulari) a livello di tupla[98]. Particolare attenzione va posta al fatto che tale tecnica, che prende il nome di *k*-anonymity, non si occupa di proteggere gli attributi oggetto di rilascio; dunque, qualora si abbia pure questa necessità, sarà indispensabile avvalersi anche di altre metodiche progettate per lo scopo. Nello specifico, "*il compito di k-anonymity è quello di garantire un certo grado di incertezza, dato che, indipendentemente da quali saranno le inferenze fatte sui dati, si otterrà come output del processo sempre una loro associazione con un gruppo di almeno k persone*"[99] [9]. Ovvero, in altri termini, possiamo considerare *k*-anonymity come uno strumento pensato per occultare un rispondente all'interno di un cluster di almeno k rispondenti.

Il principale *vulnus* di *k*-anonymity consiste nel non riuscire necessariamente a nascondere a quale rispondente sia imputabile un certo dato, malgrado renda possibile che i rispondenti si confondano tra di loro. Si rende dunque necessario avvalersi anche di un ulteriore approccio [17] che consideri le problematiche dell'omogeneità degli attributi dei valori sensibili, e della conoscenza pregressa dell'osservatore.

Nello specifico, "*la problematica dell'omogeneità dei valori degli attributi sensibili si verifica quando tutte le tuple di una tabella k-anonima, con un certo valore nel quasi-identificatore, hanno lo stesso valore in almeno un attributo sensibile*"[100] [9]. Si tratta di uno scenario potenzialmente critico poiché se il cluster in cui si confondono i rispondenti ha la caratteristica di avere lo stesso dato "sensibile" che non si vorrebbe rivelare, determinare qual è l'identità di un dato rispondente diventa un'informazione superflua per la conoscibilità del dato sensibile associato, che viene in ogni caso esposto.

Invece, la problematica della conoscenza pregressa dell'osservatore è rilevabile nello scenario in cui questo si può avvale della conoscenza pregressa di alcune informazioni esterne, aggiuntive rispetto a quelle presenti nella tabella dei microdati, per fare inferenze sul dato "sensibile" che nei *desiderata* non dovrebbe essere oggetto di rilascio. Lo scenario è particolarmente critico e complesso, dato che "*non è possibile conoscere a priori che cosa sappia l'osservatore e, dunque, quale sia il contenuto informativo in suo possesso, che potrà essere utilizzato per fare dei collegamenti atti a consentirgli di inferire le informazioni sensibili*"[101] [9].

In [17] si propone un nuovo approccio denominato *l*-diversity, basato sul cosiddetto principio non istruttivo, in cui si afferma che "*la tabella pubblicata dovrebbe fornire all'avversario poche informazioni aggiuntive oltre a quelle già precedentemente conosciute. In altre parole, non ci dovrebbe essere una grande differenza tra le convinzioni pregresse e quelle posteriori*"[102] [17]. In termini più discorsivi "*l-diversity impone che all'atto di costituire i vari cluster di dimensione l, questi vengano predisposti mettendovi delle informazioni sensibili differenti*"[103] [9].

Malgrado i benefici introdotti con questa nuova tecnica, alcune ricerche [18] hanno evidenziato che "*sebbene il principio di l-diversity rappresenta un passo importante oltre k-anonymity nella protezione dalla divulgazione degli attributi, presenta [ugualmente] diverse carenze*"[104] [18]. Il riferimento è principalmente all'attacco di asimmetria e all'attacco di somiglianza. Il primo "*si verifica quando la distribuzione in un q-blocco**[105]** è diversa dalla distribuzione nella popolazione originale*"[106] [9], mentre il secondo "*si verifica quando un q-blocco ha per l'attributo sensibile valori diversi, anche se semanticamente simili*"[107] [9].

Una soluzione utile a superare le limitazioni insite in *l*-diversity e in *k*-anonymity, è data da "*una nuova nozione di privacy chiamata t-closeness, nella quale si richiede che la distribuzione di un attributo sensibile in ogni classe di equivalenza sia vicina alla distribuzione dell'attributo nella tabella complessiva*"[108] [18]. L'idea di fondo è quella di agire affinché la distanza tra la distribuzione del mondo reale e quella presente nel blocco considerato[109] sia la minore possibile. Pertanto, "*nella costruzione del q-blocco bisognerebbe operare nell'ottica di assicurarsi che la*

distribuzione dei valori sensibili, in esso presenti, sia particolarmente vicina a quella che si ha nel mondo reale"[110] [9], pena determinare un incremento dell'intervallo di confidenza con cui è possibile dedurre il possesso di un dato "sensibile" che non si vuole rivelare.

La pseudonimizzazione e l'anonimizzazione come misure per la protezione degli archivi

Relativamente alle misure tecniche più adeguate da porre in essere nell'ambito del trattamento degli archivi sanitari, l'EAG fornisce alcune esemplificazioni su quando, e a quali condizioni, sia consigliabile procedere all'effettuazione del successivo trattamento di pseudonimizzazione. Nello specifico, nelle *EAG Guidelines* si evidenzia che "*quando si sta effettuando dell'attività di ricerca medica è importante preservare la correlazione dei differenti dati sanitari relativi a un dato paziente, anche se l'identità dello stesso paziente è irrilevante*"[111] [19], mentre "*un servizio archivistico che mantiene documenti archivistici nel pubblico interesse deve, nell'interesse degli interessati, salvaguardare l'integrità dei documenti archivistici sanitari scelti per la conservazione permanente*"[112] [19]. Dunque, solo nel primo caso i dati sanitari potranno essere sottoposti a un ulteriore trattamento di pseudonimizzazione.

Del resto le *EAG Guidelines* riconoscono che "*un servizio archivistico può anche avvalersi della pseudonimizzazione […], ma questa dovrebbe essere implementata in modo da non mettere a repentaglio il valore probatorio dei documenti archivistici*"[113] [20].

Esimi studiosi [21] hanno evidenziato un forte nesso tra il ruolo dell'archivista e il rispetto di alcuni dei principi di protezione dei dati personali di maggior interesse per la presente trattazione, dato che "*la natura stessa dell'attività dell'archivista lo spinge verso un intrinseco maggior rispetto del principio di minimizzazione dei dati*"[114] [21]. Nello specifico ciò è evincibile dal fatto che questi, "*da un lato, seleziona per la conservazione permanente i documenti di archivio contenenti dati personali solo quando ciò è effettivamente necessario in base alla legge, dall'altro, si occupa di controllare l'accesso a tale documenti, inibendo la consultazione, per il periodo temporale previsto dalla normativa vigente, dei documenti contenenti dati personali*"[115] [21].

Infine, si evidenzia come le *EAG Guidelines* riconoscano che l'utilizzo di un "*software per la descrizione archivistica che consente la creazione di due diverse versioni di un assistente di ricerca (uno con nomi reali e uno con gli pseudonimi) sia uno strumento per il rispetto dell'art. 25*"[116] del Regolamento (UE) 2016/679. Infatti, come evidenziato da Pigliapoco, "*il primo potrebbe essere utilizzato per effettuare delle ricerche finalizzate a garantire il soddisfacimento dei diritti degli interessati, mentre il secondo potrebbe essere impiegato nell'ambito delle ricerche effettuate online*"[117] [21].

Riflessioni finali

Secondo una visione tradizionalistica i principi che indirizzano la protezione dei dati personali sono alla base e ispirano l'azione del tecnologo, al quale si chiede semplicemente di operare nell'ottica di un rafforzamento delle misure di sicurezza finalizzate a minimizzare i rischi di varia probabilità e gravità per i diritti e le libertà delle persone fisiche, come a titolo esemplificativo operare affinché il dato personale risulti meno intellegibile.

Le recenti innovazioni hanno di fatto modificato tale situazione e la tecnologia ha assunto un ruolo di co-indirizzo, dato che per la loro stessa natura alcune tecnologie si oppongo radicalmente all'applicazione dei principi in materia di protezione dei dati personali. Per chiarire meglio l'affermazione, diminuendone il carattere dogmatico, è possibile considerare alcuni esempi, quali quelli della *blockchain* e della "*differential privacy*".

Per quanto concerne la tecnologia *blockchain*, questa si oppone profondamente al diritto degli interessati di ottenere la cancellazione dei propri dati personali[118] da parte del titolare del trattamento, in quanto proprio nel suo meccanismo di funzionamento interno è insita la non cancellazione del dato. Si tratta di un aspetto necessario in quanto senza la persistenza dei dati la *blockchain* cessa di funzionare.

Per quanto concerne la *differential privacy*, questa si oppone al diritto di rettifica del dato[119], secondo il quale all'interessato è riconosciuto il diritto di ottenere dal titolare del trattamento la rettifica dei dati personali inesatti che lo riguardano. Invece, tale tecnologia si caratterizza per l'utilizzo di dati che non sono esatti, bensì alterati attraverso l'aggiunta di un certo rumore statistico. Si tratta di una tecnica potente e affidabile, utilizzabile soprattutto nel rilascio di dati statistici. La privacy differenziale è già stata adottata da importanti organizzazioni come il *Census Bureau of the United States of America*, che ha scelto di avvalersene per la realizzazione integrale del censimento del 2020, riconoscendo che si tratta di un nuovo sistema di protezione della privacy di avanguardia e indispensabile "*per stare al passo con le minacce emergenti nel mondo digitale di oggi*"[120] [22].

Dunque è possibile asserire che la tecnologia cessa di essere una mera attuatrice dei principi di protezione dei dati personali, ma comincia a comprimere tali principi ponendo loro dei vincoli sempre più pressanti. Si tratta di una nuova visione nella quale è la tecnologia che determina qual è la condizione necessaria.

Ciò vuol dire che la tecnologia assurge allo stesso livello dei diritti e detta delle condizioni. Questo è uno scenario che si caratterizza per la complessità, una complessità di cui è inevitabile e necessario farsi carico, cercando di capire quali siano i dettagli in gioco e cercando di trovare una soluzione che non può in nessun modo consistere nella banalizzazione dei problemi.

Una riflessione a parte merita l'analisi delle differenze sostanziali che intercorrono tra un processo di anonimizzazione e uno di pseudonimizzazione. In questa sede non ci si vuole tanto soffermare sull'analisi degli aspetti tecnici, in quanto già precedentemente descritti, bensì sulla tipologia di tutela dei diritti e delle libertà delle persone fisiche (a cui si riferiscono i dati personali) che i due distinti processi sono in grado di offrire. Va da sé che la tutela che è insita nei due processi, lo è anche in conseguenza della tecnologia utilizzata nella loro implementazione, dato che questa inevitabilmente pone dei vincoli che si riflettono successivamente sulle tutele effettivamente implementabili.

Nello specifico, "*la tutela introdotta con la pseudonimizzazione è volta a garantire la confidenzialità del dato, non più immediatamente intellegibile, ma anche […] a garantire l'integrità contro manipolazioni anche accidentali*"[121] [8]. Mentre, "*nel caso dell'anonimizzazione la tutela è […] volta a impedire, a meno di dover ricorrere a mezzi irragionevolmente utilizzabili, la riferibilità del dato a una persona*"[122] [8]. Insomma si tratta di due distinte tipologie di trattamento che si occupano di misurare l'adeguatezza di tutele differenti; mentre la pseudonimizzazione si occupa della sicurezza del dato in esame, l'anonimizzazione si interessa alla sua riservatezza.

Un ultimo spunto di riflessione merita l'analisi dei rischi che è necessario effettuare nell'ambito del processo di gestione dei trattamenti di dati personali, anche al fine di valutare l'adeguatezza delle misure tecniche e organizzative che si vuole adottare o che si sono già adottate. L'aspetto da considerare con particolare attenzione è l'oggetto del processo di analisi e gestione del rischio. Infatti, nell'ambito del Regolamento (UE) 2016/679 il focus non è sui rischi operativi, ovvero sui rischi per i dati in sé, bensì sui rischi per i diritti e le libertà delle persone fisiche (gli interessati). "*Dunque, in questo caso, gli impatti dovranno essere valutati proprio nella prospettiva dei rischi connessi a tali diritti e libertà. Analogamente, la valutazione di impatto dovrà essere incentrata sui trattamenti, piuttosto che sui processi di business*"[123] [9]. Poiché i rischi sono quelli derivanti dal trattamento nel suo complesso e non solo dalle modalità tecniche utilizzate per garantire la protezione dei dati personali, la valutazione del rischio dovrà essere realizzata in modo olistico tenendo in considerazione questo livello di complessità. Sarà dunque utile effettuare molteplici valutazioni del rischio, tante quante sono i trattamenti in essere presso il titolare, alle quali far poi seguire una sola valutazione di impatto di tipo complessivo[124]. Quest'ultima dovrà tenere in considerazione sia le singole mappature del rischio dei trattamenti, sia il registro dei trattamenti, laddove sono evidenziati quei trattamenti che per le loro caratteristiche hanno determinato un livello di rischio di tipo alto. Sempre nell'ottica di gestire il livello di complessità insito nel processo di analisi e gestione del rischio, laddove sia necessario procedere nella valutazione dei rischi per un nuovo trattamento, sarà indispensabile prendere in considerazione tutti i rischi che possono generare un qualsiasi impatto lungo l'intera filiera del trattamento e non solo nello specifico trattamento considerato.

Bibliografia

[1] *Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/*. UE: Gazzetta Ufficiale dell'Unione europea, 2016, pp. 3–90.

[2] *Decreto Legislativo 30 giugno 2003, n. 196 (Testo coordinato con il Decreto Legislativo 10 agosto 2018, n.101)*. Italia: Garante per la protezione dei dati personali, 2018, pp. 1–66.

[3] (Garante per la protezione dei dati personali), "Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica pubblicate ai sensi dell'art. 20, comma 4, del d.lgs. 10 agosto 2018, n. 101 – 19 dicembre 2018 [9069661]." Garante per la protezione dei dati personali, Roma, 2018.

[4] (Istituto enciclopedico Treccani), "www.treccani.it – Patronimico." [Online]. Available: http://www.treccani.it/vocabolario/patronimico. [Accessed: 25-Feb-2020].

[5] (Istituto enciclopedico Treccani), "www.treccani.it – Pseudonimo." [Online]. Available: http://www.treccani.it/vocabolario/pseudonimo. [Accessed: 25-Feb-2020].

[6] (Wikipedia), “Wikipedia – Italo Svevo.” [Online]. Available: https://it.wikipedia.org/w/index.php?title=Italo_Svevo&oldid=110951696. [Accessed: 25-Feb-2020].

[7] (Article 29 Working Party), “Opinion 05/2014 on Anonymisation Techniques (WP216).” Brussels, pp. 1–37, 2014.

[8] G. D, Acquisto and M. Naldi, *Anonimizzazione Pseudonimizzazione Sicurezza*. Torino: G. Giappichelli Editore, 2017.

[9] F. Ciclosi, *La protezione dei dati e la gestione del rischio nella pubblica amministrazione. Un approccio unificato nel contesto del GDPR e del framework europeo di sicurezza cibernetica*. Santarcangelo di Romagna (RN): Maggioli Editore, 2019.

[10] (Wikipedia), “Wikipedia – Funzione crittografica di hash.” [Online]. Available: https://it.wikipedia.org/w/index.php?title=Funzione_crittografica_di_hash&oldid=107536347. [Accessed: 27-Feb-2020].

[11] (Article 29 Data Protection Working Party), “Opinion 4/2007 on the concept of personal data (WP136).” Brussels, pp. 1–26, 2007.

[12] (La Repubblica), “Scambio sacche per trapianto midollo a San Martino di Genova,” 2018. [Online]. Available: https://genova.repubblica.it/cronaca/2018/10/17/news/scambio_sacche_per_trapianto_midollo_a_san_martino_di_genova-209207054/amp/.

[13] (Il Messaggero), “Vimercate, donna muore in ospedale dopo trasfusione: sacche di sangue scambiate per omonimia,” 2019. [Online]. Available: https://www.ilmessaggero.it/AMP/italia/monza_ospedale_sacche_sangue_trasfusione_errore_donna_morta-4737967.html. [Accessed: 25-Feb-2020].

[14] S. F. V. Ciriani S. De Capitani di Vimercati and P. Samarati, “k-Anonymity,” *Secur. Data Manag. Decentralized Syst. Adv. Inf. Secur.*, vol. 33, Part I, pp. 323–353, 2007.

[15] (Commission Nationale Informatique & Libertés), “Blockchain. Solutions for a responsible use of the blockchain in the context of personal data.” pp. 1–10, 2018.

[16] (Information Commissioner’s Office), “Big Data, Artificial Intelligence, Machine Learning and Data Protection,” 2017.

[17] D. Kifer, “l-Diversity : Privacy Beyond k -Anonymity,” *Proc. 22nd Int. Conf. Data Eng.*, vol. 1, pp. 1–36, 2006.

[18] N. Li, T. Li, and S. Venkatasubramanian, “t-Closeness: Privacy Beyond k-Anonymity and l-Diversity,” in *2007 IEEE 23rd International Conference on Data Engineering*, 2007, pp. 106–115.

[19] (European Data Protection Board), “Guidelines 4/2018 on the accreditation of certification bodies under Article 43 of the General Data Protection Regulation (2016/679)-Annex 1 Version for public consultation,” Brussels, 2018.

[20] (European Archives Group), “Guidance on data protection for archive services. EAG guidelines on the implementation of the General Data Protection Regulation in the archive sector.” European Archives Group, pp. 1–37, 2018.

[21] S. Pigliapoco, “Impatto della nuova normativa in materia di protezione dei dati personali sulla gestione e conservazione degli archivi,” in *La protezione dei dati e la gestione del rischio nella pubblica amministrazione. Un approccio unificato nel contesto del GDPR e del framework europeo di sicurezza cibernetica*, F. Ciclosi, Ed. Maggioli Editore, 2019, pp. 365–375.

[22] “Statistical Safeguards.” [Online]. Available: https://www.census.gov/about/policies/privacy/statistical_safeguards.html. [Accessed: 01-Mar-2020].

---

[1] Nel prosieguo dell’articolo ci riferiremo al Regolamento (UE) 2016/679 anche con il suo titolo “*Regolamento generale sulla protezione dei dati*”, oppure con il suo acronimo in lingua inglese “*GDPR*”.

[2] Regolamento (UE) 2016/679, articolo 4(13).

[3] Regolamento (UE) 2016/679, Considerando 34.

[4] A tal riguardo, il Considerando 51, del Regolamento (UE) 2016/679 precisa che "*il trattamento di fotografie non dovrebbe costituire sistematicamente un trattamento di categorie particolari di dati personali, poiché esse rientrano nella definizione di dati biometrici soltanto quando saranno trattate attraverso un dispositivo tecnico specifico che consente l'identificazione univoca o l'autenticazione di una persona fisica*" [1].

[5] Regolamento (UE) 2016/679, articolo 4(14).

[6] Regolamento (UE) 2016/679, articolo 4(15).

[7] Regolamento (UE) 2016/679, Considerando 35.

[8] Si evidenzia come, a seconda dei casi, a tal riguardo possano rientrare nella categoria di dati riguardanti lo stato di salute dell'interessato anche il dato pseudonimizzato, nonché l'informazione aggiuntiva da aggiungere a questo per rendere l'interessato nuovamente identificato.

[9] Idem 6.

[10] Idem 6.

[11] Tali circostanze sono elencate in dettaglio nel Regolamento (UE) 2016/679, articolo 9(2).

[12] Regolamento (UE) 2016/679, articolo 89(1).

[13] Idem 10.

[14] Idem 10.

[15] Ex articolo 15 del Regolamento (UE) 2016/679.

[16] Ex articolo 16 del Regolamento (UE) 2016/679.

[17] Ex articolo 18 del Regolamento (UE) 2016/679.

[18] Ex articolo 21 del Regolamento (UE) 2016/679.

[19] Ex articolo 20 del Regolamento (UE) 2016/679.

[20] Ex articolo 19 del Regolamento (UE) 2016/679.

[21] Regolamento (UE) 2016/679, articolo 9(4).

[22] Adeguamento effettuato tramite lo strumento del d.lgs. 101/2018.

[23] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 1.

[24] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 2, lettera a).

[25] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 2, lettera b).

[26] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 2, lettera c).

[27] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 5.

[28] Idem 25.

[29] Idem 25.

[30] D.lgs. 196/2003, così come modificato dal D.lgs. 101/2018, articolo 2-septies, comma 7.

[31] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 1, comma 1.

[32] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 1, comma 2.

[33] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 1, comma 3.

[34] Con il termine “archivista” le Regole deontologiche intendono “*chiunque, persona fisica o giuridica, ente o associazione, abbia responsabilità di controllare, acquisire, trattare, conservare, restaurare e gestire archivi storici, correnti o di deposito della pubblica amministrazione, archivi privati dichiarati di notevole interesse storico*”, ma anche gli archivi privati non dichiarati di notevole interesse storico o i singoli documenti di interesse storico, per i quali i proprietari, possessori o detentori manifestano l’intenzione di applicare tali regole nella misura in cui esse sono compatibili.

[35] Con il termine “utente” le Regole deontologiche intendono “*chiunque chieda di accedere o acceda per scopi storici a documenti contenenti dati personali, anche per finalità giornalistiche o di pubblicazione occasionale di articoli, saggi e altre manifestazioni del pensiero*”.

[36] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 10, comma 1.

[37] Le regole deontologiche inseriscono tra le eccezioni anche “*i documenti di carattere riservato relativi alla politica interna ed estera dello Stato che divengono consultabili cinquanta anni dopo la loro data*”.

[38] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 10, comma 2.

[39] Idem 55.

[40] A tal riguardo il Ministro dell’Interno dovrà acquisire in via preliminare il parere del direttore dell’Archivio di Stato o del sovrintendente archivistico competenti, nonché udire la Commissione per le questioni inerenti alla consultabilità degli atti di archivio riservati istituita presso il Ministero dell’interno, secondo quanto previsto all’art. 123 del d.lgs. n. 42 del 2004.

[41] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 10, comma 4.

[42] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 10, comma 6.

[43] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 10, comma 7.

[44] Idem 60.

[45] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 11, comma 1.

[46] Idem 62.

[47] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 11, comma 2.

[48] Regole deontologiche per il trattamento a fini di archiviazione nel pubblico interesse o per scopi di ricerca storica, Allegato 1, articolo 11, comma 4.

[49] Idem 65.

[50] Regolamento (UE) 2016/679, articolo 4(5).

[51] A tal proposito si noti che, come evidenziato dal Considerando 26 del Regolamento (UE) 2016/679, “*i dati personali sottoposti a pseudonimizzazione, i quali potrebbero essere attribuiti a una persona fisica mediante l’utilizzo di ulteriori informazioni, dovrebbero essere considerati informazioni su una persona fisica identificabile*”.

[52] Idem 69.

[53] Idem 69.

[54] Regolamento (UE) 2016/679, articolo 11(1).

[55] Nella versione in lingua inglese del Regolamento (UE) 2016/679, tale concetto viene espresso con la locuzione "*privacy by design*".

[56] Regolamento (UE) 2016/679, articolo 25(1).

[57] Il riferimento è ai principi di cui all'articolo 5 del Regolamento (UE) 2016/679.

[58] Op. Cit. p. 39.

[59] Idem 75.

[60] A tal riguardo il parere 4/2007 del Gruppo articolo 29, ha evidenziato che qualora la pseudonimizzazione fosse realizzata con sistemi tracciabili, i dati così ottenuti potranno essere assimilati a delle informazioni su persone indirettamente identificabili.

[61] Op. Cit. p. 510.

[62] Idem 80.

[63] Op. Cit. p. 39.

[64] Altre tipologie di tecniche di pseudonimizzazione consistono nell'utilizzo di tabelle di transcodifica, di qrcode, oppure di barcode.

[65] In generale esistono due tipologie di funzioni hash: quelle con chiave (keyed) e quelle senza chiave. Mentre le prime garantiscono sia l'autenticazione che l'integrità di una stringa, le seconde si limitano a garantirne l'integrità.

[66] Ovvero, una sequenza di caratteri dotata di un ordine prestabilito.

[67] Ovvero la probabilità che si verifichino delle collisioni è trascurabile.

[68] Si tratta del cosiddetto "*effetto valanga*".

[69] Nello specifico l'episodio è relativo allo scambio di sacche destinate al trapianto di midollo osseo, occorso il 17 ottobre 2018 presso il centro Cellule staminali dell'ospedale San Martino di Genova.

[70] Nello specifico l'episodio è relativo allo scambio di sacche destinate a una trasfusione da realizzarsi nel corso di un intervento chirurgico al femore, occorso il 16 settembre 2019 presso l'ospedale di Vimercate, e che purtroppo ha determinato il decesso della paziente.

[71] Nella versione originale in lingua inglese: "*Released information often contains other data, such as race, birth data, sex and ZIP code, which can be linked to publicly available information to reidentify (or restrict the uncertaintly about) the data respondents, thus leaking information that was not intended for disclosure*".

[72] Op. Cit. p. 514.

[73] Nello specifico il riferimento è a *k*-anonymity, *l*-diversity e *t*-closeness.

[74] Il riferimento è al concetto di trustability.

[75] Regolamento (UE) 2016/679, articolo 11(1).

[76] Regolamento (UE) 2016/679, Considerando 57.

[77] Sono i diritti di cui al Regolamento (UE) 2016/679, articoli da 15 a 20.

[78] Idem 94.

[79] Regolamento (UE) 2016/679, articolo 11(2).

[80] In crittografia un *commitment* è una primitiva crittografica che consente a un'entità coinvolta in uno scambio di informazioni di impegnarsi su di un valore scelto (o su di un'asserzione scelta) tenendolo contestualmente celato alle altre entità coinvolte, e mantenendo la possibilità di rivelare solo successivamente qual sia il valore su cui ci si è impegnati.

[81] Traduzione originale dal testo in lingua inglese "*personal data should be registered on the blockchain preferably in the form of a commitment*".

[82] Traduzione dal testo originale in lingua inglese: "*with respect to additional personal data, in order to ensure compliance with data protection by design and by default and data minimisation obligations, the CNIL recommends solutions in which data is processed outside of the blockchain or, in which the following are stored on the blockchain, in order of preference: a) a commitment of the data; b) a hash generated by a keyed hash function on the data; c) a ciphertext of the data*".

[83] Regolamento (UE) 2016/679, Considerando 26.

[84] Idem 102.

[85] Idem 102.

[86] Op. Cit. p. 511.

[87] Op. Cit. p. 511.

[88] Op. Cit. p. 36.

[89] Idem 107.

[90] Op. Cit. p. 37.

[91] Op. Cit. p. 511.

[92] WP29 è un acronimo utilizzato per identificare il Gruppo di lavoro articolo 29.

[93] In alcuni contesti anche indicate con il termine distorsione.

[94] Op. Cit. p. 511.

[95] Citazione dal testo originale in lingua inglese: *"[Differential privacy] involves injecting 'noise' into the answers of dataset queries. The noise should be great enough to provide anonymity at an individual level, but not enough to affect the utility of the answer*".

[96] Op. Cit. p. 513.

[97] Op. Cit. p. 513.

[98] Per semplicità è possibile immaginare una tupla come una riga di una tabella.

[99] Op. Cit. p. 514.

[100] Op. Cit. p. 544.

[101] Op. Cit. p. 544.

[102] Citazione dal testo originale in lingua inglese: "*The published table should provide the adversary with little additional information beyond the background knowledge. In other words, there should not be a large difference between the prior and posterior beliefs*".

[103] Op. Cit. p. 547.

[104] Citazione dal testo originale in lingua inglese: "*while the l-diversity principle represents an important step beyond k-anonymity in protecting against attribute disclosure, it has several shortcomings*".

[105] In termini informali, possiamo considerare il q-blocco come il gruppo di tuple in cui le identità dei rispondenti si confondono.

[106] Op. Cit. p. 548.

[107] Idem 124.

[108] Citazione dal testo originale in lingua inglese: "*A novel privacy notion called t-closeness, which requires that the distribution of a sensitive attribute in any equivalence class is close to the distribution of the attribute in the overall table*".

[109] Che prende il nome di q-blocco.

[110] Op. Cit. p. 549.

[111] Traduzione dal testo originale in lingua inglese: "*when doing medical research it is important to preserve the correlation of different medical data regarding a given patient but the identity of the patient is irrelevant*". EAG guidelines, pp. 12.

[112] Traduzione dal testo originale in lingua inglese: "*an archive service that holds records in the public interest has to preserve the integrity of medical records selected for permanent preservation in the interest of the data subjects*". EAG guidelines, pp. 12.

[113] Traduzione dal testo orginale in lingua inglese *"archive services might also make use of pseudonymisation, but […] pseudonymisation should be fully reversible and should be done in a way that does not endanger the evidential value of records*". EAG guidelines, p. 13.

[114] Op. Cit. p. 369.

[115] Idem 132.

[116] Traduzione dal testo originale in lingua inglese "*software for archival description that allows the creation of two different versions of a finding aid (one with real names and one with pseudonyms) is a tool for compliance with art. 25*". EAG guidelines, p. 23.

[117] Op. Cit. pp. 372-373.

[118] A tal proposito si faccia riferimento al Regolamento (UE) 2016/679, art. 17.

[119] A tal proposito si faccia riferimento al Regolamento (UE) 2016/679, art. 16.

[120] Traduzione originale dal testo in lingua inglese "*to keep pace with emerging threats in today's digital world*".

[121] Op. Cit. p. 39.

[122] Idem 140.

[123] Op. Cit. p. 698.

[124] Del resto la DPIA o valutazione d'impatto è un'attività limitata, riguardando solo l'articolo 32 del Regolamento (UE) 2016/679.

- Bio
- Social
- Latest Posts

- By: Francesco Ciclosi

  Francesco Ciclosi è da anni professore a contratto degli insegnamenti di "Informatica" e di "Sicurezza Informatica" in corsi di laurea e master dell'Università degli Studi di Macerata. Autore di numerosi articoli e monografie, i suoi ambiti di ricerca includono la protezione dei dati, la gestione del rischio e l'evoluzione dei sistemi di gestione della sicurezza delle informazioni. Attualmente è membro del Comitato Tecnico Scientifico della Federazione IDEM

-

- See all this author's posts