Estratto della valutazione d'impatto sulla protezione dei dati (c.d. DPIA) relativa al Progetto

Gemelli Generator – Real World Data

Il Progetto **Gemelli Generator** (*GEmelli NEtwoRk for Analysis and Tests in Oncology and medical Research*) – *Real World Data* (di seguito "Gemelli Generator"), ha come obiettivo la valorizzazione clinica e scientifica di competenze, processi e dati del *Data Warehouse* (DWH) della Fondazione Policlinico Universitario A. Gemelli IRCCS (di seguito "Fondazione"), attraverso servizi avanzati di estrazione, raccolta e analisi di grandi volumi di dati.

Gli studi e le ricerche condotte nell'ambito di Gemelli Generator sono valutati e condotti considerando i potenziali benefici in riferimento al miglioramento dei percorsi di cura per il paziente, ai risultati attesi dalla comunità scientifica e al progresso sociale generato per l'intera collettività.

1) Nozione di valutazione d'impatto

"Una valutazione d'impatto sulla protezione dei dati (c.d. DPIA) è un processo inteso a descrivere il trattamento, valutarne la necessità e la proporzionalità, nonché a contribuire a gestire i rischi per i diritti e le libertà delle persone fisiche derivanti dal trattamento di dati personali, valutando detti rischi e determinando le misure per affrontarli"[1].

2) Quadro normativo

- Regolamento UE 2016/679: Articolo 9, paragrafo 2, lettera J) e Articolo 35 – Considerando C84, C89, C90, C91, C92, C93, C95;
- Decreto legislativo 196/2003 e s.m.i, Codice in materia di protezione dei dati personali: Articolo 110;

[1] Cfr. ART. 29 WORKING PARTY, *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 il 4 aprile 2017 e modificate il 4 ottobre 2017.

- Article 29 Working Party, 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 il 4 aprile 2017 e modificate il 4 ottobre 2017;
- Garante per la protezione dei dati personali, provvedimento n. 467 dell'11 ottobre 2018 – G.U. 269 del 19 novembre 2018.

3) Soggetti interessati in Gemelli Generator

L'attività di Gemelli Generator interessa il trattamento di dati riguardanti:

- pazienti assistiti o presi in cura presso la Fondazione;
- soggetti arruolati in studi clinici o progetti di ricerca condotti presso la Fondazione.

4) Valutazione preliminare

4.1 Descrizione del trattamento

La missione di Gemelli Generator è quella di contribuire a generare valore e beneficio per tutti i pazienti, operando in progetti di ricerca mediante un migliore utilizzo dei dati clinici retrospettivi e prospettici, delle informazioni relative alla qualità della vita e di quelle relative all'igiene e alla medicina sociale (*Real World Data*).

Per il raggiungimento della propria missione, Gemelli Generator opera mediante i seguenti *asset*:

- **Creazione di *database* clinico-molecolari *ad hoc*** consultabili su piattaforma *web* per una rapida visualizzazione dei dati da parte dei ricercatori secondo parametri specifici, ad esempio a seconda della patologia o del gruppo di lavoro;
- ***Data entry*:** attività di raccolta dati svolta da figure professionali con competenze tecniche, scientifiche, gestionali ed organizzative;
- ***Data registry*/archivi digitali:** sviluppo, organizzazione e gestione di registri dati/archivi digitali secondo parametri assegnati;
- ***Data Clustering*:** realizzazione di *dataset* clinici specializzati per patologia ("*Data Mart*"). Il*dataset* può includere sia dati retrospettivi che dati prospettici raccolti nell'ambito di uno studio clinico, ed integrare anche rilevazioni di sintomi e indicatori di qualità della vita, mediante dispositivi *wearable* come nel caso di *Patient Reported Outcome Measures* (PROMs) e di *Patient Reported Experience Measures* (PREMs).

4.2 Valutazione della conformità

- **Base giuridica/motivazione legittima del trattamento:** ai sensi dell'Articolo 9, paragrafo 2, lettera j) del Regolamento UE 2016/679 e dell'Articolo 110 del Decreto legislativo 196/2003 e s.m.i, il trattamento è necessario a fini di ricerca scientifica in conformità dell'Articolo 89,

paragrafo 1 del suddetto Regolamento, sulla base del diritto dell'Unione o nazionale, che è proporzionato alla finalità perseguita, rispetta l'essenza del diritto alla protezione dei dati e prevede misure appropriate e specifiche per tutelare i diritti fondamentali e gli interessi dell'interessato.

- **Soggetti che accedono ai dati:** personale interno incaricato del trattamento da parte del titolare e del responsabile.
- **Modalità di raccolta dei dati:** attraverso *software* specialistico che ha in *input* i dati pseudonimizzati contenuti nel *datawharehouse* aziendale e che fornisce in *output* dati aggregati.
- **Modalità di aggiornamento e eliminazione dei dati:** i dati sanitari pseudonimizzati trattati nell'ambito di Gemelli Generator potranno essere aggiornati e/o cancellati a seguito di apposita richiesta da parte dell'interessato.
- **Modalità di rilascio dell'informativa agli interessati:** sono affisse presso i locali della Fondazione, sia l'informativa generale sulle attività complessivamente svolte dalla Fondazione, sia l'informativa specifica dedicata alle attività eseguite nell'ambito di Gemelli Generator. Entrambe le informative sono altresì disponibili collegandosi alla sezione "*Privacy e Protezione dati*" del sito della Fondazione.
- **Trasferimento a soggetti terzi e in Paesi extra Ue:** non previsto.
- **Periodo di conservazione dei dati:** i dati sanitari pseudonimizzati verranno trattati nell'ambito di Gemelli Generator per il tempo previsto dalle normative vigenti in materia di sperimentazioni cliniche e, in particolare, per un periodo non superiore a 25 anni dal momento della raccolta dei dati.
- ***Asset* a sostegno del trattamento:** *hardware*, *software*, archivi e reti sono tutti compresi nel perimetro aziendale e gestiti da personale interno anche a livello di amministrazione dei sistemi. Gli *asset* ricadono dunque sotto le politiche di gestione e sicurezza aziendali.

4.3 Motivi della valutazione d'impatto

La DPIA è stata realizzata per valutare i potenziali rischi che possono derivare dall'attività di Gemelli Generator nei confronti degli interessati, così da consentire alla Fondazione d'intervenire preventivamente attraverso l'adozione di adeguate misure di sicurezza.

L'esecuzione della DPIA per Gemelli Generator è stata ritenuta necessaria in ragione:

- del volume e della tipologia di dati utilizzati (dati personali relativi alla salute);
- dell'ampio numero e della tipologia d'interessati coinvolti (per lo più pazienti e, dunque, suscettibili di poter essere intesi quali soggetti vulnerabili);
- della durata prolungata dell'attività di trattamento svolta nell'ambito del progetto;
- delle innovative soluzioni tecnologiche e modalità di analisi impiegate (sistemi di Intelligenza Artificiale, dispositivi *wearable*).

La DPIA è stata inoltre realizzata in ottemperanza alle previsioni dell'Articolo 110 del D.lgs. 196/2003 che individua nell'esecuzione e nella pubblicazione della DPIA uno dei requisiti la cui soddisfazione rende legittimo il trattamento dei dati relativi alla salute a fini di ricerca scientifica anche senza il consenso dell'interessato.

5) Conduzione della DPIA

5.1 Metodo adottato

Per ogni singolo trattamento vengono individuati gli asset direttamente o indirettamente ad esso collegati. Per ognuno di essi, il processo di analisi dei rischi analizza le vulnerabilità, le relative minacce, e le contromisure, dirette o indirette, attuate, fornendo il livello di rischio. Tale livello tiene anche conto della probabilità e dell'impatto che l'attuazione della minaccia avrebbe sui dati personali trattati, per mezzo degli specifici asset.

Tale metodologia può essere riassunta nella seguente funzione:

$$R_T = f (V_T , P_T , D_T)$$

Dove:

R_T = indice di rischio che insiste sul Trattamento, espresso in valori percentuali;

V_T = indice di vulnerabilità degli *asset* coinvolti nel trattamento, tenuto conto delle contromisure, dirette o indirette, attuate e del livello di criticità espresso sul singolo *asset*;

P_T = probabilità di accadimento dell'evento indesiderato sul trattamento;

D_T = gravità delle conseguenze della concretizzazione dell'evento indesiderato sul trattamento.

Il sistema può essere così rappresentato:

Scala dell'indice di rischio

La scala del livello di rischio utilizzata si configura come segue:

- Rischio molto basso
- Rischio basso
- Rischio medio
- Rischio alto
- Rischio molto alto

5.2 Identificazione e valutazione dei rischi

Di seguito un estratto del livello di rischio per tipologie di asset, calcolato anche sulla base di vulnerabilità indirette.

Rischio	Software	Luoghi	Organizzazione	Postazioni di lavoro	Server	Apparati Rete dati	Storage/Backup	disponibilità	riservatezza	integrità
Malfunzionamento e degrado di apparecchiature	\	\	basso	basso	basso	basso	basso	x		x
Uso non autorizzato di apparecchiature	\	\	\	basso	medio	basso	basso	x	x	x
Furto di apparecchiature	\	\	\	basso	basso	basso	basso	x	x	x
Indisponibilità sistema informativo	basso	\	\	medio	basso	medio	basso	x		
Ingressi non autorizzati ai locali ad accesso ristretto	\	basso	basso	\	\	\	\		x	
Mancanza di alimentazione elettrica	\	medio	medio	medio	basso	medio	basso	x		
Incendio	\	basso	basso	\	\	\	\	x		x
Allagamento	\	basso	basso	\	\	\	\	x		x
Furto di dati	basso	\	basso	medio	basso	basso	medio	x	x	

Modifica non autorizzata di dati	basso	\	basso	basso	medio	basso	basso			x
Presa visione abusiva di dati	basso	\	basso	basso	medio	basso	basso		x	
Mancata accessibilità dei dati	basso	\	basso	basso	basso	basso	basso	x		x
Mancata conservazione dati	basso	\	\	\	\	\	medio	x		x
Malfunzionamento software	basso	\	\	\	basso	\	\	x		x
Uso non autorizzato di software	basso	\	\	\	basso	\	\		x	
Azioni di software suscettibili di arrecare danno	medio	\	medio	basso	medio	medio	medio	x	x	x
Accessi non autorizzati al software	medio	\	\	\	medio	\	\		x	
Errore nello svolgimento di mansioni	\	\	basso	\	\	\	\	x	x	x
Ignoranza procedure di gestione	basso	\	basso	basso	basso	basso	basso	x	x	x
Guasto ai sistemi complementari	\	medio	medio	\	\	\	\	x		x

5.3 Mitigazione rischi

La mitigazione dei rischi sopra riportati avviene applicando le seguenti misure di sicurezza a garanzia della riservatezza, disponibilità e integrità dei dati a livello:

- **organizzativo**, tramite ad esempio: istruzioni interne, assegnazione degli incarichi a personale qualificato, formazione agli incaricati del trattamento, profilazione degli accessi a sedi e sistemi informatici;

- **fisico**, tramite ad esempio: gestione degli accessi alle sedi del trattamento, dispositivi di allarme, vigilanza, dispositivi antincendio, dispositivi di controllo di umidità e temperatura, continuità dell'alimentazione elettrica;

- **logico**, tramite ad esempio: gestione delle credenziali di accesso a sistemi e *software* con *password policy* di complessità e durata, gestione dei *log* degli accessi a sistemi e programmi, antivirus centralizzato, *firewall* perimetrali, ridondanza e virtualizzazione dell'infrastruttura informatica a supporto, *backup* dei dati, *snapshot* dei sistemi, ridondanza dei collegamenti di rete, compartimentazione logica delle reti, trasmissione cifrata dei dati, *vulnerability assessment* periodici.

6) Risultati DPIA

Tutto ciò valutato, e soprattutto

- considerata la natura aggregata dei dati presentati dai *report* di Gemelli Generator;
- considerato che gli algoritmi di I.A. sono utilizzati nelle ricerche per individuare i fattori di rischio che influenzano la fase diagnostica di una patologia **non attraverso l'individuazione di soggetti a rischio**, ma attraverso la comprensione dei fattori che influenzano l'andamento di una patologia;

si ritiene che il trattamento in esame, allo stato attuale, presenta un grado di rischio basso per i diritti e le libertà degli interessati al trattamento stesso, espresso secondo la scala di valutazione dei rischi adottata.

7) Revisione ed aggiornamento, con riesame di congruità con le esigenze di protezione dei dati

Il titolare del trattamento procede a un riesame per valutare se il trattamento dei dati personali sia effettuato conformemente alla valutazione d'impatto sulla protezione dei dati almeno quando insorgono variazioni del rischio rappresentato dalle attività relative al trattamento (Articolo 35, paragrafo 11 del Regolamento UE 2016/679). La Fondazione procederà a revisione ed aggiornamento della presente DPIA in corrispondenza dell'introduzione di nuove procedure nell'ambito di Gemelli Generator.