Digital health e protezione dei dati

Francesca Morelli - giornalista scientifica

CYBERSECURITY vantaggi e limiti per la ricerca clinica

Big Data, intelligenza artificiale, codice della privacy: la gestione dei dati sensibili impone l'adozione di misure cautelative contro potenziali attacchi cibernetici, ma non solo. A vigilare c'è il nuovo GDPR, però non sempre "vantaggioso". Il tema è stato affrontato durante una Sessione dedicata del 62° Simposio AFI

Da un lato la necessità di proteggere la sicurezza, l'integrità e la privacy del dato, nello specifico clinico-sanitario, con appositi tool e soluzioni innovative contro attacchi cibernetici; dall'altra l'arrivo di nuove norme e regolamentazioni del GDPR, Regolamento Generale sulla Protezione dei Dati, che sembrano togliere "competitività" alla ricerca clinica italiana. Qual è lo stato dell'arte? Se ne è parlato al 62° Simposio di Associazione Farmaceutici Industria (AFI) nel corso della Sessione "Privacy, Cybersecurity, nuove norme: quale impatto per la ricerca scientifica", introdotta da **Lorenzo Cottini** (AFI - Evidenze Health) e moderata da **Guido Fedele** (AFI), **Sergio Scaccabarozzi** (AFI - Arithmos) e **Francesca Vaccari** (AFI - Chiesi Farmaceutici).

Sistemi informativi e gestione dei processi

Nella ricerca clinica e scientifica, ma non solo, ciò che conta è il dato: asset fondamentale di ogni studio, esso viene giornalmente costruito, alimentato, modificato, usato, "manipolato" da più attori. Aspetti che ne denunciano la forza, ma anche la fragilità e l'alta esposizione a potenziali rischi: danneggiamento, smarrimento, furto; eventi che possono impattare anche sul business, gravati dal circuito di attori coinvolti nella gestione del dato. «In ambito di ricerca - spiega **Gian Paolo Baranzoni**, GdS Sistemi Informativi di AFI - occorre tenere in considerazione un "perimetro di riferimento" che includa soggetti e ruoli di varia forma e contenuto: aziende farmaceutiche, centri di ricerca e ricercatori, IRCCS/ospedali e organi sanitari, enti regolatori. Soggetti non omogenei che impattano sul processo di tutela e sicurezza del dato, nello specifico, con necessità organizzative, infrastrutture tecnologiche e problematiche completamente diverse. Tuttavia, realtà accomunate da una simile pluralità di fattori di rischio, dipendenti e aumentati dall'elevato numero di risorse che collaborano alla gestione, creazione, distribuzione, revisione del dato; dall'esponenziale incremento di informazioni circolanti, sia in termini di quantità - i cosiddetti Big Data - sia di qualità cui va garantita riservatezza e sicurezza e infine, ma non ultimo, i diversi supporti con cui viaggiano i dati, dalla carta al video, e le infrastrutture sempre più eterogenee e complesse che comprendono soggetti diversi, tecnologie diffe-

renti, continuità di comunicazioni trasmesse dal singolo dato di laboratorio al dossier di registrazione di un nuovo farmaco». È indubbio che in tali contesti il dato è soggetto e oggetto di attacchi *cyber*, sempre più sofisticati e complessi, e tenuto conto che resta "materia riservata" per lunghi anni, l'accesso al dato dovrebbe essere concesso solo a soggetti abilitati e protetto da sistemi IT in grado di preservare il patrimonio informativo, di gestirlo nella sua eterogeneità e limitare quanto più possibile la Data Loss Prevention: vi è evidenza che il 31% dei dati venga perso a causa di eventi correlati al cattivo controllo nell'utilizzo.

Le strategie di Data Loss Prevention

Definire il luogo e il momento di presenza del dato: questi sono i due aspetti che consentono di identificare e avviare strategie di contenimento del rischio mirate ed efficaci. «Il dato - prosegue Baranzoni - può essere presente in tre specifiche fasi: la fase di archiviazione, quella di trasferimento dal punto di archiviazione all'utilizzatore, la fase di utilizzo. Nella fase del Data at Rest Protection ci si focalizza sulla protezione del dato memorizzato. La tutela può essere garantita da strumenti quali la crittografia dell'intero disco o dispositivo o, parzialmente, a livello di file o Database, così come da altre tecnologie quali l'Information Rights Management (IRM), il MDM (Mobile Device Management), il DLP (Data Leak Prevention) e/o il CASB (Cloud Access Security Broker). Un secondo momento chiave è la fase di Data in Transit Protection in cui ci si focalizza sulla protezione del dato in transito, sia all'interno della rete privata che tramite Internet, con l'Email encryption, il Managed File Transfer (MFT), l'IRM, il DLP o CASB sono procedure di efficacia cautelativa. Infine, il terzo momento critico è Data in Use Protection dove ci si focalizza sulla protezione del dato in uso, in continua modifica o aggiornamento da parte dell'utente o delle applicazioni, in questa fase è necessaria l'attenzione all'utilizzatore e quindi agire con una serie di strumenti: l'Identity Management Tools, il Role Based Access Control e il Digital Rights Protection o l'IRM». Due nuove norme regolano la tutale del dato in ambito europeo: le *Guidelines on computerised system and electronic data in clinical trials* emesse da EMA e la Direttiva UE 2022/2555 del Par-

lamento Europeo e del Consiglio (NIS2) che impone che aziende di una certa importanza si dotino entro il 2024 di sistemi di Cybersecurity. Il mondo chimico-farmaceutico e della ricerca entrano a pieno titolo nella norma in quanto ambiti di gestione di dati sensibili.

Le best practice

L'adozione di alcuni comportamenti, fin dalle fasi di implementazione e di azioni di Cybersecurity, possono contribuire a garantirne efficacia e performance. «È necessaria la collaborazione e la partecipazione attiva al processo di tutte le aree aziendali, *stakeholders* inclusi - conclude Gian Paolo Baranzoni - mantenere un approccio metodologico alla sicurezza, applicando cioè la ISO 27000 sulla sicurezza di Dati e Informazioni (Fisiche e Logiche) e attuando i requisiti della NIS2 secondo le indicazioni del NIST; garantirsi accessi sicuri tramite strumenti e strategie dedicate quali il Segregation of Duties, IAM, IRM, RBAC, 2FA/MFA. Inoltre, è bene predisporre la crittografia dei dati con dischi crittografati, sicurezza delle chiavi, HTTPS e, infine, è raccomandato strutturare un monitoraggio proattivo con l'introduzione di processi, servizi di security quali SIEM, MDR, IPS, IDS, CSIRT». All'opposto, è da evitare il faida-te: in relazione alla complessità della materia e delle sinergie da mettere in atto, l'autonomia non è la strada per percorrere soluzioni estemporanee; occorre esser padroni delle problematiche e non riferirle per sentito dire; lasciare accessi liberi, cioè non controllati o verificati ogni 6/12 mesi, così come dati liberi o non classificati; barricarsi dietro illusorie sicurezze quali il "non è mai successo", mentre tutto può accadere, anche in assenza di precedenti. Infine, è bene porsi in una condizione di monitoraggio reattivo: in caso di Log Management o di Incident Management, il danno è già fatto.

La validazione della firma elettronica

In funzione di Cybersecurity, diversi tool sono stati sviluppati o sono in via di "progettazione". Tra questi, uno strumento-processo per la validazione della firma elettronica di documenti relativi a studi clinici, nel rispetto della *compliance* regolatoria. «È necessaria una prima distinzione - precisano **Stefano Piccoli**, di QStep, azienda specializzata in servizi di CSV e di Digital Compliance ed **Elena Ciuchini**, di CROLife, azienda specializzata in soluzioni innovative per sperimentazione clinica, regolamentazione e farmacovigilanza - fra firma elettronica semplice in cui si "dichiara" tramite l'apposizione di simbolo o di un *tick* su un documento digitale il preciso intento di firmare un documento in cui tuttavia non è possibile identificare il soggetto firmatario, dalla firma elettronica avanzata o da quella digitale. Quest'ultima, invece, consente di verificare l'identità della persona da parte di un ente terzo attraverso un "certificato" e di associare la persona firmataria al documento stesso, co-

NIST Cybersecurity Framework

Il rischio informatico viene stimato in funzione del NIST Cybersecurity Framework, un processo che si struttura in 5 fasi, dall'identificazione dell'evento di esposizione specifico alla potenziale risoluzione:

1. **Identify:** è la fase di individuazione degli eventi Cyber potenzialmente disastrosi, identificata tramite il quadro di riferimento per la gestione del rischio informatico, le diverse norme di ambito (GCP /GDPR /Nuova NIS2), le norme/guideline efficaci tra queste ENISA/ISO/NIST e tutte le aree aziendali potenzialmente coinvolte (R&D, Regulatory, Quality Assurance, IT, Purchasing, Production, Engineering, *etc.*).
2. **Protect:** è la fase in cui si valuta il contesto di rischio in relazione ad esempio a dati, classi, trattamenti/sistemi, il livello di rischio associato ai singoli eventi Cyber individuati, le probabilità di accadimento e la gravità di impatto sulle attività di business.
3. **Detect:** in questo terzo momento si predispongono e pianificano misure di prevenzione e protezione degli eventi critici di rischio individuati.
4. **Respond:** questa fase è veicolata all'implementazione

delle misure protettive individuate su processi, servizi e sistemi con azioni di mitigazione anche in caso di «Disastro» (DRP, Disaster Recovery Plan e BCP, Business Continuity Plan).

5. **Recover:** l'ultima fase richiede il monitoraggio periodico e revisione delle misure adottate, ovvero l'implementazione dei sistemi/servizi necessari, conformati con le Norme, per ridurre il rischio, impedire utilizzi anomali o verificare e anomalie, affrontare situazioni critiche, tra queste anche segnalazioni di violazione.

Il tutto (nuova fase introdotta con la versione 2) gestito all'interno di una fase Govern fondamentale per integrare la sicurezza informatica nella più ampia strategia di gestione del rischio aziendale di un'organizzazione, Govern ha l'obiettivo di orientare la comprensione del contesto organizzativo; la definizione di una strategia di sicurezza informatica e di gestione del rischio della catena di fornitura della sicurezza informatica; ruoli, responsabilità e autorità; politiche, processi e procedure; e la supervisione della strategia di sicurezza informatica.

Fonte: The NIST Cyber Security Framework 2.0, Agosto 2023

sicché in caso di manomissione la firma possa essere invalidata». Restano in materia alcune questioni aperte: è sempre richiesta la convalida di una firma elettronica in uno studio clinico? Non è *conditio sine qua non*, la necessità è determinata dall'uso e dal tipo di documento da firmare; le Predicate Rules in ambito GCP (Good Clinical Practice) richiedono l'utilizzo della firma elettronica? Le Predicate Rules non specificano se usare una firma su carta o elettronica, ma definiscono se un documento debba essere approvato. Qualora per approvare un documento richiesto dalle Predicate Rules si utilizzi una firma elettronica, questa dovrà essere validata ed essere conforme alle normative applicabili (e.g. 21CFR part 11). Da qui, la necessità di disegnare un progetto in cui definire chiaramente l'*intended use* e cioè definire i documenti dello studio clinico da approvare con lo strumento di firma, identificare i diversi workflow di firma, consentire la gestione di firmatari interni ed esterni e garantire la conformità della soluzione con i requisiti regolatori sulla firma elettronica, quali il *Code of Federal Regulation, 21 CFR - Part 11 Electronic Records Electronic Signature* e la Guidelines EMA sui sistemi computerizzati in ambito Good Clinical Practice, precedentemente citate. Norme che richiedono l'obbligo di verifica dell'identità del firmatario, ovvero che la firma sia legata in modo univoco al firmatario, che sia assicurata la non ripudiabilità della firma e l'esistenza di una connessione indissolubile tra il documento elettronico firmato e la firma stessa. Infine, che i documenti firmati contengano una specifica "anagrafica" (*audit trail*): nome e cognome del firmatario, data, ora e motivo della firma.

La scelta della soluzione

In funzione degli obiettivi e scopi sopra espressi, tra le varie opzioni disponibili per la realizzazione del progetto è stata privilegiata la firma elettronica di Adobe Acrobat Sign, un sistema Cloud che richiede l'acquisto di alcune licenze particolari per l'uso conforme alla normativa. «Abbiamo disegnato il processo di firma per i diversi documenti - prosegue Piccoli - e (con)validato le configurazioni e le funzioni che garantissero la soddisfazione dell'Intended Use, attraverso Risk Based Approach che ha consentito di identificare le possibili *failure* del processo e definire i controlli procedurali e tecnici per mitigarli. Congelati i requisiti, nella Project Phase abbiamo provveduto alla convalida e configurazione della firma, definendo utenti interni ed esterni, ciascuno provvisto delle chiavi di accesso, alla definizione dei flussi, dei motivi di firma e alla creazione delle regole di *retention* dei documenti per il trasferimento dei documenti firmati dal Cloud al *repository* finale di CROLife. Infine, è stata prevista l'inclusione dell'*audit trail* nel documento firmato». Come ultimo elemento è stato definito un processo di firma con dei passaggi fissi - Send, Sign, Track, Archive - utili a "chiudere" la firma elettronica e a poter sca-

Alcuni dei relatori che hanno partecipato alla Sessione IV del 62° Simposio AFI

ricare il documento finale, sicuro e così come richiesto dalla normativa. La Use Phase, ultimo step del progetto, è stata dedicata alla stesura di procedure per l'uso corretto della firma dei documenti, per la gestione degli utenti e del *workflow* di approvazione, l'archiviazione dei documenti e delle firme, l'amministrazione degli accessi e delle credenziali di firma, e infine la gestione delle modifiche del tool di firma.

Aspetti legali

Big Data, Intelligenza Artificiale (IA), Ricerca Clinica: in tutti questi ambiti i dati sensibili si scontrano con aspetti di privacy, oltre che di Cybersecurity. Legalmente, come comportarsi? Il Regolamento generale sulla protezione dei dati (GDPR), seppur approvato nel 2016, non sembra uno strumento attualmente sufficiente per affrontare le sfide offerte dalla mole enorme di dati e di algoritmi, esplosi esponenzialmente dopo Covid-19: occorre interrogarsi sulle problematiche e implicazioni di Big Data e IA e le possibili soluzioni. «Riguardo ai Big Data - informa l'Avvocato **Silvia Stefanelli**, GdS Salute digitale AFI - il GDPR mette a disposizione due strumenti che occorre tenere in considerazione: l'Art. 22 sulla profilazione che consente di "interrogare" correttamente i dati, e l'Art. 35 sulla valutazione di impatto (c.d. DPIA). Quest'ultimo è uno strumento molto importante, senza dubbio complesso da applicare, ma il cui utilizzo è fortemente raccomandato in quanto consente di fare emergere elementi che qualche volta possono essere "sfuggiti" a una prima analisi. La DPIA poi appare di assoluto rilievo anche in relazione a molto aspetti che già oggi sono presenti nella Proposta di Regolamento sulla AI, quali la qualità del dato, le modalità di raccolta e la convalida del software. Una delle maggiori criticità del settore è poi stabilire la corretta base giuridica per la ricerca clinica: i dati, infatti, vengono per lo più raccolti all'interno degli ospedali per diagnosi e cura e poi si apre il problema di un possibile utilizzo ulteriore (il c.d. uso secondario) per attività di ricerca. La difficoltà di questa materia e anche le diverse applicazioni del GDPR in ambito comunitario hanno portato oggi il legislatore comunitario a introdurre attraverso una proposta di Regolamento uno spazio per condividere i dati sanitari, l'Health Data Space. «La Comunità Europea - prosegue l'Avvocato - sta promuovendo un ampio *repository* di dati sanitari a cui tutti potranno andare ad attingere, dalle pubbliche amministrazioni ai privati, dalle imprese ai ricercatori, proponendo anche un regolamento per l'uso secondario dei dati sanitari elettronici. In particolare, l'Art. 34 della proposta di regolamento elenca tutti gli usi secondari, includendo fra queste attività di trattamento per il sostegno di enti pubblici o di istituzioni che operano nel settore sanitario; la produzione di statistiche ufficiali e studi di ricerca a livello nazionale, multinazionale e dell'Unione in ambito sanitario o assistenziale; attività di sviluppo e innovazione per prodotti o servizi che contribuiscono a garantire elevati livelli di qualità e sicurezza dell'assistenza sanitaria, dei medicinali o dei dispositivi medici, nonché attività di addestramento, prova e valutazione degli algoritmi, anche nell'ambito di dispositivi medici, sistemi di IA e applicazioni di sanità digitale, che contribuiscono alla sanità pubblica o alla sicurezza sociale, o che garantiscono elevati livelli di qualità e sicurezza dell'assistenza sanitaria, dei medicinali o dei dispositivi medici: infine, utilizzo secondario dei dati per l'erogazione di un'assistenza sanitaria personalizzata tale da mantenere o ripristinare lo stato di salute della persona sulla base dei dati sanitari di altre persone fisiche».

Il GDPR non sembra uno strumento attualmente sufficiente per affrontare le sfide legate a Big Data, AI e ricerca clinica

La privacy secondo il paziente

Caso emblematico sono i pazienti affetti da malattie rare: solo il 6% ha a disposizione una cura. Secondo un'ampia *survey* del 2020, condotta dalla Federazione Europea sulle malattie rare su oltre 2 mila pazienti rappresentativi di 664 diverse malattie rare sulle oltre 8.000 attualmente conosciute, il 97% condividerebbe i propri dati sanitari a scopo di ricerca e il 95% li metterebbe a disposizione anche per altri e nonostante i rischi associati ai dati, di alcuni dei quali non c'è sempre una piena consapevolezza. I principali rischi identificati sono che i propri dati sanitari vengano condivisi con terze parti senza il proprio consenso o usati per un contesto diverso da quello per cui sono stati messi a disposizione, o che vengano usati a fini discriminatori, ad esempio nel mondo del lavoro, o a propria insaputa. Aspetto rilevante: l'89% dei malati, in ambito di trattamento del dato, si fida maggiormente delle associazioni di pazienti o di ricercatori e medici non profit, piuttosto che di professionisti e aziende profit e, non ultimo, i pazienti chiedono una restituzione sull'utilizzo dei propri dati. In Italia una simile riflessione è stata portata avanti da LIRH (Lega Italiana Ricerca Huntington) in collaborazione con altri enti che conducono a evidenze simili: «È emerso - commenta **Barbara D'Alessio**, Presidente LIRH - che i pazienti con malattie rare si percepiscono più fragili rispetto ad altri, preferiscono avere un filtro tra sé e il ricercatore e ricevere una comunicazione trasparente circa i soggetti terzi coinvolti e il tipo di ricerca veicolato al proprio dato. Auspicano un ritorno, anche economico per la società, e non ultimo chiedono la garanzia dell'anonimato del loro dato in tutto il processo. Per rispondere alle attese dei pazienti e agli obiettivi di ricerca è fondamentale che il percorso del dato sia tutelato da un sistema di Governace allargata, superando cioè la visione monocratica verso una pluralità di voci, incluse quelle dei pazienti».

Un parere unanime

Il GDPR non fa chiarezza in ambito di ricerca clinica: a livello globale non rende attrattiva la ricerca in Europa e aggiunge un carico maggiore all'Italia, specie nell'esecuzione di ricerche cliniche da parte di IRCCS. Si concorda con la proposta di un approccio unitario, pragmatico, e qualora il garante si profilasse come uno scoglio duro da superare, si prospetta come possibile soluzione da adottare la definizione di un albero decisionale che a monte esprima parere favorevole per studi in determinate categorie. Il GDPR ha prodotto dichiarazioni di principio, senza arrivare a categorizzare le tipologie di ricerca; pertanto, sarebbero più che auspicabili interlocuzioni con il garante. È necessario rendere l'Italia competitiva, rivedendo le norme che riguardano la sperimentazione clinica, interagendo anche a livello Ministeriale. Si concorda, infine, sulla necessità di anonimizzazione del dato con algoritmi o metaverso. È questo il parere unanime espresso dai partecipanti alla Tavola Rotonda conclusiva (**Pietro Calamea**, Ministero della Salute; **Fabrizio Forini**, Presidente Aicro; **Gualberto Gussoni**, FADOI; **Sandra Petraglia**, AIFA; **Carlo Tomino**, AFI e IRCCS San Raffaele di Roma; **Marco Zibellini**, Farmindustria) moderata da **Lorenzo Cottini**, AFI- Evidenze Health; **Ilaria Maruti**, AFI – Astra Zeneca; **Sergio Scaccabarozzi**, AFI – Arithmos).

La ricerca nel GDPR

C'è scarsa armonizzazione del GDPR in tema di ricerca scientifica: non si è trovato ancora un accordo a livello europeo, lasciando libertà di azione agli Stati membri dell'Unione. «La disciplina italiana ha una visione consensocentrica, definita nel GDPR dagli art. 6 e 9 - spiega Silvia Stefanelli - il primo pone le basi giuridiche del trattamento e il secondo le condizioni legittimanti il trattamento di particolari categorie di dati. In tema di ricerca si incappa poi nel codice privacy regolamentato da due norme: l'Art. 107, riferito alla ricerca scientifica e Art. 110 inerente invece alla ricerca medica, biomedica e epidemiologica». Restano, tuttavia ancora molti temi aperti, tra questi:

- cosa rientri nella nozione di «ricerca scientifica» ex Art. 107 e cosa nella nozione di «ricerca medica e biomedica» ex Art. 110;
- quando un trattamento sia necessario (Corte di Cassazione ordinanza 6177/2023);
- se l'azienda Pharma che svolge la sperimentazione o l'azienda Medical Device che svolge una indagine clinica possa usare la base giuridica dell'Art. 6 (obbligo di legge) e Art. 9 lett. i) (parametri sicurezza farmaci e dm) oppure art. 9 lett. j (ricerca scientifica);
- se gli enti di ricerca, quai gli IRCSS, possano utilizzare come base giuridica Art. 6 (compito di interesse pubblico) e l'Art. 9 lett. j (ricerca scientifica);
- se non sia possibile l'anonimizzazione dei dati della ricerca.

Questa potrebbe essere una soluzione per trasportare e trattare il dato fuori dal GDPR.