Consiglio Nazionale delle Ricerche

Istituto di Fisiologia Clinica

Area della Ricerca S. Cataldo Pisa

Sez. Epidemiologia e Ricerca sui Servizi Sanitari

Technical Report / Aprile 2014

Gestione e trasmissione sicura dei dati attraverso piattaforme tecnologiche: dati sensibili e privacy

Version 1.0

Autori

L.Fortunato, A.Cutilli, S.Pieroni, S.Molinaro

Sommario

Si descrivono nel seguente documento alcune delle politiche prese in esame o adottate per la gestione dei problemi riguardanti la raccolta informatizzata di dati personali (nome, cognome, sesso, data di nascita, CF, …) e sensibili (razza, etnia, convinzioni religiose, opinioni politiche, dati che rivelano lo stato di salute, …), in linea con le norme attualmente in vigore (DL 30/06/2003 n.196 "Codice in materia di protezione dei dati personali", derivante dalla direttiva europea 95/46/CE) e con le disposizioni sul FSE (Linee Guida in tema di Fascicolo Sanitario Elettronico e di Dossier Sanitario del 16 Luglio 2009, Prescrizioni in tema di Fascicolo Sanitario elettronico, G.U. n.178 03/08/2009).

In particolare, si focalizzano gli aspetti di progettazione ed implementazione di piattaforme tecnologiche per la gestione di tali dati via web; procedure di autenticazione e tracciabilità; messa in sicurezza dell'hardware e dei database; standard crittografici; modalità di archiviazione sicura e ridondata dei dati; gestione in sicurezza di archivi offline, contenenti dati personali e sensibili. Vengono descritte le figure rilevanti che intervengono nel processo di trattamento dei dati personali e sensibili: il Titolare del trattamento, il Responsabile e l'Incaricato al trattamento; le Autorità coinvolte nel processo di trattamento del dato sensibile, il Garante per la protezione dei dati personali ed il Comitato Etico. Infine, si fa cenno alla documentazione che occorre necessariamente mettere a punto per informare il soggetto sulle finalità della raccolta ed elaborazione di dati personali/sensibili che lo riguardano, e per raccogliere il suo consenso scritto ad un determinato trattamento dei suoi dati personali.

Memorizzazione e trasmissione sicura dei dati attraverso piattaforme tecnologiche

Piattaforme web: la gestione dei dati, con operazioni di inserimento /modifica/cancellazione/consultazione/stampa è effettuata tramite sistemi web ad accesso riservato alle sole figure di riferimento, precisamente individuate (medici professionisti, infermieri, ricercatori, ..), cui viene fornita una login di accesso. Tali persone sono le stesse che dovranno essere autorizzate al trattamento dei dati sensibili dal paziente, con specifico consenso.

Profilazione delle utenze: il software applicativo deve permettere di regolamentare la visibilità dei dati, garantendo che l'accesso ad essi e le operazioni svolte siano conformi a quanto previsto per ciascuna figura professionale (profilo utente) coinvolta. Una volta gestita correttamente la profilazione delle utenze (definendo "chi può gestire/visualizzare/stampare cosa") viene garantito che i dati possano essere consultati e/o modificati dagli utenti dei sistemi solo tramite l'interfaccia web.

Database e protezione: i dati raccolti via web sono automaticamente memorizzati in un database centralizzato che risiede su un **server protetto** secondo le modalità seguenti, in accordo alla normativa vigente:

- il server che ospita il database risiede in una server farm appositamente progettata con accesso protetto (da guardia giurata, ad esempio), consentito solo a persone munite di badge e facenti parte della lista delle persone autorizzate all'accesso conservata in armadio sicuro;

- il server è protetto da firewall (hardware/software);

- il software applicativo che recepisce il dato mediante data entry non risiede sullo stesso server che ospita il database con i dati clinici. Questo prevede che ci sia una ridondanza dei sistemi hardware, oppure a livello software realizzata attraverso meccanismi di virtualizzazione, condizione ottimale che non sempre può essere soddisfatta, ma rappresenta un rafforzamento dei requisiti di sicurezza;

Archiviazione e backup dei dati: è necessario predisporre (giornalmente/settimanalmente) le procedure di export dei dati inputati; tali export vengono memorizzati su un server di backup e su supporto secondario che risiede in una struttura sicura della sala macchine. Il ripristino eventuale dei dati è affidato a procedure automatiche di recovery.

Standard crittografici

Tutti i dati che transitano dal server centrale verso la postazione dell'utente e viceversa, sono criptati mediante l'utilizzo di: **certificati sicuri** associati ai server (richiedere un certificato per un host, vuol dire richiedere la certificazione di una macchina che dovrà avere funzionalità di server) e del **protocollo HTTPS** (per connessioni sicure e crittografate, che è il risultato dell'applicazione di un protocollo di crittografia SSL al protocollo di trasferimento di ipertesti HTTP). Nel paragrafo successivo viene presentato un approfondimento sul tema dei certificati digitali e protocollo HTTPS.

Modalità di accesso al DB: al database può accedere direttamente solo l'amministratore di sistema, e tutti gli accessi al server tramite rete sono esclusivamente gestiti mediante **protocolli di cripting**:

- con **HTTPS** per gli accessi via web**;**

- con **SSH Secure Shell**, protocollo di rete che permette di stabilire una sessione remota cifrata fra un computer ed un altro host di una rete informatica. Come ulteriore misura di sicurezza, l'accesso attraverso SSH viene abilitato solo agli IP assegnati ai responsabili del database e dello sviluppo software.

Procedure di autenticazione e tracciabilità

Protocollo di autenticazione: esso prevede che a ciascun utente venga assegnata una login di accesso al sistema web; egli potrà scegliere una password composta da almeno 8 caratteri che dovranno essere sia numerici che alfanumerici.
Le password memorizzate nel database sono criptate ed hanno scadenza automatica. La password definita scade dopo 3/6 mesi e l'utente dovrà reimpostarla attraverso un'apposita funzionalità del software applicativo, scegliendo password sempre differenti. Potrà riscegliere la stessa password solo dopo il terzo cambio.

Tracing delle operazioni utente: è possibile configurare gli applicativi e/o i database server affinché i dati inseriti (a partire dai dati strategici o di principale interesse) vengano associati automaticamente all'utente che ha effettuato l'immissione/modifica. I meccanismi di tracing prevedono l'utilizzo di tabelle supplementari per ospitare lo storico dei 'movimenti' (ins/upd/del) e/o la configurazione di trigger, stored procedure ad hoc, realizzate con linguaggio SQL.

Certificati digitali e protocollo HTTPS

I **certificati digitali** vengono utilizzati per proteggere i collegamenti da/verso un server e per attestare l'identità di un sito web (così come per l'identità di un soggetto). Nel nostro Istituto (IFC-CNR Pisa) la procedura consiste nel presentare una richiesta al GARR, tramite il responsabile della rete di Istituto. Una volta ottenuto ed installato il certificato digitale valido sul server, è possibile l'accesso ai siti ospitati da quel server attraverso il protocollo **HTTPS** (ossia HTTP con l'aggiunta del protocollo crittografico **SSL, Secure Sockets Layer**). Ricorrendo all'impiego dei certificati digitali, il browser web può accertarsi che il server a cui si è connessi sia autentico, ossia che corrisponda effettivamente a quello che dichiara di essere. Se il certificato è stato firmato da un'autorità di certificazione riconosciuta, il browser provvede ad utilizzare la chiave pubblica indicata nel documento digitale per scambiare dati in modo sicuro.

Il protocollo HTTPS, insieme con un certificato digitale valido, viene quindi utilizzato da tutti i siti web che permettono la gestione di dati particolarmente importanti od informazioni sensibili (si pensi ai siti di e-commerce ed ai servizi online dei vari istituti di credito...). Abilitando l'utilizzo di HTTPS, i dati non viaggiano più "in chiaro" ma sono crittografati così da impedire l'"intercettazione" da parte di terzi di tutti i contenuti inviati e ricevuti.Le più recenti versioni dei vari browser web ben evidenziano quando si sta utilizzando una connessione HTTP e quando, invece, ci si è collegati ad una pagina web che fa uso del protocollo HTTPS. Sia Internet Explorer, sia Chrome, sia Firefox, ad esempio, espongono un **lucchetto** nella barra degli indirizzi insieme con l'indicazione **https**:ogniqualvolta si utilizzi una pagina che utilizza un certificato digitale e che quindi provvede a crittografare i dati cambiati tra client e server (e viceversa).

Scheda Anagrafica e Dati sensibili

Gli applicativi dovranno implementare una modalità di 'accesso indiretto' ai dati, tutte le volte che si vogliono richiamare gli eventi sanitari (contenenti dati sensibili) a partire dai dati anagrafici di un paziente.

Una proposta di soluzione consiste nella configurazione di una 'tabella di correlazione anagrafica', in cui vengono memorizzati: un identificativo numerico associato alla scheda anagrafica e un secondo identificativo numerico utilizzato nelle tabelle che contengono dati sensibili. Pertanto non si ha necessità di accedere ai dati identificativi ed in chiaro del paziente (nome, cognome, CF, …) ma tutte le operazioni avvengono mediante il secondo identificativo che individua lo stesso soggetto nei vari flussi informativi di dati sensibili. Attraverso il secondo ID non si può risalire all'identità dell'interessato, se non passando attraverso la tabella di correlazione e reperendo il corrispondente identificativo (primo ID) di accesso alla scheda anagrafica.

La tabella di correlazione è l'unico elemento che permette di associare dati anagrafici a dati sensibili. L'accesso alla tabella di correlazione anagrafica per l'identificazione dei soggetti deve essere espressamente autorizzato.

La tabella di correlazione deve risiedere su un differente DB rispetto ai dati, (collegamento via dblink) che talvolta induce problemi di performance dei sistemi.

A livello di fruizione dell'applicativo web, quando viene richiesta l'apertura di uno screen contenente dati sensibili, se il consenso sul trattamento dati sensibili è stato firmato dal paziente, l'applicativo ha accesso alla tabella di correlazione e si mostrano in selezione/modifica i dati sensibili. Se il consenso sul trattamento dati sensibili non è stato firmato dal paziente, a video viene mostrato un alt del sistema che nega la visualizzazione/gestione di questi dati.

Caratteristiche del dataset: l'implementazione di un ulteriore livello di sicurezza può rendersi necessario se esistono pochi elementi del set di dati (es. meno di 10 soggetti) con le stesse caratteristiche di età, patologia (diagnosi , farmaci…). Questi casi possono essere gestiti con specifici algoritmi di ricodifica della data in nascita in classi di età, ricodifica dei codici originari di diagnosi ICD9, ricodifica dei codici originari di farmaco ATC.

Archivi offline - DB Access, Excel

Un primo livello di sicurezza riguarda l'accesso alle singole postazioni di lavoro degli utenti, che deve essere protetto da password (composta da almeno 8 caratteri, sia numerici che

alfanumerici) con scadenza periodica dopo 3/6 mesi; l'utente dovrà reimpostarla scegliendo password sempre differenti.

Relativamente ai file di dati già archiviati, non più alimentati da dati applicativi online, se contenenti informazioni personali e sensibili (nome e cognome, data di nascita dei pazienti, ed altre info in chiaro), si raccomanda di non far risiedere tali file sulle singole postazioni degli utenti (operatori, ricercatori, medici..).

Dopo aver svolto una sessione di lavoro su tali file, è importante che l'utente li memorizzi unicamente sul server, nelle apposite cartelle condivise e protette, raggiungibili solo dagli utenti autorizzati.

I file devono essere protetti da password, la quale verrà comunicata solo al personale autorizzato alla consultazione/estrazione dei dati.

E' importante capire se i **dati sensibili** nei file sono memorizzati in chiaro o sono presenti codici numerici, definiti in specifiche tabelle del DB ('dizionari' contenenti codice e descrizione).

Se i dati sensibili sono in chiaro: è opportuno separare le informazioni anagrafiche da quelle sensibili su due file differenti. Il file primo file contiene le informazioni anagrafiche e la chiave univoca numerica (ID) attribuita al paziente. Il secondo file contenente i dati sensibili, dovrà contenere unicamente la chiave univoca numerica (ID) per identificare il paziente, così che non si possa risalire al nome, cognome, CF e gli altri dati che rivelerebbero l'identità del paziente.

Se i dati sensibili non sono in chiaro: si può utilizzare un unico file contenente tutte le informazioni, con l'accortezza di tenere archiviati separatamente i documenti contenenti le mappature fra 'informazioni sensibili/codici' (codebook), ovvero in una locazione differente del server accessibile in modo protetto solo dagli utenti autorizzati.

Trasmissione dei file di dati: la trasmissione dati agli enti autorizzati dal protocollo di studio deve avvenire, per ogni flusso, su 2 file differenti separando i dati anagrafici da quelli sensibili, con doppio invio, utilizzando i meccanismi descritti sopra: primo invio contenente solo i dati anagrafici dei soggetti e l'ID soggetto; secondo invio, dati sensibili con il soggetto riferito unicamente attraverso il suo ID.

Figure rilevanti nel trattamento dei dati

Titolare e Responsabile del trattamento dati

Il **Titolare** del trattamento è colui che assume le decisioni in ordine alle finalità, alle modalità del trattamento, agli strumenti utilizzati anche sotto il profilo della sicurezza. Può essere una persona fisica o una persona giuridica (es. l'ospedale). Nell'informativa deve essere dichiarato espressamente il Titolare del trattamento dei dati (es. un ospedale, un centro di ricerca). Il

titolare può affidare i compiti di gestione e controllo del trattamento dei dati ad un **Responsabile** del trattamento dei dati personali per la particolare esperienza o capacità, individuandolo all'interno della propria organizzazione o anche all'*esterno*. La designazione del Responsabile è facoltativa e il soggetto designato (società, ente, persona fisica..) è tenuto a rispettare le indicazioni e le direttive impartite dal titolare. Nell'informativa deve essere dichiarato espressamente il Responsabile del trattamento dei dati.

Incaricato al trattamento dei dati

È la persona fisica autorizzata a compiere operazioni di trattamento, dal Titolare o dal Responsabile. La nomina di uno o più incaricati è obbligatoria e l'assegnazione di una persona ad una struttura interna non soddisfa di per sé ai requisiti di legge, in quanto deve risultare chiaro a quali categorie di dati hanno accesso le persone preposte alla struttura.

La documentata assegnazione di una persona ad un'unità organizzativa per la quale sia stato individuato, per iscritto, il trattamento consentito agli addetti all'unità medesima, è equiparata alla designazione individuale e va effettuata, sempre, per iscritto.

Le norme sull'incaricato riguardano chiunque sia stato incaricato dal Titolare o dal Responsabile a raccogliere, elaborare ed utilizzare i dati.

Queste persone non verranno indicate nell'informativa, ma dovranno essere individuate nominalmente al momento dell'assegnazione delle login di accesso ai sistemi elettronici predisposti per la gestione dei dati. In questo modo, sono chiaramente definite le persone che operano sui dati, e in quali ruoli (profili di accesso).

Cenni su Autorità e Documentazione

Garante per la privacy

Il Garante per la protezione dei dati personali è un'autorità amministrativa indipendente, che deve essere informata nel caso in cui si voglia attivare una specifica raccolta di dati sensibili, attraverso l'utilizzo di uno strumento (ad esempio una piattaforma elettronica di raccolta di informazioni sensibili, riguardante lo stato patologico di alcuni soggetti). Occorre specificare in dettaglio al Garante i dati che si intendono rilevare, le modalità di raccolta, i soggetti interessati.

Comitato etico

Il comitato etico è un organismo indipendente, senza scopo di lucro, composto da personale sanitario e non, secondo criteri di interdisciplinarietà. È incaricato di garantire la tutela dei diritti, della sicurezza e del benessere di soggetti coinvolti in uno studio o sperimentazione clinica e, nello specifico, ha i seguenti compiti: la valutazione del razionale della sperimentazione, dell'adeguatezza del protocollo, la verifica della competenza e dell'idoneità dei ricercatori e, soprattutto, la valutazione degli aspetti etici della ricerca inerenti il consenso informato, la tutela della riservatezza ed benessere dei soggetti coinvolti.

Emette pareri sul protocollo di sperimentazione, sull'idoneità dello o degli sperimentatori, sulle strutture, metodi e documenti da impiegare per informare i soggetti della sperimentazione, prima di ottenere il consenso informato. In Italia l'istituzione dei Comitati Etici è prevista: nelle strutture sanitarie pubbliche e negli istituti di ricovero e cura a carattere scientifico (IRCCS) privati.

Informativa paziente e medico

Il paziente deve essere informato dal responsabile del trattamento dei dati circa:

- le finalità dello studio, descrivendo in dettaglio perché si vogliono rilevare alcuni dati personali/sensibili riguardanti la sua persona;
- quali sarebbero i benefici derivanti dal consenso al trattamento;
- quali sarebbero le conseguenze in caso di consenso negato;
- come verranno utilizzati i dati raccolti (per fini statistici, a scopo di ricerca, in forma aggregata, …);
- a chi possono essere comunicati o diffusi i dati;
- quali sono i diritti riconosciuti all'interessato all'interessato di cui all'art. 7 del Codice Privacy[1] (diritto di ottenere informazioni e comunicazioni, diritto di ottenere l'aggiornamento, la rettifica o la cancellazione dei dati, il diritto di opporsi al trattamento dei propri dati personali…);

[1] **Codice privacy: articolo 7 - Diritto di accesso ai dati personali ed altri diritti**

• L'interessato ha diritto di ottenere la conferma dell'esistenza o meno di dati personali che lo riguardano, anche se non ancora registrati, e la loro comunicazione in forma intelligibile. •

L'interessato ha diritto di ottenere l'indicazione:

1. dell'origine dei dati personali
2. delle finalità e modalità del trattamento
3. della logica applicata in caso di trattamento effettuato con l'ausilio di strumenti elettronici
4. degli estremi identificativi del titolare, dei responsabili e del rappresentante designato ai sensi dell'articolo 5, comma 2
5. dei soggetti o delle categorie di soggetti ai quali i dati personali possono essere comunicati o che possono venirne a conoscenza in qualità di rappresentante designato nel territorio dello Stato, di responsabili o incaricati

• L'interessato ha diritto di ottenere:

1. l'aggiornamento, la rettificazione ovvero, quando vi ha interesse, l'integrazione dei dati
2. la cancellazione, la trasformazione in forma anonima o il blocco dei dati trattati in violazione di legge, compresi quelli di cui non è necessaria la conservazione in relazione agli scopi per i quali i dati sono stati raccolti o successivamente trattati
3. l'attestazione che le operazioni di cui ai numeri 1) e 2) sono state portate a conoscenza, anche per quanto riguarda il loro contenuto, di coloro ai quali i dati sono stati comunicati o diffusi, eccettuato il caso in cui tale adempimento si rivela impossibile o comporta un impiego di mezzi manifestamente sproporzionato rispetto al diritto tutelato.

L'interessato ha diritto di opporsi, in tutto o in parte:

1. per motivi legittimi al trattamento dei dati personali che lo riguardano, ancorché pertinenti allo scopo della raccolta
2. al trattamento di dati personali che lo riguardano a fini di invio di materiale pubblicitario o di vendita diretta o per il compimento di ricerche di mercato o di comunicazione commerciale

- chi sono il Titolare, l'eventuale Responsabile del trattamento e gli Incaricati del trattamento (che devono essere necessariamente individuati) e dove sono raggiungibili (indirizzo, telefono, fax, ecc.).

Inoltre, viene specificato che egli dovrà successivamente compilare uno o più moduli di Consenso informato e che potrà, in qualsiasi momento, revocare uno o più specifici consensi al trattamento ed esercitare i diritti di cui all'art. 7 del Codice Privacy.

L'informativa riguarda anche il medico, qualora abbia il compito di informare preventivamente alcuni suoi pazienti per indirizzarli verso un particolare studio che richiede il trattamento di dati personali e sensibili.

Consenso al trattamento dei dati personali

E' la libera manifestazione di volontà dell'interessato con cui questi accetta espressamente un determinato trattamento dei suoi dati personali, del quale è stato preventivamente informato da chi ha un potere decisionale sul trattamento (**Informativa**). Se il trattamento riguarda dati "sensibili" occorre il consenso rilasciato **per iscritto** dall'interessato, con la sua sottoscrizione. Il consenso per essere valido ed efficace deve essere **informato**, quindi preceduto da una informativa chiara, semplice e comprensibile; **personale**, non delegabile a familiari o ad altri, in caso di minore o incapace il consenso viene esercitato da chi ha la podestà tutoria; **consapevole**, l'interessato deve essere in grado di intendere e di volere; **esplicito,** espresso e mai sottointeso; **specifico, preventivo ed attuale, libero e sempre revocabile.** Il consenso al trattamento dei dati personali è, inoltre, **specifico**, in quanto diverso ed ulteriore rispetto al consenso informato espresso per ricevere la prestazione sanitaria (anche se reso contestualmente); deve essere prestato, infine, con riferimento ad ogni comunicazione o diffusione delle informazioni raccolte a soggetti terzi (come nel caso di Dossier Sanitario e Fascicolo Sanitario Elettronico).

In alcuni casi, relativi alla **ricerca medico-scientifica, il Garante semplifica gli adempimenti e rilascia autorizzazioni generali** a università, enti di ricerca, organismi sanitari, che consentono di utilizzare dati o campioni già raccolti a fini di cura o in precedenti progetti di ricerca, di svolgere studi e ricerche su particolari patologie e terapie o sull'efficacia di determinati farmaci senza il consenso dei pazienti nei casi in cui non sia possibile fornire loro l'informativa prevista per legge *per "motivi etici" o "impossibilità organizzativa",* senza diminuire il livello di tutela per i pazienti (ad es. le Autorizzazioni n. 2/2013 - *Autorizzazione al trattamento dei dati idonei a rivelare lo stato di salute e la vita sessuale*, n. 8/2013 - *Autorizzazione generale al trattamento dei dati genetici* e n. 9/2013 - *Autorizzazione generale al trattamento dei dati personali effettuato per scopi di ricerca scientifica*).