



27 Maggio 2022



Titolo: Il futuro dei dati fra dati sintetici e dati in movimento

Sponsor: Brain-IT

Relatore: Alfredo Visconti | Presidente ANDIP, e AD di Brain-it

#cybersec_forum #csf22

La sicurezza come l'affidabilità dei processi, dei dati e i sw di riferimento, non si acquistano si progettano

START



Esistono dati senza informazioni ma non informazioni senza dati

Analisi dei dati

- Data base
- Tabelle
- Dati grezzi

Un Dato è una rappresentazione oggettiva non interpretata della realtà. I data base delle aziende sono organizzati in tabelle che contengono milioni di righe.



Business intelligence

- Fabbisogno informativo
- Raccolta delle informazioni
- Analisi, valutazione e produzione
- Utilizzazione
- Feedback

B.I. trasforma i dati in un prodotto finito ed impiegabile, ed è anche il momento nel quale è possibile orientare la ricerca delle informazioni.



Sviluppo dei sistemi

- Requisiti di sistema
- Requisiti software
- Implementazione policy privacy
- Gestione informative e consensi
-

I progetti sw iniziano con una fase di analisi che consiste in uno studio della fattibilità e una definizione dei requisiti. E' qui che il progetto viene valutato in base ai costi, ai profitti e alla realizzabilità e da oggi anche in considerazione delle politiche di privacy.

RISULTATO

Esistono informazioni scompaiono i dati che creano identificazione

Dati sintetici e dati in movimento una nuova sfida per la tutela dei dati personali

Oggi le società, che sono obbligate a lavorare con i dati e con le informazioni che ne derivano, incontrano costantemente barriere, imputabili in modo specifico alla tecnologia, alla propria struttura tecnica, e all'ordinamento giuridico.

Questi ostacoli si traducono in mancate opportunità di crescita e possono ostacolare l'innovazione generando la paura della condivisione di dati, soprattutto se sono sensibili con partner, potenziali clienti, o prospect esterni.

Diamo atto che questo passaggio da dato ad informazione è da imputarsi anche all'evoluzione del concetto di privacy.

L'art 32, tenendo conto dello stato dell'arte e dei costi di attuazione, nonché della natura, dell'oggetto, del contesto e delle finalità del trattamento, le aziende sono tenute a mettere in atto misure tali da garantire un livello di sicurezza e di riservatezza adeguato al rischio (art. 32). L'analisi dei rischi viene richiesta almeno per (art. 35):

- l'utilizzo di soluzioni tecnologiche che elaborano una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, attraverso modalità di trattamento automatiche (compresa la profilazione);
- trattamenti su larga scala di dati sensibili e/o giudiziari (ad es. numero di persone coinvolte, volume di dati trattati, trasferimento di dati) o dati di persone con maggiore necessità di protezione (es. bambini, pazienti con malattie mentali);
- sorveglianza sistematica su larga scala di una zona accessibile al pubblico



Dato grezzo



Business intelligence

A. I.



Utilizzo proattivo
corretto

Protetto

Conforme al
GDPR

Dati sintetici

Le aziende che mirano ad accedere a questi dati per sviluppare algoritmi o sistemi di machine learning, non necessariamente devono accedere a dati strettamente reali, ma devono poter accedere a set di dati realistici che simulano e mantengono le stesse proprietà statistiche dell'informazione reale. Questi dati alternativi sono noti come dati sintetici.

I dati sintetici sono dati, o il loro risultato, creati attraverso una particolare tecnica di anonimizzazione basata su modelli di machine learning di tipo generativo. Partendo da un set di dati reali, si allena un sistema di intelligenza artificiale a individuare le correlazioni e metriche statistiche del dataset originale, per poi generare un set di dati ex novo che mantiene la stessa distribuzione statistica del dataset originale, pur non condividendo alcun dato del dataset reale.

Questa tecnica detta della sintetizzazione fa sì che si possa risalire al dato reale dal dato generato, senza perdere l'informazione originale, esattamente l'opposto di quanto succede sui dati con le normali tecniche di anonimizzazione, dove il dato è privato di tutti gli elementi "personali".

I dati sintetici hanno diversi vantaggi, quali:

- il superamento delle restrizioni di utilizzo dei dati: i dati reali spesso nel loro utilizzo incontrano vincoli di utilizzo basati sulla disamina del GDPR, perchè i dati sintetici duplicano le proprietà statistiche fondamentali per l'utilizzo del dato senza esporre i dati reali;
- la generazione di dati sintetici è a regime più economica rispetto all'acquisizione di dati reali;
- la condivisione e quindi la divulgazione delle informazioni viene resa possibile dalla natura statistica delle informazioni, che non porta con se la possibilità di accoppiare l'informazione ad una persona.

I dati sintetici rappresentano una soluzione tecnologica che incomincia a trovare una strada specifica soprattutto in alcuni settori come:

- le assicurazioni, i servizi bancari/finanziari ed il retail;
- **l'healthcare, mercato dove i dati sensibili, genetici, biometrici, ma sanitari in generale, godono di un regime di protezione normativa molto complicato;**

GDPR

Ciclo del dato personale



Tipo dati raccolti

Raccogli troppi dati personali per raggiungere gli scopi prefissi?
Puoi ridurre il numero di dati raccolti ottenendo gli stessi risultati?

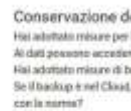
Finalità della raccolta

Quale è lo scopo della raccolta dei dati?
Lo hai specificato nell'informatica? Lo hai comunicato agli interessati?
Gli interessati sono a conoscenza di cosa accade ai loro dati quando te li conferiscono?



Aggiornamento dei dati

Hai una procedura per la verifica dell'effettività dei dati e il loro eventuale aggiornamento?



Conservazione dei dati

Hai adottato misure per la sicurezza dei dati?
Ai dati possono accedere solo le persone autorizzate?
Hai adottato misure di backup dei dati?
Se il backup è nel Cloud, il fornitore del servizio è in regola con la norma?



Tempo di conservazione dei dati

Per quanto tempo i dati vengono conservati?
Hai procedure per la verifica che i dati siano conservati solo per il periodo necessario a raggiungere gli scopi prefissi?

Diritti dell'interessato

Garantisci i diritti all'interessato (diritto di accesso, cancellazione o portabilità)?
Rispondi alle richieste dell'interessato nei tempi previsti dalla norma?



Dati in movimento

