Progettare software in linea con il nuovo regolamento GDPR

Raffaele Rialdi @raffaeler http://iamraf.net

Senior Software Architect

Presidente DotNetLiguria

Chi è Raffaele Rialdi

- Lavoro nella progettazione del software dal 1987
- Libero professionista con clienti nei settori
  - Banking / Financial, Manufacturing, Healthcare, F1 Racing, …
- Responsabile della Divisione Tecnologie Avanzate di Vevy Europe SpA
  - Settore chimico
- Microsoft MVP (Most Valuable Professional) dal 2003
  - Award riconosciuto su base annuale da Microsoft a professionisti che si sono distinti per competenza e promozione della conoscenza
  - Specializzazione: "*Developer Security*"
- Communities
  - Presidente di DotNetLiguria
  - Co-fondatore di ItalianCpp
  - Member of the board di UGIdotNET

DotNet Liguria

SDL
Security Development Lifecycle

IN PARALLELO CON IL REGOLAMENTO GDPR

SDL e il principio SD3+C

citati nella GDPR

- **Secure By Design**
  - Architettura, progettazione e implementazione devono essere pensati per essere sicuri
- **Secure by Default**
  - I componenti devono offrire minore superficie di attacco e usare il minor numero di privilegi
- **Secure in Deployment**
  - Evidenziare le condizioni necessarie in fase di installazione
  - Documentazione per evitare errori nell'infrastruttura
- **Communications**
  - Architetti e sviluppatori devono essere pronti ad eseguire modifiche a seguito della scoperta di vulnerabilità
  - Organizzare frequenti riunioni periodiche e documentare vulnerabilità e contromisure

GDPR
Le prescrizioni 'tecniche'

GENERAL DATA PROTECTION REGULATION

REGOLAMENTO (UE) 2016/679

http://eur-lex.europa.eu/legal-content/it/TXT/?uri=celex:32016R0679

Articolo 5

I dati da proteggere

- Devono essere adeguati, pertinenti e limitati alla finalità
  - La limitazione implica la segregazione dei dati nel caso siano condivisi
    - Esempio: Due applicazioni (con utenze diverse) condividono lo stesso database, ma solo una usa il numero cellulare
- Deve essere possibile cancellare o rettificare i dati inesatti rispetto alle finalità
- L'utente deve poter disporre dei dati in tempo utile per le finalità indicate nel trattamento
  - Attenzione ai Denial of Service, mancanza di backup, corruzione dei dati

Articolo 6

Possiamo registrare dati di privacy se …

- L'utente ha espresso il consenso per determinate finalità
- I dati sono necessari all'esecuzione di un contratto con l'utente
- Se esiste un obbligo legale per registrare i dati
  - Fatture, DDT, etc.
- Se è necessario per interessi vitali di qualcuno
- Nel caso manchi un esplicito consenso per una data finalità
  - È necessario fornire garanzie aggiuntive come la cifratura o la pseudonimizzazione

DotNet Liguria

Cifratura

- Si usa per rendere illeggibile il dato nello storage
  - Chi apre il database non riesce a riconoscere i dati dell'utente
- La chiave crittografica deve essere custodita separatamente
- La crittografia può essere fatta da:
  - Database sull'intero file: relativamente semplice, inefficace per la protezione dagli amministratori di sistema
  - Database solo sulle anagrafiche: complessa, richiede modifiche all'app
  - Sistema operativo (bitlocker): protegge solo dal furto fisico del disco

Pseudonimizzazione

- È un processo applicativo che usa pseudonimi al posto dei dati veri
- Deve essere prevista/gestita a livello applicativo
- Può comportare problemi nelle ricerche/filtri
- Per essere efficace i dati veri dovrebbero essere conservati su un database diverso
  - Questo complica di parecchio la logica applicativa
  - Gli ORM non sono in grado di gestire questa casistica (a meno di custom provider)

Articolo 7.1

Condizioni per il consenso

- Il titolare del trattamento **deve dimostrare** che l'interessato ha prestato il proprio **consenso al trattamento** dei propri dati personali.

- Interpretazione garante:
  - http://www.garanteprivacy.it/fondamenti-di-liceita-del-trattamento

- Considerazioni
  - Nel **consenso "digitale"**, il problema è dimostrare che il consenso sul database non sia stato alterato
    - Una possibile soluzione potrebbe essere un "timestamp" la cui chiave viene conservata altrove

DotNet Liguria

Articoli 13-22

Trattamento dei dati

- Esiste un **periodo di conservazione** dei dati personali
  - Viene chiesto all'utente o ricavati tramite un 'criterio'
- L'utente può chiedere di ottenere i dati e di essere cancellato (diritto all'oblio)
- I dati possono essere trattati con finalità diverse rispetto a quella per cui sono stati raccolti, ma l'utente deve essere informato
  - L'utente potrà chiedere come conseguenza di essere cancellato
- Se i dati sono stati ottenuti da terzi, è necessario informare l'utente sulla loro origine prima di poterli trattare
- L'utente può opporsi a diverse tipologie di trattamento
  - alla cancellazione dei dati e richiederne invece una limitazione di utilizzo.
  - al cambio di soggetto del trattamento
  - **ad una o più finalità del trattamento**
  - …

DotNet Liguria

Data protection by design and by default

Articolo 25

Protezione dei dati fin dalla progettazione e protezione per impostazione predefinita

- «… 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 …»
- «Il titolare del trattamento mette in atto **misure tecniche e organizzative** adeguate per garantire che siano trattati, per impostazione predefinita, solo i dati personali necessari per ogni specifica finalità del trattamento»
  - Si applica a quantità, portata del trattamento periodo di conservazione e accessibilità
- Considerazioni
  - È una analisi di rischio e tutte le scelte devono essere verbalizzate con relative motivazioni
  - In termini di threat model queste sono buoni esempi di "mitigazioni"

DotNet Liguria

Articolo 32.1

Sicurezza del trattamento

- «**Tenendo conto dello stato dell'arte** e dei **costi di attuazione**, nonché della **natura**, dell'**oggetto**, del **contesto** e delle **finalità** del trattamento, come anche del rischio di varia probabilità e gravità per i diritti e le libertà delle persone fisiche, il titolare del trattamento e il responsabile del trattamento mettono in atto misure tecniche e organizzative adeguate per garantire un livello di sicurezza adeguato al rischio, che comprendono, tra le altre, se del caso …»
- Considerazioni:
  - Anche questi sono concetti tipici delle analisi di rischio
  - Non applicare le security updates può essere considerata una violazione
  - Le mitigazioni devono avere un «level of security appropriate to the risk»
  - La mancanza di verbalizzazione o la sottovalutazione del rischio sono un'aggravante

DotNet Liguria

Articolo 32.1 a

Sicurezza dei dati personali

- «…, tra le altre, se del caso:»
  - la pseudonimizzazione e la cifratura dei dati personali
    - La cifratura dei file dell'intero DB è relativamente semplice
    - Uso di Bitlocker
    - Uso di TLS (es: HTTPS) per movimentare i dati

Articolo 32.1 b-d

Sicurezza dei dati personali

- «la capacità di assicurare su base permanente la riservatezza, l'integrità, la disponibilità e la resilienza dei sistemi e dei servizi di trattamento;»
  - riservatezza ➔ "Confiendentiality"
  - integrità ➔ "Integrity"
  - disponibilità, resilienza ➔ "Availability"

} Le categorie di rischio usate nei threat models
C, I, A

- Considerazioni: l'articolo continua evidenziando la necessità di poter
  - Ripristinare tempestivamente da eventuali malfunzionamenti
    - Backup? Backup in luoghi fisici diversi!
  - Un processo per testare regolarmente l'efficacia delle procedure tecniche

DotNet Liguria

Articoli 32.2

Sicurezza dei dati personali

- «Nel valutare l'adeguato livello di sicurezza, si tiene conto in special modo dei rischi presentati dal trattamento che derivano in particolare dalla **distruzione, dalla perdita, dalla modifica, dalla divulgazione non autorizzata o dall'accesso, in modo accidentale o illegale**, a **dati personali trasmessi, conservati** o comunque trattati.»

- Considerazioni:
  - Si parla di hacking o perdita di dati (assenza di backup)
  - I boundary dei diagrammi DFD nel Threat Model evidenziano:
    - Dove sia **necessaria la protezione del trasporto (HTTPS)** ➔ la gestione dei certificati è importante!
    - L'implementazione di una **adeguata policy sulle password**
    - Ma **non** parlano esplicitamente di pratiche "IT" (ma SDL raccomanda i backup)

DotNet Liguria

Valutazione d'impatto sulla protezione dei dati

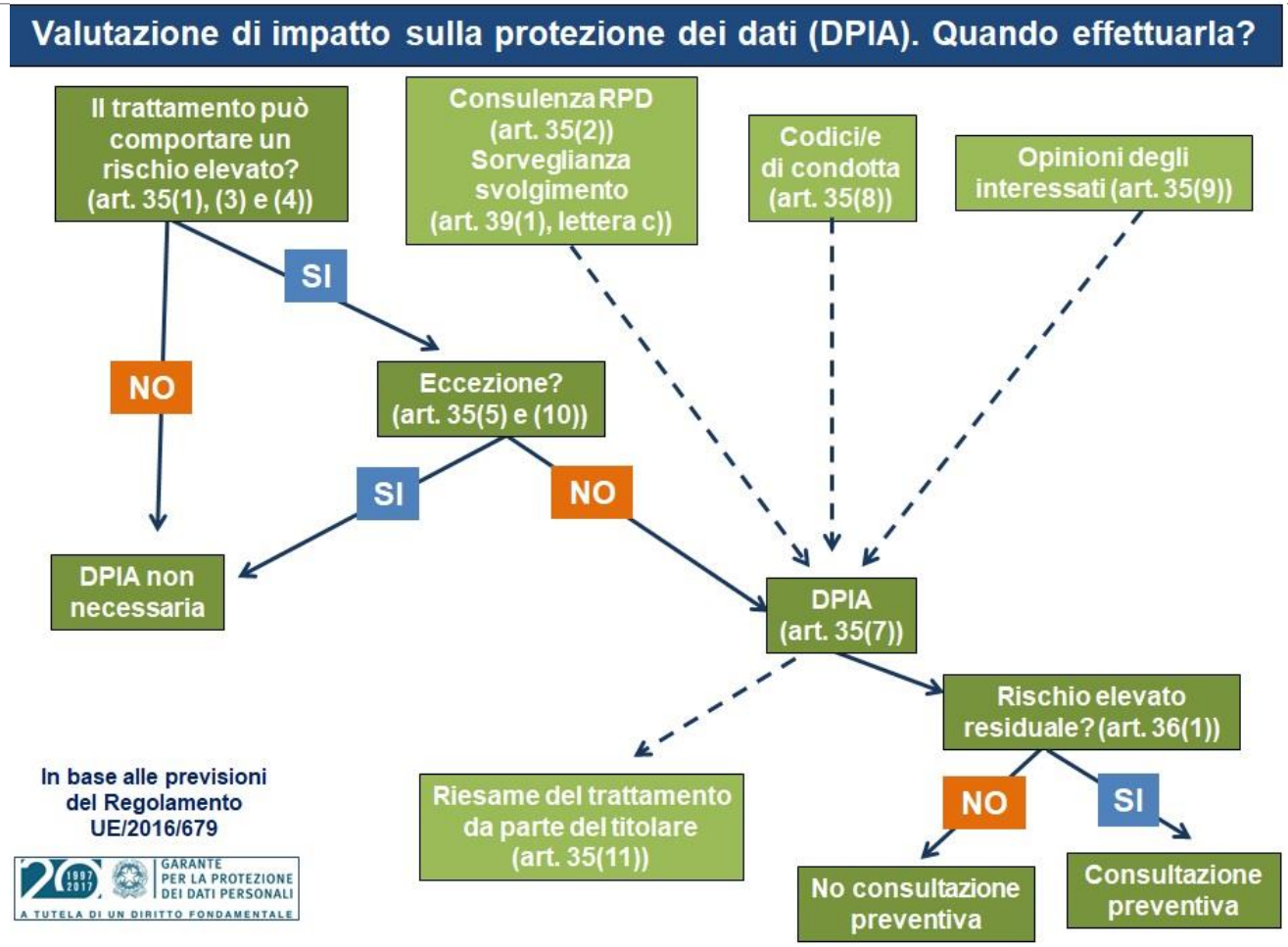
Articolo 35

- «Quando un tipo di trattamento, …, può presentare un rischio elevato per i diritti e le libertà delle persone, …, il titolare del trattamento effettua, …, **una valutazione dell'impatto** dei trattamenti previsti sulla protezione dei dati personali»

- Considerazioni
  - **La conservazione di password in chiaro** può consentire ad un intruso di entrare in possesso di dati anche su altri siti/banche dati/etc. da cui può derivare un "rischio elevato" e una lesione della libertà del soggetto.
  - La conservazione di un hash delle password è un elemento fondamentale per evitare questo rischio
  - L'uso di un sistema di 2FA (eventualmente fatto scegliere all'utente) è auspicabile
  - Se la valutazione di impatto è necessaria, è auspicabile imponga meccanismi di protezione **superiori** a quelle dei dati non a rischio.

Valutazione di impatto secondo il garante

http://www.garanteprivacy.it/DPIA



DotNet Liguria

Valutazione d'impatto sulla protezione dei dati

Articolo 35.7

- «La valutazione contiene almeno:
  a) una **descrizione sistematica dei trattamenti** …
  b) una valutazione della **necessità e proporzionalità** dei trattamenti …
  c) una **valutazione dei rischi** per i diritti e le libertà degli interessati …
  d) le **misure previste per affrontare i rischi**, includendo le garanzie, le **misure di sicurezza** e i **meccanismi per garantire la protezione dei dati personali** …»
- Considerazioni:
  - Il **Threat Model** copre quasi tutti questi punti
  - Resta fuori la finalità del trattamento e il punto (b)

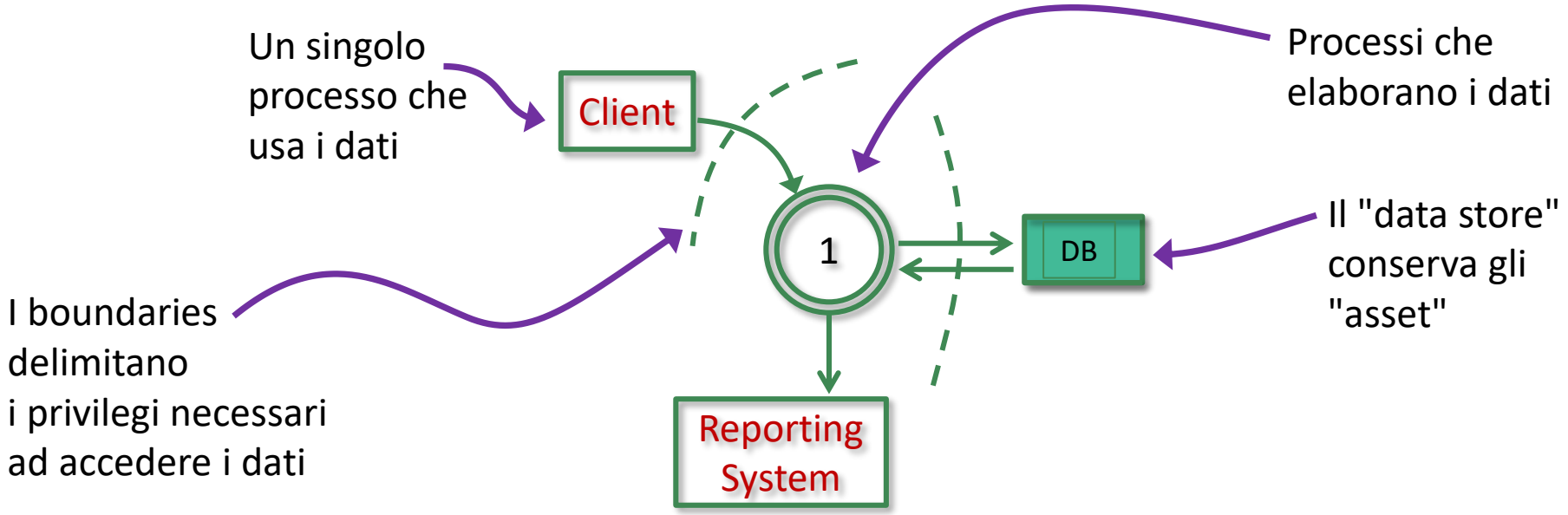
DotNet Liguria

Il Threat Model secondo la SDL

- Il Threat Model (modello delle minacce) è un processo che consente di comprendere le potenziali vulnerabilità di un sistema informatico
  - Guida all'analisi dell'architettura dell'applicazione
  - Evidenzia i rischi
  - Suggerisce le mitigazioni
- Gli speciali diagrammi DFD aiutano a visualizzare i flusso di dati
  - DFD = Data Flow Diagram
- È paragonabile ad una classica "analisi di rischio"
  - Fornisce una classificazione comprensibile
  - Al minimo otteniamo dei numeri oggettivi: Rischio = Probabilità * Impatto

Esempio di threat modeling

- Esistono diversi tool.
- Il TMT di SDL fa uso di 6 semplici simboli
  - Descrive i flussi all'interno di un sistema
  - I flussi più complessi possono essere "esplosi" in altri TM



DotNet Liguria

SDL: Le categorie delle minacce STRIDE

• Threat	Descrizione	Esempi di mitigazione	Categoria C, I, A
Spoofing	Uso di falsa identità per entrare nel sistema	• Autenticare • Filtrare gli IP non validi	Integrity
Tampering	Modifica dei dati tra due sistemi (man in the middle)	• Proteggere il trasporto dei dati	Integrity
Repudiation	Esecuzione di azioni nel sistema senza poterne provare la paternità	• Autenticare • Adottare politiche di auditing	Integrity
Information Disclosure	Rivelazione di informazioni sensibili	• Autorizzazione (SSL, etc.) • Hash credenziali	Confidentiality
Denial of Service	Rendere il sistema inutilizzabile (anche se con richieste legittime)	• Filtrare le richieste per IP • Controllare l'uso della banda	Availability
Elevation of Privilege	Eseguire azioni con privilegi maggiori di quello assegnati	• Eseguire i servizi con il minimo privilegio	Integrity

SDL: Quantificare il rischio

Si parte assegnando i valori DREAD

- **D**amage potential [Impatto]
  - Quanto danno provoca il successo di un attacco?
- **R**eproducibility [Probabilità]
  - Quanto è facile replicare l'attacco?
- **E**xploitability [Probabilità]
  - Quanto è semplice compiere l'attacco?
- **A**ffected users [Impatto]
  - Quanti utenti (percentualmente) sono colpiti dall'attacco?
- **D**iscoverability [Probabilità]
  - Quanto è semplice scoprire la vulnerabilità?

Nota: La scala è Alto, Medio, Basso

Si calcola il fattore di rischio

- Metodo 1

  Rischio = Probabilità * Impatto

- Metodo 2

  Rischio = Min(D, (D+R+E+A+D) / 5)

Damage Potential

Da dove si inizia

- Verbalizzare un «data protection impact assessment» (art 35)
  - È una analisi di rischio per valutare se fare una analisi di rischio più dettagliata
- Verificare (con un legale) i moduli di raccolta dati (finalità, etc.)
  - Audit o altre garanzie di inalterabilità
- Se possibile, separare i dati sensibili in un database differente
  - Gestire questi dati dentro un token come "Claim" è una buona soluzione
- Imbastire tutta la gestione del trattamento: categorie, oblio, storia, ..
  - Eventualmente chiedere una conferma dei dati alla successiva logon

Se il software esiste già, iniziare da modifiche meno invasive

- Analizzare tutte le possibili path rispetto agli entry-point
  - Una analisi statica del codice aiuta in modo significativo
- Redigere un threat model
  - Ordinare i rischi e mitigare da subito quelli più gravi
  - Produrre il documento di analisi dei rischi tramite i tool (TMT, etc.)
- Continuare a mantenere il threat model aggiornato ad ogni rilascio
  - Incrementare il dettaglio del threat model man mano che si va avanti
- Fare l'enforcing da codice delle modalità di deploy (https, …)
  - I container sono una soluzione ottimale

Cosa ne esce?

- Il Threat Model evidenzia molte delle possibili mancanze:
  - https, input validation, output encoding, repudiation dei dati, etc.
- In alcuni casi serve di più:
  - Le deleghe della gestione dei dati sono delicate: la login con 2FA o fingerprint aiutano
- Nei casi peggiori, se il software è vecchio, può non esserci soluzione
  - Es: VB6, privilegi troppo elevati
  - Pianificare una migrazione a nuove tecnologie
- Se il problema è di vulnerabilità del canale e la fix è difficile
  - Mitigare con una VPN, anche all'interno della LAN

I cambiamenti all'architettura possono essere dolorosi

- Evitare la login con username-password a favore di un IP esterno
- Crittografia e pseudonimizzazione sono complesse da implementare
- La separazione del DB utenti è valida ma non sempre facilmente conciliabile con l'architettura
- Implementare la segregazione dei dati è complesso, ma aiuta
  - L'audit è funzionale alla buona riuscita
- Il diritto all'oblio può essere molto complesso
  - Sovrascrivere i dati utente in modo irreversibile è accettabile
    - Nel caso di fotografie o video ovviamente è più complesso o non realizzabile
- L'utente deve poter fare il download di tutto ciò che lo riguarda

Domande?

DotNet Liguria