Contents

[Introduzione 2](#_Toc466721012)

[1 MyData, Big Data 3](#_Toc466721013)

[2 Mobility as a Service e SMAll 4](#_Toc466721014)

[2.1 Mobility as a Service 4](#_Toc466721015)

[2.1.1 Mobility Profile e Journey Planner 4](#_Toc466721016)

[2.2 Smart Mobility for All 5](#_Toc466721017)

[3 Architettura MyData 5](#_Toc466721018)

[3.1 Entità fondamentali 5](#_Toc466721019)

[3.2 Service Registry, Service Linking 6](#_Toc466721020)

[3.3 Autorizzazioni e Consent 6](#_Toc466721021)

[3.4 Personal Data Storage 6](#_Toc466721022)

[4 Analisi e Design 7](#_Toc466721023)

# Introduzione

Negli ultimi anni, con la diffusione degli smartphone e l’aumento dell’accesso a Internet, la tecnologia è diventata sempre più legata all’utente dietro lo schermo, con esempi che possono spaziare dai social network ai dispositivi di monitoraggio della salute e dell’attività fisica. La quantità di informazioni raccolte tramite questi servizi non si limita a ciò che l’utente sceglie di condividere (foto, email, transazioni online), ma comprende anche dati osservati (abitudini di navigazione, dati di geolocalizzazione) e deduzioni (*targeted advertising*, previsioni sul flusso del traffico).

L’importanza dei dati personali è diventata tale da richiedere l’esistenza di leggi che ne regolamentassero l’utilizzo, come il regolamento generale sulla protezione dei dati (GDPR, *General Data Protection Regulation- Regolamento UE 2016/679)*.

In questo contesto si sviluppa il modello di *MyData*, fondato sull’idea secondo cui ogni utente ha il controllo sui propri dati personali, ed è a conoscenza dell’utilizzo che ne viene fatto, in modo trasparente. D’altra parte, l’architettura *MyData* vuole offrire al mercato e alle imprese un contesto di sviluppo di applicazioni software in cui vengono rispettate le leggi in materia di protezione dei dati sensibili, favorendo l’interoperabilità fra di esse.

L’approccio *MyData*, seppur innovativo, non è senza precedenti. Molti dei suoi concetti fondamentali sono anche alla base dei *personal cloud*, come ad esempio la possibilità di contenere in modo sicuro i dati personali dell’utente e offrire interoperabilità fra diversi servizi, utilizzando concretamente i dati memorizzati.

Questa tesi si propone di studiare i concetti fondamentali alla base del modello di *MyData*, e di implementare un prototipo di gestore di dati personali che ne rispetti le specifiche. A titolo di esempio, si è scelto di utilizzare come servizio “consumatore” di dati personali un sistema di previsione del viaggio più probabile.

L’obiettivo è quello di realizzare un sistema che permetta all’utente di monitorare l’utilizzo dei suoi dati personali in tempo reale ed in modo trasparente, attraverso l’implementazione di una politica di controllo degli accessi. Inoltre, il gestore opera in un contesto in cui il rispetto di politiche di sicurezza è fondamentale: anche se la realizzazione di un sistema sicuro non rientra negli obiettivi del lavoro di tesi, sono presenti alcuni accorgimenti che puntano in questa direzione.

# 1 MyData, Big Data

Il progetto MyData nasce in Finlandia, all’università di Aalto, con alla base l’idea di rafforzare i diritti digitali dell’individuo. All’interno del contesto vi è anche la stringente legislazione europea sulla protezione dei dati personali, e la volontà di offrire alle aziende un nuovo modo di approcciarsi al cliente, basato non su incomprensibili “Termini e Condizioni” ma sulla fiducia reciproca.

Attualmente, grandi quantità di dati vengono continuamente raccolti, senza che se ne sfrutti pienamente la ricchezza dell’informazione. Il loro percorso e le operazioni di processamento a cui vengono sottoposti sono sempre più numerose, e aumenta la difficoltà nel tenerne traccia. Inoltre, l’utilizzo di software proprietari limita le possibilità di analisi dei dati a causa della scarsa interoperabilità che si crea inevitabilmente. Infine, va anche considerato l’impatto che questo modus operandi ha sugli utilizzatori finali: secondo un sondaggio riportato dal World Economic Forum, “Fully 78% of consumers think it is hard to trust companies when it comes to use of their personal data” (Orange, The Future of Digital Trust, 2014).

MyData cerca di trovare una risposta a questi temi attraverso un cambio di paradigma che ponga l’utente al centro, rendendolo capace di gestire i propri dati personali e l’uso che ne viene fatto. In questo modo potrebbe ottenere migliori suggerimenti su servizi e opportunità, descrizioni più approfondite del proprio comportamento (*self tracking*), o ottenere un compenso per il processamento dei suoi dati (*data monetization*).

L’approccio a livello infrastrutturale permetterebbe l’indipendenza da specifici settori (sanità e salute, finanza, ecc.), favorendo una completa portabilità dei dati. Il rispetto e l’aderenza alle leggi verrebbe realizzato dall’infrastruttura stessa, permettendo alle aziende di sviluppare i software in maniera meno vincolata, acquisendo al contempo la fiducia del cliente grazie alla trasparenza e all’affidabilità garantite da MyData.

# 2 Mobility as a Service e SMAll

## 2.1 Mobility as a Service

Nonostante siano state evidenziate le caratteristiche dell’architettura MyData sviluppate in misura più rilevante nel progetto di tesi, non bisogna tralasciare il contesto in cui si sviluppa questo modello. MyData è infatti un progetto dell’università finlandese Aalto, all’interno del più ampio contesto del “Mobility as a Service”.

Il paradigma MaaS descrive un nuovo utilizzo delle tecnologie, in particolare dei mezzi di trasporto, che elimina l’auto di proprietà per passare al trasporto pubblico, o in generale a un mezzo condiviso. Questo cambiamento permetterebbe di aumentare l’efficienza con cui i mezzi condivisi sono usati, eliminando gli sprechi che inevitabilmente derivano dal possesso di un autoveicolo e dal suo inutilizzo. L’accesso del privato ai mezzi di trasporto avverrebbe, in questa nuova ottica, attraverso un software in grado di calcolare una ottimizzazione per i mezzi condivisi, ad esempio raggruppando gli utenti per fasce orarie, tratte comuni e generiche preferenze.

### 2.1.1 Mobility Profile e Journey Planner

Come *proof of concept*, sono state realizzate dall’università di Aalto due applicazioni, rilasciate su piattaforma Android e iOS, Mobility Profile e Journey Planner.

Mobility Profile raccoglie i dati personali dell’utente, li mantiene in un database relazionale e svolge le operazioni di calcolo del prossimo viaggio più probabile. Questa applicazione funziona come base di appoggio per Journey Planner, che raccoglie i risultati forniti e restituire un feedback al processo sottostante, in modo da rendere più precisi i suggerimenti. Le API di comunicazione fra le due sono requestSuggestions(), requestTransportModePreferences(), e sendSearchedRoute(Place startLocation, Place destination).

In questo esempio, Mobility Profile non rispetta precisamente le specifiche dettate per MyData, ma è comunque possibile riconoscere alcune sue caratteristiche che rientrano nella filosofia sottostante. Ne sono un esempio la richiesta esplicita di un permesso (revocabile in ogni momento) per l’utilizzo di dati personali come gli impegni del calendario e lo storico delle posizioni GPS, e anche lo sviluppo di una applicazione separata per l’utilizzo dei risultati (Journey Planner) rispetto a quella che raccoglie i dati e li processa.

## 2.2 Smart Mobility for All

L’idea centrale su cui si fonda la piattaforma SMAll è di costruire un insieme di servizi di mobilità, potenzialmente anche per mezzi di trasporto molto diversi fra loro, che gestisca l’intero ciclo di vita di un viaggio, dalla prenotazione all’effettivo spostamento. Si tratta di Smart Mobility poiché attraverso la raccolta e l’analisi dei dati di un utente è possibile fare inferenze sui prossimi spostamenti, proponendo anche l’acquisto dei titoli di viaggio più adatti alle abitudini e alle necessità dell’individuo.

Anche in questo contesto risulta evidente l’importanza della gestione dei dati personali, non solo in termini di protezione della privacy ma anche con uno sguardo all’interoperabilità e al corretto funzionamento di un insieme di servizi eterogenei.

# 3 Architettura MyData

## 3.1 Entità fondamentali

L’architettura di MyData si costruisce su quattro componenti di base: l’utente finale, detto anche *Account Owner*, l’operatore MyData, o *Operator*, e due generiche entità che, da una parte, “producono” dati, e, dall’altra, li “consumano”. Queste sono dette rispettivamente *Source* e *Sink*. Mentre i ruoli di *Account Owner* e di *Operator* sono chiari e generalmente statici, quelli di *Source* e *Sink* sono fortemente variabili nel tempo e si possono applicare anche ad entità molto diverse fra loro, poiché definiti con un alto livello di astrazione. Convenzionalmente, si può identificare un servizio come “consumatore” di dati, mentre l’account dell’utente può essere un “produttore” di dati personali. In MyData invece, è possibile che un servizio occupi entrambi i ruoli, o anche che l’*Operator* stesso rientri in questa classificazione quando si trova a compiere operazioni sui dati.

Il ruolo di *Operator* comprende operazioni di vario genere, fra i quali vi sono la gestione degli utenti, dei servizi e delle interazioni che avvengono fra le due parti.

## 3.2 Service Registry, Service Linking

Viene indicata con *Service Registry* quella parte dell’*Operator* che contiene un database di tutti i servizi registrati presso quell’operatore, insieme ad un servizio detto *Service Discovery*, utilizzato dagli utenti per trovare nuovi servizi secondo certe specifiche. Ogni nuovo servizio che vuole essere utilizzabile all’interno dell’architettura MyData deve quindi sottoporsi ad una procedura di registrazione, al termine della quale, in caso di successo, viene inserito all’interno del *Registry*.

L’iscrizione di un utente presso uno di questi servizi avviene tramite un processo chiamato *Service Linking*, in cui l’Operatore MyData si occupa di realizzare una identificazione mutua delle parti. Non è possibile utilizzare il servizio se non è precedentemente avvenuta questa fase.

## 3.3 Autorizzazioni e Consent

Come specificato dal GDPR, ogni tipo di operazione svolta sui dati personali di un utente deve essere stata autorizzata dallo stesso tramite un permesso (*Consent*), che quindi acquista in questo contesto una valenza legale.

Il ruolo del *Consent* è quindi particolarmente rilevante: esso definisce quali dati possono essere utilizzati e identifica le entità *Source* e *Sink* fra le quali avviene lo scambio, specificando come avviene il processamento dei dati.

Per dare la possibilità ad un utente di ritirare il permesso accordato ad un servizio, un *Consent* ha tre stati possibili: *Active*, *Disabled* e *Withdrawn*. Il primo è lo stato standard di funzionamento, in cui l’accesso ai dati è consentito; si hanno poi gli stati “disabilitato” e “ritirato”, in cui l’accesso è impedito. Nel caso in cui il permesso sia “ritirato” è necessario provvedere all’emissione di un nuovo permesso.

## 3.4 Personal Data Storage

Nonostante non venga dato un peso rilevante a questa componente all’interno dei documenti di MyData, non è possibile prescindere dall’esistenza di un database che mantenga tutti i dati relativi ad un *Account Owner*. Non si tratta infatti solo di dati personali, ma di ogni dato utilizzato da un generico servizio o ad esempio inserito volontariamente dall’utente.

Non è specificato se il Personal Data Storage (spesso chiamato anche Personal Data Vault) faccia parte dell’ecosistema dell’operatore: ciò è possibile, ma non sono da escludere implementazioni alternative che prevedono il salvataggio delle informazioni presso il dispositivo dell’utente.

# 4 Analisi e Design

Ad alto livello, per realizzare il servizio di gestione dei dati personali è richiesta la presenza delle seguenti entità:

* Un utente finale, che disponga di un account presso l’Operatore MyData e di un account presso il servizio di calcolo del prossimo viaggio più probabile;
* Il servizio che realizzi la logica di business;
* L’Operatore MyData;
* Un Service Registry, presso cui il servizio è registrato;
* Un gestore dei permessi per l’utilizzo del servizio e l’accesso ai dati dell’utente;
* Un Personal Data Vault per la memorizzazione dei dati.