



# Analisi Tecnica Preliminare OTP-Italy

OTP-Italy, un progetto che punta a migliorare la **transizione dagli standard *de facto* a quelli *de jure* dei dati sulla mobilità**, potenziando l'**interoperabilità** e migliorando l'**accessibilità** del trasporto pubblico.

<b>1. Introduzione</b>	<b>1</b>
<b>2. Mapping tra i formati dati ed approccio selezionato dei converter</b>	<b>3</b>
2.1 Macro categorie di informazioni	3
2.2 Mapping	8
2.2.1 Analisi delle corrispondenze tra il profilo italiano di NeTEx e GTFS	8
2.2.2 Analisi delle corrispondenze tra il profilo italiano di SIRI e GTFS-RT	13
2.3 Definizione dell'approccio per la conversione dei formati dati	14
<b>3. Miglioramento ed ottimizzazione dei dati in formato GTFS e GTFS-RT</b>	<b>16</b>
3.1 Creazione di file shapes.txt	16
3.2 Implementazione di route-type extension	18
3.3 Miglioramento della granularità delle informazioni	19
3.4 Miglioramento nella gestione delle tariffe	21
3.5 Integrazione dati Pathways e Levels	23
<b>4. Sviluppo del fork di OpenTripPlanner</b>	<b>26</b>
<b>5. Bibliografia</b>	<b>29</b>

# 1. Introduzione

Questo documento presenta un'analisi tecnica preliminare svolta nel contesto di **OTP-Italy**, progetto svolto in collaborazione fra OpenMove e Cefriel e finalizzato allo sviluppo di un **convertitore dati da GTFS<sup>1</sup> a NeTEx profilo italiano<sup>2</sup> e da GTFS-RT<sup>3</sup> a SIRI profilo italiano<sup>4</sup>** (di seguito *converter*) e all'implementazione di un **fork del motore di calcolo del percorso *OpenTripPlanner* (OTP)**. In particolare, questo documento rappresenta la **prima milestone di progetto** ed andrà a delineare brevemente il **lavoro preparatorio** svolto nelle prime 12 settimane di progetto.

Il Capitolo 2 descrive il lavoro preparatorio svolto per l'analisi delle macro categorie di informazioni esprimibili nei profili italiani degli standard NeTEx e SIRI. Viene presentata la mappatura con i formati GTFS e GTFS-RT, evidenziando le corrispondenze e le discrepanze tra i diversi standard. Inoltre, il capitolo illustra l'approccio definito per la conversione dei dati, basato su un modello di riferimento che garantisce interoperabilità e scalabilità nell'integrazione delle informazioni di trasporto pubblico.

Il Capitolo 3 approfondisce le aree di miglioramento e le strategie di ottimizzazione dei dati GTFS e GTFS-RT prima della loro conversione, con l'obiettivo di arricchire e rendere più accurate le informazioni per il sistema di routing. Vengono analizzati diversi aspetti chiave, tra cui la creazione del file *shapes.txt* mediante tecniche di map-matching per migliorare la rappresentazione delle tratte, l'implementazione della *route-type extension* per una classificazione più dettagliata dei mezzi di trasporto e il miglioramento della granularità delle informazioni relative alle fermate. Inoltre, il capitolo tratta l'ottimizzazione della gestione tariffaria attraverso l'evoluzione da GTFS-Fares V1 a V2 e l'integrazione dei dati

---

<sup>1</sup> <https://gtfs.org/documentation/schedule/reference/>

<sup>2</sup> <https://github.com/5TsrI/netex-italian-profile>

<sup>3</sup> <https://gtfs.org/documentation/realtime/reference/>

<sup>4</sup> <https://github.com/5TsrI/siri-italian-profile>

sulla struttura fisica delle stazioni tramite i file *pathways.txt* e *levels.txt*, migliorando l'accessibilità e la navigazione negli hub di trasporto.

Il capitolo 4 introduce il motore di calcolo del percorso OpenTripPlanner (OTP) e spiega l'utilità di questa soluzione software all'interno dell'ecosistema di mobilità. Il capitolo approfondisce anche ad alto livello la modalità di funzionamento di OTP e le modalità di gestione dei flussi dati provenienti dai convertitori.

L'obiettivo del documento è di definire una **roadmap tecnica** che permetta di ottimizzare il processo di conversione dei dati e di migliorarne l'usabilità, con particolare attenzione all'affidabilità e alla qualità delle informazioni per la pianificazione multimodale di offerte agli utenti finali.

## 2. Mapping tra i formati dati ed approccio selezionato dei converter

La prima parte del lavoro preparatorio si è concentrata prevalentemente su tre aree: **l'identificazione delle macro categorie di informazioni** esprimibili mediante i profili italiani degli standard NeTEx e SIRI, lo svolgimento di attività di **mapping tra GTFS e NeTEx e tra GTFS-RT e SIRI** e la definizione dell'**approccio selezionato per l'implementazione dei converter**.

### 2.1 Macro categorie di informazioni

Per quanto attiene la prima area, le attività svolte nell'analisi dei profili italiani di NeTEx e SIRI sono state relative a identificare le macro categorie di informazioni rappresentabili adottando queste due specifiche. Questa analisi, oltre ad essere rilevante ai fini dello sviluppo del converter, è propedeutica all'identificazione di quali informazioni risultano utili e gestibili da OTP, sia con riferimento alle **informazioni statiche** (orari, percorsi, fermate, etc.) che a quelle **dinamiche** (ritardi, cancellazioni, variazioni di percorso, etc.).

**NeTEx** (Network Timetable Exchange) è uno standard europeo, basato su XML, concepito per la gestione e lo scambio dei **dati statici relativi al trasporto pubblico** come orari, itinerari, fermate, strutture tariffarie e servizi. Lo standard **NeTEx** viene promosso a livello europeo per garantire maggiore interoperabilità tra diversi operatori della mobilità, facilitare la condivisione e l'interpretazione corretta dei dati. Lo standard è stato definito grazie al lavoro del Comitato Europeo di Normazione (CEN) ed è basato sul più generico modello dati **Transmodel**<sup>5</sup> che garantisce l'interoperabilità anche con gli altri standard

---

<sup>5</sup> <https://transmodel-cen.eu/>

derivati dallo stesso modello, tra cui SIRI. Lo standard prevede la possibilità di specificare dei *profili* che definiscono il subset di informazioni da considerare e la rappresentazione che si intende adottare conformemente allo standard.

Il **profilo italiano di NeTEx** è caratterizzato dall'adattamento del modello generale alle specifiche esigenze del sistema di trasporto pubblico italiano, includendo dettagli come le caratteristiche dei servizi, le informazioni sui biglietti e le regole di validazione dei dati specifiche per il mercato italiano. Esso è strutturato in 5 diversi livelli.

Il **Livello 1** raggruppa le informazioni fondamentali per l'alimentazione dei servizi di informazione ai passeggeri del trasporto pubblico. Il livello 1 è conforme alla norma UNI CEN/TS 16614-4:2021 "Trasporto pubblico - Scambio di rete e orario (NeTEx) - Parte 4: Profilo europeo relativo alle Informazioni sui passeggeri" che definisce profilo europeo EPIP del NeTEx<sup>6</sup>. In particolare definisce informazioni relative a:

- Nodi di accesso delle fermate;
- Calendario di operatività;
- Topologia della rete e itinerari/linee;
- Orari di operatività;
- Operatori dei trasporti.

Il **Livello 2** integra il livello 1 con informazioni per la correlazione del servizio di trasporto pubblico con gli aspetti contrattuali di interesse delle pubbliche amministrazioni, nello specifico definisce:

- Contratti di servizio;
- Tipologie di veicoli;
- Servizi a bordo;

---

<sup>6</sup> <https://github.com/NeTEx-CEN/NeTEx-Profile-EPIP>

Il **Livello 3** integra i livelli precedenti con le informazioni di tariffazione del servizio, in particolare definisce informazioni relative a:

- Parcheggi (Tariffe);
- Titoli di viaggio;
- Tariffe e Zone Tariffarie;
- Punti e Piattaforme di Vendita;
- Classi di passeggeri.

Il **Livello 4** integra i livelli precedenti con una descrizione dei servizi di mobilità alternativi al trasporto pubblico locale, come car sharing, bike sharing o altri servizi di mobilità condivisa:

- Dotazione e caratteristiche dei Veicoli;
- Flotta dei veicoli di bike o car sharing.

Il **Livello 5** si concentra su dati specifici che descrivono le caratteristiche di accessibilità delle infrastrutture e dei servizi di trasporto pubblico

- Accessibilità delle fermate e delle stazioni;
- Caratteristiche dei veicoli in termini di 'accessibilità';
- Descrizione di servizi di assistenza specifici disponibili presso fermate, stazioni o a bordo dei mezzi;
- Informazioni aggiuntive per passeggeri con esigenze speciali riguardo modalità di prenotazione per servizi dedicati e segnalazione di eventuali limitazioni o condizioni di utilizzo.

Il **protocollo SIRI** (Service Interface for Real Time Information) è uno standard europeo, basato su XML, volto a supportare lo scambio di **dati dinamici in tempo reale**, tra sistemi potenzialmente eterogenei, relativamente ai servizi e ai veicoli del trasporto pubblico. SIRI prevede in totale 10 servizi funzionali basati sullo stesso framework, ognuno dei quali

designato allo scambio di tipi specifici di dati. Come per NeTEx, lo standard è stato definito grazie al lavoro del CEN ed è basato sul più generico modello dati **Transmodel**.

Il **profilo italiano di SIRI** considera esclusivamente 4 dei 10 servizi funzionali per lo scambio in tempo reale dei dati. L'analisi delle macro categorie di informazioni tiene conto dunque dei 4 servizi inclusi nel profilo italiano.

Il servizio **SIRI-VM (Vehicle Monitoring)** permette di comunicare la posizione dei mezzi in tempo reale e/o gli orari di passaggio (effettivi o previsti) dei mezzi a specifiche fermate. Le macro categorie di informazioni incluse sono:

- Posizione del mezzo in tempo reale (latitudine, longitudine, direzione del mezzo);
- Se il mezzo è in fermata o meno;
- Orario teorico di partenza dalla fermata;
- Orario effettivo di partenza dalla fermata;
- Ritardo del mezzo;
- Stato occupazionale del mezzo.

Il servizio **SIRI-ET (Estimated TimeTable)** permette di comunicare gli orari di transito registrati durante lo svolgimento di una corsa specifica e informazioni relative ai passeggeri saliti, discesi, a bordo per ciascuno dei transiti in fermata già registrata. In particolare, include le seguenti macro categorie di informazione:

- Orari teorici/effettivi di partenza/arrivo del mezzo (d)alla fermata;
- Orari previsti di partenza/arrivo del mezzo alle prossime fermate (nel caso l'operatore posseda un previsore);
- Percentuale di occupazione del mezzo;
- Numero di passeggeri saliti/discesi/a bordo veicolo dopo la partenza dalla fermata
- Numero di sedili estraibili/posti per carrozzine/posti riservati usati dopo la partenza dalla fermata.

Il servizio **SIRI-SX (Situation Exchange)** permette di comunicare gli eventi che impattano sulla regolarità dei servizi di trasporto pubblico. Le macro categorie di informazioni che supporta sono:

- Stato dell'evento;
- Causa dell'evento;
- Quali componenti influenza (linee, corse, fermate, e relativa posizione);
- Conseguenze (data/ora della conseguenza, gravità, se l'evento impedisce o meno la discesa/salita dei passeggeri, ritardo causato).

Il servizio **SIRI-FM (Facility Monitoring)** permette di comunicare lo stato di specifiche infrastrutture, e si concentra sulle seguenti macro categorie di informazione:

- Stato occupazionale dei parcheggi (disponibile, rimosso, etc.) e conteggio/percentuale degli stalli per veicoli/biciclette disponibili/fuori servizio/occupati/riservati;
- Disponibilità e stato degli stalli di ricarica dei veicoli elettrici in un parcheggio e conteggio/percentuale degli stalli disponibili/fuori servizio/occupati/riservati;
- Disponibilità e stato delle biciclette di bike sharing e conteggio/percentuale degli stalli disponibili/fuori servizio/occupati/riservati.

## 2.2 Mapping

Con riferimento alla seconda area, è stata svolta una prima attività di **mapping** (ovvero **l'analisi dei punti di corrispondenza**) tra il profilo italiano di NeTEx e GTFS, e tra il profilo italiano di SIRI e GTFS-RT. Questa analisi è propedeutica all'implementazione dei converter ma è stata svolta anche al fine di **identificare eventuali discrepanze e lacune informative** negli standard GTFS e GTFS-RT rispetto ai profili italiani di NeTEx e SIRI. I risultati presentati in questa sezione definiscono quindi il campo di applicazione dei

converter che verranno sviluppati, ma forniscono anche un input all'analisi svolta da OpenMove (Sezione 3) per garantire che le informazioni rilevanti siano presenti nei GTFS e GTFS-RT di input e per identificare ulteriori fonti dati aggiuntive necessarie ai fini della conversione. Le informazioni non presenti o non rappresentabili tramite GTFS e GTFS-RT devono infatti essere recuperate da ulteriori fonti dati per permettere il popolamento delle relative strutture dati nei profili italiani NeTEx e SIRI. In questa direzione, l'analisi sui requisiti minimi per fornire un input adeguato ai converter prenderà in considerazione i vincoli dei profili italiani di NeTEx e SIRI (campi obbligatori e raccomandati) ma anche le tipologie di dati maggiormente propedeutici ad un miglior funzionamento del software OTP.

### 2.2.1 Analisi delle corrispondenze tra il profilo italiano di NeTEx e GTFS

Per quanto riguarda il profilo italiano di **NeTEx**, è stato portato a termine un confronto preliminare con **GTFS**, che è lo standard *de facto* per lo scambio di informazioni statiche relative alla rete di trasporto pubblico.

Ciascun livello del NeTEx profilo italiano definisce diverse strutture, dette "Frame", che a loro volta contengono campi e sottostrutture che racchiudono le diverse categorie di informazioni descritte nella sezione precedente. Il formato GTFS è invece definito da una serie di file CSV, ognuno dedicato ad uno specifico tipo di informazioni.

Riguardo al **Livello 1** del profilo italiano di NeTEx:

- Il *ResourceFrame* contiene dettagli sulle organizzazioni coinvolte nel servizio di trasporto pubblico che trovano una corrispondenza nel GTFS (file), ma potrebbe essere necessario un arricchimento dei dati per soddisfare completamente i requisiti del profilo italiano di NeTEx. Il *ResourceFrame* definisce anche le tipologie di

veicoli utilizzati nel servizio, i modelli e i dettagli sui singoli veicoli utilizzati nel servizio. GTFS non fornisce informazioni a questo livello di dettaglio sui veicoli.

- Il *ServiceCalendarFrame* contiene informazioni sui calendari di validità dei servizi di trasporto nel contesto NeTEx. Queste informazioni sono rappresentate in GTFS attraverso la combinazione dei file **calendar.txt** e **calendar\_dates.txt**.
- Il *SiteFrame* include informazioni sulla struttura fisica delle località di trasporto pubblico (e.g., fermate, stazioni, parcheggi e altre infrastrutture correlate). In GTFS, le fermate sono rappresentate nel file **stops.txt**, dove ogni fermata ha un identificativo univoco, un nome e coordinate geografiche. Tuttavia, NeTEx offre una struttura più gerarchica e dettagliata per le fermate, per esempio permette di rappresentare i punti specifici dove i veicoli si fermano per il carico e scarico dei passeggeri, come banchine o marciapiedi. Alcune informazioni aggiuntive sulle stazioni possono essere ottenute in GTFS se i files **pathways.txt** e **levels.txt** sono presenti. Il *SiteFrame* permette di rappresentare anche entità geografiche come città, quartieri o altre aree rilevanti per il servizio di trasporto non disponibili in GTFS.
- Il *ServiceFrame* rappresenta informazioni su linee, percorsi, fermate, collegamenti tra fermate e pattern di viaggio. Le reti di trasporto possono essere rappresentate nel file GTFS **networks.txt**. Le rotte e le linee dei servizi di trasporto sono rappresentati nel file **routes.txt** e possono essere mappate alle reti e raggruppati mediante il file **route\_networks.txt** se disponibile. I punti di fermata sono rappresentate in GTFS nel file **stops.txt**. Il percorso della linea può essere costruito combinando l'informazione contenuta nei file **stops.txt**, **stop\_times.txt** and **shapes.txt**. L'assegnazione delle fermate ai vari servizi o linee e eventuali zone tariffarie sono presenti nel file **stops.txt** del GTFS. Per molte tipologie di dati del *ServiceFrame* esiste un mapping in GTFS ma l'informazione esprimibile è molto più limitata rispetto a NeTEx.

- Il *TimetableFrame* rappresenta gli orari programmati dei servizi di trasporto. I viaggi dei veicoli secondo l'orario programmato possono essere estratti combinando le informazioni nei file GTFS **trips.txt** e **stop\_times.txt**. Le possibili coincidenze o interscambi tra diversi viaggi all'interno della rete di trasporto pubblico sono esprimibili in GTFS tramite il file **transfers.txt**.

Il **Livello 2** del NeTEx profilo italiano integra le informazioni del livello 1 con dettagli sui veicoli e sui servizi a bordo:

- Il *GeneralFrame* viene utilizzato per incapsulare informazioni generali che possono includere contratti di servizio, politiche operative e altre informazioni di alto livello che non hanno una diretta rappresentazione in GTFS.
- Il *ResourceFrame* viene esteso con la descrizione dei consorzi di aziende di trasporto pubblico e dei ruoli dei vari attori del sistema all'interno di ogni contratto di servizio. Entrambe queste informazioni non sono disponibili in GTFS.
- Il *SiteFrame* viene esteso per rappresentare le aree di parcheggio correlate al sistema di trasporto pubblico, come parcheggi di interscambio o strutture per biciclette. GTFS non include informazioni sulle aree di parcheggio.
- Il *ServiceFrame* viene esteso con le informazioni visualizzate sui pannelli informativi relativi alle destinazioni dei servizi, estraibili dai file **stop\_times.txt** e **trips.txt** del GTFS, e i raggruppamenti di più fermate, definiti nei files **areas.txt** e **stop\_areas.txt**.
- Il *TimeTableFrame* viene esteso con la descrizione delle tipologie di servizio non prevista in GTFS.

Il **Livello 3** descrive la tariffazione e la gestione economica dei servizi di trasporto tramite l'introduzione del *FareFrame*. Le possibilità di mapping tra GTFS e NeTEx profilo italiano variano a seconda che sia utilizzato GTFS-Fares v1 o GTFS-Fares v2.

Nel caso di GTFS-Fares v1 le tariffe applicate ai servizi di trasporto e i parametri di utilizzo delle tariffe come restrizioni o validità temporale sono accessibili attraverso il file

**fare\_attributes.txt.** Ulteriori parametri per determinare una tariffa, per esempio se dipendente da specifiche stazioni o zone di origine/destinazione, sono definiti tramite il file **fare\_rules.txt.**

Alcuni elementi del *FareFrame* definiti dal NeTEx profilo italiano possono essere popolati solo nel caso il GTFS-Fare v2 sia utilizzato. GTFS-Fare v2 introduce una struttura più flessibile che consente di rappresentare ed estrarre informazioni su tariffe complesse, come abbonamenti, tariffe basate su fasce orarie, sconti per categorie specifiche e trasferimenti tra operatori diversi.

GTFS non permette però di rappresentare informazioni sui canali di distribuzione dei titoli di viaggio, come vendite online o presso rivenditori fisici, e la loro combinazione con particolari condizioni di vendita. Similmente, GTFS non copre la possibilità di specificare tariffe relative ai parcheggi associati ai servizi di trasporto.

Il **Livello 4** del NeTEx profilo italiano integra le informazioni dei livelli precedenti con i dati relativi a servizi di mobilità alternativi al trasporto pubblico locale, come car sharing, bike sharing o altri servizi di mobilità condivisa:

- *MobilityServiceFrame* definisce diverse strutture per descrivere la flotta dei veicoli disponibili, le modalità operative, le condizioni e le zone di utilizzo.
- Estende il *ResourceFrame* per dettagliare i profili dei modelli di veicoli utilizzati nei servizi di trasporto.

GTFS non contiene informazioni riguardo i servizi di mobilità alternativa e per questo non presenta elementi di corrispondenza con il Livello 4 del profilo NeTEx profilo italiano. La specifica *General Bikeshare Feed Specification* (GBFS)<sup>7</sup> può essere considerata per la definizione di un converter dedicato che però non è oggetto delle attività di questo progetto.

---

<sup>7</sup> <https://gbfs.org/>

Il **Livello 5** si concentra su dati specifici che descrivono le caratteristiche di accessibilità delle infrastrutture e dei servizi di trasporto pubblico. Non è definito nessun nuovo frame per questo livello, ma è definita una struttura *AccessibilityAssessment* che è richiamata all'interno di altri frame:

- Nel *SiteFrame*, permette di modellare informazioni di accessibilità per le fermate e limitate informazioni sono reperibili nel file **stops.txt**.
- Nel *ServiceFrame*, è usata per modellare l'accessibilità di una linea. Il GTFS non prevede un campo specifico per valutare l'accessibilità a livello di linea.
- Nel *TimetableFrame*, permette di esprimere l'accessibilità delle singole corse. Nel GTFS, il file *trips.txt* specifica l'accessibilità delle corse.

Per tutti i casi considerati tuttavia, le informazioni di accessibilità modellabili tramite GTFS risultano estremamente limitate rispetto a quelle definite dal NeTEx profilo italiano.

In ultima considerazione, recentemente GTFS ha incorporato al suo interno la specifica GTFS Flex<sup>8</sup>, precedentemente proposta come estensione, che permette di modellare le informazioni relative ai servizi Demand Responsive Transportation (DRT). Il NeTEx profilo italiano non permette attualmente di modellare questa tipologia di servizi ma è previsto che il supporto venga introdotto tramite nuovi livelli del profilo definiti in futuro. Nel contesto di questo progetto considereremo eventuali specifiche GTFS Flex per il loro diretto caricamento in OTP senza l'intermediazione dei converter.

### 2.2.2 Analisi delle corrispondenze tra il profilo italiano di SIRI e GTFS-RT

Per quanto riguarda il profilo italiano di **SIRI**, è stato portato a termine un confronto preliminare con **GTFS-RT**, che è lo standard *de facto* per lo scambio di informazioni in tempo reale relative alla rete di trasporto pubblico.

---

<sup>8</sup> <https://gtfs.org/community/extensions/flex/>

In particolare, le funzioni di SIRI del profilo italiano trovano una corrispondenza sui messaggi di GTFS-RT come segue:

- **SIRI-VM (Vehicle Monitoring)** ha una corrispondenza sui campi inclusi nel messaggio GTFS-RT **VehiclePosition**. Alcuni campi non sono direttamente estraibili dal messaggio di GTFS-RT, ma possono essere recuperati da messaggi GTFS, per esempio, il nome della fermata a cui si sta facendo riferimento e l'orario programmato di partenza del mezzo dalla fermata.
- **SIRI-ET (Estimated Timetable)** corrisponde a campi inclusi nel messaggio **TripUpdate**. In GTFS-RT risultano però mancanti le informazioni, presenti in SIRI, relative al numero di passeggeri che sono saliti, discesi o sono a bordo dopo la partenza da una fermata, e il numero di sedili estraibili, posti per carrozzine e posti riservati usati dopo la partenza. In GTFS-RT è unicamente possibile specificare in maniera qualitativa lo stato occupazionale del mezzo, scegliendo in una lista di valori (per esempio, *empty*, *few\_seats\_available*, etc.). Inoltre, come per Vehicle Monitoring, alcuni campi non si mappano direttamente su GTFS-RT, ma possono essere estratti da GTFS.
- **SIRI-SX (Situation Exchange)** può essere valorizzato da campi inclusi nel messaggio **Alert** di GTFS-RT. In GTFS-RT risultano tuttavia mancanti un numero significativo di informazioni previste in SIRI, ovvero: stato dell'evento (per esempio, chiuso, aperto, in attesa di approvazione, etc.), dettagli sul fornitore dell'evento (come nazionalità e azienda), tipologia della fonte (per esempio, via mail, fax, telefono, etc.), dettagli sulle conseguenze dell'evento (inizio e fine delle conseguenze, durata dell'eventuale ritardo causato dall'evento, possibilità o meno di scendere/salire alla fermata su cui agisce l'evento), al di fuori dei valori della lista *Effect* (per esempio, *accessibility\_issue*, *significant\_delays*, etc.). Inoltre, come per i due messaggi precedentemente trattati, alcuni campi del messaggio Alert non si mappano direttamente su GTFS-RT, ma possono essere recuperati da GTFS (per esempio, le coordinate della fermata coinvolta dall'evento).

- **SIRI-FM (Facility Monitoring)** non ha una corrispondenza in GTFS-RT, in quanto quest'ultimo non prevede la modellazione di informazioni relative ai parcheggi. Alcune delle informazioni previste da SIRI-FM sono modellate in GBFS in relazione ai servizi di sharing mobility ma non sono oggetto del presente lavoro.

A completamento dell'attività di mapping va notato come alcuni valori siano estraibili solo considerando più messaggi GTFS-RT per ottenere una singola funzione SIRI. Per esempio, l'orario di partenza effettivo del mezzo dalla fermata non è recuperabile da VehiclePosition per SIRI-VM, ma dal messaggio TripUpdate, che si mappa invece su un'altra funzione SIRI (SIRI-ET). Inoltre, il dato sulla percentuale di occupazione del mezzo dopo la partenza dalla fermata indicata non è presente in TripUpdate per SIRI-ET, ma si mappa su un campo del messaggio VehiclePosition.

## 2.3 Definizione dell'approccio per la conversione dei formati dati

Diversi approcci possono essere applicati per consentire la conversione tra diversi formati di dati di mobilità. Tali approcci spaziano da soluzioni ad-hoc mirate a uno scenario specifico a soluzioni più generali e scalabili che supportano più stakeholder e formati di dati. L'approccio any-to-one semantic mapping, basato su [1], consiste nell'utilizzo di un modello di riferimento (e.g., Transmodel<sup>9</sup>) e sulla definizione di mappature da una specifica rappresentazione di dati (e.g., GTFS) al modello di riferimento e viceversa.

Un'applicazione specifica di questo approccio si basa sull'utilizzo di tecnologie del Semantic Web ed è stato convalidato in [2][3] e nel contesto del progetto SNAP<sup>10</sup>. Nelle implementazioni considerate il modello concettuale viene definito tramite un'ontologia e le

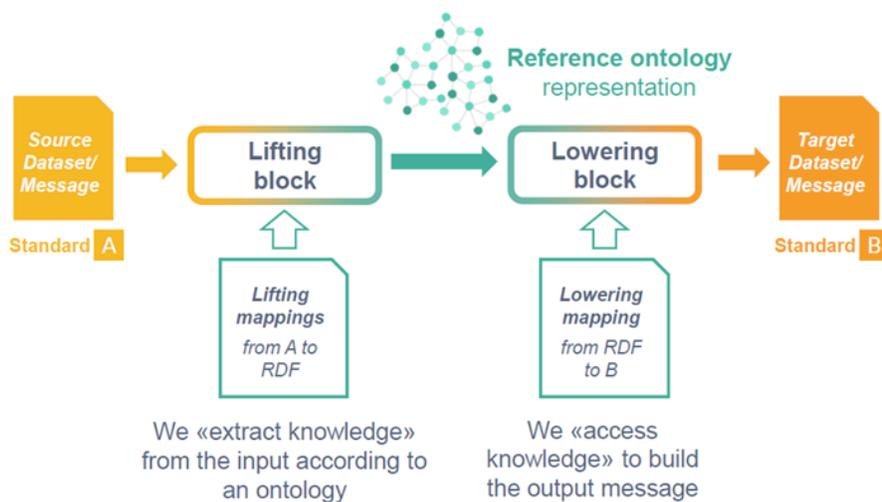
---

<sup>9</sup> <https://www.transmodel-cen.eu/>

<sup>10</sup> <https://snap-project.eu/>

soluzioni di mappatura permettono di ottenere e convertire dati eterogenei tramite l'utilizzo di un Knowledge Graph (KG) interoperabile basato sull'ontologia di riferimento. La soluzione open-source Chimera<sup>11</sup> [5], basata su Apache Camel<sup>12</sup>, consente, ad esempio, di realizzare l'approccio any-to-one mediante la definizione di pipeline di trasformazione dei dati mediante diversi componenti per la costruzione, trasformazione, convalida e utilizzo di Knowledge Graph. In particolare, Chimera abilita la trasformazione dei dati mediante mapping dichiarativi [4][6], che permettono il disaccoppiamento delle regole di mapping dal software di conversione, migliorando la scalabilità e manutenzione delle soluzioni di conversione.

La Figura 1 dimostra l'approccio presentato nel generico caso di conversione tra uno Standard A e uno Standard B. Il blocco di *lifting* esegue un set di regole di mapping per estrarre le informazioni rilevanti dall'input e ottenere una rappresentazione tramite l'ontologia di riferimento in un KG. Il blocco di *lowering* accede l'informazione nel KG e la utilizza per costruire un output secondo lo Standard B che rappresenta l'obiettivo della trasformazione.



**Figura 1:** Approccio any-to-one semantic mapping

<sup>11</sup> <https://github.com/cefriel/chimera>

<sup>12</sup> <https://camel.apache.org/>

Nel contesto di questo progetto, applicheremo l'approccio descritto per la definizione di due converter: il primo da GTFS a NeTEx profilo italiano e il secondo da GTFS-RT a SIRI profilo italiano. Il TANGENT Conceptual Model definito in [3] verrà riutilizzato ed eventualmente esteso come ontologia di riferimento per supportare le trasformazioni. I converter verranno implementati tramite specifiche Chimera pipelines configurate per eseguire le conversioni tra gli standard menzionati. Le pipeline saranno integrate e deployate nel contesto della soluzione sviluppata dal progetto per permettere la comunicazione con OTP. La flessibilità fornita da Apache Camel nella configurazione delle pipeline definite tramite Chimera sarà utilizzata per definire un converter che possa essere invocato su richiesta per i dati statici, e servizi di conversione in continua esecuzione per processare i dati dinamici.

## 3. Miglioramento ed ottimizzazione dei dati in formato GTFS e GTFS-RT

La seconda parte del lavoro preparatorio si è concentrata sull'identificazione delle potenziali aree di miglioramento e sulle modalità di **arricchimento** e di **ottimizzazione dei dati** in formato GTFS e GTFS-RT prima della loro conversione. Questo processo di arricchimento ed ottimizzazione sarà infatti finalizzato all'**integrazione di nuove fonti dati** e all'**ampliamento delle informazioni disponibili**, al fine di garantire una maggiore accuratezza e completezza delle informazioni per il sistema di routing e di conseguenza **fornire risultati più accurati ed utili** all'utenza viaggiante. Alcune delle aree di miglioramento sono discusse più in dettaglio qui di seguito.

### 3.1 Creazione di file *shapes.txt*

L'integrazione di strumenti di **map-matching** rappresenta un elemento chiave per migliorare la qualità e l'affidabilità dei dati di mobilità, in particolare per quanto riguarda la rappresentazione delle tratte all'interno del dataset GTFS. Il **map-matching** è una tecnica che permette di associare dati di localizzazione grezzi (spesso derivati da GPS) a una rete stradale o ferroviaria predefinita, correggendo eventuali imprecisioni nelle coordinate e generando percorsi più realistici. Nel contesto di questo progetto, tale funzionalità assume particolare rilevanza nei casi in cui i dati sulle **shapes** (cioè la rappresentazione geometrica delle rotte seguite dai mezzi di trasporto pubblico) siano assenti o incompleti all'interno del dataset GTFS fornito dagli operatori di trasporto.

Grazie all'uso di algoritmi avanzati di map-matching, il sistema sarà in grado di elaborare le coordinate fornite dai dispositivi GPS montati sui veicoli e confrontarle con la rete stradale

esistente, determinando con maggiore accuratezza il percorso effettivamente seguito dai mezzi. Questo consentirà di generare shapes precise e dettagliate, migliorando la qualità del dataset GTFS e, di conseguenza, la rappresentazione grafica delle linee di trasporto su mappe digitali. Una rappresentazione più accurata delle tratte non solo facilita la comprensione visiva per gli utenti finali, ma ottimizza anche il calcolo dei percorsi all'interno del **motore di routing** di OpenTripPlanner. Infatti, percorsi ben definiti riducono il rischio di discrepanze tra i dati teorici e quelli reali, migliorando l'affidabilità delle informazioni fornite ai viaggiatori in termini di orari stimati, distanze percorse e possibili connessioni tra diversi mezzi di trasporto.

Un ulteriore vantaggio dell'integrazione di map-matching è la possibilità di adattare il sistema a contesti in cui i dati di trasporto pubblico non sono sempre disponibili in formati standardizzati. Ad esempio, in alcune aree rurali o in sistemi di mobilità emergenti (come servizi di bus a chiamata o linee flessibili), i percorsi possono variare dinamicamente e non essere predefiniti nei dataset forniti. In questi casi, il map-matching può servire per costruire le shapes delle tratte basandosi sui dati storici e rilevazioni GPS, offrendo così una soluzione adattabile anche a sistemi di trasporto non convenzionali.

Dal punto di vista tecnico, il progetto potrà sfruttare framework open-source di map-matching come **Pfaedle**<sup>13</sup> o **Open Source Routing Machine (OSRM)**<sup>14</sup>, che offrono algoritmi efficienti per l'allineamento delle tracce GPS alla rete stradale e ferroviaria. Inoltre, l'integrazione con **dati cartografici di OpenStreetMap (OSM)**<sup>15</sup> permetterà di ottenere una mappatura accurata della rete viaria, migliorando ulteriormente la precisione del sistema. La combinazione di questi strumenti con OTP consentirà di ottenere un motore di routing più affidabile e una visualizzazione più fedele della realtà, garantendo

---

<sup>13</sup> <https://github.com/ad-freiburg/pfaedle>.

<sup>14</sup> <https://project-osrm.org/>.

<sup>15</sup> OpenStreetMap (<https://wiki.openstreetmap.org/>) è un database geografico collaborativo open source che fornisce dati cartografici utilizzabili in applicazioni GIS.

agli utenti un'esperienza di viaggio più fluida e coerente con le effettive condizioni del trasporto pubblico.

## 3.2 Implementazione di *route-type extension*

Il GTFS utilizza una classificazione standard per i tipi di percorso (**route\_type**), che assegna un valore numerico predefinito a ciascuna tipologia di trasporto pubblico, come autobus, tram, treni o traghetti. Tuttavia, questa categorizzazione può risultare limitante in alcuni contesti, poiché non sempre permette di rappresentare con sufficiente granularità la varietà di mezzi di trasporto esistenti.

Per superare questa limitazione, Google ha introdotto una estensione delle tipologie di percorso (**route\_type extension**<sup>16</sup>), che consente di definire in modo più dettagliato i diversi veicoli utilizzati nell'ambito del trasporto pubblico. Questo approccio si rivela particolarmente utile nella creazione di **VehicleModels**, poiché permette di ipotizzare le caratteristiche dei veicoli in base alla loro categoria di appartenenza, migliorando così la precisione nella modellazione dei servizi di mobilità.

In alternativa, un'altra opzione per rappresentare le caratteristiche specifiche dei veicoli è l'introduzione di un file **vehicles.txt** all'interno del feed GTFS. Questo file, ispirato al modello **VehicleType** definito nello standard NeTEx, consentirebbe di descrivere dettagliatamente i diversi veicoli in termini di capacità, dimensioni, accessibilità e altre caratteristiche tecniche<sup>17</sup>.

Tuttavia, è importante sottolineare che l'adozione di **vehicles.txt** come standard ufficiale è ancora in fase di discussione. Di conseguenza, mentre l'estensione delle tipologie di percorso offre una soluzione già operativa e compatibile con GTFS, l'utilizzo di **vehicles.txt**

---

<sup>16</sup> <https://developers.google.com/transit/gtfs/reference/extended-route-types>.

<sup>17</sup> Questo approccio è stato proposto nella issue <https://github.com/google/transit/issues/458>.

potrebbe rappresentare un'evoluzione futura, garantendo una modellazione ancora più dettagliata del parco mezzi impiegato nei sistemi di trasporto pubblico.

### 3.3 Miglioramento della granularità delle informazioni

Uno degli aspetti chiave nella gestione e nell'organizzazione dei dati di trasporto pubblico è il livello di dettaglio fornito agli utenti. La qualità e la granularità delle informazioni non solo migliorano la coerenza e l'organizzazione dei dati ma hanno anche un impatto diretto sull'usabilità e sulla precisione delle informazioni per gli utenti finali.

Attualmente, in molti dataset relativi ai trasporti pubblici, le fermate vengono spesso indicate in maniera generica, senza fornire dettagli aggiuntivi che potrebbero essere fondamentali per la navigazione degli utenti. Ad esempio, indicare "Stazione di Trento" può risultare insufficiente per chi necessita di informazioni più precise, come il numero di binario o la posizione esatta della fermata. Aggiungere questo livello di dettaglio non è solo una questione di completezza informativa, ma anche di usabilità: più informazioni precise e organizzate vengono fornite, più semplice sarà per gli utenti pianificare i loro spostamenti.

Un metodo efficace per migliorare l'organizzazione delle fermate consiste nel controllare i nomi o le descrizioni degli stop adiacenti e raggrupparli in modo logico.

Si può immaginare il caso di una stazione ferroviaria in cui esistano tre fermate distinte nei dati:

- **"Stazione Trento - Binario 1"**
- **"Stazione Trento - Binario 2"**
- **"Stazione Trento - Binario 3"**

Invece di trattare queste fermate come entità completamente separate, è possibile organizzare queste informazioni tramite una struttura logica di tipo gerarchica; per fare ciò risulta necessario:

- Creare una fermata principale **“Stazione di Trento”** che racchiuda tutte le sottostazioni.
- Trattare i singoli binari come sotto-entità della stazione principale, identificandoli come **“Binario 1”**, **“Binario 2”**, ecc.

Questo approccio consente di mantenere un'informazione strutturata e facilmente interpretabile, migliorando sia l'organizzazione dei dati sia l'esperienza dell'utente che consulta la piattaforma.

Un altro esempio pratico riguarda fermate degli autobus o tram situate sulla stessa strada ma su lati opposti. Ad esempio, potremmo avere:

- **“Via Roma (direzione Centro)”**
- **“Via Roma (direzione Periferia)”**

Ancora una volta, invece di considerare queste fermate come due entità scollegate, si può rappresentarle come due fermate distinte collegate a un'unica fermata principale **“Via Roma”**, migliorando così la leggibilità e la gestione delle informazioni.

Nello standard NeTEx questa logica di organizzazione si traduce nell'utilizzo di due concetti fondamentali situati nel SiteFrame:

- **StopPlace:** rappresenta un'area di fermata complessa, come una stazione ferroviaria o una fermata principale di un autobus.
- **Quay:** rappresenta punti di accesso più specifici all'interno di uno StopPlace, come un binario specifico o una banchina.

Nell'esempio della Stazione di Trento, la mappatura in NeTEx potrebbe essere strutturata così:

- **StopPlace “Stazione di Trento”**
  - **Quay “Binario 1”**
  - **Quay “Binario 2”**
  - **Quay “Binario 3”**

Questo approccio consente di mantenere una gerarchia logica e di fornire dati più dettagliati senza sacrificare la coerenza dell'informazione.

## 3.4 Miglioramento nella gestione delle tariffe

Come già menzionato nella sezione 2.2.1, il Livello 3 di NeTEx descrive la tariffazione e la gestione economica dei servizi di trasporto tramite l'introduzione del FareFrame. Le possibilità di mapping tra GTFS e NeTEx profilo italiano variano a seconda che sia utilizzato GTFS-Fares V1<sup>18</sup> o GTFS-Fares V2<sup>19</sup>.

Nel caso di GTFS-Fares V1, le tariffe applicate ai servizi di trasporto e i parametri di utilizzo delle tariffe, come restrizioni o validità temporale, sono accessibili attraverso il file *fare\_attributes.txt*. Ulteriori parametri per determinare una tariffa, per esempio se dipendente da specifiche stazioni o zone di origine/destinazione, sono definiti tramite il file *fare\_rules.txt*. Tuttavia, questa versione presenta limitazioni nella rappresentazione di scenari tariffari complessi, rendendo necessaria l'adozione di GTFS-Fares V2 per una maggiore granularità e flessibilità.

---

<sup>18</sup> <https://gtfs.org/documentation/schedule/examples/fares-v1/>

<sup>19</sup> <https://gtfs.org/documentation/schedule/examples/fares-v2/>

GTFS-Fares V2 introduce una struttura più avanzata che consente di rappresentare e gestire informazioni su tariffe articolate, come abbonamenti, tariffe basate su fasce orarie, sconti per categorie specifiche e trasferimenti tra operatori diversi. L'elemento chiave di questa evoluzione è l'introduzione dei *fare products*, che permettono di definire in modo esplicito le tariffe associate alle diverse tipologie di biglietto, e delle *fare leg rules*, che stabiliscono le condizioni di utilizzo per ciascuna tariffa.

L'integrazione dei dati tariffari esistenti può avvenire attraverso due strategie principali: se le tariffe non sono già disponibili, è necessario procedere alla loro raccolta e modellizzazione; se invece le tariffe sono presenti in formato GTFS-Fares V1, è possibile effettuare una migrazione al modello Fares V2, a condizione che siano disponibili i dati richiesti per una rappresentazione più dettagliata. Tale transizione permette di ottenere una migliore qualità dell'output, rendendo più trasparente e strutturata la gestione delle informazioni tariffarie.

Nonostante i miglioramenti introdotti con GTFS-Fares V2, alcune limitazioni permangono. Ad esempio, GTFS non consente di rappresentare informazioni sui canali di distribuzione dei titoli di viaggio, come vendite online o presso rivenditori fisici, né la loro combinazione con particolari condizioni di vendita. Inoltre, non è prevista la possibilità di specificare tariffe relative ai parcheggi associati ai servizi di trasporto, aspetto che potrebbe risultare rilevante per una pianificazione tariffaria completa.

L'adozione del profilo italiano di NeTEx può contribuire a superare queste carenze, fornendo un framework più ricco e dettagliato per la gestione delle tariffe e dei relativi metadati. NeTEx permette di includere informazioni su politiche di sconto, tariffe integrate tra più operatori, e condizioni di vendita specifiche, offrendo un livello di dettaglio superiore rispetto a GTFS. In questo contesto, il *FareFrame* diventa l'elemento centrale per organizzare e strutturare i dati tariffari in un formato standardizzato e interoperabile.

L'evoluzione da GTFS-Fares V1 a V2 e l'eventuale adozione di NeTEx rappresentano quindi passi fondamentali per migliorare la qualità dei dati tariffari, rendendo possibile una gestione più efficace e dettagliata delle tariffe di trasporto pubblico.

## 3.5 Integrazione dati Pathways e Levels

Come accennato nel paragrafo 2.2.1, ulteriori informazioni aggiuntive sulla struttura fisica delle stazioni e sulle connessioni tra le varie aree al loro interno possono essere ottenute in GTFS attraverso l'utilizzo dei file **pathways.txt**<sup>20</sup> e **levels.txt**<sup>21</sup>. Questi file sono infatti particolarmente utili per sistemi complessi come metropolitane, hub multimodali o stazioni ferroviarie con più livelli, scale mobili, ascensori e percorsi pedonali interni.

Nei feed GTFS, queste informazioni sono generalmente assenti. Per colmare questa lacuna, verrà sviluppato un applicativo in grado di elaborare una mappa **OpenStreetMap** (OSM) e un feed GTFS come input. L'applicativo analizzerà i nodi **OSM** insieme alle fermate del feed GTFS. Il risultato sarà un feed GTFS arricchito, comprensivo dei file *pathways.txt* e *levels.txt*.

In particolare, il file *pathways.txt* viene utilizzato per descrivere i collegamenti tra diverse fermate, ingressi o aree all'interno di una stazione. Ogni riga di questo file definisce un percorso specifico tra due punti, specificando il tipo di connessione e i relativi attributi. Questo è particolarmente utile per aiutare i sistemi di pianificazione dei percorsi a fornire istruzioni più dettagliate per la navigazione all'interno di una stazione complessa.

Il file *pathways.txt* contiene i seguenti campi:

- **pathway\_id**: identificativo univoco del percorso;

---

<sup>20</sup> <https://gtfs.org/documentation/schedule/reference/#pathwaystxt>

<sup>21</sup> <https://gtfs.org/documentation/schedule/reference/#levelstxt>

- **from\_stop\_id** e **to\_stop\_id**: identificativi delle fermate o punti della stazione collegati dal percorso;
- **pathway\_mode**: indica il tipo di percorso (per esempio, scala mobile, ascensore, tunnel, percorso pedonale);
- **is\_bidirectional**: specifica se il percorso è utilizzabile in entrambe le direzioni o solo in una;
- **length**: lunghezza del percorso in metri;
- **traversal\_time**: tempo stimato per percorrere il percorso in secondi;
- **stair\_count**: numero di gradini se il percorso include una scala;
- **max\_slope**: pendenza massima del percorso, utile per valutazioni di accessibilità;
- **min\_width**: larghezza minima del percorso in metri;
- **signposted\_as**: nome del percorso come indicato nei cartelli fisici della stazione;
- **reversed\_signposted\_as**: stessa descrizione di signposted\_as ma nel caso il percorso sia bidirezionale e venga usato da to\_stop\_id a from\_stop\_id.

Grazie a queste informazioni, il sistema di routing può offrire indicazioni dettagliate all'interno di una stazione, suggerendo ad esempio il percorso più veloce per passare da un binario all'altro o fornendo alternative accessibili per utenti con mobilità ridotta.

Il file **levels.txt** permette invece di definire la struttura verticale delle strutture, specificando i vari livelli o piani presenti e permette l'uso di ascensori in congiunta all'utilizzo delle Pathways . Questo è particolarmente utile per le stazioni su più piani, come quelle ferroviarie o metropolitane, dove i collegamenti tra livelli sono cruciali per una navigazione efficiente.

Il file include i seguenti campi:

- **level\_id**: identificativo univoco del livello;
- **level\_index**: numero che indica l'altezza relativa del livello rispetto agli altri (per esempio, 0 per il piano terra, 1 per il primo piano, -1 per un piano sotterraneo);

- **level\_name**: nome descrittivo del livello, ad esempio "Piano binari" o "Atrio principale".

L'integrazione di **levels.txt** con **pathways.txt** permette di costruire una rappresentazione dettagliata delle connessioni all'interno delle stazioni, rendendo possibile mostrare agli utenti una navigazione dettagliata all'interno di quest'ultime. Un sistema di routing avanzato può così calcolare non solo il tragitto più breve tra due fermate, ma anche indicare il livello in cui si trovano e fornire dettagli sulla presenza di scale mobili o ascensori per l'accessibilità.

## 4. Sviluppo del fork di OpenTripPlanner

OpenTripPlanner (di seguito OTP) è un pianificatore di viaggi multimodale open-source, progettato per facilitare gli spostamenti utilizzando i mezzi di trasporto pubblico e integrandoli con altre modalità di trasporto, come la camminata, il bike sharing e il car sharing, etc. Lanciato nel 2009, il progetto ha sin da subito attirato una attivissima comunità di utenti e sviluppatori, ricevendo il supporto di pubbliche amministrazioni, centri di ricerca ed aziende private.

OTP è progettato per supportare la mobilità sostenibile e l'integrazione di diverse modalità di trasporto all'interno di un unico sistema. Il software è particolarmente utile per enti di trasporto pubblico, aziende di mobilità e servizi di smart city che desiderano fornire agli utenti soluzioni di viaggio efficienti e integrate.

Il componente principale di OTP è sviluppato in Java e si basa su dati geografici e di trasporto pubblico provenienti da fonti aperte, principalmente OpenStreetMap (OSM) per le mappe e GTFS e GTFS-RT per gli orari e le informazioni statiche ed in tempo reale sui mezzi pubblici. Recentemente, il sistema ha introdotto anche il supporto per NeTEx.

Il funzionamento di OTP si articola in diversi passaggi:

1. **Creazione del grafo di transito:** OTP costruisce un grafo dai dati della mappa OSM e dati riguardanti il trasporto pubblico nel formato GTFS;
2. **Avvio del server:** OTP copre il ruolo di server rimanendo in ascolto di richieste da parte degli utenti;
3. **Elaborazione delle richieste degli utenti:** l'utente interroga il server OTP e specifica il punto di partenza, la destinazione e le preferenze di viaggio. OTP calcola il percorso più efficiente combinando mezzi pubblici, percorsi pedonali, ciclabili e automobilistici;

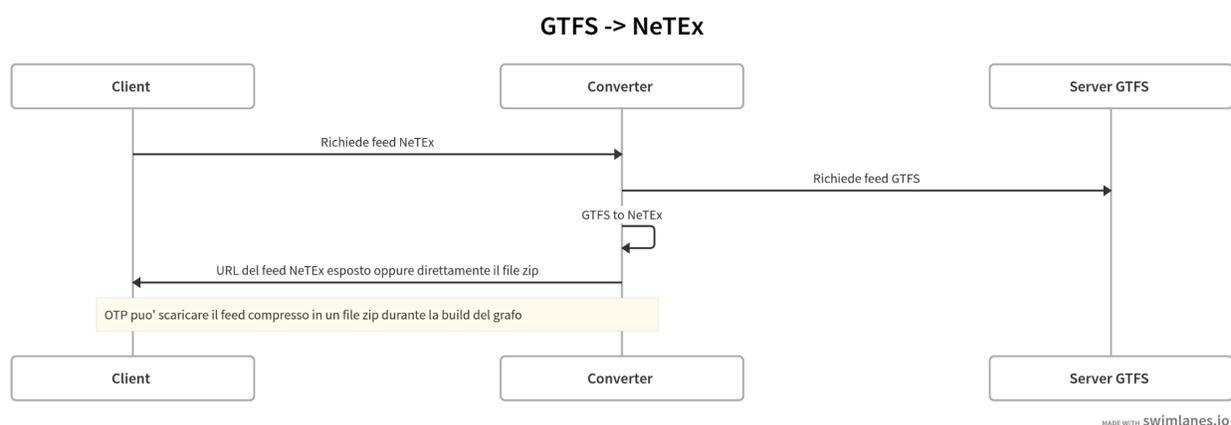
4. **Presentazione dei risultati:** il sistema restituisce una lista di opzioni di viaggio con indicazioni dettagliate, tempi di percorrenza e intermodalità disponibili.

Grazie a queste caratteristiche, OTP è utilizzato in numerosi progetti di mobilità urbana, rendendo più accessibili ed efficienti i trasporti pubblici.

Il fork di OTP nasce dall'esigenza di migliorare l'integrazione del software con gli standard e le specificità del sistema di trasporto pubblico italiano. In particolare, l'obiettivo principale dello sviluppo è ottimizzare il supporto ai profili italiani dei formati NeTEx e SIRI, garantendo una maggiore compatibilità con i dati forniti dalle aziende di trasporto italiane e migliorando l'accuratezza delle informazioni fornite agli utenti sugli itinerari di viaggio.

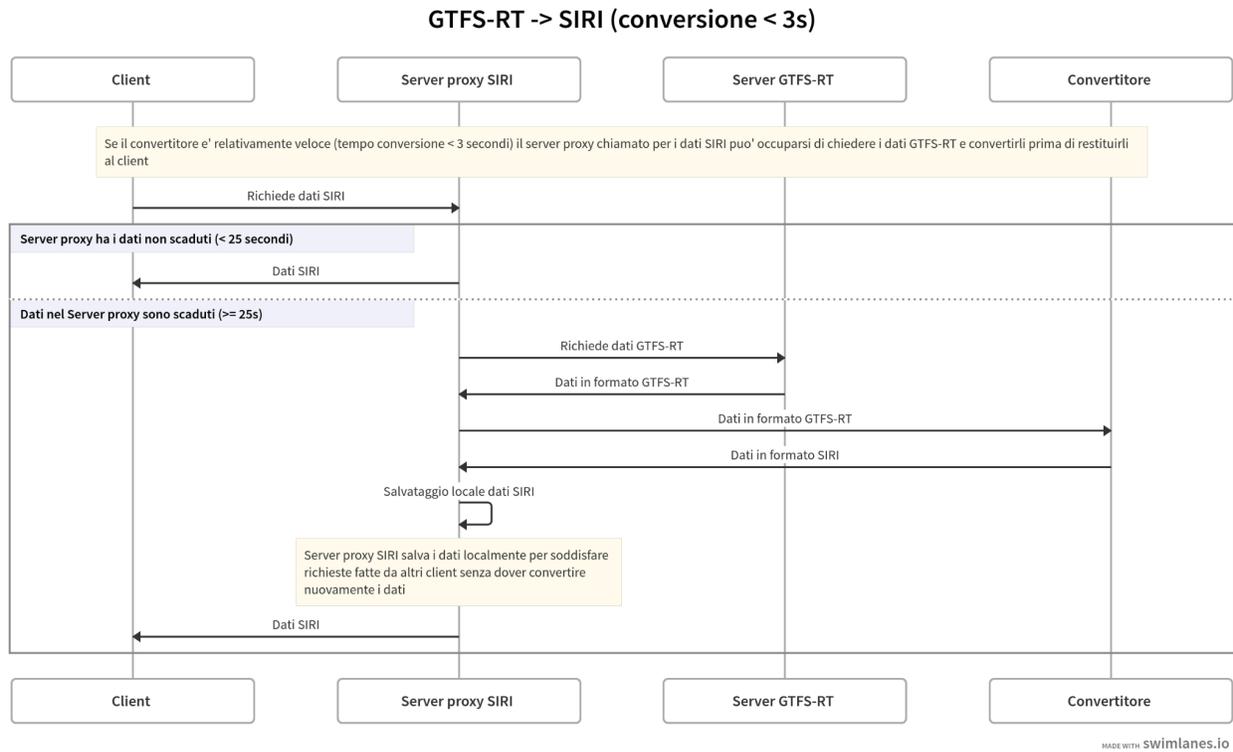
Per integrare i dati convertiti in formato NeTEx e SIRI all'interno di OTP, verranno adottate due modalità distinte di gestione dei flussi informativi, una per NeTEx e una per SIRI.

La figura 2 illustra il caso di NeTEx, in cui il processo di integrazione con OTP prevede l'utilizzo dei dati convertiti dal converter nel formato NeTEx conforme al profilo italiano. Questi dati vengono poi impiegati per costruire il grafo di transito all'interno di OTP, rappresentando così il primo passo fondamentale nel funzionamento del sistema.



**Figura 2:** Gestione flusso dati NeTEx

La figura 3, invece, descrive il caso di SIRI, in cui il metodo di integrazione segue un approccio differente. In questo scenario, OTP può accedere direttamente agli URL esposti dal converter, che forniscono i dati in formato SIRI. Durante l'elaborazione della richiesta, il converter assume il compito di trasformare le informazioni dal formato GTFS-RT a SIRI profilo italiano.



**Figura 3:** Gestione flusso dati SIRI

Vale comunque la pena sottolineare come il fork di OTP che verrà sviluppato sarà comunque capace di gestire qualunque flusso di dati in formato NeTEx e SIRI profilo italiano, senza la necessità che questi debbano necessariamente essere prima sottoposti alla procedura di conversione.

## 5. Conclusioni e prossimi passi del progetto

Questo documento di analisi tecnica preliminare ha descritto le attività svolte nelle prime 12 settimane del progetto **OTP-Italy**, focalizzandosi su quattro aree principali: il mapping tra i diversi formati dati e la definizione dell'approccio più adatto per la realizzazione dei converter; l'impostazione delle strategie di conversione per NeTEx e SIRI; il miglioramento e l'ottimizzazione dei dati GTFS e GTFS-RT e l'avvio dello sviluppo del fork di *OpenTripPlanner* per l'adattamento al contesto italiano.

I prossimi passi del progetto prevedono l'avvio delle attività di sviluppo vero e proprio, con la realizzazione di una prima versione alpha dei converter, il progressivo miglioramento dei processi di integrazione dei dati, incluse nuove fonti da importare nel formato GTFS ed ulteriori attività relative allo sviluppo del fork di OTP, con l'obiettivo di ottenere una versione del motore di calcolo pienamente compatibile con i profili dati italiani e in grado di supportare casi d'uso reali.

## 6. Bibliografia

- [1] Vetere, G., Lenzerini, M.: *Models for semantic interoperability in service-oriented architectures*. IBM Systems Journal 44(4), 887–903 (2005). <https://doi.org/10.1147/sj.444.0887>
- [2] Scrocca, M., Comerio, M., Carenini, A., Celino, I.: *Turning transport data to comply with EU standards while enabling a multimodal transport knowledge graph*. In: Proceedings of the 19th International Semantic Web Conference. vol. 12507, p.411–429. Springer (2020). [https://doi.org/10.1007/978-3-030-62466-8\\_26](https://doi.org/10.1007/978-3-030-62466-8_26)
- [3] Scrocca, M., Grassi, M., Comerio, M., Carriero, V., Delgado Dias, T., Vieira Da Silva, A., Celino, I.: *Intelligent Urban Traffic Management via Semantic Interoperability across Multiple Heterogeneous Mobility Data Sources*. In Proceedings of the 23rd Intern. Semantic Web Conference, (2024). Pre-print version available at: <http://arxiv.org/abs/2407.10539>
- [4] Van Assche, D., Delva, T., Haesendonck, G., Heyvaert, P., De Meester, B., Dimou, A.: *Declarative RDF graph generation from heterogeneous (semi-)structured data: A systematic literature review*. Journal of Web Semantics (2022), publisher: Elsevier
- [5] Grassi, M., Scrocca, M., Carenini, A., Comerio, M., Celino, I.: *Composable Semantic Data Transformation Pipelines with Chimera*. In: Proceedings of the 4th, International Workshop on Knowledge Graph Construction. CEUR Workshop Proceedings (2023).
- [6] Scrocca, M., Carenini, A., Grassi, M., Comerio, M., Celino, I.: *Not Everybody Speaks RDF: Knowledge Conversion between Different Data Representations*. In: Proceedings of the 5th, International Workshop on Knowledge Graph Construction. CEUR Workshop Proceedings (2024).