



# *Sii-Mobility*

## **Supporto di Interoperabilità Integrato per i Servizi al Cittadino e alla Pubblica Amministrazione**

**Trasporti e Mobilità Terrestre, SCN\_00112**

**Deliverable ID: DE4.4a**

**Titolo: Servizi per gestione interfacce di acquisizione  
dati**

<b>Data corrente</b>	M16, Aprile 2017
<b>Versione (solo il responsabile può cambiare versione)</b>	0.4
<b>Stato (draft, final)</b>	final
<b>Livello di accesso (solo consorzio, pubblico)</b>	Pubblico
<b>WP</b>	4
<b>Natura (report, report e software, report e HW..)</b>	Software e processi
<b>Data di consegna attesa</b>	M16, Aprile 2017
<b>Data di consegna effettiva</b>	M16, Aprile 2017
<b>Referente primario, coordinatore del documento</b>	Laura Cocone laura.cocone@swarco.com
<b>Contributor</b>	Roberto Di Vincenzo roberto.divincenzo@swarco.com, Silvia Capato silvia.capato@swarco.com
<b>Coordinatore responsabile del progetto</b>	Paolo Nesi, UNIFI, paolo.nesi@unifi.it

# Sommario

1	Executive Summary .....	3
2	Introduzione .....	4
2.1	Il Progetto Sii-Mobility .....	4
2.2	Scopo del documento .....	6
2.3	Struttura del documento .....	6
3	Acquisizione dati.....	7
4	Integrazione con sistemi degli enti gestori, TPL.....	8
4.1	Fonte e tipologia dati .....	8
4.2	Modalità di integrazione.....	9
5	Integrazione con enti gestori traffico .....	11
5.1	Fonte e tipologia dati .....	11
5.2	Modalità di integrazione.....	12
6	Integrazione con Open Data, OD comuni via ETL.....	14
6.1	Fonte e tipologia dati .....	14
6.2	Modalità di integrazione.....	14
7	Integrazione con LOD.....	14
7.1	Fonte e tipologia dati .....	14
7.2	Modalità di integrazione.....	14
8	Integrazione con centrali meteo e altre informazioni ambientali.....	14
8.1	Fonte e tipologia dati .....	14
8.2	Modalità di integrazione.....	15
9	Acronimi .....	15

# **1 Executive Summary**

Il seguente documento costituisce il deliverable DE4.4 “Servizi per gestione interfacce e acquisizione dati”.

Lo scopo del documento è quello di descrivere le interfacce usate per l’integrazione dei dati provenienti da:

1. Sistemi degli enti gestori, TPL, AVM, etc.;
2. Open Data, OD dei vari comuni via ETL;
3. LOD vari per arricchimento dati per big data reasoning;
4. Sensori e reti di sensori;
5. Centrali per il meteo e altre informazioni ambientali;
6. Centrale di autostrade e sensori;
7. Centrale di ferrovie;
8. Kit di privati diversi da quelli Sii-Mobility (modello generico);
9. Centrale del 118 ed emergenze.

## 2 Introduzione

### 2.1 Il Progetto Sii-Mobility

**Sii-Mobility** (Supporto all'interoperabilità integrata per i servizi ai cittadini e alla pubblica amministrazione) è un progetto dello Smart City Nazionale in area Trasporti e Mobilità Terrestre, cofinanziato dal MIUR (Ministero dell'Istruzione dell'Università e della Ricerca).

Sii-Mobility promuove nell'ambito della mobilità urbana su gomma e/o su rotaia, lo sviluppo di nuove tecnologie e soluzioni ICT innovative finalizzate a migliorare l'interoperabilità dei sistemi di gestione informativi e di infomobilità urbana e metropolitana, in attuazione delle disposizioni della normativa comunitaria vigente in materia, in sinergia con quelle regionali e con le azioni nazionali ed internazionali.

I sistemi di trasporto per la mobilità terrestre si trovano sempre più spesso in stati di congestione dovuti a condizioni difficilmente prevedibili. Il sistema di trasporto locale presenta alti costi sociali legati al disagio dei cittadini per quanto riguarda le soluzioni di mobilità a causa della scarsa interoperabilità e intelligenza tra:

- 1) gestione e monitoraggio di sistemi di trasporto;
- 2) servizi per la mobilità;
- 3) servizi e sistemi per il trasporto di merci;
- 4) ordinanze e servizi pubblici (quali ospedale, centri, musei, ...);
- 5) trasporto privato, trasporto ferroviario, parcheggi e le persone che si muovono, a causa della limitata capacità del sistema di incorporare e reagire ai cambiamenti della città e dei cittadini.



**Figura 1: Concept di Sii-Mobility**

Per ridurre i costi sociali si devono ottimizzare i servizi per esempio sfruttando informazioni integrate, una maggiore interoperabilità e integrazione fra sistemi ed instaurando un colloquio con le comunità e i singoli cittadini tramite sistemi di informazione e raccolta umori ed opinioni dei cittadini (vedere Figura 1: Concept di Sii-Mobility). Le soluzioni a tali problemi non sono semplici perché il sistema della città e/o delle metropoli è un ecosistema sociale complesso.

Sii-Mobility propone:

- Soluzioni di guida connessa (supporto alla guida intelligente con e senza mezzi): servizi personalizzati, controllo del traffico, invio di messaggi, comandi e informazioni personalizzate;
- Piattaforma partecipata e di informazione: per fornire informazioni al cittadino, il cittadino come un sensore intelligente, informare ed educare i cittadini attraverso totem, applicazioni mobili, web, ecc.;
- Gestione dei criteri di accesso personalizzati: utilizzo di politiche di incentivo e dissuasione del veicolo, mobilità dei crediti, flussi di monitoraggio;
- Interoperabilità e integrazione dei sistemi di gestione: contributo alla validazione di test e dati, riconciliazione, norme, ecc.;
- Integrazione di metodi di pagamento e di identificazione: politiche di pay-per-use, monitoraggio del comportamento dell'utente;
- Gestione dinamica dei confini ZTL: cambiamenti nei confini di aree riservate (inquinamento, congestione, ...), dinamiche dei prezzi per categoria di veicoli;
- Gestione di una rete condivisa per lo scambio di dati tra operatori: valutazione dell'affidabilità dei dati, separazione delle responsabilità, interfacciamento, dati aperti;
- Monitoraggio in tempo reale della domanda e dell'offerta di trasporto pubblico: soluzioni di integrazione e di elaborazione dati.

Nell'architettura del progetto **Sii-Mobility** si possono notare le interfacce per la connessione con altri sistemi di Smart city, con il sistema di mobilità nazionale, la rilevazione dati ambientali, le ordinanze, etc.

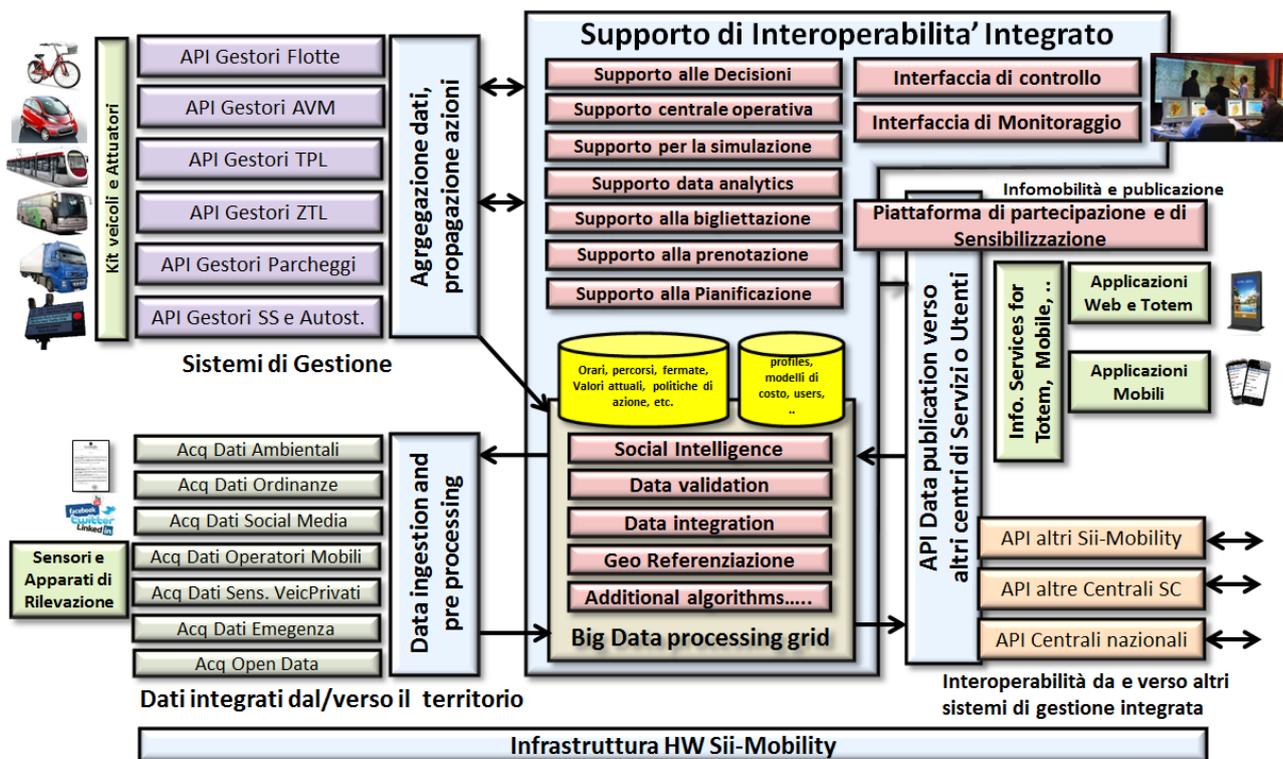


Figura 2: L'architettura di Sii-Mobility, con i dettagli interni di del SII, Supporto di Interoperabilità Integrato

## 2.2 Scopo del documento

Il seguente documento costituisce il deliverable DE4.4a “Servizi per gestione interfacce e acquisizione dati”.

## 2.3 Struttura del documento

Il presente documento è organizzato in capitoli e paragrafi, la cui struttura è descritta qui di seguito:

- **Capitolo 1 – Executive Summary**; costituisce una linea guida per il lettore con lo scopo di presentare il punto di partenza del lavoro e gli obiettivi prefissati.
- **Capitolo 2 – Introduzione**; questo capitolo è strutturato in tre sotto paragrafi con l’intento di orientare velocemente il lettore all’interno del documento; include una breve descrizione del progetto Sii-Mobility (paragrafo 2.1 – Il Progetto-Sii-Mobility), descrive lo scopo del documento (par. 2.2 – Scopo del documento) e si conclude illustrando la struttura del documento (par. 2.3 - Struttura del documento).
- **Capitolo 3 - Acquisizione dati**; presenta brevemente il processo ETL, usato per l’acquisizione dei dati dei diversi sistemi collegati alla piattaforma Sii-Mobility.
- **Dal Capitolo 4 al Capitolo 13** sono presentati più nel dettaglio i processi di integrazione per i sistemi integrati nella piattaforma Sii-Mobility, questi sono:
  - Sistemi degli enti gestori, TPL;
  - Enti gestori di traffico;
  - Open Data, OD dei vari comuni via ETL;
  - LOD;
  - Sensori e reti di sensori;
  - Centrali meteo e altre informazioni ambientali;
  - Centrale di autostrade e sensori;
  - Centrale di ferrovie;
  - Kit di privati diversi da quelli Sii-Mobility (modello generico);
  - Centrale del 118 ed emergenze.

Questi capitoli presentano la stessa struttura, descrivendo prima la fonte e tipologia di dati acquisita e poi dettagliando la modalità di integrazione.

- Infine, **Capitolo 14 – Acronimi**, dove viene riportato il significato degli acronimi usati nel documento.

### 3 Acquisizione dati

L'acquisizione dei dati presenti sui differenti sistemi informativi collegati alla piattaforma di Sii-Mobility viene effettuata attraverso dei processi ETL (Extraction, Transformation and Loading – estrazione, trasformazione e caricamento).

In base alla fonte informativa utilizzata, in particolare alla forma in cui è disponibile il dato ed alle modalità di connessione per la sua fruizione, i processi ETL di acquisizione sono realizzati attraverso differenti modalità/tecnologie.

I processi ETL realizzati effettuano l'acquisizione di nuovi OpenData da diverse sorgenti e li trasformano in Triple RDF, per poterli integrare all'interno di un sistema già esistente. A tal proposito i processi ETL si basano sulle seguenti tre specifiche fasi:

- Fase 1 -> Data Ingestion
- Fase 2 -> Quality Improvement
- Fase 3 -> Triplification

La fase di **Data Ingestion** prevede l'acquisizione dei file sorgente dai diversi portali/sistemi integrati con la piattaforma Sii-Mobility, la loro memorizzazione in una specifica cartella e il caricamento sulle specifiche tabelle di appoggio (solitamente su database HBase).

La fase di **Quality Improvement** ha come obiettivo quello di migliorare la qualità dei dati acquisiti nella fase di ingestion, operando delle opportune modifiche sullo specifico dato in esame e memorizzando il contenuto in una nuova tabella HBase. Ad esempio, un'attività di Quality Improvement che può essere effettuata è quella di normalizzazione degli indirizzi.

La fase di **Triplification** ha l'obiettivo di generare un insieme di triple RDF partendo dai dati acquisiti e migliorati nelle due fasi precedenti e "mappandoli" su un modello costruito sulla base delle relazioni definite all'interno di una specifica ontologia di riferimento.

## 4 Integrazione con sistemi degli enti gestori, TPL

I sistemi di gestione delle flotte del trasporto pubblico locale (TPL), saranno integrati con la piattaforma Sii-Mobility al fine di inviare le informazioni relative alle previsioni di arrivo in fermata dei mezzi pubblici.

I sistemi di gestione delle flotte che saranno oggetto dell'integrazione sono quelli relativi ai 5 bacini identificati, ovvero Arezzo, Pisa, Pistoia, Prato, Siena.

### 4.1 Fonte e tipologia dati

I messaggi inviati dai sistemi di gestione delle flotte conterranno informazioni relative alle previsioni sui tempi di arrivo in fermata (Bus Stop Forecasts). Tali informazioni saranno inviate, ovviamente, in anticipo rispetto all'orario reale di arrivo dell'autobus.

I messaggi saranno inviati in formato JSON; di seguito un esempio di messaggio inviato dai sistemi di gestione flotte, con la descrizione del modello dati.

```
{
  "BusStops": [
    {
      "AreaId": "1",
      "timestamp": "2017-07-06 14:57:07",
      "BusStopId": 1,
      "Forecasts": [
        {
          "TheoreticalPassageTime": "2017-07-06 15:28:00",
          "TripCode": "001536",
          "LineLabel": "1",
          "LineDescription": "CECINA - DONORATICO - (VENTURINA) - PIOMBINO",
          "RealPassageTime": "2017-07-06 15:40:31",
          "IsDeleted": "0",
          "Destination": "001 - CECINA"
        }
      ]
    },
    {
      "AreaId": "1",
      "timestamp": "2017-07-06 14:57:07",
      "BusStopId": 3,
      "Forecasts": [
        {
          "TheoreticalPassageTime": "2017-07-06 15:28:00",
          "TripCode": "001536",
          "LineLabel": "1",
          "LineDescription": "CECINA - DONORATICO - (VENTURINA) - PIOMBINO",
          "RealPassageTime": "2017-07-06 15:39:58",
          "IsDeleted": "0",
          "Destination": "001 - CECINA"
        }
      ]
    }
  ]
}
```

**Tabella 1 - Modello dei dati - Bus Stop**

Campo	Tipo	Descrizione	Chiave
AreaId	intero	Identificativo univoco dell'area (bacino)	Sì
BusStopId	stringa	Identificativo della fermata	Sì
timestamp	data/ora	Data di aggiornamento del dato	Sì

**Tabella 2 - Modello dei dati - Bus Stop Forecasts**

Campo	Tipo	Descrizione	Chiave
TheoreticalPassageTime	data/ora	Tempo di passaggio alla fermata teorico. Per le fermate normali, coincide con il tempo di arrivo. Per le fermate di tipo Terminal, coincide con l'orario di partenza	Sì
TripCode	stringa	Identificativo della corsa	
LineLabel	stringa	Identificativo della linea	
LineDescription	stringa	Descrizione della linea	
RealPassageTime	data/ora	Previsione di arrivo in fermata calcolata. Per le fermate normali, coincide con il tempo di arrivo. Per le fermate di tipo Terminal, coincide con l'orario di partenza	
IsDeleted	booleano	0: La previsione può essere pubblicata 1: La previsione deve essere eliminata	
Destination	stringa	Destinazione visualizzata dal veicolo mentre sta raggiungendo la fermata	

## 4.2 Modalità di integrazione

L'integrazione tra i sistemi di gestione delle flotte e la piattaforma Sii-Mobility è realizzato attraverso l'utilizzo di un servizio web di tipo REST/JSON, esposto su rete pubblica da parte di UniFi.

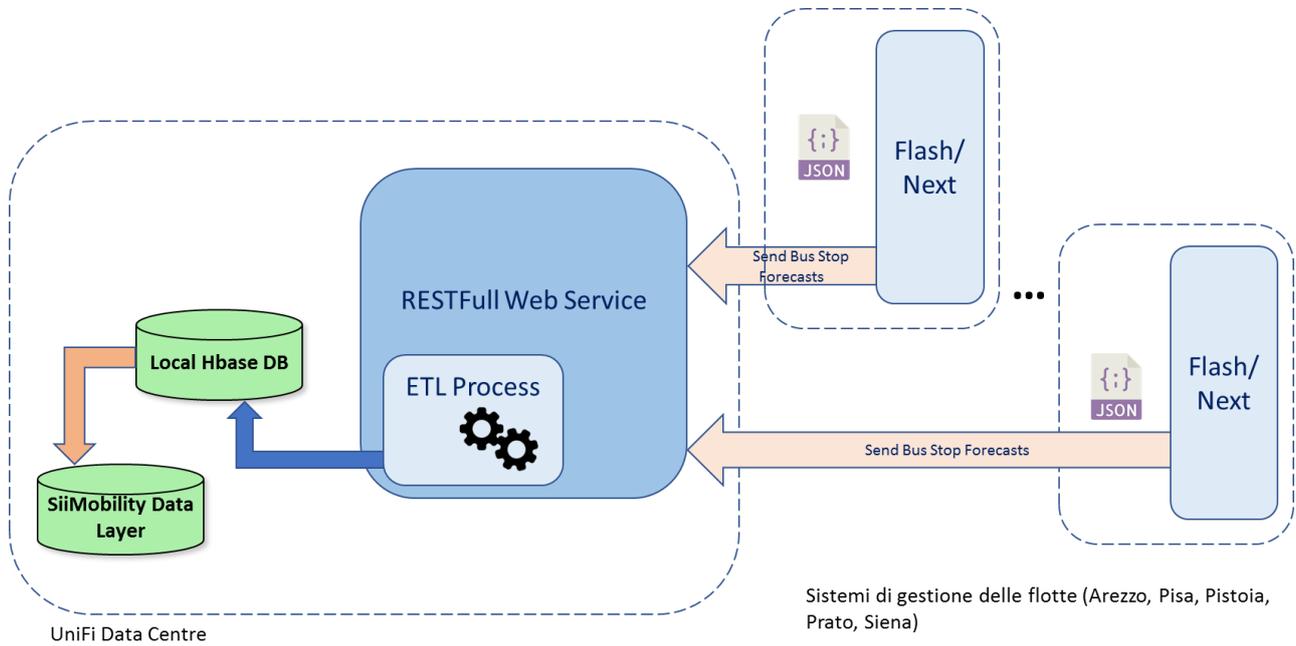
I sistemi di gestione delle flotte, attraverso un modulo software specificatamente realizzato, invocano, ad intervalli di tempo regolari, il servizio REST esposto da UniFi per inviare i dati aggiornati di previsione di arrivo in fermata. La cadenza di invio dei dati da parte dei sistemi di gestione flotte è configurabile; al fine di fornire una visione aggiornata dei dati di previsione, ma senza sovraccaricare i sistemi, la cadenza di invio sarà configurata con valori compresi tra 1 e 5 minuti.

I dati vengono inviati in formato JSON, secondo il modello definito al paragrafo precedente, attraverso una chiamata HTTP POST, così come definito solitamente per i servizi di tipo REST.

Il servizio REST esposto permette di acquisire i dati da tutti e 5 i bacini individuati; le informazioni ricevute saranno memorizzate su un database locale HBase prima di essere acquisite dalla

piattaforma Sii-Mobility. Il database HBase locale memorizza, per ogni previsione di arrivo in fermata, l'informazione più aggiornata ricevuta dai sistemi di gestione delle flotte.

La Figura 3 mostra l'architettura relativa all'integrazione sopra descritta.



**Figura 3: Architettura integrazione con sistemi degli enti gestori**

## 5 Integrazione con enti gestori traffico

I sistemi di gestione del traffico, in particolare il sistema di Traffic Management della Città Metropolitana di Firenze, sarà integrato con la piattaforma Sii-Mobility al fine di inviare le informazioni relative agli eventi di traffico ed alle strategie eventualmente adottate a fronte di eventi definiti.

### 5.1 Fonte e tipologia dati

Il sistema di Traffic Management della Città Metropolitana di Firenze espone, su due pagine web distinte, i seguenti dati:

- Eventi di traffico correntemente attivi;
- Strategie di traffico eventualmente attive.

Le pagine web sono rese accessibili ai soli sistemi della rete dell'Università di Firenze.

I dati sono esposti in formato XML; le relative pagine hanno un periodo di aggiornamento di 5 minuti (configurabile).

Di seguito un esempio di pubblicazione di eventi di traffico in formato XML, e la relativa descrizione del modello dati.

```
<?xml version="1.0" encoding="UTF-8"?>
<EventList
xmlns="http://schemas.datacontract.org/2004/07/EventsAndStrategiesMonitor.Models"
xmlns:i="http://www.w3.org/2001/XMLSchema-instance">
  <Events>
    <Event>
      <arcs>
        <Arc>
          <arcId>2</arcId>
        </Arc>
      </arcs>
      <eventCode>EVT:20170612150745858</eventCode>
      <eventInstanceId>28108</eventInstanceId>
      <lat>43.7586173304</lat>
      <lon>11.2064766337</lon>
      <notes i:nil="true" />
      <severity>5</severity>
      <source>Da operatore</source>
      <startTime>2017-06-12T15:07:00</startTime>
      <stopTime i:nil="true" />
      <subTypeId>100</subTypeId>
      <typeId>2</typeId>
    </Event>
    <Event>
      <arcs>
        <Arc>
          <arcId>3</arcId>
        </Arc>
      </arcs>
      <eventCode>EVT:20170612150829051</eventCode>
      <eventInstanceId>28109</eventInstanceId>
      <lat>43.7284625001</lat>
```

```

<lon>11.021979002</lon>
<notes i:nil="true" />
<severity>5</severity>
<source>Da operatore</source>
<startTime>2017-06-12T15:08:00</startTime>
<stopTime i:nil="true" />
<subTypeId>495</subTypeId>
<typeId>1</typeId>
</Event>
</Events>
</EventList>

```

**Tabella 3 - Modello dei dati – Eventi di traffico**

Campo	Tipo	Descrizione	Chiave
eventInstanceId	intero	Identificativo univoco dell'evento	Sì
eventCode	stringa	Codice univoco dell'evento	No
typeId	stringa	Identificativo del tipo di evento	No
subTypeId	stringa	Sotto identificativo del tipo di evento	No
Source	stringa	Sorgente che ha generato l'evento (es. operatore)	No
severity	intero	Gravità dell'evento	No
startTime	data/ora	Data di inizio dell'evento	No
stopTime	data/ora	Data di fine dell'evento	No
Arcs	lista	Insieme degli archi della rete coinvolti dall'evento	No
Lat	decimale	Coordinate geografiche identificative dell'evento (latitudine)	No
Lon	decimale	Coordinate geografiche identificative dell'evento (longitudine)	No
Notes	stringa	Campo note libero	No

## 5.2 Modalità di integrazione

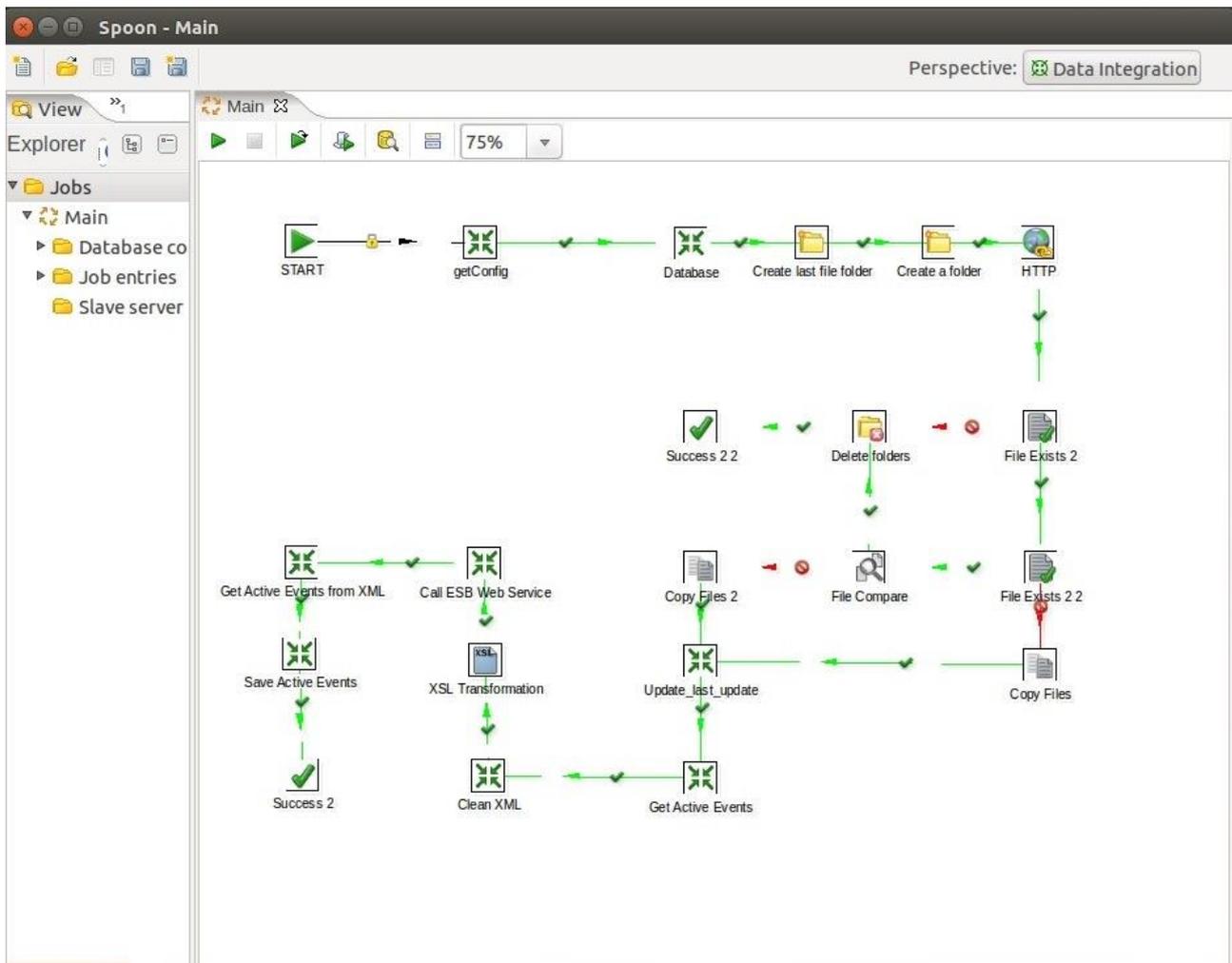
L'integrazione viene realizzata attraverso la piattaforma di integrazione messa a disposizione dall'Università di Firenze, basata sul tool Pentaho Data Integration (Kettle).

Pentaho Kettle è un progetto open source che comprende un potente tool ETL, il quale facilita la gestione dei processi di estrazione, trasformazione e caricamento in un sistema informatico.

I processi ETL sono realizzati mediante l'utilizzo di un IDE (Integrated Development Environment) che permette la creazione di un file XML descrittore, attraverso un editor grafico.

Il file XML descrittore viene poi eseguito, ad intervalli regolari secondo quanto configurato, da un applicazione Java che rappresenta il motore di esecuzione della piattaforma Pentaho; il motore di esecuzione esegue le istruzioni contenute nello script XML seguendo il flusso logico definito.

La figura seguente mostra, in modalità grafica, lo script realizzato per l'acquisizione degli eventi di traffico dal sistema di Traffic Management della Città Metropolitana di Firenze.



Ogni azione definita nello script (che prende il nome di step nell'ambito di Pentaho), si occupa di effettuare una determinata azione, semplice o complessa (ovvero costituita a sua volta da più step).

I primi step dello script si occupano di recuperare la configurazione dello stesso, all'interno della quale sono presenti i riferimenti all'indirizzo web delle sorgenti dati.

La fase successiva del processo ETL si occupa di accedere alle sorgenti dati (le pagine XML) per recuperarne il contenuto informativo. Attraverso alcuni step, il processo estrae dal file xml gli eventi di traffico non ancora processati precedentemente; gli eventi di traffico già elaborati in esecuzioni precedenti vengono scartati.

Segue dunque la fase di trasformazione, nella quale i dati vengono manipolati in modo da risultare compatibili per il sistema di destinazione. I dati così elaborati vengono memorizzati su un database HBase locale.

La fase di triplification viene effettuata, a partire dal database locale HBase, da UniFi che provvede, sulla base dell'ontologia definita, a salvare le relative triple RDF sulla piattaforma Sii-Mobility.

La figura seguente mostra l'architettura relativa all'integrazione sopra descritta.

## Traffic ETL Process

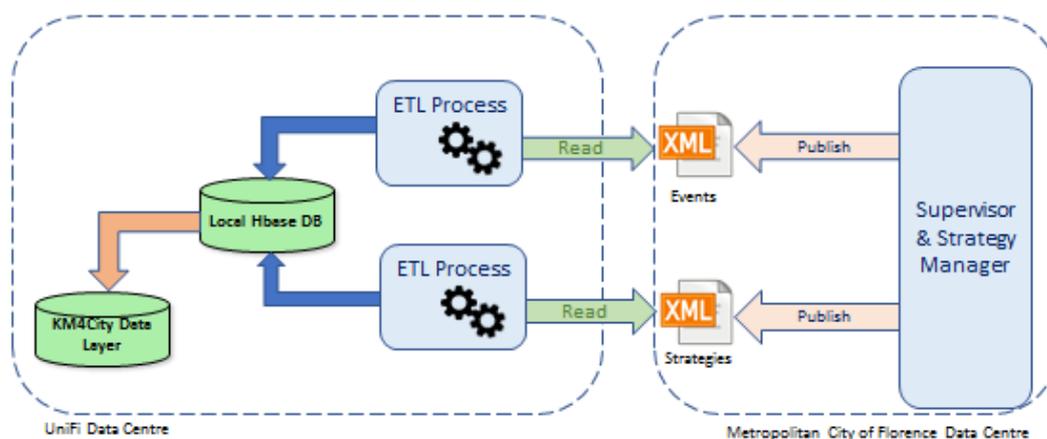


Figura 4: Architettura integrazione con enti gestori di traffico

## 6 Integrazione con Open Data, OD comuni via ETL

Si vedano i DE 4.2 e 4.3

### 6.1 Fonte e tipologia dati

Si vedano i DE 4.2 e 4.3

### 6.2 Modalità di integrazione

Si vedano i DE 4.2 e 4.3

## 7 Integrazione con LOD

Si vedano i DE 3.18, 4.2 e 4.3

### 7.1 Fonte e tipologia dati

Si vedano i DE 3.18, 4.2 e 4.3

### 7.2 Modalità di integrazione

Si vedano i DE 3.18, 4.2 e 4.3

## 8 Integrazione con centrali meteo e altre informazioni ambientali

Si vedano i DE 3.18, 4.2 e 4.3

### 8.1 Fonte e tipologia dati

Si vedano i DE 3.18, 4.2 e 4.3

## 8.2 Modalità di integrazione

Si vedano i DE 3.18, 4.2 e 4.3

## 9 Acronimi

- API: Application Program Interface
- AVL: Automatic vehicle location
- AVM: Automatic Vehicle Monitoring
- BDaaS: Big Data as a Service
- CAP principle: Consistency Availability Partition Tolerance principle
- CBB: Content Based Billing
- CBB: Content Based Billing
- CEN: European Committee for Standardization
- DBMS: database management system
- FCD: Floating Cellular Data
- GPRS: General packet radio service
- GPS: Global positioning System
- GSM: Global System for Mobile
- ICT: Information and Communication Technologies
- ITS: Intelligent Transport Systems
- LCD: liquid-crystal display
- LOD: linked open data
- MC: Mobile Collector
- MMS: Multimedia Messaging Service
- NLP: Natural Language Processing
- NoSQL: no SQL database
- OD: open data
- OD: Open Data
- OGC: Open Geospatial Consortium
- OWL: Web Ontology Language
- PA: Pubblica Amministrazione
- PMI: Piccola e Media Impresa
- PMS: Private Mobile Systems
- POS: part-of-speech
- RDF: Resource Description Framework
- RFID: Radio Frequency IDentification o Identificazione a radio frequenza
- RTTI: Real-time Travel & Traffic Information
- SDI: Spatial Data Infrastructures
- SII: sistema di interoperabilità integrato
- SIMONE: progetto Simone
- SMS: Short Message Service
- SN: social networking, oppure sensor network
- SOA: Service Oriented Architecture
- SOAP: Simple Object Access Protocol
- SSAMM: Agenzia per la Mobilità Metropolitana strumenti di supporto, TOSCANA
- TPEG: Transport Protocol Experts Group
- TPL: gestore trasporto pubblico locale
- UML: Unified Modeling Language
- UMTS: Universal Mobile Telecommunications System
- UTC: Urban Traffic Control

- UUDI: Universal Description Discovery and Integration
- V2I: Vehicle-to-Infrastructure
- V2V: vehicle-to-vehicle
- VMS: Variable Message Sign
- VWSN: Vehicular Wireless Sensor Networks
- W3C: World Wide Web Consortium
- WSD: Word Sense Disambiguation
- WSDL: Web Services Description Language
- WSN: Wireless Sensor Networks
- XMI: XML Metadata Interchange standard di OMG
- XML: Extensible Markup Language
- ZTL: Zona a Traffico Limitato