



SII-Mobility

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

Trasporti e Mobilità Terrestre, SCN_00112

Deliverable ID: DE5.13a

**Titolo: Risultati dei test sull'integrazione con gestori
parcheggi ed altre tipologie**

Data corrente	<u>M18, Giugno 2017</u>
Versione (solo il responsabile può cambiare versione)	<u>3</u>
Stato (draft, final)	<u>Draft</u>
Livello di accesso (solo consorzio, pubblico)	<u>Pubblico</u>
WP	<u>OR5</u>
Natura (report, report e software, report e HW)	<u>Report & Software</u>
Data di consegna attesa	<u>M18, Giugno 2017</u>
Data di consegna effettiva	<u>M18, Giugno 2017</u>
Referente primario, coordinatore del documento	Leonardo Rigutini (QuestIT) rigutini@quest-it.com , Alessandro Gronchi (QuestIT) gronchi@quest-it.com
Contributor	TIME, Midra, Liberologico, UNIFI DISIT
Coordinatore responsabile del progetto	Paolo Nesi, UNIFI, paolo.nesi@unifi.it

Sommario

Sommario	2
1. Introduzione ed obiettivi	3
1.1 Descrizione	3
1.2 Progetto e realizzazione: DE5.11	4
1.3 Pubblicazione e messa in produzione: DE5.13a	6
2. Triplification, deploy e screenshot	7
2.1 ETL di interfacciamento con SienaParcheggi	7
2.1.1 Dati	7
2.1.2 Dati real time: Triplification	8
2.1.3 Deposito sul sistema di hosting pubblico “GitHub”	10
2.1.4 Deploy in produzione	10
3. Conclusioni	11
Acronimi	12

1. Introduzione ed obiettivi

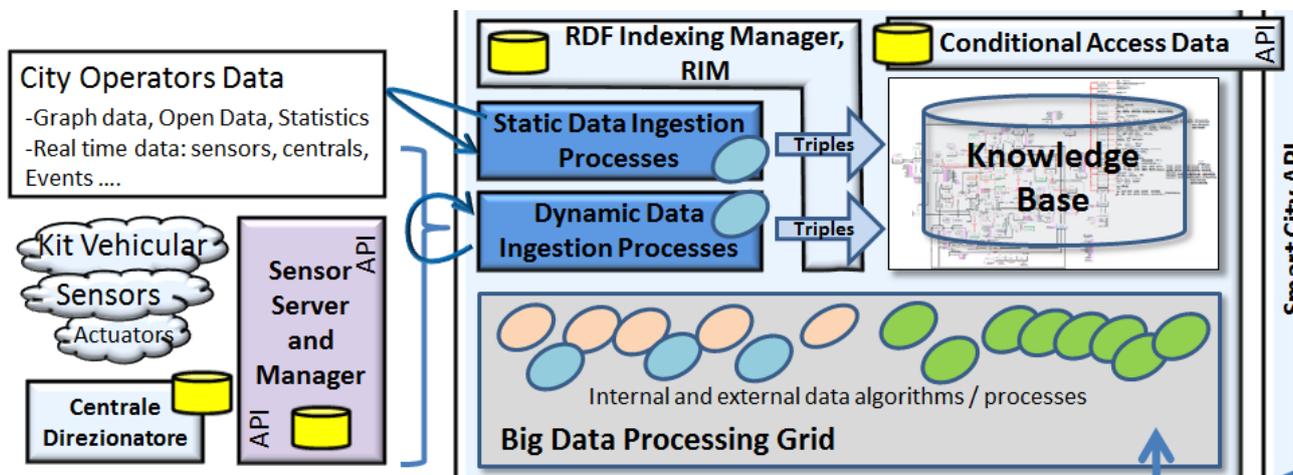
Il progetto SII-Mobility consiste nella realizzazione di un sistema centralizzato di data storage di dati relativi alla infomobilità. In particolare, il progetto prevede tre layer funzionali:

- un sistema centrale di data storage;
- un sistema periferico di data ingestion (acquisizione dati da gestori);
- un sistema periferico di utilizzo dei dati memorizzati nel sistema centrale (app, web, ecc...).

In questo schema, l'**Obiettivo Realizzativo OR5 - Integrazione ed Ottimizzazione dei sottosistemi** riguarda il layer di acquisizione dati da gestori (layer n. 2) ed in particolare l'**Attività 5.4 - Integrazione con parcheggi, ed altri gestori servizi** si concentra sull'acquisizione dei dati dalle seguenti tipologie specifiche di gestori:

- gestori di parcheggi;
- gestori di car and bike sharing;
- gestori taxi e car service;
- gestori nettezza urbana e servizi.

La seguente figura mostra i moduli dell'architettura generale di SII-Mobility interessati dalla attività 5.4 :



1.1 Descrizione

L'interconnessione tra i sistemi periferici (ovvero i gestori dei servizi) ed il sistema centrale è realizzata attraverso appositi software chiamati ETL (Extract, Transform, Load). Tali software costituiscono un layer di astrazione intermedio che permette di uniformare e normalizzare la enorme variabilità dei dati dei gestori. In particolare, gli ETL consistono in appositi software che permettono di accedere ai dati dei differenti gestore ed attuare una serie di operazioni intermedie volte ad ottenere un output standard, compatibile con quello gestito dal sistema centrale:

1. selezione dei dati che sono di interesse per il sistema centrale
2. normalizzazione dei dati (per esempio eliminando i duplicati)

3. traduzione dei dati codificati
4. derivazione di nuovi dati calcolati
5. accoppiamenti (join) tra dati recuperati da differenti repository
6. raggruppamento dei dati

La figura seguente posiziona il livello di ETL all'interno della architettura generale del sistema:



1.2 Progetto e realizzazione: DE5.11

La progettazione e la realizzazione degli ETL previsti nel task 5.4 è riportata nel deliverable DE5.11. In tale deliverable vengono analizzati, al fine di una corretta progettazione, gli aspetti legati al tipo di dato (cosa rappresenta) e alla modalità con cui tale dato deve essere recuperato.

In particolare, nel report, sono dapprima identificate tre macro-categorie di tipologia dei dati:

- dati statici;
- dati dinamici;
- dati statistici.

Tali famiglie hanno, ovviamente, caratteristiche molto diverse tra loro e richiedono quindi differenti modalità di interrogazione. Ad esempio, i dati statici sono dati con tasso di variabilità molto ridotto e, quindi, le richieste al gestore di tali dati possono essere ridotte al minimo; i dati dinamici, invece, rappresentano informazioni “real-time” sullo stato dei servizi e necessitano di un continuo processo di sincronizzazione degli stessi con il database del gestore. I dati statistici, invece, identificano le informazioni aggregate che possono essere derivate dai dati real-time nel recente passato, come ad esempio il tasso di disponibilità del servizio, andamento temporale, ecc... Come tali essi possono essere interrogati dal sistema con tempistiche periodiche ma non con vincoli di allineamento real-time tipici dei dati dinamici.

Successivamente, il DE5.11 analizza ed individua due tipologie distinte per il recupero dati (ingestion):

1. recupero dati manuale - in cui i dati del gestore vengono inseriti nel sistema attraverso una procedura manuale da parte di un operatore, come il caricamento di file in formato CSV/EXCEL o l'utilizzo di form web dedicati.
2. recupero automatizzato - i dati del gestore vengono trasferiti sul sistema centrale tramite interazione diretta con i sistemi proprietari del singolo gestore.

Le due differenti modalità di aggiornamento non risultano ovviamente adatte per tutte le tre tipologie dei dati. Per esempio, pensare ad una procedura manuale per l'ingestion dei dati real-time è impossibile mentre un approccio manuale è fattibile per i dati statici con un rate di variabilità temporale molto basso (dell'ordine di pochissime volte all'anno).

Il report DE5.11 riporta poi la progettazione e la realizzazione degli ETL per tutte le tipologie di

servizio previste nel OR5.4 ed in particolare:

- Task 5.4.1 - parcheggi;
- Task 5.4.2 - servizi di car & bike sharing;
- Task 5.4.3a - taxi e car-service;
- Task 5.4.3b - operatori della logistica (corrieri);
- Task 5.4.5 - nettezza e servizi vari (es. spazzamento strade).

Per ogni tipologia di servizio sono stati forniti contributi dai partner previsti come illustrato nella seguente tabella:

		TIME	liberologico	Midra	QUESTIT
5.4.1	integrazione con parcheggi, ed altri gestori servizi: parcheggi vari		X		X
5.4.2	integrazione con parcheggi, ed altri gestori servizi: car and bike sharing				X
5.4.3.a	integrazione con parcheggi, ed altri gestori servizi: taxi, car service			X	X
5.4.3.b	integrazione con car service e servizi di logistica merci dell'ultimo miglio	X			
5.4.4	integrazione con parcheggi, ed altri gestori servizi: nettezza e servizi			X	X
Tabella 1: Attività 5.4: integrazione con parcheggi, ed altri gestori servizi Tempistica: inizio attività mese: M7, M21 fine attività mese: M17, M28 Coordinatore: QUESTIT Partner: come da tabella					

In particolare, ogni partner si è adoperato per la ricerca ed il coinvolgimento nel progetto di operatori e gestori di servizi collegati alla tipologia di competenza. In molti casi, purtroppo, i gestori contattati non hanno manifestato il loro interesse al progetto e non hanno manifestato l'intenzione di rilasciare la liberatoria all'uso dei dati in loro possesso né il supporto necessario alla loro integrazione (interfacciamento con gli ETL da progettare e realizzare). Il motivo di tale mancato interesse è vario: in alcuni casi la disponibilità del dato è vista come contro-produttore all'azienda stessa in quanto fornisce un indicatore sensibile della copertura del servizio (es TAXI).

In altri casi la mancata partecipazione al progetto è legata ad una particolare situazione interna. Ad esempio, nel caso della società SEI Toscana (gestore del servizio di Nettezza Urbana per il distretto Toscana Sud), pur avendo manifestato telefonicamente un forte interesse al progetto, non è stato possibile arrivare ad un accordo in quanto la società è stata commissariata ad inizio 2017 dal ministero).

La ricerca ed il coinvolgimento di gestori di servizi urbani ha portato alla individuazione di 13 gestori di servizi, dei quali ben 8 sono stati completamente integrati nel sistema SIIMobility.

In particolare:

1. Parcheggi in struttura – Sono stati realizzati gli ETL per i parcheggi in struttura di SistemaGrosseto (4 parcheggi) e Siena Parcheggi (10 parcheggi);

2. Parcheggi lungo strada – Sono stati realizzati gli ETL per i parcheggi lungo strada di Pisa e Firenze.
3. BikeSharing – Sono stati integrati i servizi di bike sharing di Siena (SIPedala) e di Pisa.
4. Taxi – Sono stati contattati i servizi Taxi di Siena e Firenze ma entrambi i gestori non si sono dimostrati interessati alla partecipazione al progetto.
5. Logistica Merci – Per tale categoria sono stati integrati 2 gestori realizzando i relativi ETL.
6. Nettezza Urbana – Sono stati contattati 2 gestori (SEI Toscana e Quadrifoglio) ma in entrambi i casi non è stato possibile integrare le stesse nel sistema. In particolare, SEI Toscana, pur manifestando un forte interesse, si è trovata in una situazione di commissariamento che ha impedito il suo coinvolgimento nel progetto.

1.3 Pubblicazione e messa in produzione: DE5.13a

Il presente deliverable è conforme alle prescrizioni del manuale della qualità del progetto ([SIIM 3]) e riporta le informazioni relative al deploy ed alla prima fase di test degli ETL progettati e sviluppati nel DE5.11. In particolare, il documento riporta la fase di triplicazione, la fase di deposito del codice sul sistema di versioning “github” utilizzato dal consorzio SIIMobility ed alcuni esempi di primo utilizzo tramite screenshot del ServiceMap.

Dato che la fase di triplicazione e deposito del codice su “GitHub” è comune a tutti gli ETL realizzati nel DE5.11, per evitare ridondanze ed un inutile appesantimento del documento, in questo documento, viene riportata la procedura di un singolo ETL.

2. Triplification, deploy e screenshot

L'ETL selezionato è stato quello riguardante l'integrazione dei dati dei parcheggi in struttura gestiti da SienaParcheggi. Come descritto in precedenza, la procedura seguita per tutti gli altri ETL descritti nel DE5.11 è la stessa.

2.1 ETL di interfacciamento con SienaParcheggi

SienaParcheggi è società in-house del comune di Siena che gestisce gli 8 parcheggi in struttura del territorio comunale: Parcheggio La Stazione, Parcheggio Il Campo, Parcheggio Stadio-Fortezza, Parcheggio Eliporto, Parcheggio FastPark, Parcheggio San Francesco, Parcheggio Santa Caterina e Parcheggio Il Duomo.

I sistemi di automazione dei parcheggi sono forniti dalla società Designa, una società esterna da cui SienaParcheggi ha acquistato il servizio e che si occupa anche della gestione dei dati.

I contatti avuti con il direttore della società hanno permesso la integrazione delle informazioni real-time sullo stato di occupazione dei posti auto (capienza totale, posti liberi ed occupati) di tali parcheggi con il sistema SII-Mobility, come descritto nel DE5.11, in modo da poter essere utilizzati da applicazioni di terze parti connesse al sistema SII-Mobility.

2.1.1 Dati

Il relativo ETL progettato e realizzato nel task 5.4 e descritto nel DE5.11 si occupa di recuperare i dati dal database interno del gestore ed renderli disponibili in maniera normalizzata al sistema SIIMobility. Tecnicamente parlando, Siena Parcheggi espone i dati tramite un web-service REST all'indirizzo:

http://95.110.228.237/siena_parcheggi/sp_data.php?s=siimobility&f=json

Il servizio fornisce in output un json contenente le informazioni su tutti i parcheggi in struttura gestiti da SienaParcheggi. Il json restituito include in un'unica risposta sia i dati **statici** sia i dati **dinamici** (real-time) dei parcheggi ed è strutturato come un array associativo di elementi. In particolare, le principali informazioni recuperabili per ogni parcheggio sono:

- id – identificativo
- nome e descrizione
- georeferenziazione
- informazioni sui posti totali e quelli occupati (da cui derivare il numero di posti disponibili)
- stato del parcheggio – un codice numerico che indica se il parcheggio è aperto (codice 80) o meno (es. chiuso, riservato, lavori, ecc...).
- timestamp dell'ultimo aggiornamento
- alcuni campi proprietari relativi all'eventuale impostazione manuale del parcheggio
- un campo che fornisce informazioni sulla visibilità o meno del parcheggio

Dal DE5.11 riportiamo di seguito un'estratto di json relativo ad un singolo parcheggio:

```

1. {
2.   "10": {
3.     "id": "10",
4.     "name": "Il Campo",
5.     "description": "Via di Fontanella n.4",
6.     "lat": "43.3142722222222",
7.     "lng": "11.3335388888889",
8.     "total": "544",
9.     "used": 335,
10.    "status": "80",
11.    "updated": "2017-03-23 10:55:54",
12.    "type": "auto",
13.    "current_mode": "auto",
14.    "free_m": "0",
15.    "used_m": 0,
16.    "total_m": "0",
17.    "status_m": "0",
18.    "visible": "yes"
19.  },
20.  ...
21. }

```

Come descritto nel DE5.11, per ogni gestore sono in verità progettati e realizzati due ETL distinti: uno per i dati statici ed un altro per i dati dinamici. Sebbene questa distinzione produca due ETL diversi, anche in questo caso è possibile restringere la trattazione solamente ad un caso ed estenderla facilmente all'altro. Per tale motivo è possibile concentrarsi sull' ETL realizzato per SienaParcheggi relativo ai dati dinamici.

2.1.2 Dati real time: Triplification

La descrizione dettagliata dello schema dell'ETL è riportata nel DE5.11. In particolare, per le due fasi di recupero sono stati progettati e realizzati due ETL distinti:

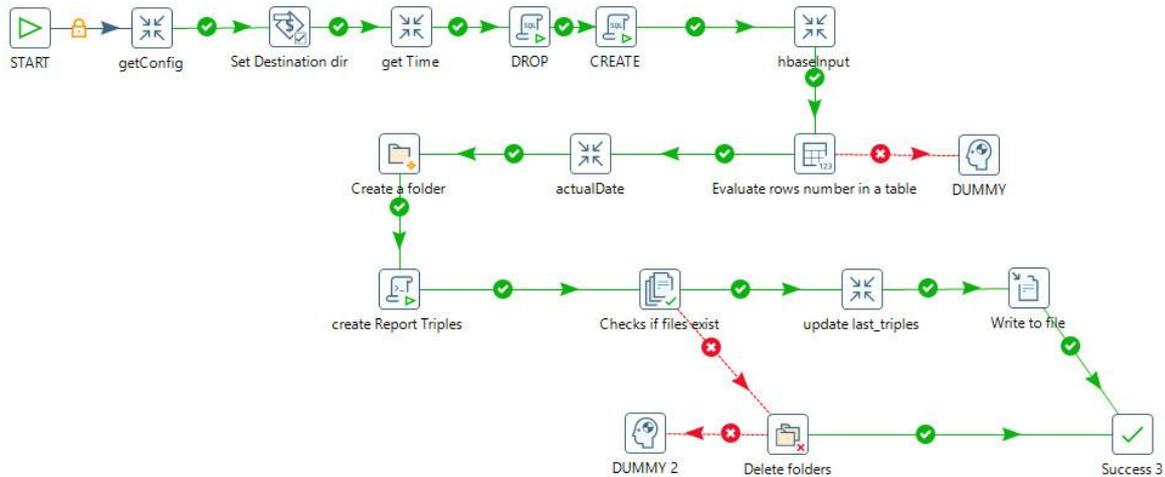
- Ingestion – recupero del dato real time attraverso API del gestore, selezione, processing, normalizzazioni ed infine memorizzazione interna;
- Triplification – trasformazione del dato in una tripla conforme alla ontologia specifica di SIIMobility (Km4City).

La fase di ingestion è descritta in dettaglio nel DE5.11. Di seguito vengono invece riportati dettagli relativi alla fase di triplification.

Il modulo di Triplification è costituito da quattro sotto-moduli principali:

- **getConfig**: lettura dei parametri di configurazione (database MySQL, database HBase, ecc.) da file csv
- **Database**: preparazione di una tabella dedicata del database MySQL per la scrittura dell'output finale
- **HBaseInput**: lettura dei dati real time dal database HBase (tabella dei dati raw) e scrittura degli stessi sulla tabella MySQL
- **create Report Triples**: modulo di generazione delle triple a partire dall'ontologia Sii-Mobility

Si riporta di seguito lo schema di dipendenza tra i sotto-moduli estratto da Spoon:



[Per i dettagli dei singoli moduli ed i parametri di configurazione si rimanda all'analisi del codice del progetto.]

Come step propedeutico all'implementazione del ETL, è stato necessario generare la mappatura del set di dati collezionati nelle fasi precedenti sull'ontologia di riferimento di KM4City (<http://www.disit.org/5606>), utilizzando Karma per la creazione di modelli ad-hoc (uno per ogni ETL).

Lo strumento di integrazione dei dati Karma consente agli utenti di modellare semanticamente i dati in un ambiente visivo. Karma rende più facile capire il modo migliore per mappare i dati sulle ontologie date e per creare un modello basato su relazioni semantiche. La procedura completa utilizzata per la creazione dei modelli è quella indicata nel capitolo 6 del DE4.2.

Nello schema a blocchi del modulo ETL di triplificazione, dopo aver creato la tabella MySQL con i campi conformi al dato di cui è stato effettuato l'ingestion e il QI, si procede alla copia dei dati dalla tabella HBase alla tabella MySQL.

Si occupa della triplificazione vera e propria il blocco "Create Report Triples": questo blocco esegue uno shell script che utilizza il modello Karma per creare le triple RDF. Tali triple RDF sono salvate in file con formato ".n3".

Esempio di shell script per generazione di triple RDF:

```
mvn exec:java -Dexec.mainClass="edu.isi.karma.rdf.OfflineRdfGenerator" -
Dexec.args="--sourcetype DB --modelfilepath
"/media/Triples_RT/Models/Siena_Parcheggi.ttl" --outputfile
${DestinationDir}/${processName}.n3 --dbtype MySQL --hostname 192.168.0.01
--username x --password x --portnumber 3306 --dbname Mob --tablename
${processName}" -Dexec.classpathScope=compile
```

2.1.3 Deposito sul sistema di hosting pubblico "GitHub"

I sorgenti degli ETL prodotti e descritti in dettaglio nel DE5.11 sono stati pubblicati sulla piattaforma di hosting per progetti software GitHub, basata sullo strumento di controllo versione distribuito Git.

GitHub è utilizzato dagli sviluppatori open, che caricano il codice sorgente dei loro programmi e lo rendono scaricabile dagli utenti.

Repo GitHub per gli ETL di Sii-Mobility:

<https://github.com/disit/smart-city-etl>

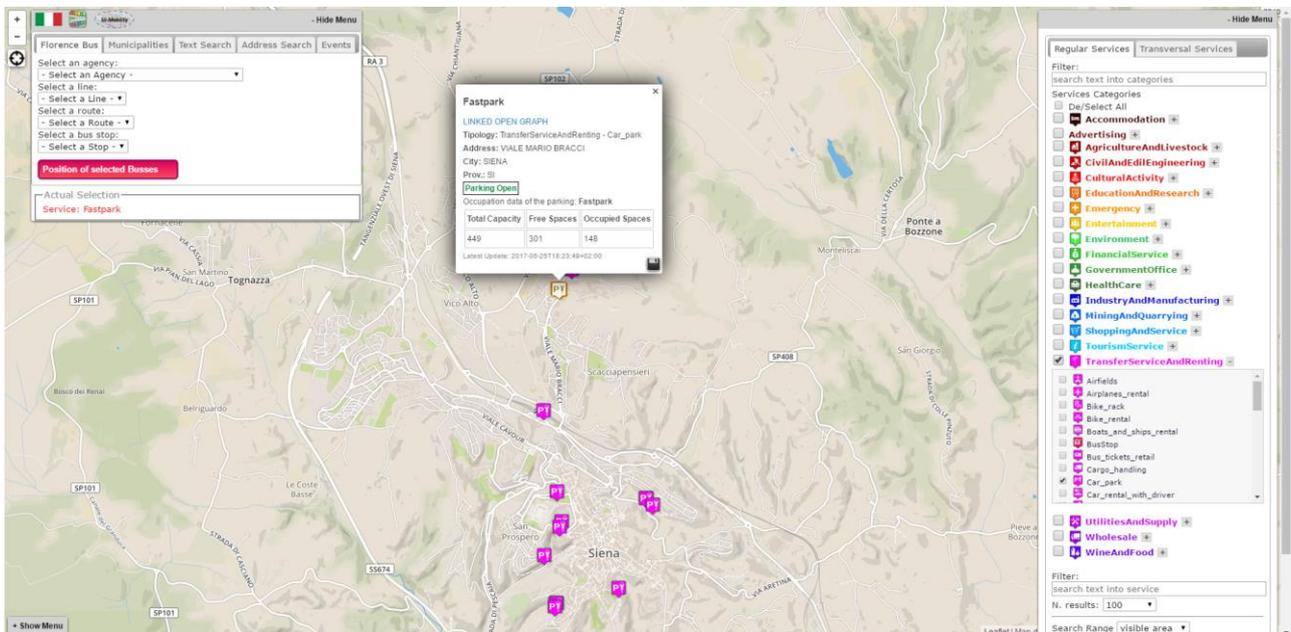
2.1.4 Deploy in produzione

I dati risultanti delle tre fasi (ingestion, QI, triplification) per i gestori di parcheggi in struttura di Siena (SienaParcheggi) sono già disponibili sul ServiceMap di Sii-Mobility:

<http://servicemap.disit.org/WebAppGrafo/mappa.jsp>

E' possibile visualizzare i dati su mappa selezionando dal menu "Regular Services" > "TransferServiceAndRenting" > "Car_park" e centrando la mappa sull'area di interesse di Siena.

Per tutti i parcheggi gestiti è possibile ottenere i dati statici e i dati real-time sull'occupazione dei posti auto, e utilizzare le piene funzionalità di ServiceMap come la generazione di API per ottenere i dati in formato XML/JSON.



3. Conclusioni

Il presente deliverable descrive le attività di triplification, deplly e primi screenshot dei servizi collegati all'integrazione dei gestori di servizi con il sistema SII-Mobility il cui progetto e realizzazione sono descritti nel DE5.11 .

I servizi considerati sono stati:

- parcheggi;
- servizi di car & bike sharing;
- taxi e car-service;
- operatori della logistica merci;
- nettezza e servizi vari (es. spazzamento strade).

Dato che tali fasi sono comuni per le diverse tipologie di servizio, è stato scelto di riportare le attività per un singolo ETL evitando di appesantire il documento con informaizioni ridondanti e praticamente doppione tra di loro. L'ETL selezionato è stato quello relativo ai dati dinamici per il gestore SienaParcheggi realizzato da QuestIT per l'integrazione dei dati relativi al monitoraggio dei parcheggi in struttura del comune di Siena (8 parcheggi).

Il documento riporta quindi la descrizione delle fasi di triplification e deploy di tale modulo di interfacciamento ed i relativi screenshot estratti dal ServiceMap di SIIMobility dei parcheggi integrati. La procedura illustrata per il gestore di riferimento è facilmente estendibile a tutti gli ETL progettati e realizzati per il Task 5.4 e descritti nel DE5.11.

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