



Enterprise Information Management Infrastructure (EIM) & Linked Open Data(LOD)

Infrastruttura Enterprise Information Management (EIM) & Linked Open Data (LOD)



UNIONE EUROPEA
Fondo Sociale Europeo
Fondo Europeo di Sviluppo Regionale



*Agenzia per la
Coesione Territoriale*



Presidenza del Consiglio dei Ministri
**Dipartimento della
Funzione Pubblica**



Dipartimento
dell'Amministrazione Generale,
del Personale e dei Servizi



**CLOUDIFY
NoiPA**

Abstract

The document provides more details on the Enterprise Information Management (EIM) Infrastructure, i.e. on all those solutions aimed at the optimal use of information and supporting the decision-making processes related to project activities, which require the availability of knowledge. The document also deals with the topic of Linked Open Data (LOD) - which represent structured data, linked and published in order to allow semantic queries. Finally, the text summarizes the main properties of open data:

- Not related to individual people
- Non-proprietary format, correlated by related metadata
- Associated license
- Available for free

Abstract

Il documento fornisce maggiori dettagli sull'Infrastruttura **Enterprise Information Management (EIM)**, ossia su tutte quelle soluzioni finalizzate ad un uso ottimale dell'informazione e al supporto dei processi decisionali, inerenti le attività progettuali, che richiedono disponibilità di conoscenza.

Viene, inoltre, affrontato il tema dei **Linked Open Data (LOD)** – che rappresentano i «*linked data* o *dati collegati*», ossia dati strutturati, collegati e pubblicati in modo da consentire interrogazioni semantiche.

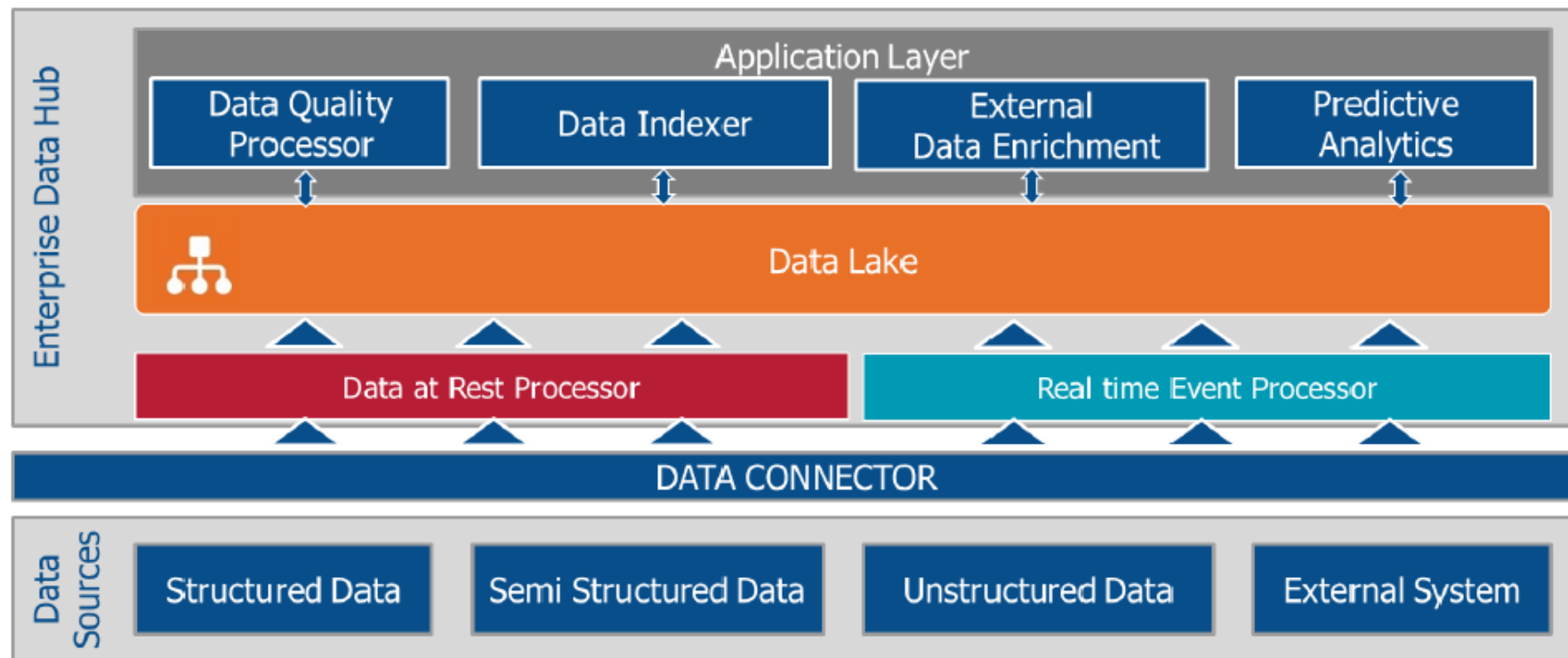
Vengono, infine, riassunte le principali proprietà di un dato aperto:

- *Non riferito a singole persone*
- *Formato non proprietario, correlato da relativi metadati*
- *Licenza associata*
- *Disponibili gratuitamente*

Agenda

- Infrastruttura EIM
- Open Data – Stato dell'arte
- Linked Open Data – Infrastruttura
- Linked Open Data – Processo creazione LOD
- Linked Open Data – Demo

EIM – Architettura Logica



Cloudera Manager

cloudera MANAGER

Clusters ▾

Hosts ▾

Diagnostics ▾

Audits

Charts ▾

Backup ▾

Administration ▾



Search

Support ▾

admin ▾

Home

Status All Health Issues Configuration  5 ▾ All Recent Commands

Add Cluster

Cluster 1 (CDH 5.13.0, Parcela)

Hosts  5

HDFS

Hive

Hue

Impala

Kafka

Oozie

Solr

Spark

YARN (MR2...)

ZooKeeper

Charts

30m 1h 2h 6h 12h 1d 7d 30d 

Cluster CPU



Cluster 1, Host CPU Usage Across Hosts 1.4%

Cluster Disk IO



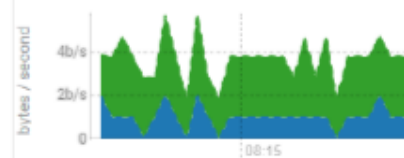
Total Disk Bytes Read 0 Total Disk Bytes Written 397K/s

Cluster Network IO



Total Bytes Read 10.8K/s Total Bytes Written 11.4K/s

HDFS IO



Total Bytes Read 1b/s Total Bytes Written 2.8b/s



Linked Open Data

Linee guida AGID per la valorizzazione dei dati

Obiettivi

- Pubblicare basi di dati di interesse nazionale
- Pubblicare in formato *Open Data*
- Utilizzare vocabolari controllati e modelli dei dati



Proprietà di un dato aperto

- Non riferito a singole persone
- Formato non proprietario, correlato da relativi metadati
- Licenza associata
- Disponibili gratuitamente

Linee guida pubblicate dall'AgID nel Piano triennale per l'informatica nella pubblica amministrazione 2017-2019

Classificazione Open Data

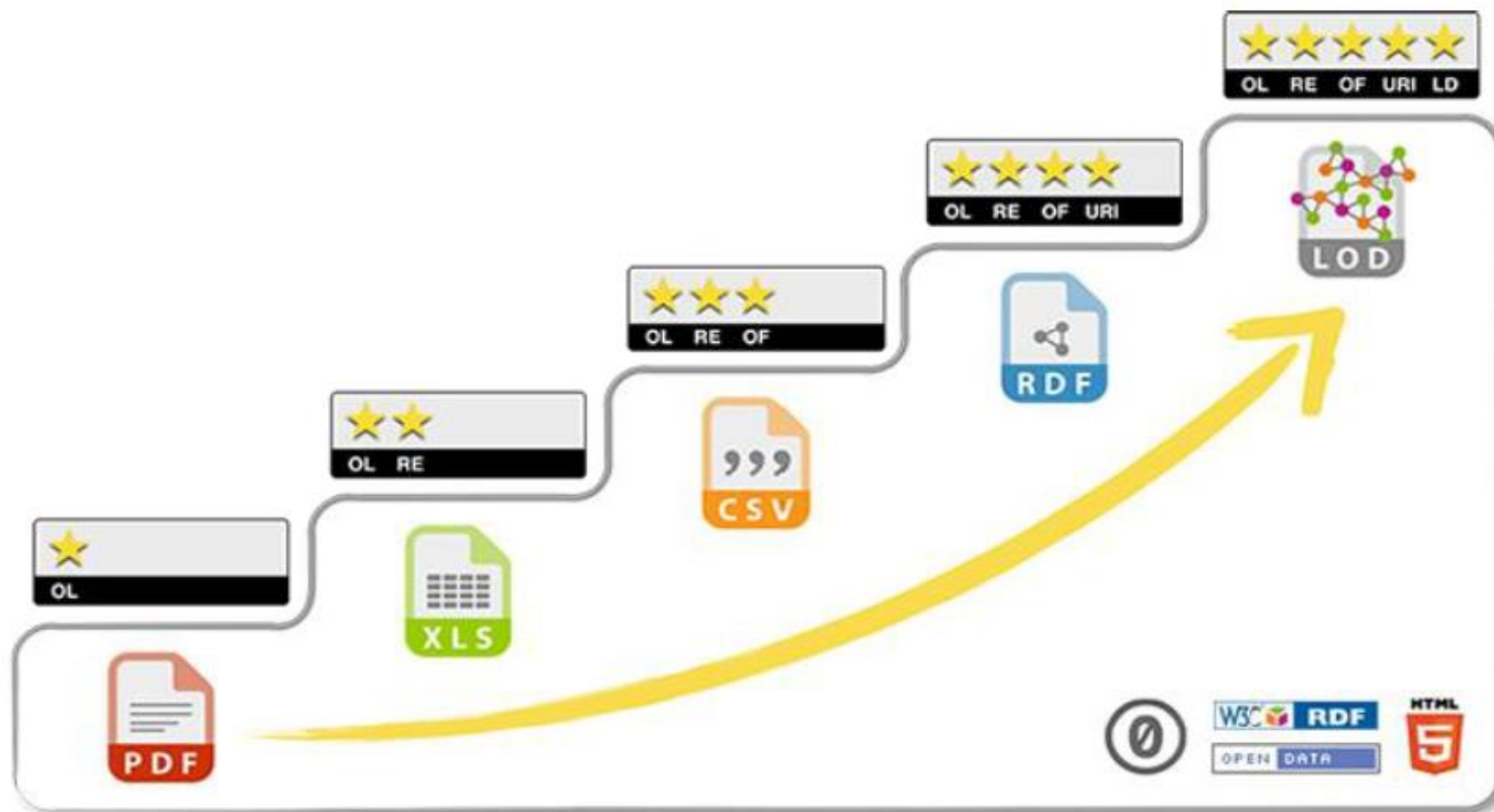
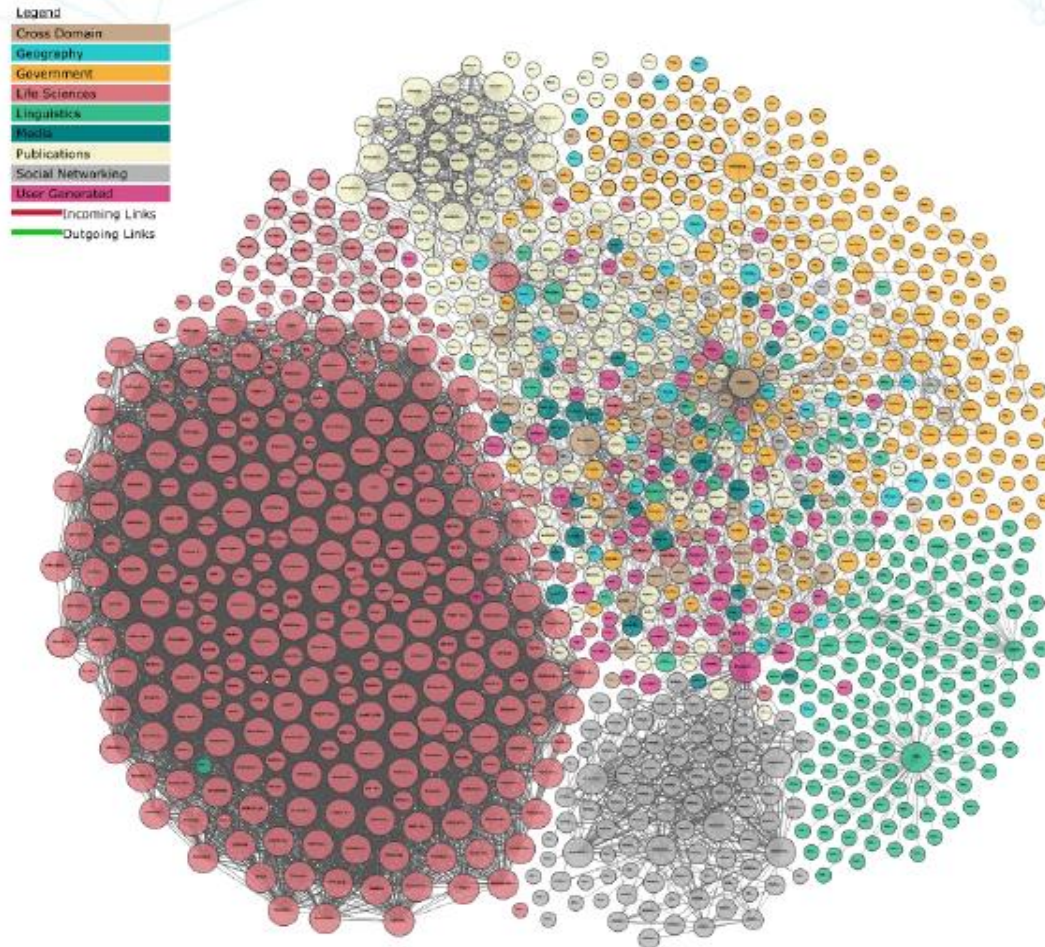


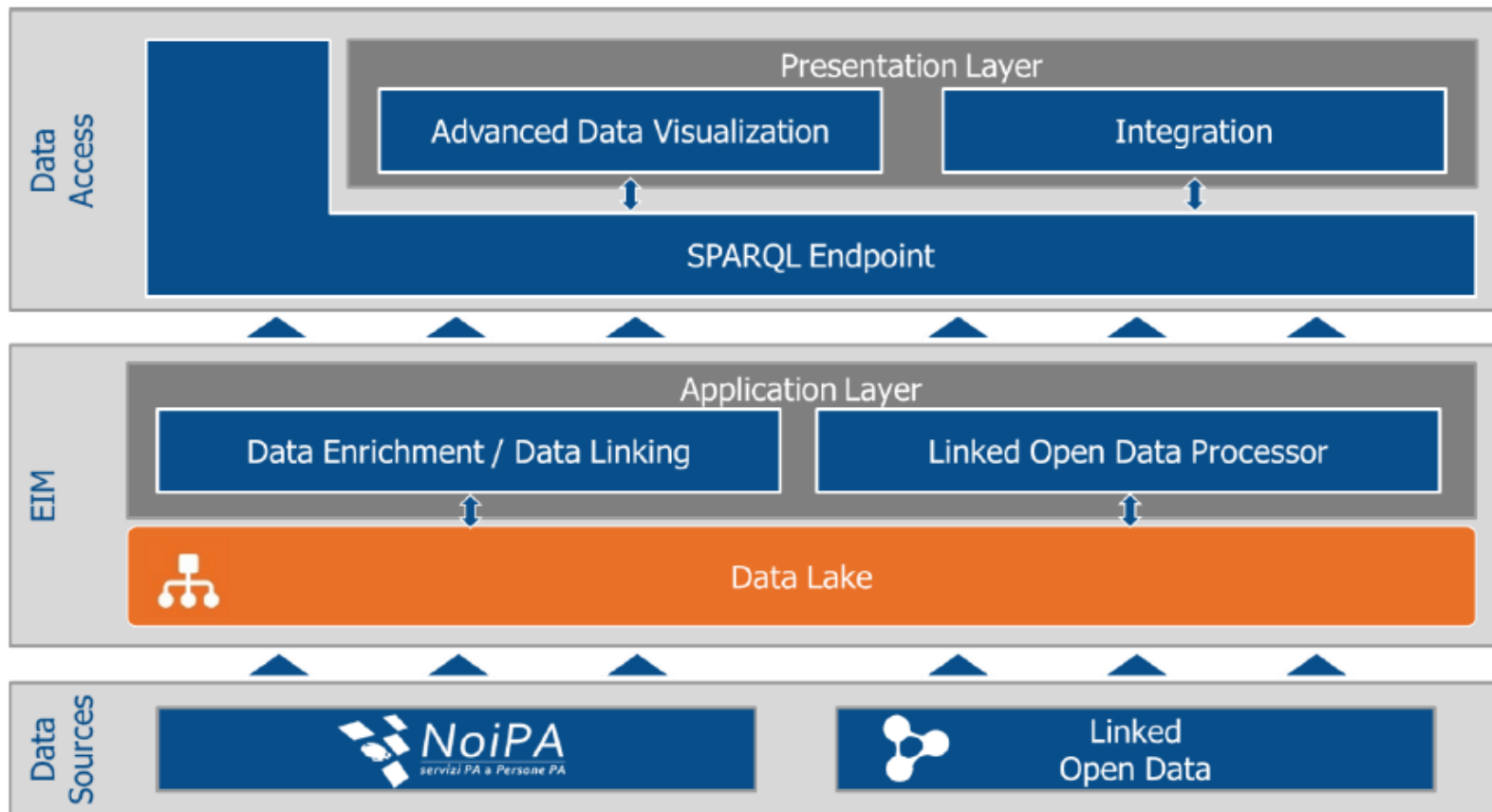
Diagramma LOD Cloud



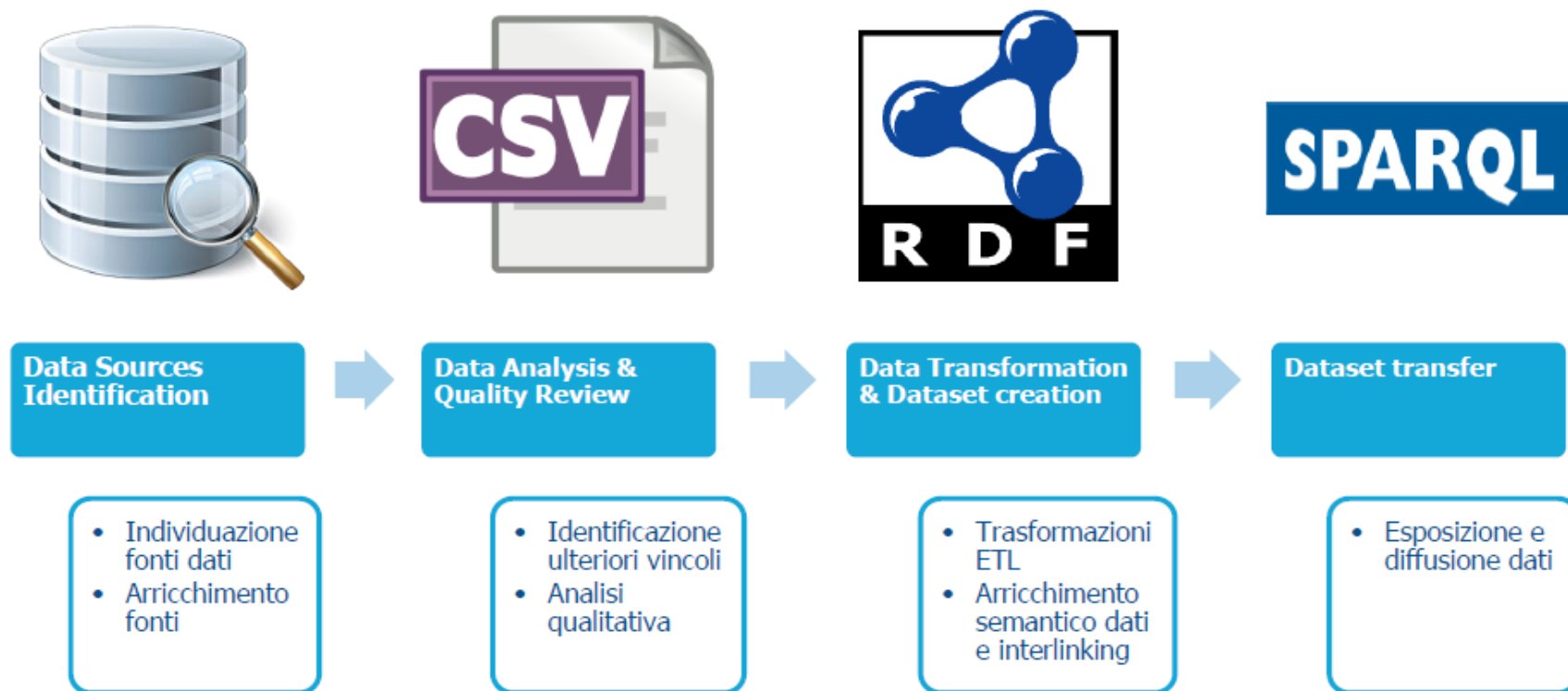
1163 dataset pubblicati

Source: <http://lod-cloud.net/>

Architettura logica



Processo creazione e pubblicazione LOD



Dataset di input



Dataset portale Open Data NoiPA

- Distribuzione amministratori per comune dell'Unità Organizzativa
- Distribuzione amministratori per comune di residenza
- Inquadramento amministratori per comune dell'Unità Organizzativa
- Retribuzione dipendente
- Risparmi di gestione per riduzione part time per comune dell'Unità organizzativa

Anagrafiche

- Enti
- Unità organizzative

Collegamenti esterni



- [DBpedia – Risorse](#)
- [DBpedia – Ontologia](#)



- [Istat – Territorio](#)
- [Istat – Variabili Censuarie](#)



- [RDFs – RDF Schema](#)
- [vCard – Persone e Organizzazioni](#)

Mappatura dati – Esempio



Proprietà

Descrizione del Comune ➡ `dbo:city | vcard:locality | ter:comune_di_REG`

```
<rdf:Description rdf:about='rdf-syntax-ns#descrizione_del_comune'>
  <rdfs:label>Descrizione del comune</rdfs:label>
  <rdfs:subClassOf rdf:resource='rdf-syntax-ns#descrizione_della_provincia' />
  <owl:sameAs rdf:resource='http://dbpedia.org/ontology/city' />
  <owl:sameAs rdf:resource='https://www.w3.org/2006/vcard/nslocality' />
  <owl:sameAs rdf:resource='http://datiopen.istat.it/odi/page/ontologia/territorio/comune_di_REG' />
</rdf:Description>
```

Valore

Comune di Alessandria ➡ `dbr:Alessandria`

```
<descrizione del comune rdf:resource='http://dbpedia.org/resource/Alessandria' />
```

dbr: <http://dbpedia.org/resource/>

ter: <http://datiopen.istat.it/odi/page/ontologia/territorio/>

vcard: <https://www.w3.org/2006/vcard/ns>

dbo: <http://dbpedia.org/ontology/>

Evoluzioni future

- Implementazione template ambienti EIM e LOD
(sw Cloudera, MongoDB, Virtuoso)
- Definizione e Pubblicazione nuovi dataset
(es. Amministrati per area territoriale, serie storica numero amministrati, dati spesa per funzione, etc.)
- Linking con ulteriori fonti esterne
(es. INAIL, Camera dei deputati, etc.)
- Perfezionamento ontologia interna
(es. utilizzo di ulteriori vocabolari controllati e ontologie standard)
- Analisi modello ontologico Cloudify NoiPA
- Avvio in esercizio



Grazie