

OpenData: un Framework di sviluppo

di T.Nicosia

```
<myo:Framework rdf:resource="http://neuron4web.palermo.enea.it/opendata/_twitterendpoint">  
  < dct:author> T.Nicosia </dct:author>  
</myo:Framework>
```

`<rdf:Description rdf:about=`” **Infografica di dati georeferenziati** “`>`

Il sistema è composto da un Database no sql

- Procedura temporizzata di raccolta da Waze
- Procedura asincrona di raccolta dati da Twitter
- Procedura asincrona di compilazione catalogo
- Demone di ascolto per richieste in sparql
- Pagina in javascript/HTML per il frontend utenza

`<rdf:Description rdf:about="` Sorgente dati `">`

Raccolti da fine marzo 2017,

- dati provenienti catturati dell'account Twitter della PM del comune di palermo, da demone bot in python. (861)
- Dati dall'applicativo Waze, di informazioni del traffico, raccolti ciclicamente ad intervalli di 20 minuti.(3911)

`<rdf:Description rdf:about="` Base dati `">`

- Apache Couchdb

Database nosql sviluppato per applicazioni IoT



Single Node Database

CouchDB is a terrific single-node database that works just like any other database behind an application server of your choice.

Most people start with a single node CouchDB instance. More demanding projects can seamlessly upgrade to a cluster.



Cluster

CouchDB is *also* a clustered database that allows you to run a single logical database server on any number of servers or VMs.

A CouchDB cluster improves on the single-node setup with higher capacity and high-availability without changing any APIs.



HTTP/JSON

CouchDB makes use of the ubiquitous HTTP protocol and JSON data format and is compatible with any software that supports them. CouchDB also works great with external tools like HTTP proxy servers, load balancers.



Offline First Data Sync

CouchDB's unique Replication Protocol is the foundation for a whole new generation of "Offline First" applications for Mobile applications and other environments with challenging network infrastructures.



Ecosystem

CouchDB is built for servers (from a Raspberry Pi to big cloud installations), while PouchDB is built for mobile & desktop web-browsers and Couchbase Lite is built for native iOS & Android apps.

And all of them can seamlessly replicate data with each other.



Reliability

CouchDB is serious about data reliability.

Individual nodes use a crash-resistant append-only data structure. A multi-node CouchDB cluster saves all data redundantly, so it is always available when you need it.

<rdf:Description rdf:about=” **Trattamento dei dati raw** “*>*

I dati sono immagazzinati 'as-is', ed elaborati da applicazione esterne per la trasformazione.

Sono rappresentate solo le informazioni di interesse:

- i tracciati geojson di tipo LineString, le velocità per il traffico e la data di pubblicazione.
- Il testo dei messaggi twitter, e l'interpretazione delle strade interessate in geojson tipo 'LineString'.

I dati twitter contengono inoltre un riferimento (owl:sameAs) all' URL bitly e all'URL del servizio di geoRSS del comune di palermo.

(<https://mobilitasostenibile.comune.palermo.it/news.php>)

`<rdf:Description rdf:about="` **Accessibilità dei dati** `>`

- Ogni documento nella base dati, contiene un identificativo unico (UUID) generato al momento dell'inserimento.
- Le serializzazioni sono create al momento della richiesta da URI, con dereferenziazione per: 'application/rdf+xml', 'text/plain', 'application/x-turtle'.
- URL di accesso con prefisso
“<http://neuron4web.palermo.enea.it/opendata/resource/>”

<rdf:Description rdf:about=" Interrogazioni*">*

- Sono forniti due endpoint, per la sottomissione delle query SPARQL (HTTP GET) sui dataset.
- Fornisce come output MIME 'application/sparql-result+xml', o serializzazioni JSON, XML, TURTLE su MIME 'text/plain'.

<rdf:Description rdf:about=" **Applicazione Web** *>*

- Interrogazioni in SPARQL, agli endpoint dei dati interessati.
- Contiene una sezione dedicata alla visualizzazione del grafo dei singoli dataset.
- Una sezione per la visualizzazione dei dati su mappa.
- Una sezione per le interrogazioni manuali.

SPARQL query editor

Waze traffic dataset Query Text

```
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dct:<http://purl.org/dc/terms/>
PREFIX tns:<http://neuron4web.palermo.enea.it:8080/opendata/_trafficResource/>
PREFIX myo:<http://neuron4web.palermo.enea.it:8080/opendata/_vocabulary/myowl#>
PREFIX geo:<http://www.w3.org/2003/01/geo/wgs84_pos#>
SELECT DISTINCT ?s ?p ?o WHERE {?s ?p ?o} LIMIT 100
```

SPARQL Result format

Twitter dataset Query Text

```
PREFIX owl:<http://www.w3.org/2002/07/owl#>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dct:<http://purl.org/dc/terms/>
PREFIX tns:<http://neuron4web.palermo.enea.it:8080/opendata/_twitterResource/>
PREFIX myo:<http://neuron4web.palermo.enea.it:8080/opendata/_vocabulary/myowl#>
PREFIX geo:<http://www.w3.org/2003/01/geo/wgs84_pos#>
SELECT DISTINCT ?s ?p ?o WHERE {?s ?p ?o} LIMIT 100
```

SPARQL Result format



2017-04-10T07:05:00

2017-04-10T06:22:33 -- lavori stradali
via ugo la malfa 86, limitazione della
circolazione veicolare e pedonale per
lavori di scavo.

2017-04-10T06:17:11 -- incidente
stradale via d. scina- via f. crisci
incidente stradale, possibili
rallentamenti

2017-04-10T06:17:10 -- lavori via
tranchina altezza via salvatore
minutilla limitazione temporanea della
circolazione per lavori di scavo.

2017-04-10T05:22:13 -- incidente
stradale piazzale einstein dir. ct lato
monte incidente stradale, possibili
rallentamenti.

2017-04-10T05:17:09 --
manifestazione piazza verdi ore 10.30
manifestazione celebrativa del 165°
anniversario della fondazione della
poli...

2017-04-10T05:07:20 -- lavori via c.
camilliani ang. via pozzo. chiusura al
transito veicolare per lavori urgenti.



<rdf:Description rdf:about=" **Librerie** *">*

- Javascript:
 - D3 , bootstrap , mapbox , jquery
- Python:
 - rdflib, requests, tweepy, json.

<rdf:Description rdf:about=" **Risorse** *">*

- http://neuron4web.palermo.enea.it/opendata/_twitterendpoint
 - Endpoint sparql per risorse Twitter
- http://neuron4web.palermo.enea.it/opendata/_trafficeendpoint
 - End point sparql per risorse Waze
- <http://neuron4web.palermo.enea.it/sparqlendpoint>
 - Interfaccia web

</rdf:Description>

Grazie per l'attenzione.