

OpenData: un Framework di sviluppo

di T.Nicosia

```
<myo:Framework rdf:resource="http://neuron4web.palermo.enea.it:8080/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. (3911)
- Dati dall'applicativo Waze, di informazioni del traffico, raccolti ciclicamente ad intervalli di 20 minuti.(861)

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

- Apache Couchdb

Database nosql sviluppato per applicazioni lo



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:8080/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 XML Run Query Reset

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 XML Run Query Reset

1660-a1485e16817c3f557b0bfda9b1655606
1621-ecae0306f0c621435f7319a697b3dc9c
1673-33572533f99594518a32b64659d671f4
1645-1cb2b4a371dd8ae9b40105a1ce76355e
1666-c519f3b281f31b966f58f9da3a3d93d5
1625-314f32103fee60a36b5dab09e4fa4c6c
1648-008c294c904bc6553b21f95f71e65ab6
1664-fd9aaaa4f63f7b434ee20fd50f379e6f
1652-f5fa116f54352e638f94598d1a78161d
1634-edeeafa563000f45668b97ea7425463a
1629-4deea05db00f6b19c76bed4d0157cc
1682-d3ac088fd93a48c4ec33c77eb9b41a10
1646-39fba66711c447026d55a4f982aa72f
1696-8a503cc861242e1d03c0cb3b639a2b12
1693-ed8066d8c22b6c520f92a598a00aa190
1690-fee031adde613bd76b799a7c51cddee
1675-3b04a519cf4c2858658aed918241ae5d
1671-66ed5ee75c838731b82138178f50569c
1649-6c25e1c344349a9d12817ece47c99b43

hashem

line034
line015
line016
line043
line021
line012
line018
line038
line002
line023
line011
line024
line030
line029
line008
line039
line022
line026
line014
line032
line019
line006
line028
line031
line020
line010
line013
line000
line033
line025
line027
line003
line004
line009
line044
line045

type

Resource

issued

2017-04-11T13:57:00

type

issued

onStreet

hgpsw/jyn

sameAs: http://nc=1&id=1585

geometry

Resource

equivalentClass

match

twitter

twittertext

248-771121717688e912ec263e445cfa8134
245-321b38c0fc71ac9e0d465d5c2d03aact
243-2c1246e30995800d41ade0361584fa36
249-973d791994b6038e6fee3d530049d3be
247-e032b30b1c2e730d1b4185f548d5eb79
244-2ec7211552a64e2a6f0cd23a560c78d7

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. crispi
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>

Grazie per l'attenzione.