LOD4SDI: LINKED OPEN DATA FOR GEOGRAPHICAL INFORMATION IN SPATIAL DATA INFRASTRUCTURES

planetek italia

LINKED OPEN SPATIAL DATA INFRASTRUCTURE

The publication of spatial data as Linked Open Data (LOD) facilitates their aggregation, processing and analysis with other data and maximizes the value of geographic information. Thanks to their open and standardized format, LOD can be accessed and reused by any computer application and can be crossed and linked automatically with other heterogeneous information, without any human intervention, but simply with machine to machine operations.

To ensure the web publication of its own geospatial data as open and linkable data, ISPRA (Institute for Environmental Protection and Research) adopted LOD4SDI, a solution developed by Planetek Italia that automatically generate LOD according to the standard RDF/XML and using the cartographic web OGC services (WFS, CSW etc) provided by Spatial Data Infrastructure as input data.

Thanks to LOD4SDI it's possible: a full standardization of open data, their indexing on search engines, their connection with other open data published on the Web and their full accessibility for anyone who wants access to publish data, extract them, download them in various formats and cross them with other linked data available.

REFERENCES

The reference data used by ISPRA in its first implementation are the European Corine Land Cover 2006 and 2012. The ontologies considered are those already known and widely accepted of the harmonISA project, which define the semantic descriptions of use of the territory and categories of land cover. LOD4SDI is also designed in view of the update the Corine using the Copernicus data.

Author

Mauro Casaburi
casaburi@planetek.it
Planetek Italia s.r.l.
http://www.planetek.it

Valter Sambucini
valter.sambucini@isprambiente.it
ISPRA (Institute for Environmental Protection and Research)
http:// www.isprambiente.gov.it

DOWNLOAD the pdf version on your mobile

LOD4SDI

A *plug'n'play* component, fully compatible with the INSPIRE SDI, that can transform and provide all the spatial data as Linked Open Data.

LOD4SDI: Benefits

Citizens: Search and download as Open Data all the INSPIRE Spatial Data.

Developers: Build new app that can use the INSPIRE Spatial Data in real time connecting through the OGC geoSPARQL End-Point.

Administrations: Provide all the INSPIRE spatial data as LOD with no SDI upgrade and preserving the previous investments.

Connect the SDI to the Open Data Cloud improving the environmental awareness and knowledge.
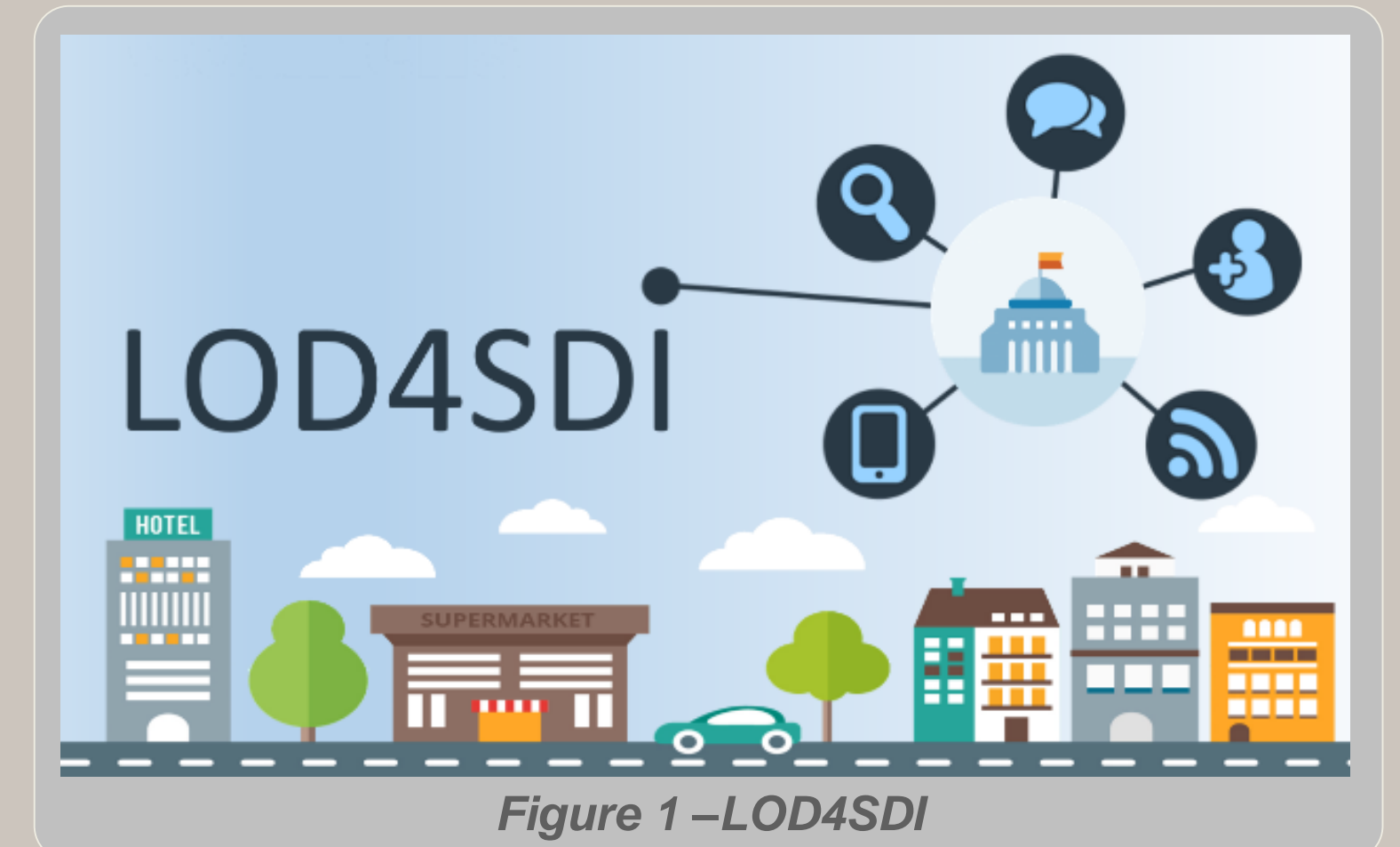

Figure 1 –LOD4SDI

LOD4SDI: How it works

LOD4SDI is a stand-alone component that receives in input the data through the standard services (CSW and WFS) of any INSPIRE SDI, transforms the metadata and the data in Linked Open Data (RDF) by dedicate ontologies, store its in the Triple Store Data Base and expose the LOD through an endpoint OGC geoSPARQL .
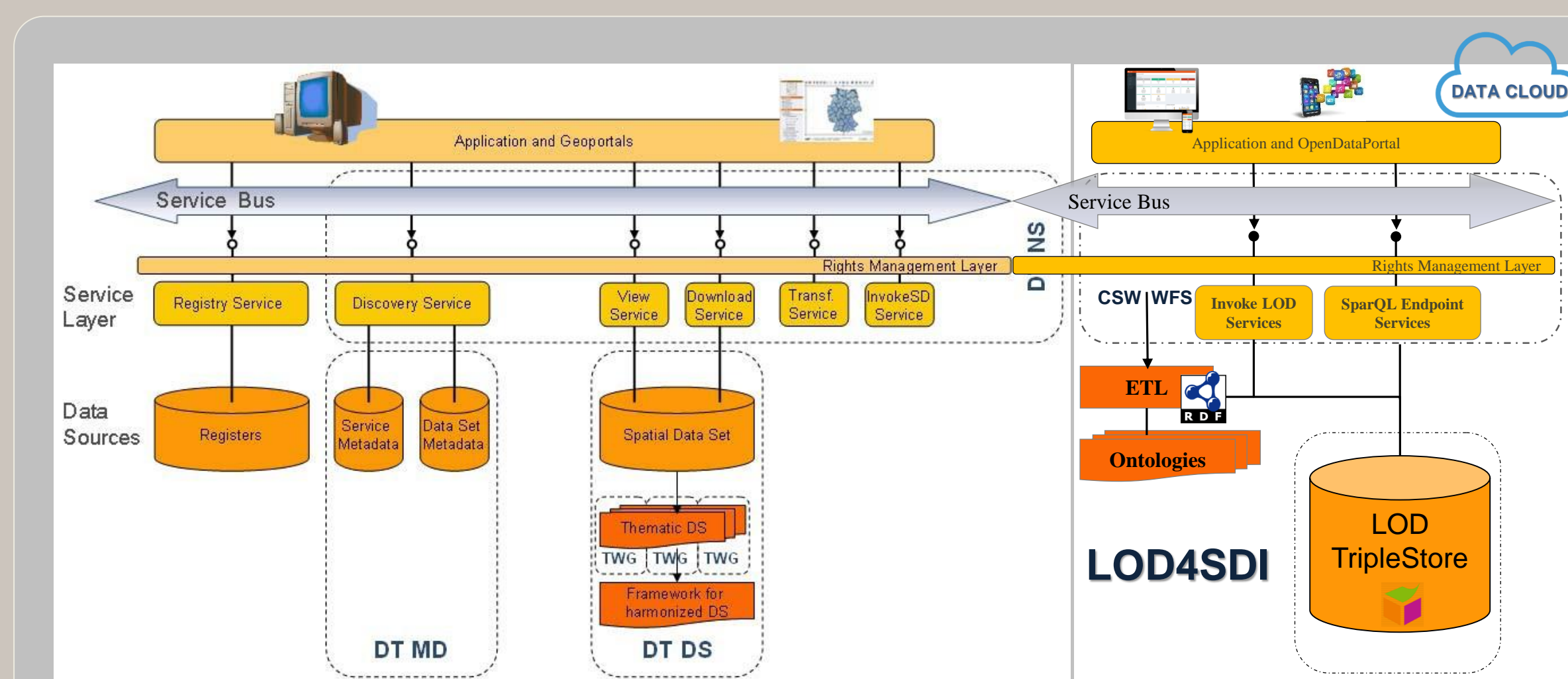

Figure 2 – INSPIRE SDI Architecture extended with LOD4SDI component.

LOD4SDI is easy to install and manage. Thanks to a user friendly inteface, the administrator can manage it in few steps:

A. Insert the URL of the INSPIRE services (CSW;WFS) as data source.

B. Choose wich metadata/data he wants to provide as Linked Data.

C. Shedule the Extract-Transform-Load job. For dinamic data could be necessary to schedule the job as repetitive in time to update the Linked Data.
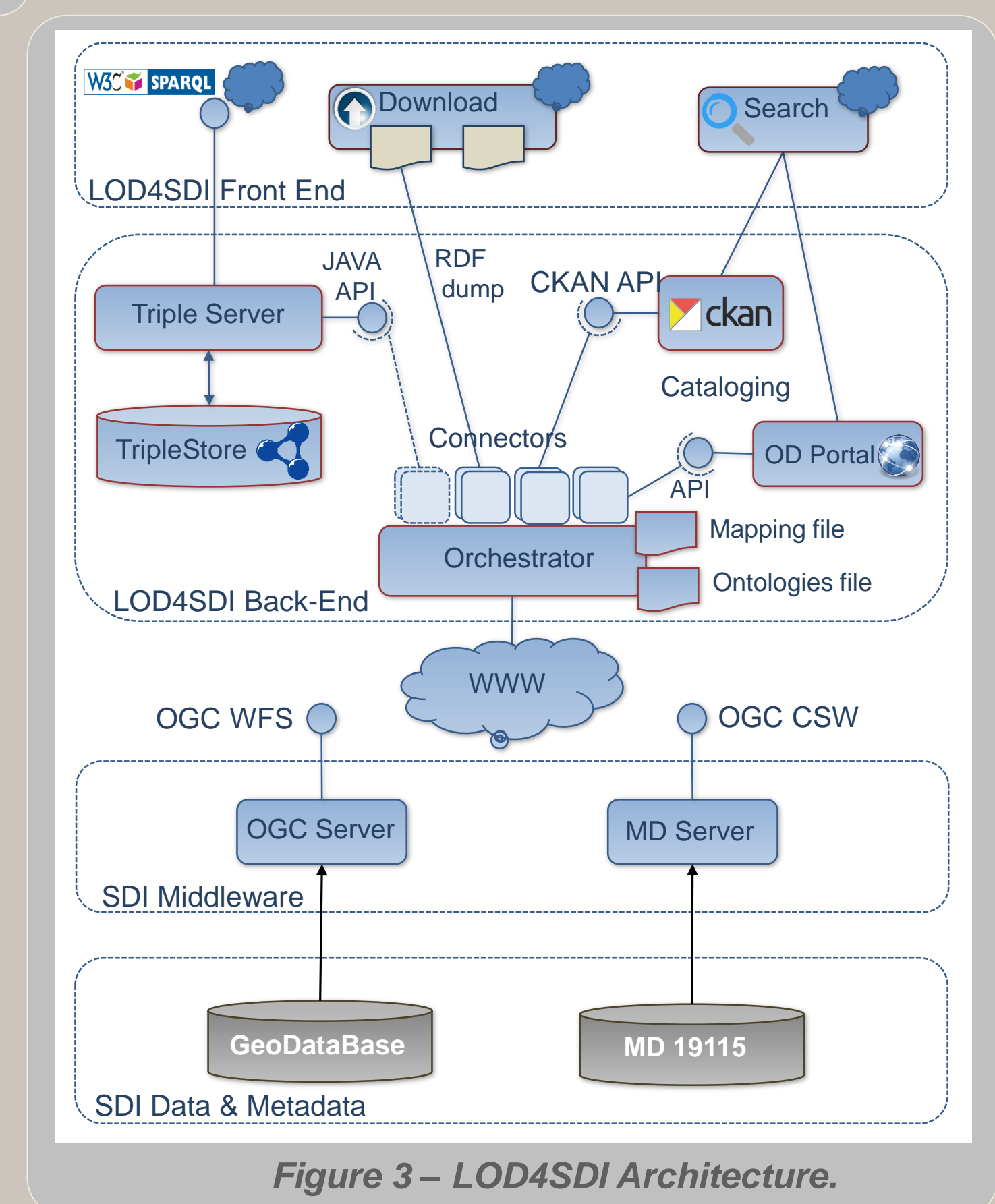
LOD4SDI: Software Architecture

All the LOD4SDI system is based on several components combining together according to the modern modular software architectures. This is guarantee of capability in terms of scalability and extensibility. These components are based on open source libraries and frameworks. The core of the LOD4SDI component is the Orchestrator, that manages When transform the spatial data in linked data, How to do that and on Which data. The Orchestrator take in input the Metadata/Data from the OCG services and, using GeoKettle workflow as geoETL engine, execute and transform the XML/GML (Metadata/Data) in RDF. To obtain the RDF it takes in input the mapping file that relates the ontologies and the geofeatures. The Orchestrator has also API interfaces to provide the RDF files to the Triple Server directly connected to the TripleStore based on Strabon and to provide the RDF to the other Open Data Software like Open Data Portals, CKAN Portal, Search and Download interface. The LOD4SDI Front End provides the possibility to the Admin user to manage all the system and expose the data through the search and download interface and through the geoSPARQL End Point.

Figure 3 – LOD4SDI Architecture.

Semantic INSPIRE Knowledge base

- Improve the INSPIRE environmental knowledge-base in worldwide interoperability data cloud

- Spatial Data harmonized and interconnected through standard services

- Improving the environmental awareness end knowledge

- Machine to Machine (M2M) connection to spatial information

#OPEN DATA