



Idra - Open Data Federation Platform

A single entry point for Open Data

Engineering R&D lab - Palermo



www.eng.it

Open Data: fragmentation and interoperability

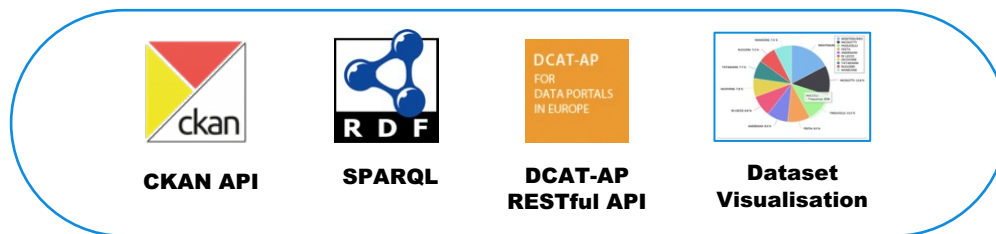
- One of the most relevant problems that limits to adoption and usage of open data is related to technology and standard fragmentation: every entity that wants to provide open data (e.g. Public Administration) uses different platforms and formats for their publication.
- This issue represents a barrier for the user that wants to search for open data, and for third-party applications that are not able to simply access and reuse it.
- Idra is an open source platform that provides a **unique access point to open datasets** coming from heterogeneous Open Data management systems, such as portals provided by Municipalities and PAs.
- Idra, allowing a unique and standard access to the open data, is in line with **Freedom of Information** laws, to foster and simplify the access to data held by national governments.

Idra - Open Data Federation Platform

- Idra is an “Open Data Federation Platform” composed by a back-end module and a web portal that allows to search for open data and linked open data, in a federated and multilanguage way
- The datasets remain in the original portals, the platform imports and manages only the metadata, updating it periodically in order to assure that the information provided is up to date
- The platform is able to access to open data portals based on different technologies even that they don't provide API



Idra overview



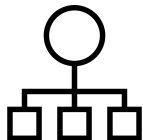
idra



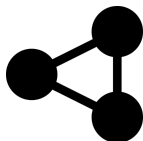
Idra: main features



Web interface to search for open data in a federated and multi-language way



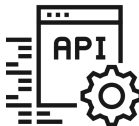
Unique access to open data regardless the **heterogeneity** of the underlying technologies



Ease of performing **SPARQL queries** on 5 stars RDF linked open data



Integration of third party applications for the open data visualisation and federation



Public and open APIs to build smart and innovative applications



Compliance to **DCAT-AP** standard and **European Data Portal** specifications

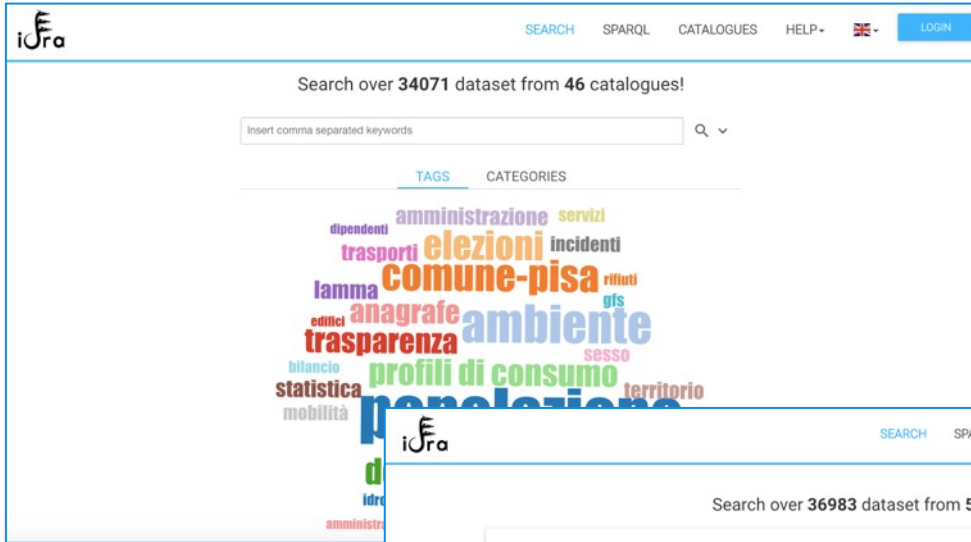


Web scraping technology to federate open data portals that do not provide APIs

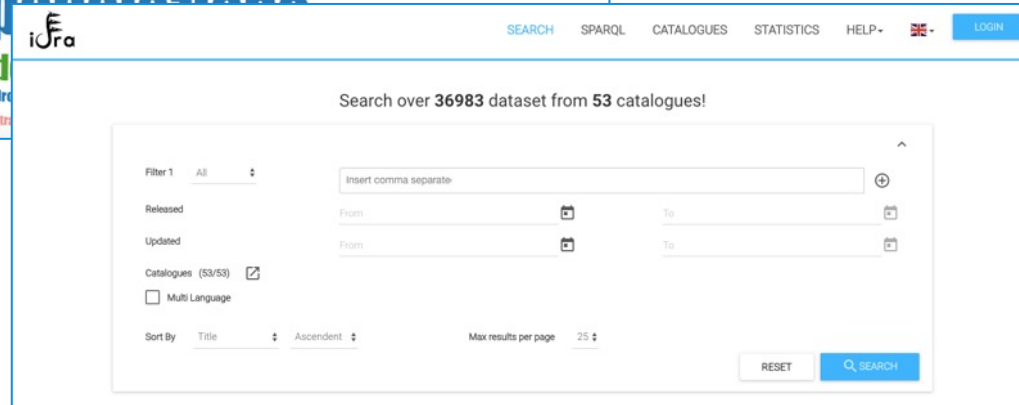


Connectors to simplify the federation of new portals reducing effort to the minimum

Search interface (1)



Filters and ordering : it is possible to search and order open data by:
Keywords related to title, description, tags, data of creation and of modification etc.



Selection of federated ODMS: it is possible to select federated ODMS of interest

Search interface (2)

Filter 1 All

water

Insert comma separated keywords

Released

From

To

Updated

From

To

Catalogues (53/53)

☒ Multi Language

Source

Targets

English

Italiano

Sort By Title Ascendent

Tags

ambiente (39)
acqua (36)
idrografia (33)
cover (25)
land (25)
surface (25)
cartografia di base (21)
ctr (21)
fiumi (20)
gfs (20)
Show all Tags

Formats

zip (90)
csv (69)
wms (44)
html (27)
json (27)
xml (18)
rdf+xml (13)
shp (12)
xlsx (11)
wfs (9)
Show all Formats

BACK

Sort By Title

Results 25

Datasets found: 186

PTCP 2013 - Quadro Conoscitivo QC08.3 Beni paesaggistici. Aree tutelate per legge - Città Metropolitana di...

ODN - Città Metropolitana di Firenze

Piano Territoriale di Coordinamento Provinciale approvata con Delib. C.P. del 10/01/2013 (BURT n.11 del 13/03/2013). La Carta rappresen...

WMS (1)

ZIP (1)

Uso delle acque comunali - Treviso - Anno 2014

Open Data Veneto

contiene i dati relativi all'utilizzo delle acque nell'anno 2014 nel Comune di Treviso

CSV (1)

Uso delle acque comunali - Treviso - Anno 2016

Open Data Veneto

contiene i dati relativi all'utilizzo delle acque nell'anno 2016 nel Comune di Treviso

CSV (1)

SWE (Snow Water Equivalent) 2017

Open Data Lombardia

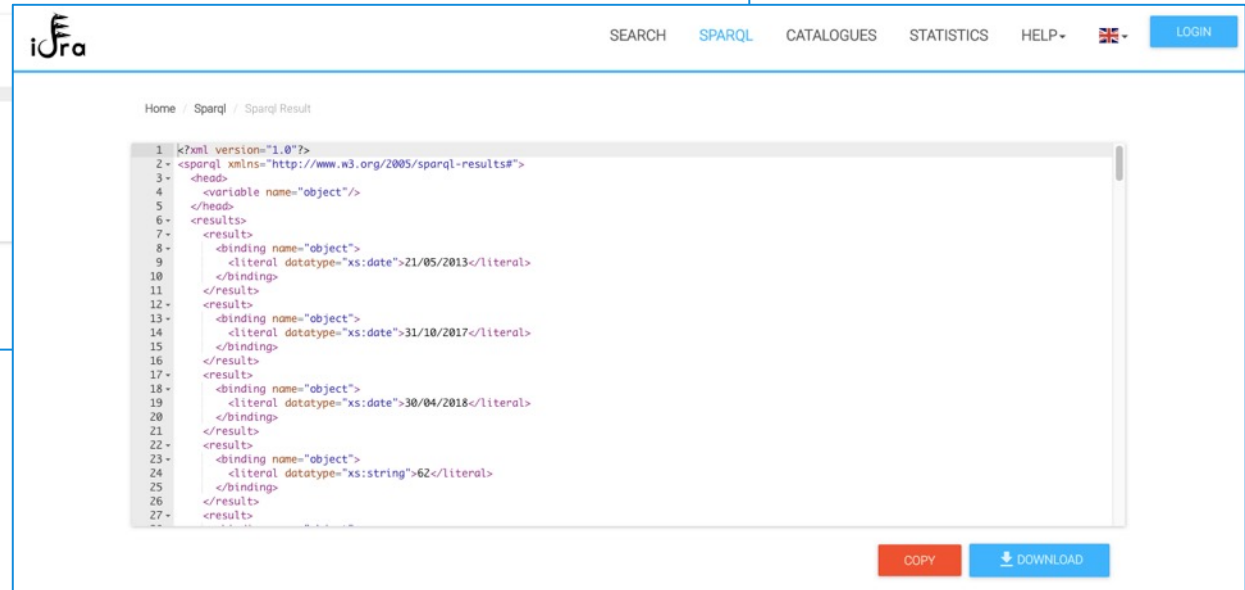
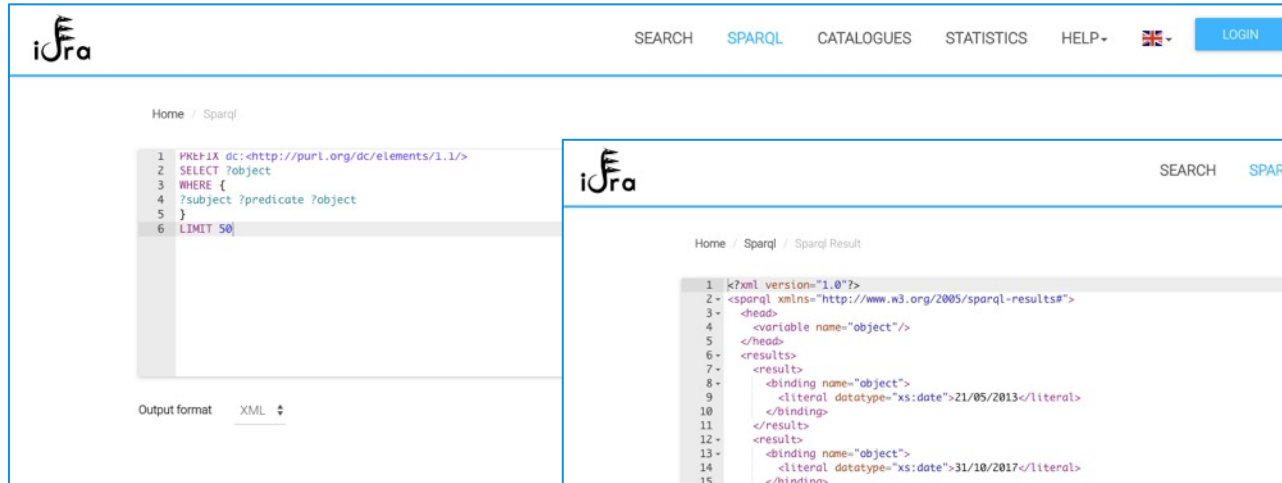
Mapa spazializzata dell'equivalente idrico della neve (SWE), derivano dall'integrazione dei dati rilevati dalle stazioni nivometeorologiche e ...

ZIP (1)

EuroVoc: it is possible to execute multilanguage searches based on the European thesaurus [EuroVoc](#)

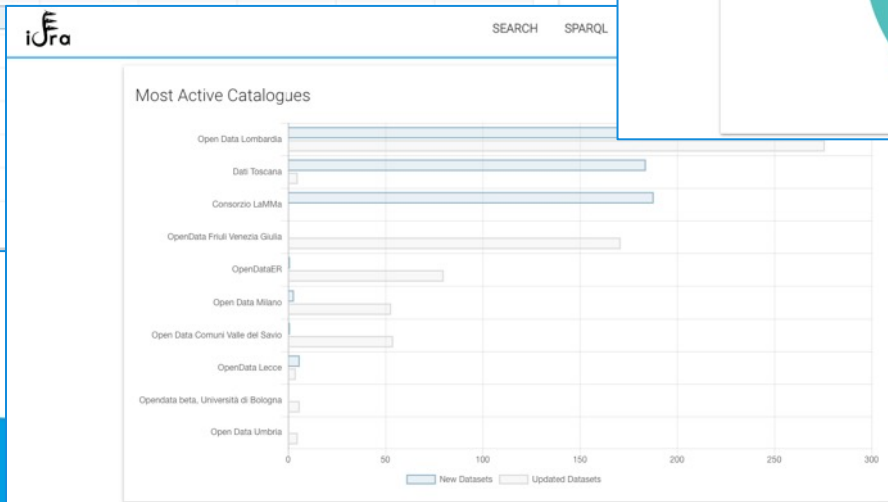
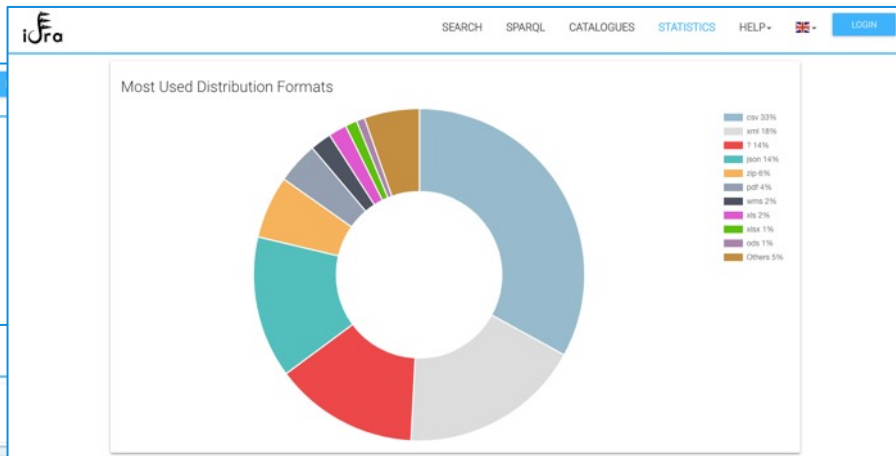
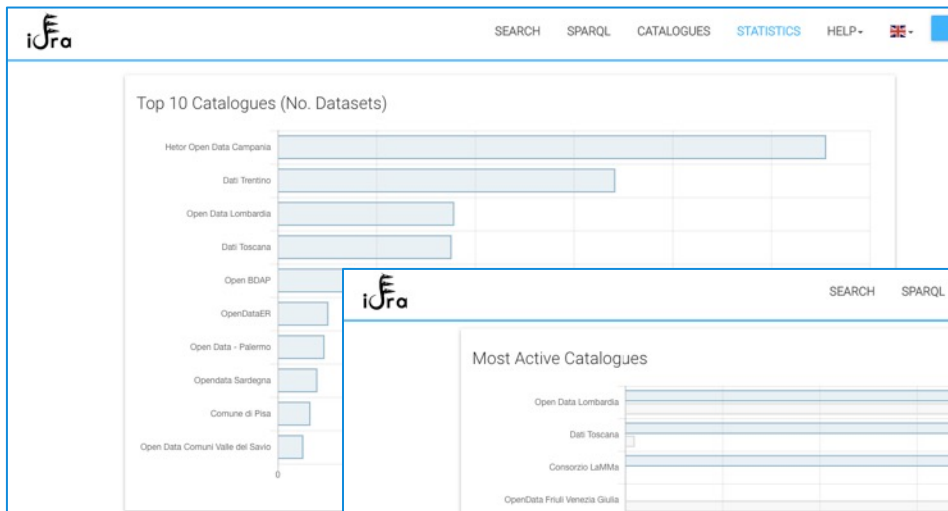
Linked Open Data SPARQL Editor

SPARQL queries: Idra provides a SPARQL editor to perform queries on Linked Open Data



Statistics about federated catalogues

Idra shows comparative statistics about federated catalogues: number of datasets, distribution format types, most active catalogues etc.



Integration with external apps : Datalet creator

Home / Datasets / Cognomi di tutti i residenti di Matera aggiornati al 31/08/2013

Cognomi di tutti i residenti di Matera aggiornati al 31/08/2013

OpenData Matera

Il file CSV contiene i cognomi di tutti i residenti nel comune di Matera. Nella seconda colonna è visualizzato il numero totale di individui aventi il relativo cognome. L'aggiornata al 31 Agosto 2013.

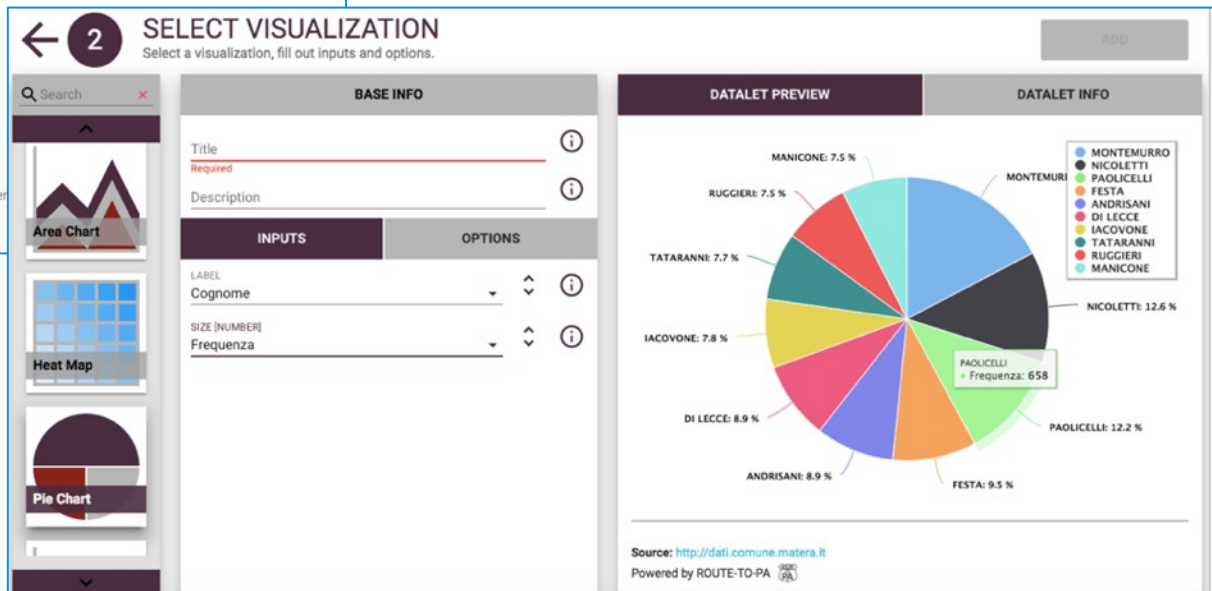
Tags

cognomi residenti statistica

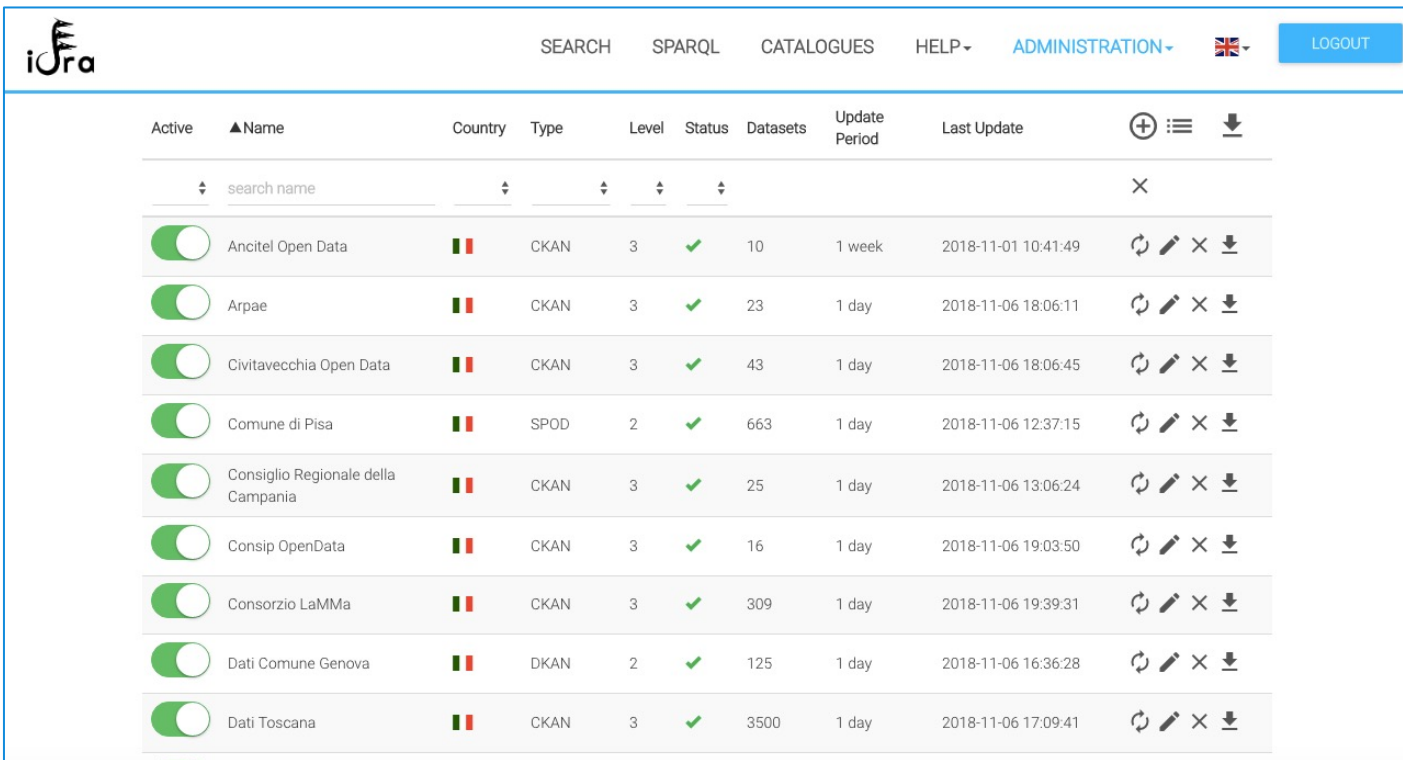
Resources

Format	File Name	Type	Description
JSON	cognomimatera.json	JSON	file in formato json
CSV	cognomimatera.csv	CSV	I cognomi di tutti i residenti di Matera

- The platform is integrated with an external component (Datalet creator) able to generate “on the fly” diagrams from open data (developed by Route2PA project)
- Further external components can be integrated to add specific visualisation and data analysis functionalities



Administration console: federation management

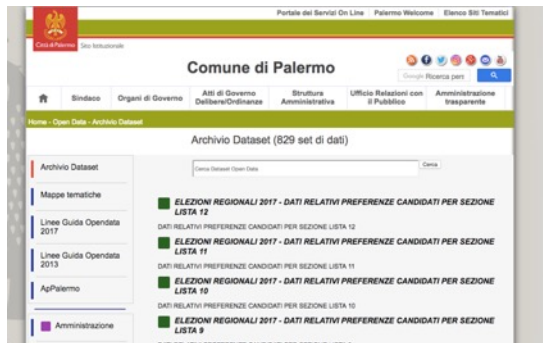


Active	Name	Country	Type	Level	Status	Datasets	Update Period	Last Update	
<input type="checkbox"/>	search name								X
<input checked="" type="checkbox"/>	Ancitel Open Data	Italy	CKAN	3	✓	10	1 week	2018-11-01 10:41:49	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Arpae	Italy	CKAN	3	✓	23	1 day	2018-11-06 18:06:11	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Civitavecchia Open Data	Italy	CKAN	3	✓	43	1 day	2018-11-06 18:06:45	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Comune di Pisa	Italy	SPOD	2	✓	663	1 day	2018-11-06 12:37:15	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Consiglio Regionale della Campania	Italy	CKAN	3	✓	25	1 day	2018-11-06 13:06:24	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Consip OpenData	Italy	CKAN	3	✓	16	1 day	2018-11-06 19:03:50	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Consorzio LaMMa	Italy	CKAN	3	✓	309	1 day	2018-11-06 19:39:31	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Dati Comune Genova	Italy	DKAN	2	✓	125	1 day	2018-11-06 16:36:28	↻ ✎ ✕ ⬇
<input checked="" type="checkbox"/>	Dati Toscana	Italy	CKAN	3	✓	3500	1 day	2018-11-06 17:09:41	↻ ✎ ✕ ⬇

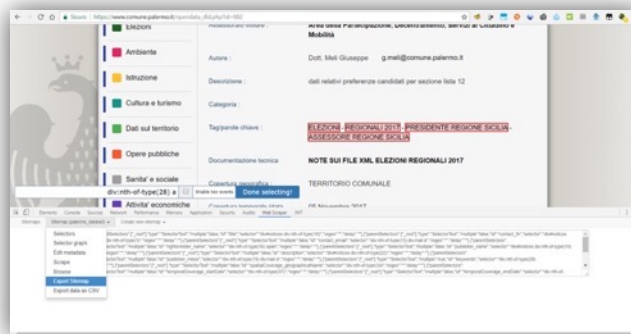
Idra provides administration functionalities to manage federated ODMS and platform configuration

Federation of custom open data portals

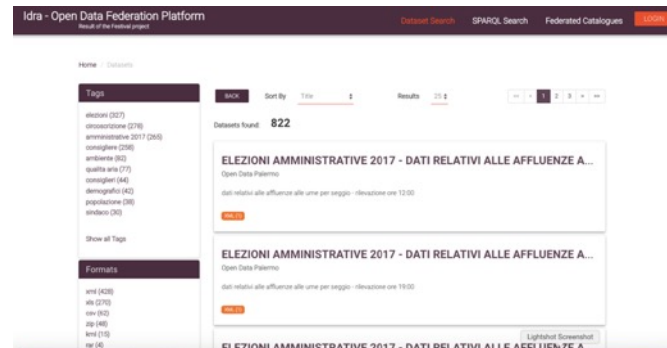
Custom open data portal without API
and no standard metadata



Visual mapping of
the webpage
elements to DCAT-AP
model using a
Chrome web
browser plugin

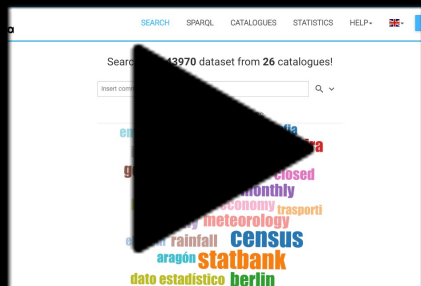


Idra uses sitemap and web scraping to federate the
portal: the datasets are now DCAT-AP compliant

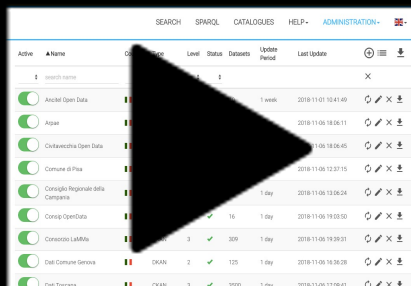


Import the mapping
description file
(sitemap) on Idra

Idra demo videos



User functionalities



Administration console

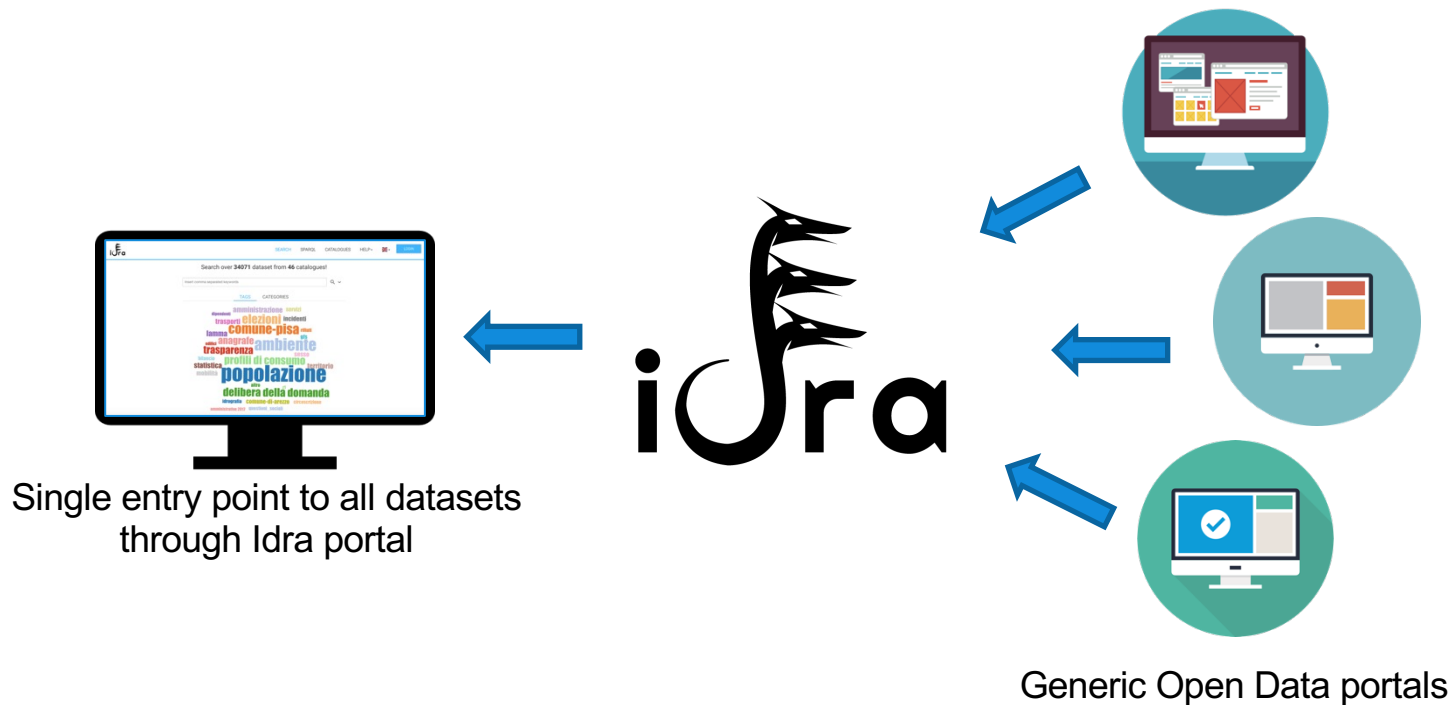


Web scraping

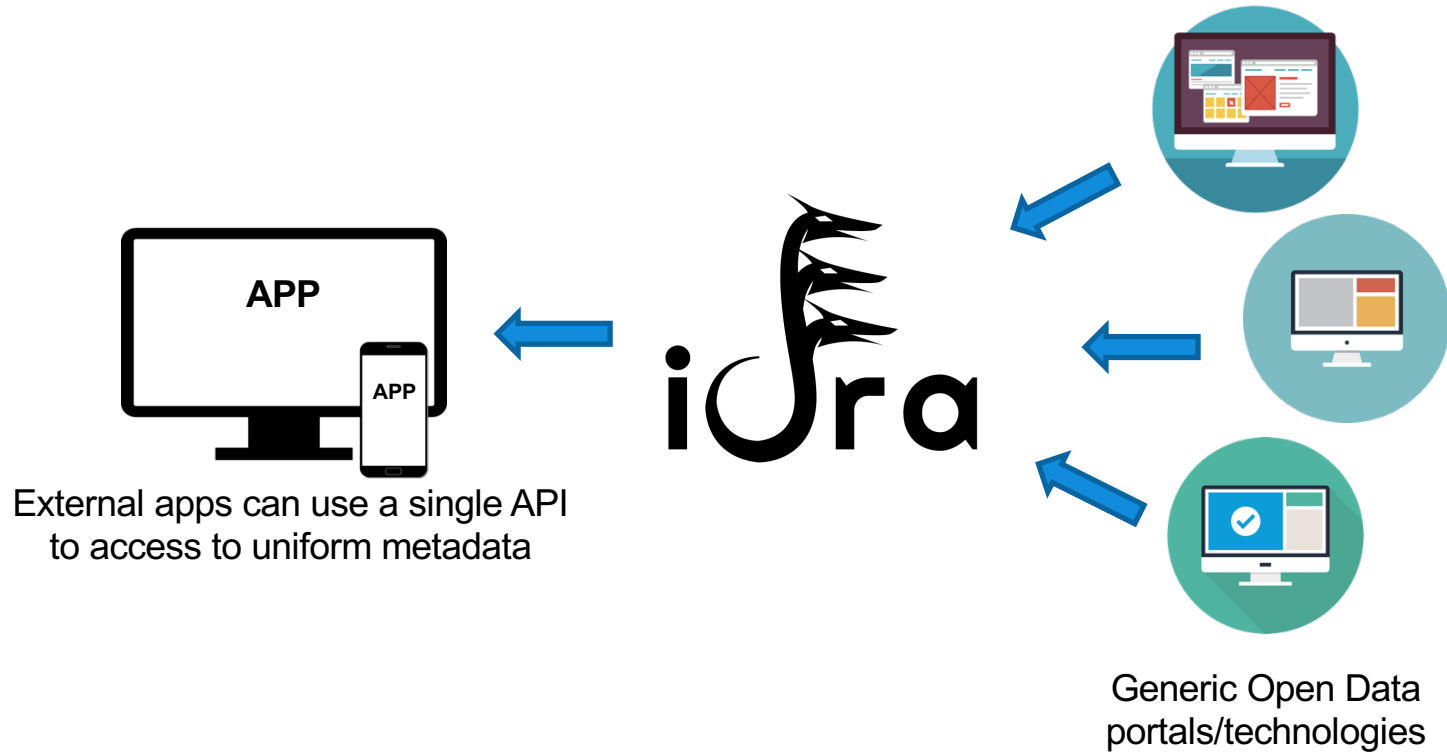
Idra connectors

Connector	Supported version	Description
CKAN	API v3	Connector for CKAN portals
DCATDUMP	DCAT AP v1.1 , DCAT-AP IT v1.0	Connector to import metadata via dump file compliant with DCAT-AP or DCAT-AP_IT
DKAN	DCAT Catalog endpoint	Connector for DKAN portals
Junar	API v2.0	Connector for Junar portals
NATIVE	API v1.0	Connector for Idra Federation API
OpenDataSoft	Search API v2	Connector for OpenDataSoft portals
ORION	NGSI v2 , NGSI-LD	Connector that allows to create datasets with NGSIv2/NGSI-LD query-based distributions
SOCRATA	DCAT Catalog endpoint	Connector for Socrata portals
SPARQL	v1.1	Connector that allows to create datasets with SPARQL query-based distributions
SPOD	CKAN API v1&2	Connector for Stat Portal Open Data (SPOD)
WEB	v1.0	Connector to import sitemaps in order to scrape and federate custom websites

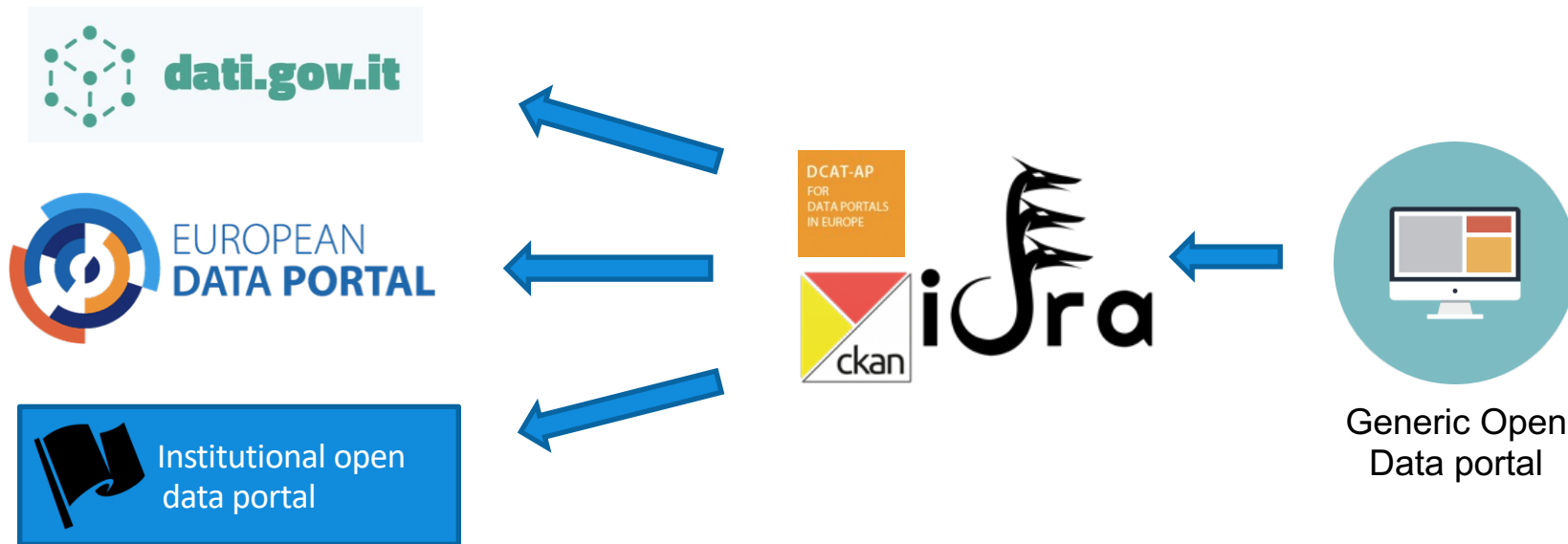
Create a meta-catalogue as single entry point to different open data portals



Provide a single API to access to heterogenous open data technologies



Enable the possibility to be harvested by European/national open data catalogues



Idra: an open source project

- The first version of Idra has been developed during the EU-Japan collaborative project [FESTIVAL](#)
- The platform has been later extended and it is **currently improved and maintained by Engineering**
- Idra is an open source project released under the **AGPLv3 license**
- Idra is part (incubated Generic Enabler) of the European open platform [FIWARE](#), a suite of open source components to support different applicative domains, in particular **IoT e Smart City**



Developer resources

- GitHub repository : <https://github.com/OPSILab/Idra>
- DockerHub: <https://hub.docker.com/u/idraopendata/>
- API documentation: <https://idraopendata.docs.apiary.io/>
- User manual: <http://idra.readthedocs.io/>
- Online demo: <https://idra.eng.it/>
- Information and support: idra@eng.it
- Developers Italia (Italian P.A. open source catalogue):
<https://developers.italia.it/en/software/opsilab-idra-015316.html>



Idra: benefits for the stakeholders (1)

Benefits for Public Administrations

- **Standardise their open data** exposing it through open API and standard models (DCAT_AP) compliant with European guidelines. The API could be accessed by other open data federation portals (e.g. European Data portal)
- **Provide to the citizen**, with minimum effort, **new functionalities** not existent in their native portals (e.g. SPARQL queries, graphical visualisation of datasets etc)
- **Improve visibility** because their data are part of a federation potentially accessible by a larger number of users

Benefits for the citizens

- Use a **single access point** for the open data of different entities avoiding problems related to find and use heterogeneous open data portals
- Access to advanced search and visualisation functionalities not existent in the original portals

Idra: benefits for the stakeholders (2)

Benefits for companies

- Access to a large number of open datasets in a unique and standard way together with data visualisation and analytics functionalities for different scopes:
 - Statistics, market analysis, sector-specific data research
 - Data reuse to build innovative and added-value application

Benefits for software developers

- Access to Open API to develop innovative applications based on a large set of open data in a simple and unique way
- Contribute to the development of the platform producing an added value for the international open data community

Thank you



ENGINEERING

Martino Maggio

Research & Development Laboratory
Viale Regione Siciliana, 7275 – Palermo - Italy

 martino.maggio@eng.it



www.eng.it



@EngineeringSpa



Engineering Ingegneria Informatica Spa



gruppo.engineering

