

ENGAGE Project Information

Acronym ENGAGE

Title An Infrastructure for Open, Linked Governmental Data Provision towards Research Communities and Citizens

Website <http://www.engage-project.eu>

Platform <http://www.engagedata.eu>



e-infrastructure

Research Infrastructures

Contract no RI-283700

Project type CP-CSA

Start date 01/06/2011

Duration 36 months

Partners 9



**Framework Programme 7
(2007-2013)**

Project participants

NTUA (Coordinator)	GR
TU-DELFT	NL
MIC-GR	GR
IBM-ISRAEL	IL
INTRASOFT	LU
STFC	UK
FhG-FOKUS	DE
AEGEAN	GR
EUROCRIS	NL



Ministries / local public agencies websites



National Statistical Offices



Unstructured / "Semi-structured"

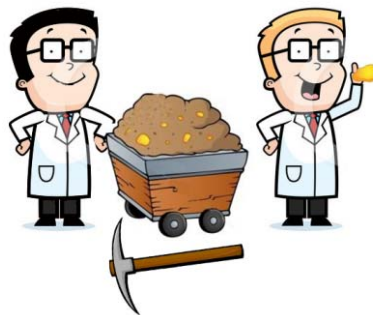


ENGAGE traverses across distributed and diverse public sector information resources



ENGAGE provides a single point of access to PSI sources as well as relevant tools in order to cover the needs of researchers and citizens

- o Integration of original PSI data and derived / curated datasets created, maintained and extended by users (researchers, citizens, journalists, computer specialists) in a **collaborative environment**. A research / data curation community platform with focus on the SSH domain.
- o The vision of the ENGAGE infrastructure is to extract, highlight and **enhance the RE-USE value** of PSI data.
 - o HOW: Moving slowly from low-structured, isolated, difficult to find PSI data to high-structured, easy to link , easy to process datasets => Crowd-sourcing.



for Scientists Citizens Journalists

Unstructured / Semi-structured / Structured



Public data sources



Discovery and Context Metadata



ENGAGE



Crowdsourcing

Moving from low structured, low value datasets to highly structured and / or derived datasets

Low

Re-Use Value / Quality structure / metadata



High

Re-Use Value / Quality structure / metadata



`<\/>`



Conversion
Data Enrichment
Metadata Enrichment
Cleansing
"Snapshots"



ENGAGE 2.0

- On top of ENGAGE basic functions (catalog, search, visualizations, API)

Researchers / Citizens / Journalists:

- Extend other datasets (official or already extended - derived datasets)
 - Conversions (e.g. HTML- PDF to xls, PDF to RDF)
 - Data Cleansing (e.g. duplicate records, empty rows, errors)
 - Metadata Enrichment (missing metadata, Linked Data Enablers!)
 - Data Enrichment (enrich datasets with more information)
 - Snapshots of real-time data (e.g. Diavgeia_decisions_10_2012_to_12_2012.xls)
 - Mash-ups / Interlinking (e.g. Combine Election results to UV radiation levels!)
- View the version tree of official – derived datasets (clean solution - easy to understand and manage the contributions / versions)

ENGAGE 2.0

Researchers / Citizens / Journalists:

o Data Requests

- o Looking for a dataset (e.g. I can't find it elsewhere. Does it exist?)
- o Looking for a curation / conversion / enrichment (e.g. I am looking for the election results in Greece in XLS.)
- o Looking for data verification (e.g. Do you think this dataset is valid?)
- o Freedom of Information Requests

o Integration of tools

- o Google Refine
- o ScraperWiki
- o Visualizations

ENGAGE 2.0

Data Providers:

- Maintainers of Official Datasets
- Work as a group
- Bring the community which works on their data closer to them/
direct communication
- See and take advantage of ENGAGE Data Curation Community
work (e.g. cleansing, better formats)
- Easy to see / gather all the Applications that are based on their
official datasets.
- See the impact of their datasets.
- Understand which datasets have RE-USE value for users.
- Community Help in the process of Digitalization and Opening of
current or older Public Data (history dimension)

Search for a dataset...

The screenshot shows the Engage website interface. The browser address bar displays www.engagedata.eu. The website header includes the Engage logo, navigation links (HOME, DATASETS, REQUEST DATA, COMMUNITY, OPEN DATA SITES), and user options (Sign in, Register). A search bar labeled "SEARCH FOR PUBLIC DATA" is prominently displayed. A language dropdown menu is open, showing options: English, Deutsch, Nederlands, Français, فارسی, Bahasa Indonesia, Ελληνικά, and עברית. The search bar and the language menu are circled in red.

...use your own language

Check dataset information...

Excel - Civil Service People Survey 2011-CF

The Civil Service People Survey (CSPS) is an annual survey open to all Civil Servants and those that work for Civil Service organisations. In 2011 299,410 Civil Servants across 97 organisations participated.

Category: Arts, Recreation and Travel, Banking, Finance and Insurance

Submitted by:
smouzakitis

Excel - Civil Service People Survey 2011-CF

Uploaded by smouzakitis

Extend

★★★★★ No rating yet

The Civil Service People Survey (CSPS) is an annual survey open to all Civil Servants and those that work for Civil Service organisations. In 2011 299,410 Civil Servants across 97 organisations participated.

Resources

Detailed Metadata

Excel Conversion

Size 214.0 KB

csp2011_allorganisations_20120201.xls



view

download

visualize

Basic Information

Revised dataset



Conversion to other format

Maintainer: [smouzakitis](#)

Publisher: None specified

Author: None specified

Release Date: March 29, 2013, 7:52 a.m.

Last Update: March 29, 2013, 7:52 a.m.


Country: None specified

Licence: None specified

Extension Graph

...and download it

Faceted search available...



The screenshot displays the data.gouv.fr search interface. On the left, a sidebar contains filter sections: Country (France 3927), Categories (Financials 2121, Education 543, Arts, Recreation and Travel 271, Agriculture 133, Census 125), Publisher (data.gouv.fr 3925, regardscitoyens.org 1), and License (French Open License 3925, Licence Ouverte (LO) 1, French Open Licence 1). A red oval highlights these filter sections. The main content area features a search bar, a 'Clear all filters' button, and a list of 3927 datasets found. Three dataset entries are visible, each with a title, description, update date, and the data.gouv.fr logo.

Country

France (3927)

Categories

Financials (2121)
Education (543)
Arts, Recreation and Travel (271)
Agriculture (133)
Census (125)

Publisher

data.gouv.fr (3925)
regardscitoyens.org (1)
(1)

License


French Open License (3925)
Licence Ouverte (LO) (1)
French Open Licence (1)

Search bar: 

✖ Clear all filters

3927 datasets found

Plan de Relance 2009 - Crédits budgétaires par région Updated:16-06-13
Détail des crédits budgétaires allouées dans le cadre du plan de relance de l'économie 2009.Sont distingués les autorisations d'engagement et les crédits de paiement, ainsi que leur répartition par secteur et par région....


Programme du séminaire prospective info "L'aménagement du territoire à l'international : de l'identification de pratiques innovantes à leur appropriation" Updated:16-06-13
Programme détaillé du séminaire précisant le sujet de la journée et la qualité des intervenants....


Parc naturel régional - PNR Updated:16-06-13
Les parcs naturels régionaux (PNR) concourent à la politique de protection de l'environnement, d'aménagement du territoire, de développement économique et social et d'éducation et de formation du public. Ils constituent un cadre privilégié ...


...with several filters

Extend the datasets...

Party positions for the May and June 2012 parliamentary elections in Greece

Uploaded by smouzakitis

★★★★★ No rating yet

This dataset includes expert estimates on the policy positions of 14 Greek political parties on more than 40 policy issues as well as scaled positions on the left-right and pro/anti-memorandum dimensions. Expert estimation was conducted through the Del

Extend

Basic Information

Maintainer: [smouzakitis](#)
Publisher: Data Archiving and Networked Services
Original url: [Original Dataset page](#)
Author: Gemenis, Dr. K. (University of Twente) Trika, Dr. V.

Extend Dataset

Revision type:

- Please choose an extension typ ▾
- Please choose an extension type..
- Conversion to other format
- Data Cleansing
- Metadata Enrichment
- Data Enrichment
- Snapshots of real-time data
- Combination of Datasets / Mash-ups

...in several ways...

...and keep the provenance information

Uploaded by smouzakitis

Extend

★★★★★ No rating yet

The Civil Service People Survey (CSPS) is an annual survey open to all Civil Servants and those that work for Civil Service organisations. In 2011 299,410 Civil Servants across 97 organisations participated.

Resources

Detailed Metadata

UTF-8 Encoding changed

Size 6.9 KB

cspcs2011_allorganisations_20120202.csv

view

download

visualize

Revised dataset

Combination of Datasets / Mash-ups

Maintainer: [smouzakitis](#)

Publisher: None specified

Author: None specified

Release Date: April 3, 2013, 7:31 a.m.

Last Update: April 29, 2013, 10:19 a.m.

Country: None specified

Licence: None specified

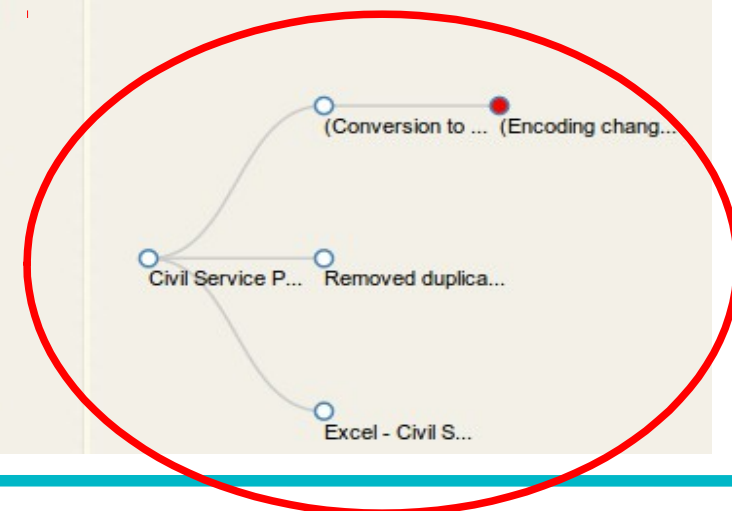
Items based on this dataset

+ Submit your own Item

Discussions

Leave a reply

Extension Graph



BETA

engage

HOME DATASETS REQUEST DATA COMMUNITY

Home | Datasets | Civil Service People Survey 2011

Civil Service People Survey 2011

Uploaded by smouzakitis

★★★★☆ Score: 3.0 / 1 votes

The Civil Service People Survey (CSPS) is an annual survey open to all Civil Servants and those that work for Civil Service organisations. In 2011 299,410 Civil Servants across 97 organisations participated.

Resources Detailed Metadata

Size 204.4 KB

Civil-Service-People-Survey-2011-Civil-Service-benchmark-final.pdf

view visualize

Extend Dataset

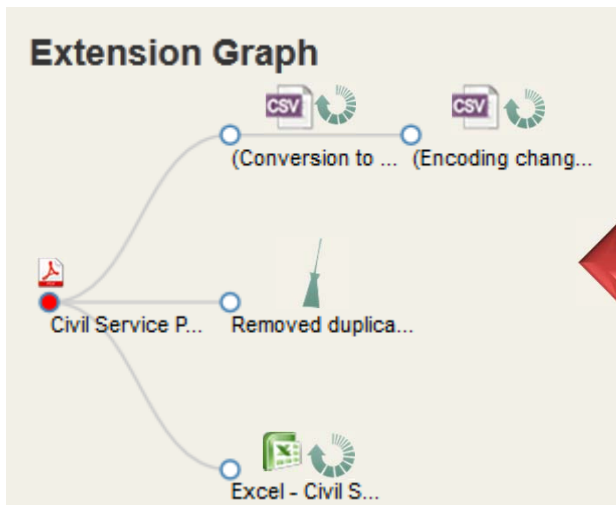
Revision type:

- Data Cleansing
- Please choose an extension type..
- Conversion to other format
- Data Cleansing**
- Metadata Enrichment
- Data Enrichment
- Snapshots of real-time data
- Combination of Datasets / Mash-ups

Data Cleansing

Data cleansing, data cleaning or data scrubbing is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database. Used mainly in databases, the term refers to identifying incomplete, incorrect, inaccurate, irrelevant, etc. parts of the data and then replacing, modifying, or deleting this dirty data

You can read more about this [here](#)



FaceFilter undo/redo

1 - 10 of 19 rows

Refresh Browse by rows records Reset All Remove All change

Lang Detect

2 choices sorted by name count

facet by choice counts

	Name	Facebook ID	Lang Detect	Facebook URL	Fans
1	7 Nesnáám soyč ská7u z baráku a dojde mi pávu?na ai 1 19 ?	105911579453004		http://www.facebook.com/pepsi-nesnam-soy-z-baraku-a-dojde-mi-pavu-na-ai-1-19-?	2805
2	Ano, šéf!	53837383556	["confidence":0.092383504,"language":"cs","isReliable":false]	http://www.facebook.com/anoasef	123163
3	Blene Maja	34901000985	["confidence":0.51053035,"language":"de","isReliable":true]	http://www.facebook.com/pages/Blene-Maja/45015000985	5490
4	Big Brother 10 - Die officiele Fanpage	202560057443		http://www.facebook.com/BioBrother10Cieffah	14971
5	Bob a Bobek	4002423573		http://www.facebook.com/pages/Bob-a-Bobek/5022423573	19120
6	Comeback	16853810046	["confidence":0.8317397,"language":"cs","isReliable":true]	http://www.facebook.com/comebackfan	156430
7	Comedystreet	44287783709	["confidence":0.048393045,"language":"de","isReliable":false]	http://www.facebook.com/pages/Comedystreet/44287783709	14305
8	Fraggles	58353972410	["confidence":0.51778406,"language":"de","isReliable":true]	http://www.facebook.com/pages/Fraggles/58353972410	17505
9	Galleo	57541544534	["confidence":0.08777907,"language":"de","isReliable":false]	http://www.facebook.com/Galleo	264725
10	GAME ONE	50207717075		http://www.facebook.com/m/gamesone	112297
11	Hyde Park	207375037633	["confidence":0.6087329,"language":"cs","isReliable":true]	http://www.facebook.com/hydeparkcz	8785
12	Im Sick of cash For Gold Advert!	150005551182		http://www.facebook.com/pages/Im-Sick-of-cash-For-Gold-Advert!/150005551182	15512
13	Konec? funk?i zpřesb jak sledovat všechny televizní stanice a v7a zdarma	12217094479850	["confidence":0.41852877,"language":"cs","isReliable":false]	http://www.facebook.com/pages/Konec-funk-i-zpřesb-jak-sledovat-všechny-televizni-stanice-a-v7a-zdarma/12217094479850	70458
14	Krkonoské pohádky	39671570005	["confidence":0.252422,"language":"cs","isReliable":false]	http://www.facebook.com/pages/Krkonoske-pohadky/39671570005	22652
15	Lindenastraße	177512484076	["confidence":0.2849542,"language":"de","isReliable":false]	http://www.facebook.com/Lindenastrasse	18294
16	MTV Home	402549730374	["confidence":0.2833688,"language":"de","isReliable":false]	http://www.facebook.com/pages/MTV-Home/402549730374	91299
17	MTV Switzerland	13303517911	["confidence":0.8096947,"language":"de","isReliable":true]	http://www.facebook.com/MTV-Switzerland	6900
18	NewKids	12354548431329	["confidence":0.35466942,"language":"de","isReliable":false]	http://www.facebook.com/newKidsOfficial	48063
19	Showtime (ABS CBN)	173116234940		http://www.facebook.com/pages/Showtime-ABS-CBN/173116234940	982584

Describe the metadata...

Add new dataset

Add new dataset

Add new dataset

Basic Information

Coverage

Scientific Context

Scientific domain of the dataset
Electrical engineering, electro... ▼

Data Collection type

Data Collection Description

Statistical Software package used

Analysis Unit

Statistical Methodology

Title

Description

Data

Basic Information

Organizations

Coverage

Scientific Context

Data

Basic Information

Organizations

Coverage

Scientific Context

Data

Basic Information

Organizations

Coverage

Scientific Context

Join the community...

Community Users



Zoe Christidou
Archivist / Librarian

Other natural sciences
Greece

View Profile



vb
Citizen

Other humanities
United Kingdom (Great Britain)

View Profile

Home | **valcas**

Profile summary

My Groups

Inbox

Show me:



All groups



Groups I administrate



Groups I am a member of

Create a new group

Group Name

Members

Role

Data for Public Administration Corruption

xalexopoulos (Charalampos Alexopoulos)
a.zuiderwijk.eijk (Anneke Zuiderwijk - Van Eijk)

Not A Member

ENGAGE Developers

skous (Sotiris Koussouris)
smouzakitis (Spiros Mouzakitis)
johncon
mpetyx2

Not A Member

French community

valcas (Valerie Brasse)

Owner

SQUID-Applications

WilliWichtig

Not A Member

..and create your groups

Rate the datasets...

Civil Service People Survey 2011

Uploaded by smouzakitis


★★★★☆ **Score: 3.0 / 1 votes**

The Civil Service People Survey (CSPS) is an annual survey those that work for Civil Service organisations. In 2011 299,4 organisations participated.


[Resources](#) [Detailed Metadata](#)

Discussions

1 reply

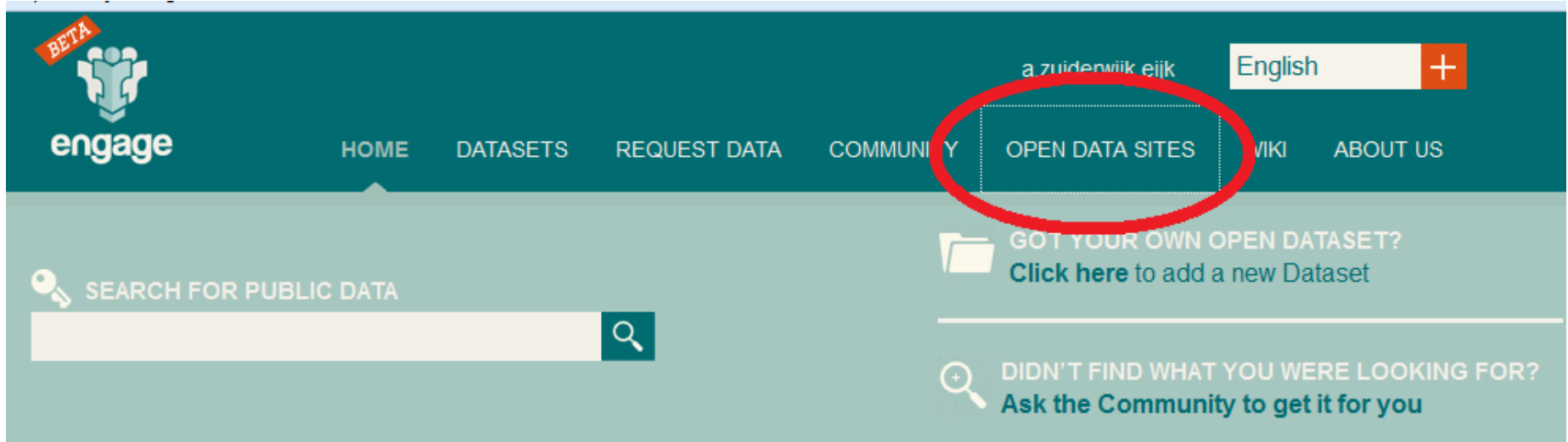
 smouzakitis replied 3 weeks, 1 day ago

The file is in SPSS format



..and share your thoughts

Find out about Open Data sites...



The image shows the homepage of the Engage website. The top navigation bar is dark teal and contains the Engage logo (a stylized group of people) with a 'BETA' badge, the text 'engage', and navigation links for HOME, DATASETS, REQUEST DATA, COMMUNITY, OPEN DATA SITES (circled in red), WIKI, and ABOUT US. To the right of the navigation bar, there is a language selector showing 'English' and a plus sign. Below the navigation bar, the main content area is light teal. On the left, there is a search bar with the text 'SEARCH FOR PUBLIC DATA' and a magnifying glass icon. On the right, there are two promotional boxes. The first box has a folder icon and the text 'GOT YOUR OWN OPEN DATASET? Click here to add a new Dataset'. The second box has a magnifying glass icon with a plus sign and the text 'DIDN'T FIND WHAT YOU WERE LOOKING FOR? Ask the Community to get it for you'.

...per country

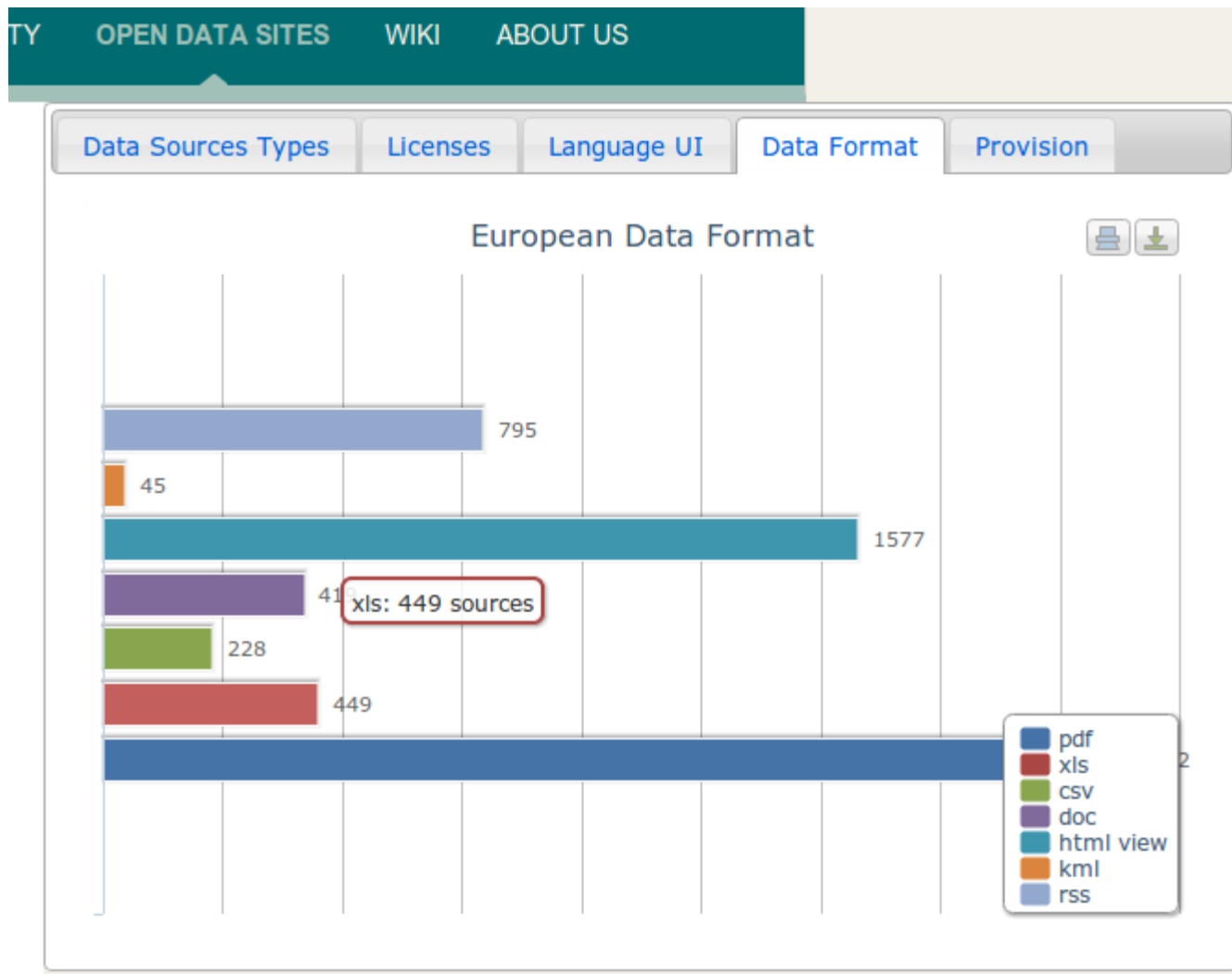
The screenshot shows the Engage website interface. At the top, there is a navigation bar with the Engage logo (a stylized 'E' with three figures) and a 'BETA' badge. The navigation menu includes: HOME, DATASETS, REQUEST DATA, COMMUNITY, OPEN DATA SITES, WIKI, and ABOUT US. On the right side of the navigation bar, there are links for 'Sign in' and 'Register', and a language selector set to 'English' with a plus sign icon.

The main content area features a map of Europe. A dropdown menu is open over the map, listing countries: Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Czech republic, Denmark, Estonia, European Union, Finland, France, German, Germany, Greece, Hungary, Iceland, Ireland, Israel, and Italy. Below the map, there are several filter dropdowns: Country, DataSource Type, Publisher, License, Category, and State of Data.

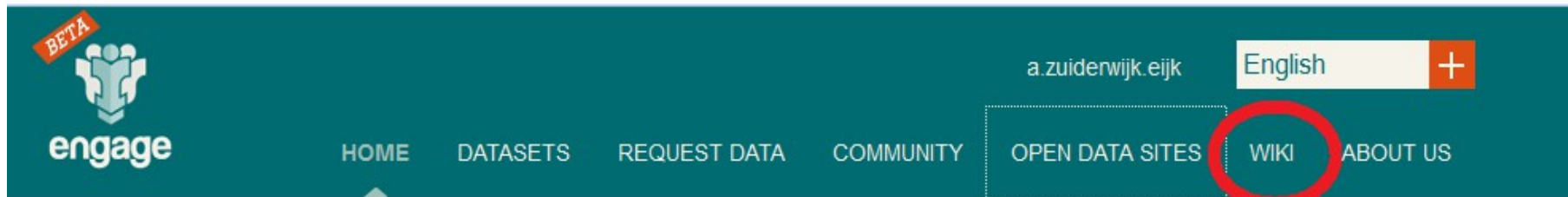
On the right side, there is a panel titled 'European Data Sources T'. It contains a horizontal bar chart showing the number of data sources for various countries. The chart has a 'Feedback' button on the right side. The data points from the chart are as follows:

Country	Number of Data Sources
Austria	61
Belgium	716
Bulgaria	65
Cyprus	8
Czech Republic	168
Czech republic	54
Denmark	16
Estonia	45
European Union	574
Finland	34
France	129
Germany	15

...or other criteria



Learn more about...



Wiki Home

/ ENGAGE DATA Platform / Data Curation Methods /



Data Curation Methods data_curation

[View](#)

[ENGAGE Extension / Curations types](#)

[The market of data curation and wrangling](#)

[Google Refine](#)

[Import a dataset](#)

[Import a dataset from the ENGAGE portal](#)


[Navigation and Access to common Operations](#)

[Adding a column based on another column](#)

[Data transformations](#)

[Data cleansing](#)

[How to fix string data type issues](#)

 **GOT YOUR OWN OPEN DATASET?**
[Click here](#) to add a new Dataset

 **DIDN'T FIND WHAT YOU WERE LOOKING FOR?**
[Ask the Community](#) to get it for you

...ENGAGE,
the ENGAGE API
and Data Curation
Methods

When downloading a dataset, you should always do a quality check against a number of p
claims that its quality is high and there are no quality issues with the contents. Apart from

Value Proposition through individual tools

- Search in diverse and dispersed data sources in EU supported by ENGAGE
- Be able to transform your datasets keeping the valuable information with the ENGAGE external tools (Open Refine, Scrapperwiki etc.)
- See your results through visualisation tools
- Structure your data according to your needs – control all the levels of your dataset (data, metadata, format)
- Refine existing datasets by metadata enrichment

Value Proposition through collaboration

- Create your community(ies) with members of mutual interests
- Each community will be able to increase the value of its data sets by applying their own perspectives based on its unique needs
- Upload your work and share it with your community
- Find other data sets, valuable for your work, uploaded by your community (Collaborate / Exchange / Ask / Provide)
- Combine their results with yours – make new datasets

ENGAGE Technology Stack

