

The Widukind project

A unique database for national
and international economic data

July 2016

CEPREMAP

CENTRE POUR LA RECHERCHE ECONOMIQUE ET SES APPLICATIONS

The birth of the project

Originally a tool to improve data handling processes in macroeconomic research, complementary to DYNARE (a free software to build economic models), with three objectives:

- ▶ Simplifying retrieval of economic data
- ▶ Automatically providing updated data
- ▶ Allowing reproducible results

Increasingly, objectives go beyond academia:

- ▶ Public/Private Sector: economic indicators, panels, graphs, ...
- ▶ Journalism: data journalism, fact-checking, ...
- ▶ Civil society: professors, associations, bloggers, ...

A partnership between the CEPREMAP and France Stratégie (and the financial support of the Investments for the Future Programme)

Objectives

Create a free, open-source server to aggregate publicly-available data series provided by national and international statistical institutions.

Value added: a unique economic database with wide, systematic coverage of economic data (70 million series at present)

Four important principles of the project:

- ▶ Data series are taken directly from providers and kept unchanged
- ▶ Data series are stored in a tree similar to the provider's
- ▶ Data series are automatically updated via provider-specific functions
- ▶ Archive system: each revision of the data series is archived

A free platform designed for every user

Why use it ?

- ▶ A website: <http://widukind.cepremap.org>
with different search tools depending on needs
- ▶ An API: <http://widukind-api.cepremap.org>
to easily access each data series from statistical software (R, Python, Julia, Octave, Matlab, LibreOffice Calc, Excel, SAS, Eviews, Stata)
- ▶ Free and Open source (GNU Affero General public License) to allow
the creation of a Widukind community

Providers already included

- ▶ BIS - Complete : 324,199 series
- ▶ ECB - Complete : 475,607 series
- ▶ Eurostat - Economy; population; labor : 60,741,121 series
- ▶ INSEE - BDM : 159,549 series
- ▶ IMF - Complete : 2,203,735 series
- ▶ ESRI - Japanese National Accounts : 1,597 series
- ▶ Fed - Board of Governors - Complete : 28,440 series
- ▶ OECD - MEI and EO : 81,154 series
- ▶ World Bank - GEM and WDI : 337,944 series
- ▶ BEA - NIPA and Fixed Assets : 33,434 series

Plugins already included

Three available formats : SDMX; JSON; HTML
(API: <http://widukind-api.cepremap.org>)

- ▶ Python plugin:
<https://github.com/Widukind/widukind.py.pandaSDMX>
- ▶ R plugin : <https://github.com/Widukind/widukind.R.sdmx>
- ▶ Matlab plugin : <https://github.com/Widukind/widukind.m>
- ▶ Excel, LibreOffice Calc, and Eviews plugins

Planning

Providers:

- ▶ Next step: Banque de France
- ▶ Planned: BLS, UN, ILO, Germany, United Kingdom, Italy, Canada, Australia, Spain, Sweden, Belgium, China, India...

Plugins :

- ▶ Next step: Julia
- ▶ Planned: Octave, Stata, SAS

Questions

- ▶ Differences in targeted users (i.e. not only within academia)
- ▶ Governance of the project
- ▶ Maintenance and financing of the (open-source) project in the middle/long term

Technical details and contacts

- ▶ Website : <http://widukind.cepremap.org>
- ▶ API REST : <http://widukind-api.cepremap.org>
- ▶ Code : <https://github.com/Widukind>
- ▶ Questions : <https://github.com/Widukind/widukind/issues>
- ▶ Examples : <http://obsmacro.cepremap.fr/>