

# Open data at Statistics Netherlands

**Michiel Eijkhout**

**2016 Workshop on the Dissemination and Communication  
of Statistics, 5-7 October, The Hague**



**Statistics  
Netherlands**

# Overview presentation

- How do we define Open data and how did we get started?
- History of the database of Statistics Netherlands
- Implementation of Open data
- Involvement of users of Open data
- Products and examples
- Lessons learned
- Issues and discussion

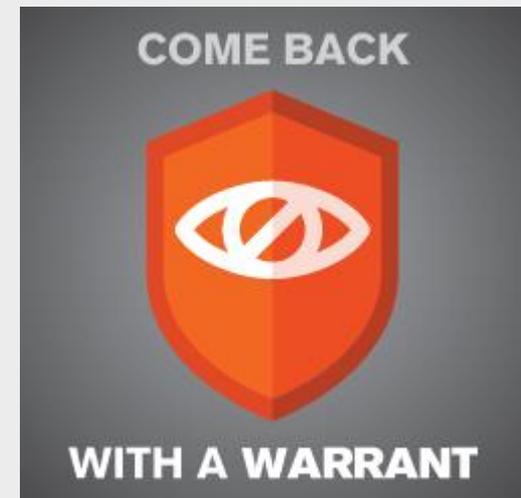


# What is Open data?

Open Data is (government) data made available electronically in such a way that programmers can use them directly for various applications.

## Open data are

- Public data
- Free of licence
- Funded by public means
- Machine-readable
- Using open IT-standards
- Respecting privacy



# Timeline StatLine database

<b>Circa 1990</b>	First statistics database and interface
<b>1990-2000</b>	Slow expansion of StatLine database
<b>2000</b>	StatLine becomes primary channel for all figures
<b>2001-2009</b>	Rapid expansion of StatLine database
<b>2010</b>	First experiment with web services; iPhone app
<b>2011</b>	Improvement output database: metadata as xml files; data as text files; no SQL
<b>2012</b>	Decision to implement Open data for StatLine
<b>2012-2014</b>	Development of Open data

# Project approach

Demand for web services had existed for several years.

Starting impulse: government policy to have Open data in 2015.

Development by IT department in close contact with some key users:

- Research institutions;
- Small IT companies working on data visualisation tools.

Some key users selected for feedback during first stages.

Input before, during and after introduction of Open data.

LinkedIn group established for contact with users.

Organisation of user meetings twice a year.



# Road to Open data

- 2012**     **Proof of concept**  
Microsoft Azure Data market  
Open standard for Open data queries
- 2013**     **Beta version**  
Version with 150 small and popular tables  
Positive feedback from customers
- 2014**     **Final release**  
All StatLine tables (3,700 at the time)  
App: beta-version interface using Open data
- After**     Fine-tuning and expansion of Open data facility



# LinkedIn group

The screenshot shows a LinkedIn group page for 'Centraal Bureau voor de Statistiek; Open Data' with 626 members. The group is led by Erik Willems, a developer at LocalFocus. A post from Erik Willems, dated 2 months ago, is titled 'Publieke demo koppeling CBS > LocalFocus'. The post text reads: 'De koppeling tussen LocalFocus en API CBS draait nu ruim een jaar stabiel. Petje af! Hier een openbare demo: <http://playground.localfocus.nl/?public> en druk op CBS-datasets. Met ons interne dashboard kunnen we in minuten nieuwe koppelingen met andere datasets maken.'

The post includes a screenshot of a dashboard titled 'LocalFocus Playground'. The dashboard features a search bar and a grid of 12 data visualization charts. The charts display various statistical data, including line graphs, bar charts, and a map of the Netherlands. The dashboard interface includes a navigation menu on the left with options like 'Nieuw', 'Mijn profiel', and 'Groepen'. The top of the dashboard shows the user's name 'Erik Willems' and the group name 'Overseen CBS dataset'.

On the right side of the LinkedIn page, there is a section titled 'ABOUT THIS GROUP' with the description: 'Platform voor uitwisseling van kennis over Open Data.' Below this, there is a 'MEMBERS' section showing 626 members and a row of profile pictures. A blue button labeled 'Invite others' is visible below the member list.

[link](#)

7



# Open data community Netherlands

User meetings at Statistics Netherlands, twice a year.

Central government site for all Open data sets:

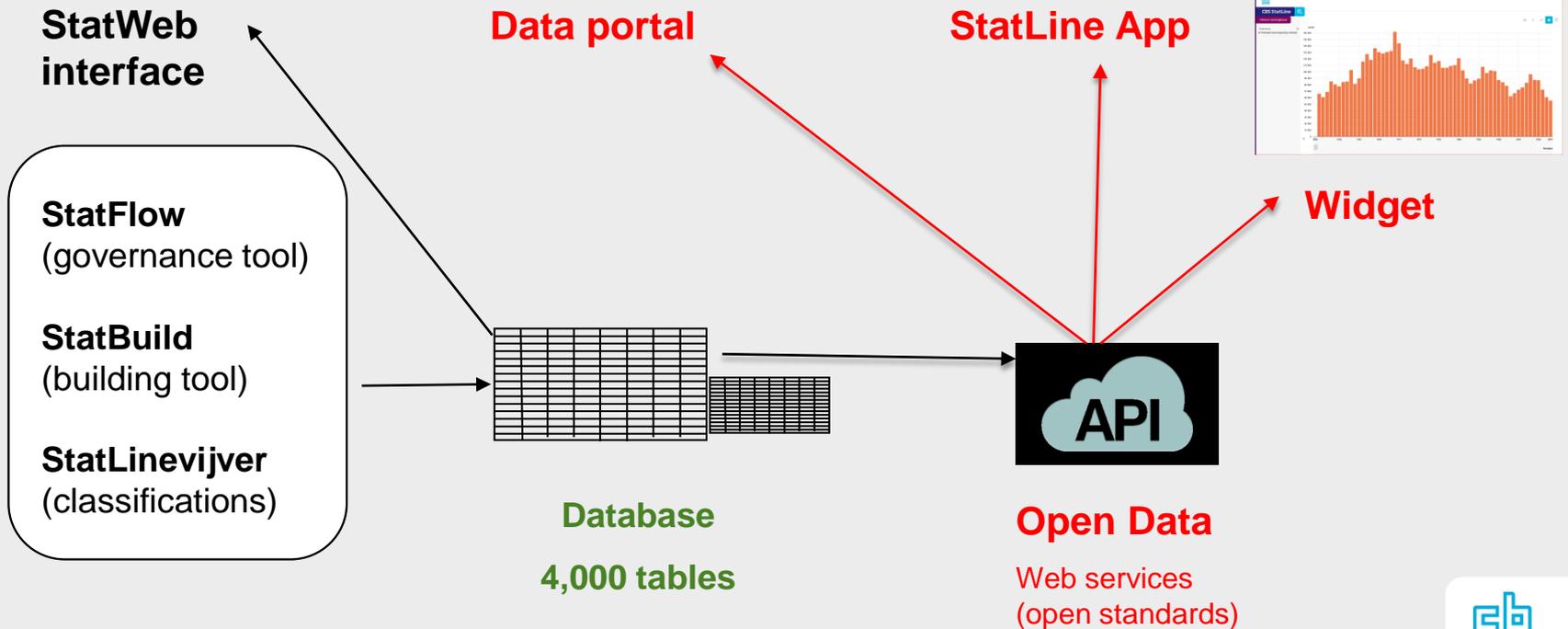
<https://data.overheid.nl/>

This site gives access to all data sets of Statistics Netherlands. Owned by the Ministry of the Interior.

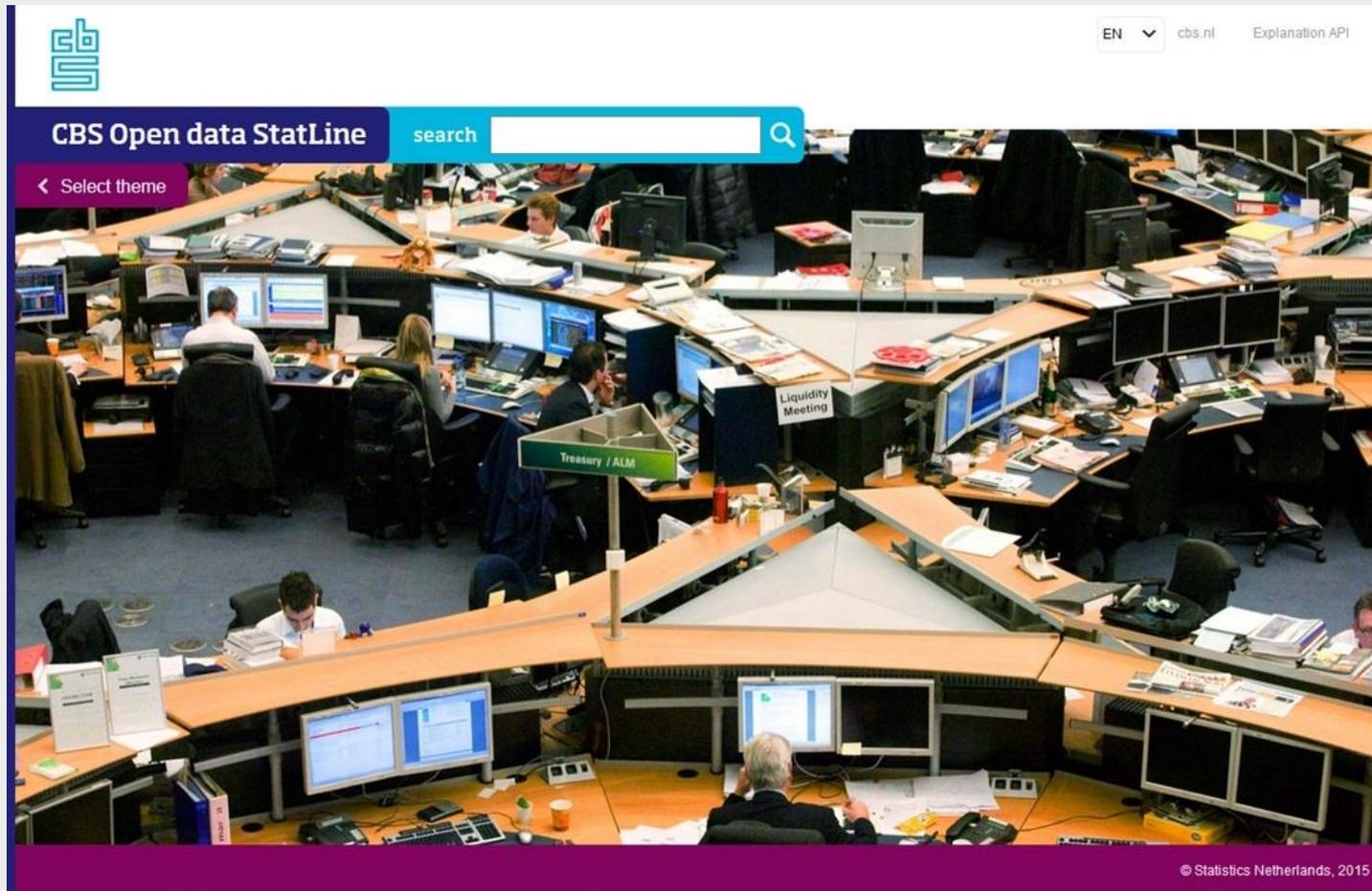
Universities, data brokers, VNG, municipalities, IT companies, financial institutions, data scientists, health institutions and so on.

Many other meetings on Open data: congresses, workshops, hackatons, app-building contests, etc.

# System setup



# Data portal (homepage)

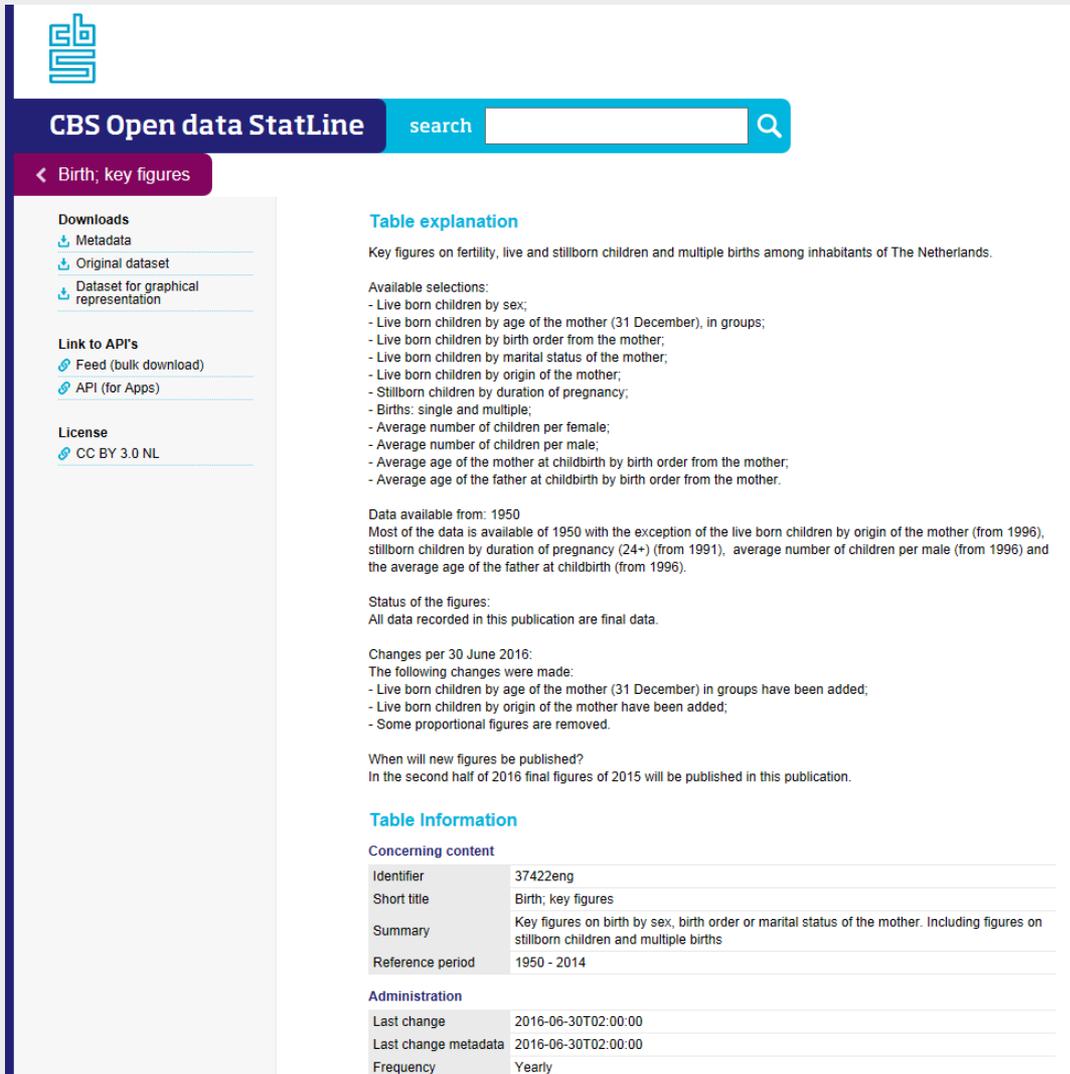


[link](#)

10



# Data portal (table)



The screenshot shows the CBS Open data StatLine interface. At the top left is the CBS logo. Below it, the text 'CBS Open data StatLine' is displayed next to a search bar. A navigation breadcrumb shows '< Birth, key figures'. On the left side, there are sections for 'Downloads' (Metadata, Original dataset, Dataset for graphical representation), 'Link to API's' (Feed (bulk download), API (for Apps)), and 'License' (CC BY 3.0 NL). The main content area is titled 'Table explanation' and contains the following text:

**Table explanation**

Key figures on fertility, live and stillborn children and multiple births among inhabitants of The Netherlands.

Available selections:

- Live born children by sex;
- Live born children by age of the mother (31 December), in groups;
- Live born children by birth order from the mother;
- Live born children by marital status of the mother;
- Live born children by origin of the mother;
- Stillborn children by duration of pregnancy;
- Births: single and multiple;
- Average number of children per female;
- Average number of children per male;
- Average age of the mother at childbirth by birth order from the mother;
- Average age of the father at childbirth by birth order from the mother.

Data available from: 1950  
Most of the data is available of 1950 with the exception of the live born children by origin of the mother (from 1996), stillborn children by duration of pregnancy (24+) (from 1991), average number of children per male (from 1996) and the average age of the father at childbirth (from 1996).

Status of the figures:  
All data recorded in this publication are final data.

Changes per 30 June 2016:  
The following changes were made:

- Live born children by age of the mother (31 December) in groups have been added;
- Live born children by origin of the mother have been added;
- Some proportional figures are removed.

When will new figures be published?  
In the second half of 2016 final figures of 2015 will be published in this publication.

**Table Information**

Concerning content

Identifier	37422eng
Short title	Birth; key figures
Summary	Key figures on birth by sex, birth order or marital status of the mother. Including figures on stillborn children and multiple births
Reference period	1950 - 2014

Administration

Last change	2016-06-30T02:00:00
Last change metadata	2016-06-30T02:00:00
Frequency	Yearly

[link](#)



# Data portal (Central government)



Ministerie van Binnenlandse Zaken en  
Koninkrijksrelaties

## Dataportaal van de Nederlandse overheid

[Home](#) [Data](#) [Dataverzoeken](#) [Ondersteuning](#) [Over Ons](#) [Contact](#)

### Dataportaal van de overheid

U vindt in dit portaal een overzicht van alle beschikbare en herbruikbare data van de Nederlandse overheid.

#### Actueel:

Meest bekeken datasets (05/07/16)  
Kamerbrief open data (22/06/16)  
High Value Datasets (11/05/16)  
Trendrapport Open Data (24/04/16)  
TenderNed Datasets (26/03/16)



#### Voor hergebruikers:

Zoek beschikbare datasets  
Doe een dataverzoek  
Bekijk openstaande verzoeken

### Zoek datasets

Zoeken

**Totaal aantal datasets: 7885**

met open data licentie 7881  
met andere licentie 4

**Kwaliteit**

Laatste aanpassing 09-09-2016

### Blader in datasets op thema (thema's nog niet compleet)

Bestuur (469)

Cultuur en recreatie  
(247)

Economie (729)

Financiën (44)

Huisvesting (21)

Internationaal (10)

Landbouw (122)

Migratie en integratie  
(1)

Natuur en milieu (1445)

Onderwijs en  
wetenschap (86)

Openbare orde en  
veiligheid (8)

Recht (12)

Ruimte en infrastructuur  
(52)

Sociale zekerheid (174)

Verkeer (348)

Werk (7)

Zorg en gezondheid (55)

[link](#)

12



# Examples of Open data (1)



The screenshot shows the website 'De Staat van het MKB'. At the top, there is a navigation bar with links for 'Nieuws', 'Feiten & cijfers', 'Database', and 'Over Staat van het MKB'. Below this is a large banner image of a man in a blue shirt looking up at a pink flower. The main content area is titled 'Nieuws' and includes a 'Filters:' section with buttons for 'Publicaties', 'Onderzoek', 'Internationaal', 'Gerelateerd', and 'Achtergrond'. There are two featured articles:

- Article 1:** 'De nieuwste MKB data in een overzichtelijk dashboard'. It features a speedometer icon and a 'Dashboard' button. The text below reads: 'Op de Staat van het MKB vindt u naast relevante artikelen over het MKB ook een dashboard met de nieuwste CBS-data over het midden- en kleinbedrijf.'
- Article 2:** 'Directe en indirecte rol MKB bij export beter in kaart gebracht'. It features a lightbulb icon, the date '14 juni 2016', and an 'Onderzoek' button. The image below shows two workers in red uniforms in a factory setting.

[link 1](#)

[link 2](#)

# Examples of Open data (2)

Blog Contact Products & services ▾ Tutorials ▾ Login ▾ NL EN



## LocalFocus

Turn datasets into stories

Get new insights from datasets and share them with colleagues and your audience. Publish graphs and maps on your website and use them for print. Simple and fast.

Get to work with LocalFocus



What can you create with LocalFocus?



[link](#)

# StatLine App

**CBS StatLine**

EN Feedback cbs.nl Help

**In the spotlight**

**Search by theme**

**Search by search term**

**StatLine** is the database of Statistics Netherlands and offers a wealth of data on the economy and Dutch society. You can compile, adjust and download your own graphs and tables

© Statistics Netherlands, {}

[link](#)

15



# Open data for third parties

Statistics Netherlands has a second web service for third parties, such as other government organisations.

Most of these are financed by one of the Ministries.

- JM (Youth monitor)
- IV3 (Financial information from local governments)
- MKB (Information aimed at small and medium-sized businesses)
- MLZ (Monitor long-term care)
- RIVM (Public health and health care)

Custom-made open data service. Some only have a web service, some have a data portal, some have the app or a widget.



# Lessons learned: practical feedback

- Tables need to be up 24/7.
- Re-design of a table has an unpredictable effect on existing Open data applications. Adding or changing topics of a table in particular turns out to be disruptive.
- Codes of topics and categories become more important. 'Machine-readable' means using codes is standard, not free-format texts, as is common with a user interface. Standardising classifications gets more priority: same label, same code, same description.
- Explanation of API (application programming interface) necessary. User manuals are important, e.g. for Excel Power Pivot.



# Lessons learned

- The possibilities for combining data are huge
- Open data has become a buzz word
- Using your own service is a good way of learning
  
- Open data services are meant for developers, not for the average internet surfer
- Open data user questions are technical, not content-related
- Data from Statistics Netherlands are mainly used as a basic layer, to which data from other sources may be added
  
- Developing a business model for Open data products is difficult;
- General applications are rare; usually a commissioning party hires an IT company for a well-defined product or project



# A look into the future

What could lie  
around the  
corner?



Open data road

# Issues

- Business models → no money, no investments
- Goal for Open data in 2015 has not been attained
- Not all government institutions (such as municipalities) are ready for Open data
- Financial investment is for government institutions, but the financial revenues are harvested by Open data users
- Intangible revenues for Statistics Netherlands are plenty: better use of available data, more brand awareness
- Use of Open data is difficult to measure. One query on a table may lead to thousands of data views.... or none.
- Government data are open, but actual users are closed; this may result in a *take-the-data-and-run* approach to Open data.
- Metadata of tables are not machine-readable.



# Discussion

OPEN DATA

