Open data, API

Annegrete Wulff
awu@dst.dk

STATISTICS
DENMARK

Open data

• Accesability

| On-line |
| Free |
| Detailed |
| Metadata |
| Commercial and non-commercial |

… and machine readable

STATISTICS
DENMARK

Nothing new …

- 2003 PSI (Public Sector Information) Directive
  - *The PSI Directive requires that re-use is subject to   fair and non-discriminatory conditions.*
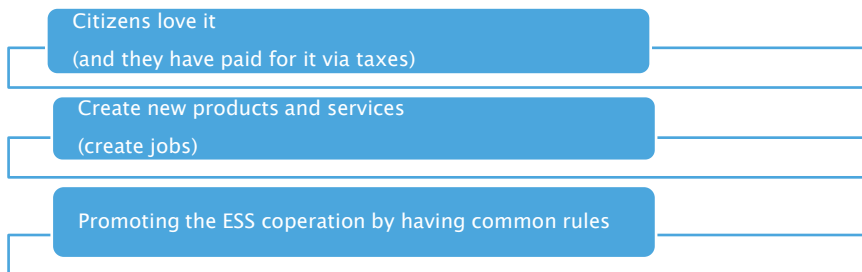 - *commercial or non-commercial re-use.*

Statistics Denmark's policy

- Free re-use of data from www.dst.dk and the StatBank is permitted as long as *the source* is acknowledged.

- The commercial re-use of complete publications and newsletters is permitted only upon written agreement from Statistics Denmark.

- Statistics Denmark shall not be held responsible for any errors or delays and neither for direct or indirect losses or consequential damage to the user.

STATISTICS DENMARK

Why open data ?

Citizens love it
(and they have paid for it via taxes)

Create new products and services
(create jobs)

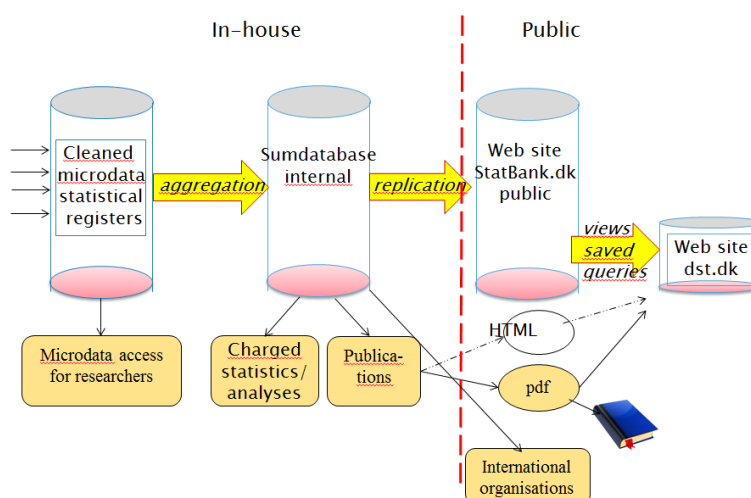Promoting the ESS coperation by having common rules

What's in?

- Data (aggregated statistical data) available from the web site

- Statistical data retrieved automatically in machine readable form (e.g. API, web services, bulk download)

- Metadata

- Descriptions on the website (e.g. texts, graphics)

STATISTICS
DENMARK

What's not in?

- Microdata (regardless of confidentallity)

- Personal data (e.g. personal addresses)

- Third party data (e.g. maps, photos, geographiccal coordinates)

- Tailored retrievals

- Publications

Information flow

STATISTICS DENMARK

Solutions since 2005

Register
- Register as a user in the StatBank (free) (All official statistics are there)

Select
- Make your selection from tables

Save
- Save the retrieval (the query) for later use

Re-use
- The query gets a unique id, that can be referenced later. E.g an id, 85858, can be used in xml

- www.statbank.dk/xml/85858 data in xml
- www.statbank.dk/xls/85858 gives the same data in Excel.
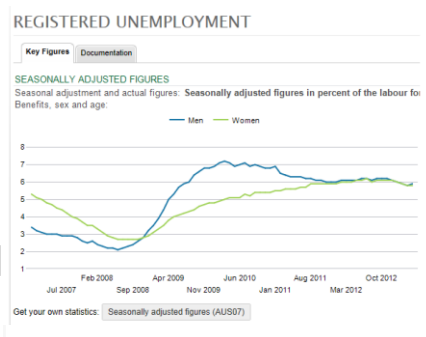

In-house use #1

- 200 topic sites built and automatically maintained via 300-400 saved queries and web service

EMPLOYED PERSONS
Area of workplace: All Denmark | Unit: Number

	2008	2009	2010	2011	2012
Commuting, total	2 857 565	2 801 519	2 684 992	2 667 424	2 673 552
No commuting	231 267	185 395	179 754	179 537	179 933
Up to 5 km	900 641	893 927	850 497	839 326	831 435
5-10 km	516 490	504 525	480 940	475 630	475 064
10-20 km	534 814	526 920	502 036	497 832	501 216
20-30 km	265 069	266 368	256 558	257 684	261 358
30-40 km	138 313	139 127	135 943	137 145	140 002
40-50 km	77 755	78 539	76 789	78 331	80 459
More than 50 km	185 939	199 832	195 522	195 315	197 210
Unknown	7 277	6 886	6 953	6 624	6 875

Footnote ≛

In connection with the publishing of register-based labour force 2012 the statistics have been reviewed for the years 2009-2011. This matrix will also be reviewed for 2009-2011 as soon as possible.

Get your own statistics: Employed persons (AFSTA1)

REGISTERED UNEMPLOYMENT

Key Figures | Documentation

SEASONALLY ADJUSTED FIGURES
Seasonal adjustment and actual figures: Seasonally adjusted figures in percent of the labour for Benefits, sex and age:

— Men — Women

Get your own statistics: Seasonally adjusted figures (AUS07)

In-house use

- 42 indicators

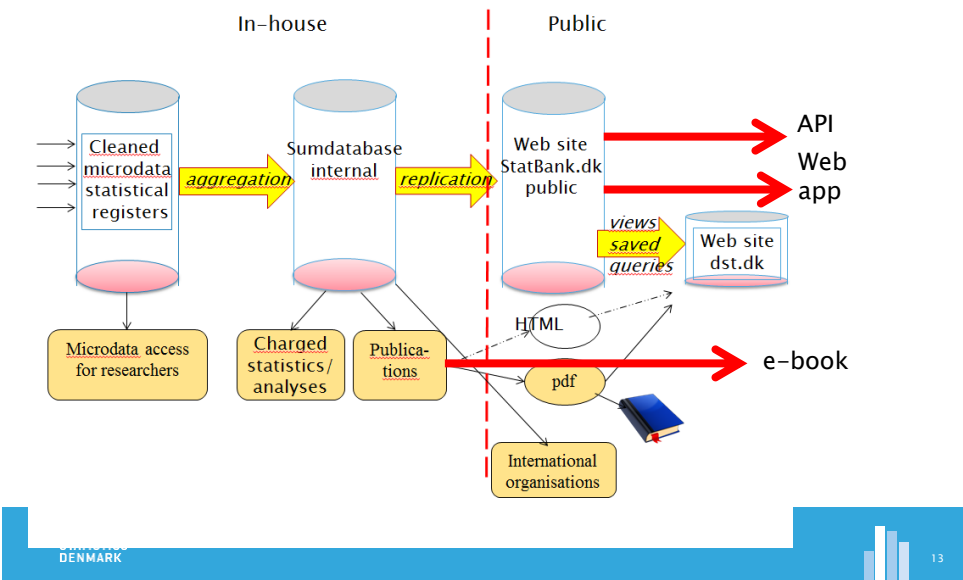**SHORT TERM TRENDS**

	Period	Value	Dev.
National accounts			t/(t-1)
Gross domestic product, real growth (DKK million, 2005-price level chain figures)	Q1/13	387,164	0.2
Private consumption, real growth (DKK million, 2005-price level chain figures)	Q1/13	194,411	0.1
Government consumption, real growth (DKK million, 2005-price level chain figures)	Q1/13	107,152	-2.1
Investments in buildings and civil engineering works, real growth (DKK million, 2005-price level chain figures)	Q1/13	29,814	-2.1
Other gross fixed capital formation, real growth (DKK million, 2005-price level chain figures)	Q1/13	42,097	0.7
Imports of goods and services, real growth (DKK million, 2005-price level chain figures)	Q1/13	204,686	1.6
Exports of goods and services, real growth (DKK million, 2005-price level chain figures)	Q1/13	214,802	-0.4
Employment. National accounts (1000 persons)	Q1/13	2,788	-0.3
Labour market			t/(t-4)
Jobs. WTA (number)	Q4/12	2,890,593	-0.1
Jobs. Hours worked, WTA (1000 hours)	Q4/12	924,467	-0.9
Indices of average earnings in the private sector (1. qtr. 2005=100)	Q4/12	123.7	1.4
Indices og average earnings in the public sector, central government (1. qtr 2005=100)	Q1/13	126.6	2.0
Indices of average earnings in the public sector, local government (1. qtr. 2005=100)	Q1/13	126.1	0.6
Registered unemployment			t/(t-12)
Gross unemployment (converted into full-time) per cent of the labour force [2]	Apr/13	5.9	-0.1

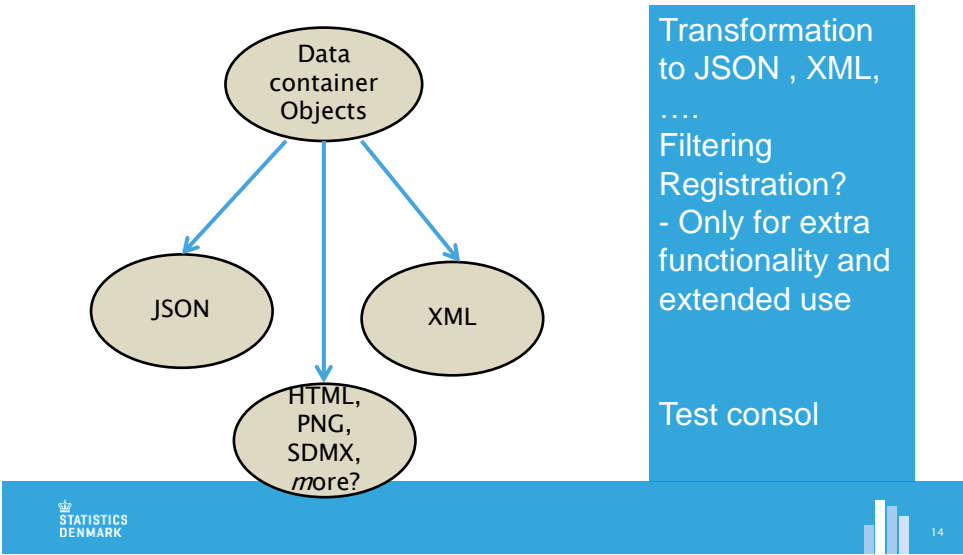## ... new possibilities are coming up

- API (Application Programming Interface)

Access to all StatBank data via the API will ease the use and re-use

Information flow

API

Use case – several filters – an example

- 1 The client sends HTTP request to the API

  2 Has the client specified any filtering?

  - 2.1 If the client did not specify filtering

    - 2.1.1 Make a selection of all the tables that have not been marked as discontinued

  - 2.2 If the client has specified filtering

    - 2.2.1 Start with all the tables

    - 2.2.2 If the client has specified filtering on topic

      - 2.2.2.1 Exclude tables that do not exist at any level in the hierarchy under the specified topics

Usecase #2

- 2.2.3 If the client has specified filtering on update date

  - 2.2.3.1 Exclude tables that are not updated within the specified number of days back from the expressed date

- 2.2.4 If the client specified filtering on discontinuation status

  - 2.2.4.1 If the client has not expressed filtering on discontinuation status

    - 2.2.4.1.1 Exclude tables, which are marked as discontinued

  - 2.2.4.2 If the client has specified filtering on discontinuation status

    - 2.2.4.2.1 Exclude tables that do not have the specified discontinuation status

3 Return the complete selection of tables

1) Vælg API-funktion:

Hent tabeller ▼

Henter tabeloversigt. Filtrering på emner gøres ved brug af koder fra funktionen "subjects".

2) Vælg parametre for funktion:

Type af kald POST ▼
Sprog Engelsk ▼
Format JSON ▼
Filtrér på emne (kommasepareret liste) _____
Hent kun tabeller opdateret indenfor et antal dage 30
Inkludér afsluttede tabeller ☐

3) Kald API'et med de valgte parametre:

http://api.statbank.dk/v1/tables

```
{
    "lang": "en",
    "pastdays": 30,
    "format": "JSON"
}
```

Udfør

17

4) Se svaret fra API'et:

Formatér JSON ☐

[{"id":"FOLK1","text":"Population at the first day of the
quarter","updated":"2/11/2014","firstPeriod":"2008Q1","latestPeriod":"2014Q1","active":true,"vari
ables":["region","sex","age","marital status","ancestry","country of
origin","citizenship","time"]},{"id":"FOLK2","text":"Population 1.
January","updated":"2/11/2014","firstPeriod":"1980","latestPeriod":"2014","active":true,"variable
s":["age","sex","ancestry","citizenship","country of origin","time"]},
{"id":"FOLK3","text":"Population 1.
January","updated":"2/20/2014","firstPeriod":"2008","latestPeriod":"2014","active":true,"variable
s":["day of birh","birth month","year of birth","time"]},{"id":"FT","text":"Population figures
from the
censuses","updated":"2/11/2014","firstPeriod":"1769","latestPeriod":"2014","active":true,"variabl
es":["national part","time"]},{"id":"BEF5","text":"Population 1.
January","updated":"2/11/2014","firstPeriod":"1990","latestPeriod":"2014","active":true,"variable
s":["sex","age","country of birth","time"]},{"id":"BEF5F","text":"People born in Faroe Islands
and living in Denmark 1.
January","updated":"2/11/2014","firstPeriod":"2008","latestPeriod":"2014","active":true,"variable
s":["sex","age","parents place of birth","time"]},{"id":"BEF5G","text":"People born in Greenland
and living in Denmark 1.
January","updated":"2/11/2014","firstPeriod":"2008","latestPeriod":"2014","active":true,"variable
s":["sex","age","parents place of birth","time"]},{"id":"KM1","text":"Population at the first

18

9

4) Se svaret fra API'et:

Formatér JSON ☑

```
      "text": "Population at the first day of the quarter",
      "updated": "2/11/2014",
      "firstPeriod": "2008Q1",
      "latestPeriod": "2014Q1",
      "active": true,
      "variables": [
         "region",
         "sex",
         "age",
         "marital status",
         "ancestry",
         "country of origin",
         "citizenship",
         "time"
      ]
   },
   {
      "id": "FOLK2",
      "text": "Population 1. January",
      "updated": "2/11/2014",
```

Statistics Denmark will use the API

- Mobile StatBank (Smart phones)
- Apps (web Apps)

- Who else will use it?
- Which statistics will be interesting?
- Do we loose control?

STATISTICS DENMARK

Spreading use and knowledge

- Pre-announce potential API users
  - We know some of them
- Agreements with test users of a beta version
- Information on the web site with examples