The Danish Basic-Data Initiative and the Data-Distributor in Service

Presentation for the
Joint PCC and EuroGeographics CLRKEN webinar
Copenhagen, June 16th 2020

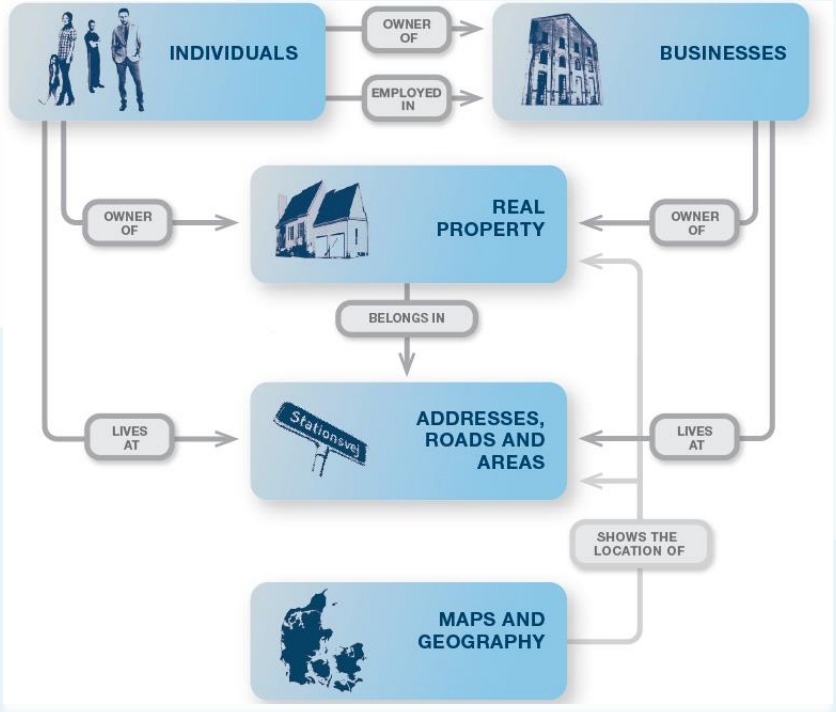
Agency for
Data Supply and
Efficiency

Leif Hernø, Chief-Consultant and
Senior Project Manager
Data Distributor Project

Introduction

- The Danish public authorities register core information about
  - individuals,
  - businesses,
  - real properties and buildings,
  - addresses, maps and geography.
- The information, called Basic-Data, is used every day throughout the public sector.
- High-quality Basic-Data is an essential basis for the authorities to perform their tasks properly and efficiently across units and sectors.
- Basic-Data is used in a wide range of areas also in the private sector: i.e. in real estate, insurance, financial and telecom sectors.

The relational model of the Basic-Data registers:



What was the problem?

- The need for modernization:

  Data structure and content reflected the needs of the public sector at the time of development,

- Poor data quality:

  Data was not necessarily coherent and could be redundant from one register to the other,

- Non-consistent modelling and documentation of data:

  Difficult to combine and exchange data between registers,

- Complex data-distribution:

  Request- and service-structures differed from one register to the other and it was difficult to gain an overview of the infrastructure and the integrations,

- Complex data-management:

  Data was distributed through a series of channels owned by the public sector or outsourced to private IT-companies,

- Costly data usage:

  The use of Basic-Data was in most cases payable.

What were the aims of the Basic-Data Initiative?

The Initiative aimed at securing:

- free, fast and reliable access to public Basic-Data for authorities, companies and citizens,
- high-quality and coherent basic-data updated once and for all at one place,
- increased effectiveness, modernization and high-quality public administration through standardized data and interfaces,
- innovation, growth and job creation in the private sector through opening up for new ways of using and combining public authoritative data,
- a reduction of costs through the use and the sharing of public data.

In this sense the Basic-Data Initiative aimed at more than just opening up to the public data collections.

The BC of the Basic-Data Initiative:

Public investment:

- The Basic-Data Initiative is a joint venture of the Danish Government, the municipalities (Local Government Denmark) and the Danish Regions.
- The investment amounts to 135 million Euro.

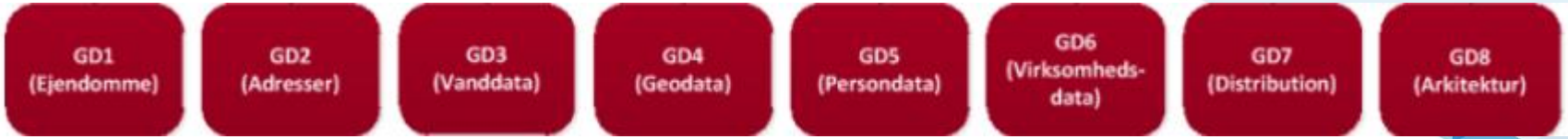
Financial benefits:

- The Basic-Data Initiative is expected to deliver economic benefits of 35 million Euro a year in the public sector.
- Private businesses are given the opportunity to use free Basic-Data in developing new solutions and products. Basic-Data Initiative thus contributes to growth and innovation in the private sector. Expected economic benefits: 70 million Euro a year.

The 4 pillars of the Danish Basic-Data Initiative:

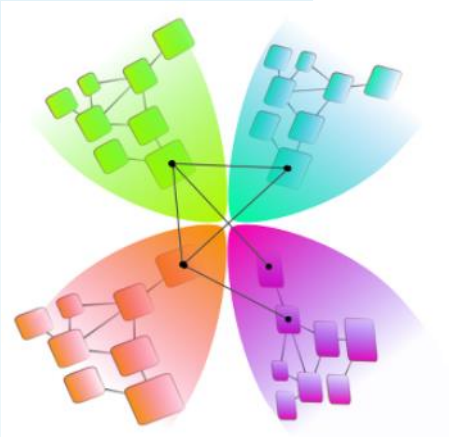
1. Making Basic-Data freely available for everyone.
2. Standardizing Basic-Data:
   - The Datamodelling Project (GD8)
3. Improving the quality of Basic-Data and adding new data, where needed:
   - The Real Property and Address Data Projects (GD1/GD2)
4. Improving the availability of Basic-Data through a shared and efficient distribution platform:
   - The Data-Distributor Project (GD7)

The objectives of the Initiative was achieved through a series of subprojects:



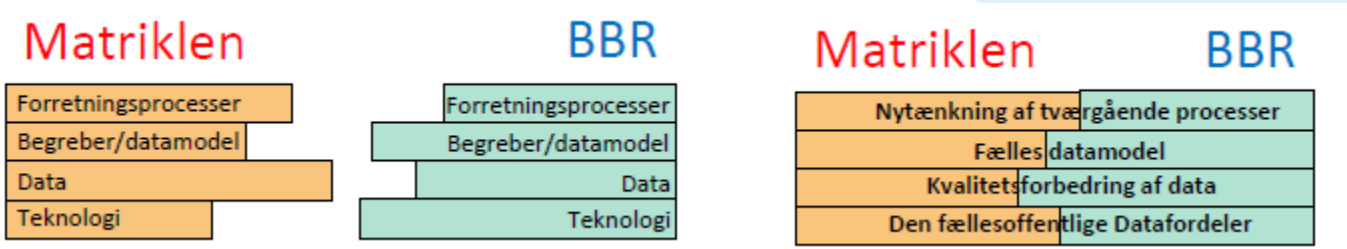
The Datamodelling Project:

- The project has established a coherent common data-model for all Basic-Data.
- The Basic-Data model was established through close cooperation between the Danish Agency for Digitisation and the Basic-Data authorities.
- The model gives an overview over data and relations between data.
- The model has made it possible to link data across the sectors.
- It is a characteristic feature of the model that objects and attributes are unique and can only be found in one Basic-Data register.
- The model is public.
- For more information look here: http://data.gov.dk/model/



The Real Property Data Project:

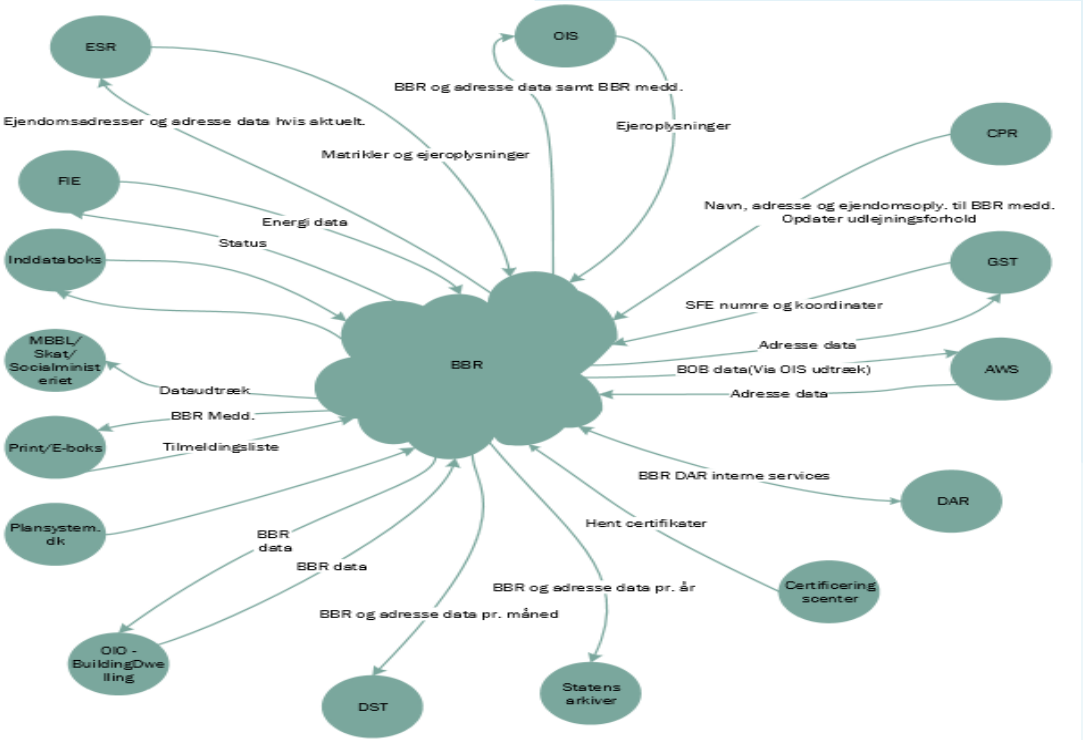
- Until now three registers contained data on real property: The Cadastre (Matriklen), the Building and Dwelling Register (BBR) and the Land Register (Tinglysningen). However, data was inconsistent and there was no unique identifier for identical objects.
- Through the Real Property Data project data was enhanced and made consistent, a.o. through the introduction of a unique ID for real property.
- At the same time the basic registration of all types of real property was made coherent and recorded in one register: The Cadastre.



The Data-Distributor Project:

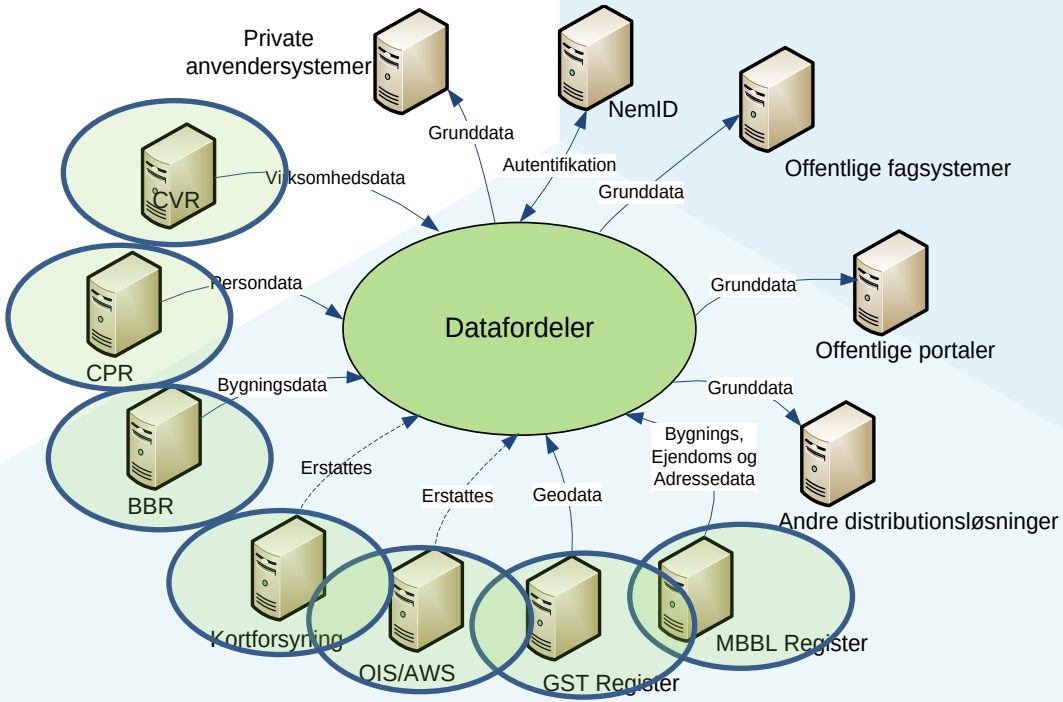
What is the Data-Distributor?

The illustration below shows as example the multiple and complex interfaces and integrations of just one Basic-Data-register before the Data-Distributor-Project:



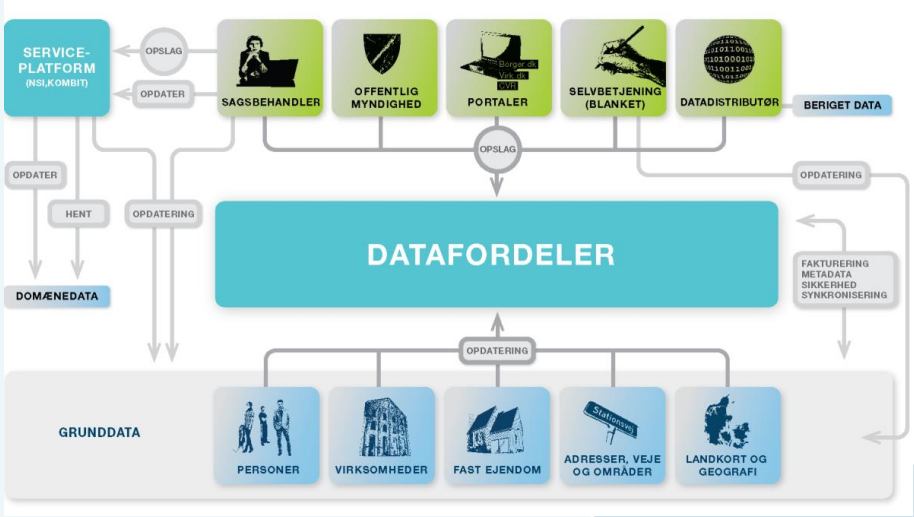
The Data-Distributor Project:

The project aims at establishing the Data-Distributor as the core single-point infrastructure element for the distribution of Basic-Data:



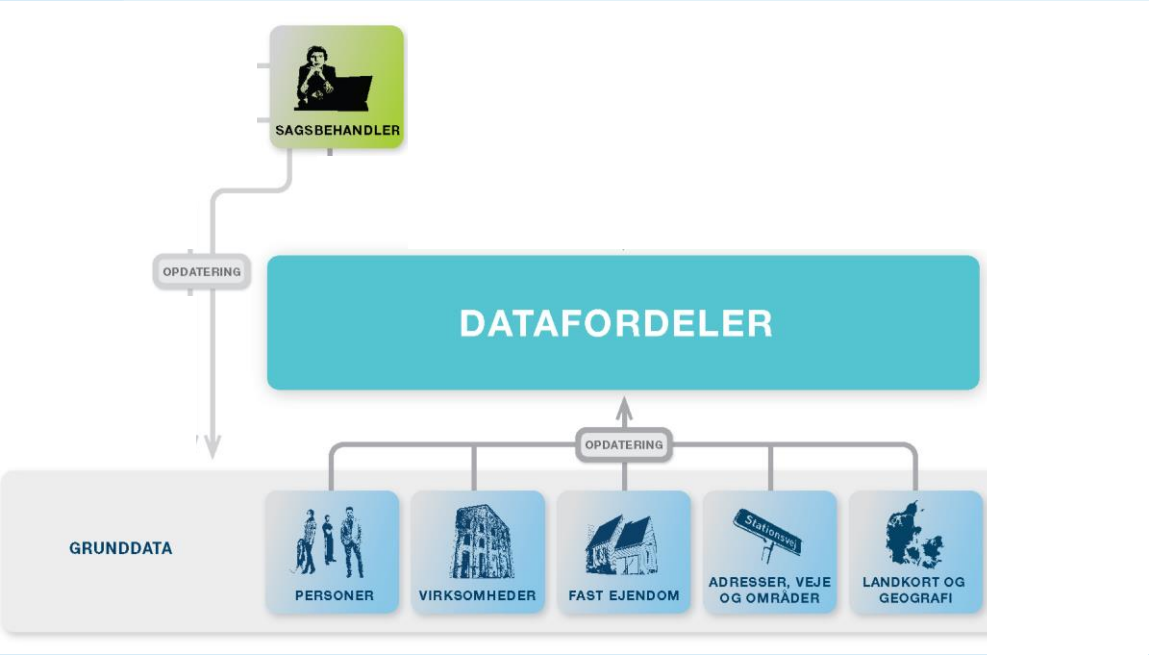
What can a user expect from the Data-Distributor?

- A secure and stable, scalable platform,
- High availability (99,9%) and high performance,
- Easy access to data through standardized interfaces,
- The possibility to compile data from different registers in a single service,
- Standardized service-types: online-requests, file-download,
- Distribution of events by changes of data (create, update, delete),
- Integration of data and the use of data,



The logic model of the Data-Distributor functionality

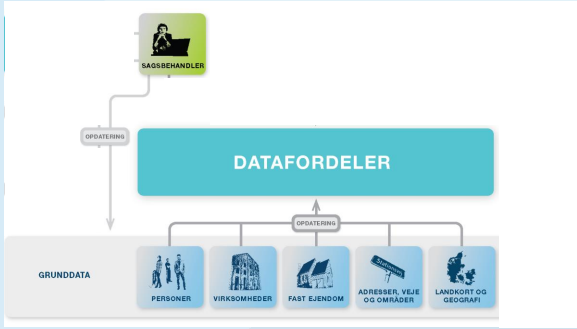
Part 1. The Basic-Data registers:



How does the Data-Distributor function?

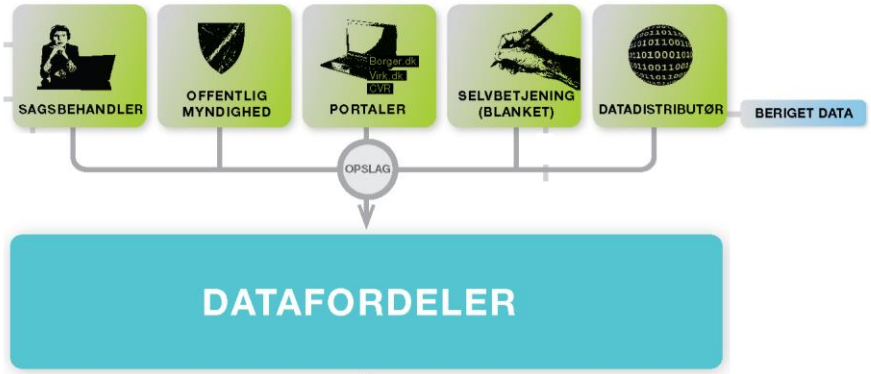
1. Registers, data and legal responsibility for the data:

- The administrative IT-systems of the public authorities (CPR, CVR, Matriklen, BBR etc.) continue to exist and the legislation behind them remains as a whole unchanged,
- The authorities own the registers and are responsible for the data, their collection and validity,
- The authorities are responsible for the operation, the maintenance and the support of the administrative systems,
- The authorities authorize the access to the data,
- As previously data is created, updated and deleted through the interfaces of the administrative systems, i.e. through updates by public servants,
- The authorities save all data as masterdata in their own systems,
- The authorities develop a dedicated interface for replicating data to the Data-Distributor,
- The authorities specify for the Data-Distributor to configure:
    - Datamodel,
    - Replication and synchronization of data,
    - Terms of use and security,
    - Services (rest-services, events, filedownload),
    - Metadata



The logic model of the Data-Distributor functionality

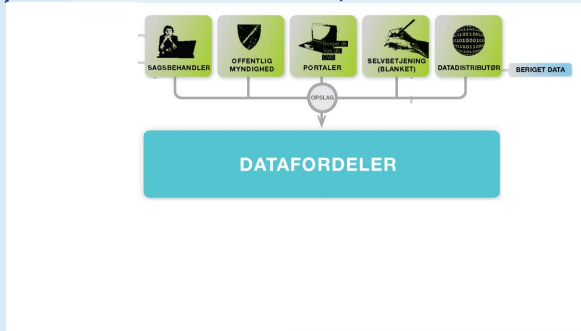
Part 2. The users:



How does the Data-Distributor function?

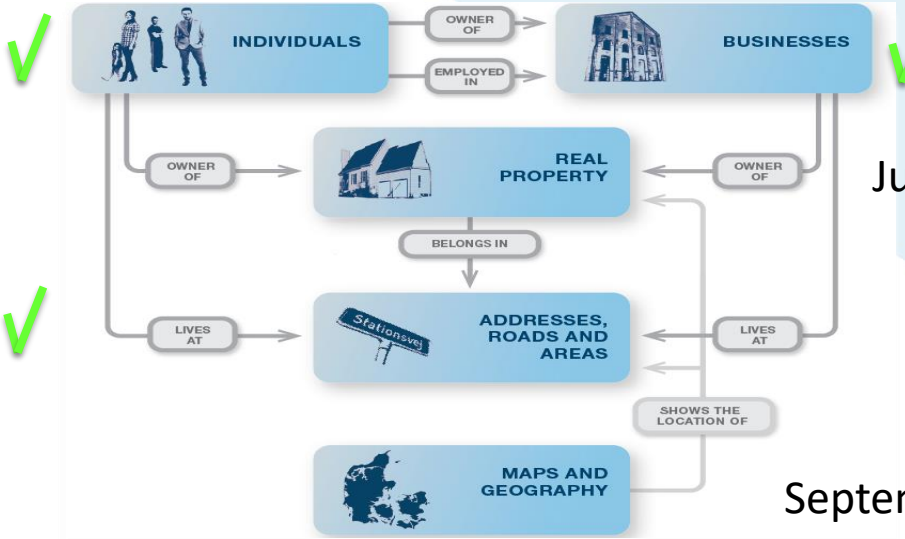
2. The users:

- The users have to register as users of the Data-Distributor,
- For access to certain data-collections the users will need authorization by the authority responsible for the data,
- To gain access to data, where access for legal reasons is restricted, the user will have to identify himself through i.e. password, certificates and/or IP-whitelisting,
- The users access data through the services, filedownloads and events, which the registers have specified,
- The users gain access to data through standardized web-requests,
- All users can – if authorized – get access to data, but communication with the Data-Distributor is “encoded”, as the Data-Distributor primarily is meant for system-to-system-communication,
- It is part of the genetic code of the Data-Distributor, that Basic-Data can be used by all users, among other things for the development of clients, user interfaces, intelligent solutions,
- For more information see here: http://datafordeler.dk



What is the status of the Data-Distributor development:

- The Data-Distributor has been developed and is in production displaying all planned Basic-Data-registers
- Data and services have been configured within all Basic-Data-areas
- The Data Distributor is now the one-stop distribution channel, giving access to a wide range of interoperable basic public data, ranging from information about individuals, businesses, real property, to geospatial data on buildings.



June 2019

September 2019

The Data-Distributor in production – status May 2020:

USAGE

SERVICES:	Oct.	Nov.	Dec.	Jan.	Feb.	March	April	May
Webservice-requests, mill.	*37,3*	*35,1*	*41,5*	*59,6*	*73,7*	*83,0*	*67,3*	*96,3*
Events, thous.	210.714	213.747	154.252	-	-	-	-	-
- Events Pull				122.919	112.875	120.052	114.218	153.848
- Events Push				71.050	61.603	73.948	63.938	73.782
Filedownloads, thous.	784	5.974	9.818	10.833	5.077	-	-	-
- Userdefined Small						6.610	6.951	7.445
- Userdefined Medium						2.156	2.447	2.922
- Userdefined Large						797	845	921
- Predefined						1.832	1.660	1.786

USERS:	Oct.	Nov.	Dec.	Jan.	Feb.	March	April	May
Unique, Active users	271	281	265	328	371	382	357	395

DATAIMPORTS, packages, thous.	Oct.	Nov.	Dec.	Jan.	Feb.	March	April	May
Real-time-updates	357.781	362.738	254.509	266.189	284.190	335.294	250.031	248.138
Batch-updates	31.840	32.436	27.024	35.906	37.895	40.546	28.356	35.299

The challenges?

- Time to market:
  - The IT-world changes fast and with this also the business needs of the registers – and of the users. As the core infrastructure element for the distribution of Basic-Data the Data-Distributor has to adjust to the development and meet the needs.
  - Changes to the data-structure of the registers can be complex and time-consuming because of the tight knit integration and the centralized distribution through the Data-Distributor.
- Governance:
  - The Basic-Data Initiative implies a change of business for the data-responsible authorities. They will have to find new ways of cooperating on inter-data questions. A strong governance-forum is necessary.
- Calibration:
  - Calibration includes the calibration of workflows in the integration between the registers, the calibration between the registers and the Data-Distributor, between the registers and the users as well as between the Data-Distributor and the users.
- Closing down the existing distribution channels:
  - The BC of the Basic-Data Initiative relies on the closing down of the existing distribution platforms. However: This has always been an extremely difficult job…

Questions?