STATISTICS DENMARK

EU Twinning Project JO 21 ENI ST 01 22

Component 1: Roadmap for the development of an integrated administrative data system in Jordan

Activity 1.2.1 Strategy for cooperation with administrative data owners for establishing a SBR 16-19 January 2023

Søren Andersen (ssa@dst.dk)
Director, Business Statistics
Statistics Denmark

Content

STATISTICS DENMARK

- Statistics Denmark’s business model and strategy
- The register based statistical system
- Zooming in on the Statistical Business Register
- The National Basic Data Initiative
- Usage of admin data in the production – an overview

STATISTICS DENMARK

Statistics Denmark's business model and strategy

Statistics Denmark's organisation

STATISTICS DENMARK

BOARD
Chair, Mrs Birgitte Anker

NATIONAL STATISTICIAN
DG, Mrs Birgitte Anker

MINISTRY
of Digitalisation

SECRETARIAT
Mr Simon Svith

ADMINISTRATION & IT Mr Torben Søborg	SOCIAL STATISTICS Mr Niels Ploug	BUSINESS STATISTICS Mr Søren S. Andersen	ECONOMIC STATISTICS Mrs Kirsten Balling	COMMUNICATION & SALES Mr Carsten Zangenberg
Finance	Population & Education	Business Dynamics	National & Environmental accounts	Communication
Service	Labor market & Income	Short Term Statistics	Public finance	Consult
HR	Social affairs & Health	Science, Tech. & Culture	Prices & Consumption	Survey
IT	Research services	Food Industries	External Economy	Method, analysis & Data Science
		Data Collection and SBR	Economic models	

Historical landmarks/milestones

1801: First Danish census

1849: First Danish constitution and Parliament

1850: Statistical Bureau established

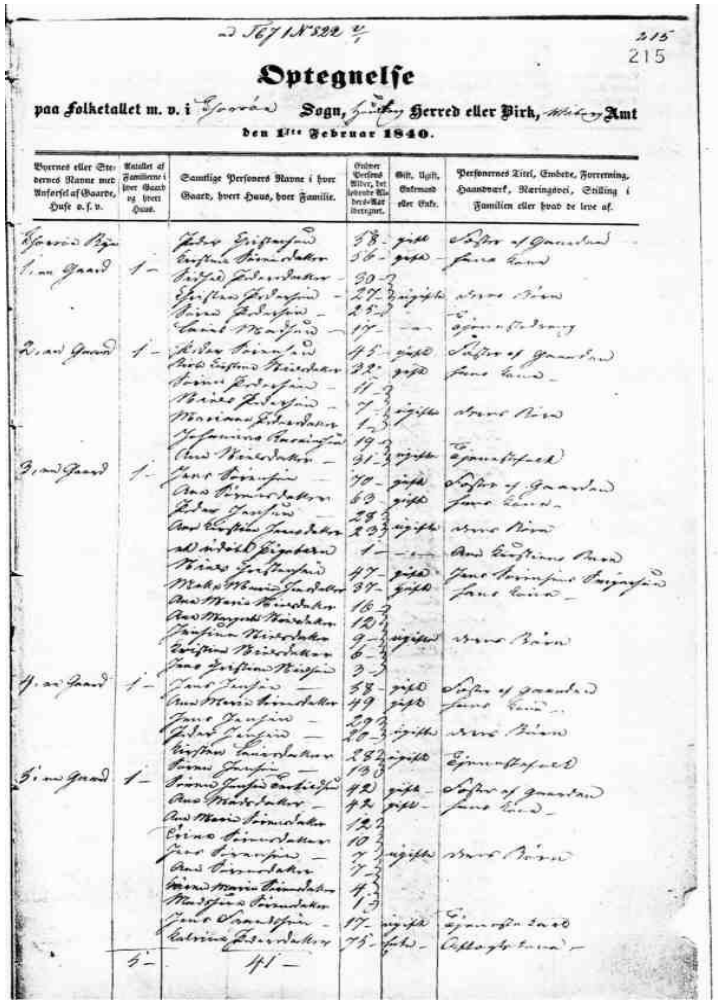
1966: Present law on SD adopted

1976: First register-based population census

2001: First ‘StatBank Denmark’ on the internet free of charge

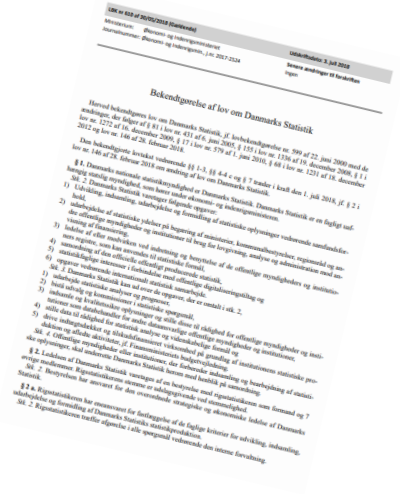
2001: Online internet access to microdata for researchers

2018: Latest revision of the Law on Statistics Denmark



The law on Statistics Denmark

STATISTICS DENMARK



a. **The purpose and obligation of SD (Art. 1)**
b. **The governance structure of SD (Art. 2)**
    a. the National Statistician, incl. for appointment and dismissal
    b. the Board and it's appointment and composition
    c. the Minister and government appropriations
c. **Compliance with EU legislation (Art. 3.a)**
d. **Digital reporting (Art. 3.b)**
e. **Publicly produced statistics and Official Statistics – i.e. coordination of the National Statistical System (Art. 4)**
f. **Access to administrative data (Art. 6)**
g. **Data confidentiality (Art. 7)**
h. **Mandatory data collection from businesses (Art. 8-12)**
i. **Legal basis for fining in case of refusal to provide data (Art. 13)**

Tasks of Statistics Denmark

STATISTICS DENMARK

Shall:

- Develop and produce statistics;
- Provide statistical services for lawmaking and public administration;
- Manage or assist in the design and use of the registers of public authorities and institutions that can be used for statistical purposes;
- Co-ordinate the official statistics;
- Safeguard professional interests in the statistical domain in connection with public digitalisation initiatives;
- Represent Denmark in international fora.

Can:

- Produce statistical analyses;
- Assist committees and commissions;
- Collect and perform quality assurance of data for public authorities;
- Make de-identified data available for statistical analyses and scientific purposes;
- Carry out commercial and grant-financed activities.

Statistics Denmark's Strategy 2025

STATISTICS DENMARK

MISSION

"We are the national supplier of independent statistics and data for society"

VISION

"We are the preferred source of reliable knowledge on Danish society"

FIVE CORPORATE VALUES

"Trustworthiness, transparency, user focus, adaptability and data accountability"

NINE STRATEGIC OBJECTIVES

The strategy pin-points three areas of strategic focus with nine goals in total ...

STATISTICS

We provide new and improved statistics for the development of Danish society

Added relevance

We will match the present-day needs of users – and for that reason, we will cover new areas of relevance to society, develop fast indicators and use new channels of communication.

High quality

We will ensure that the quality of our statistics, communication and documentation is of the highest international standard.

Stronger development

We will set an ambitious and bold development agenda, experiment and test new opportunities whenever the potential to do better presents itself.

DATA

We offer the most convenient and secure access to data on Danish society

One common entry point

We will create one common entry point to all data on society for researchers, analysts and authorities.

New data sources

We will find new data sources, secure their interaction with existing data, and translate them into new knowledge of relevance to society.

High data security

We are and will remain a leading, recognized and certified data environment, where security is always a primary consideration.

ORGANISATION

We are a competent and efficient organisation enabling others to succeed

Closer cooperation

We will initiate and enter into more mutually beneficial partnerships and strengthen our relationship with stakeholders and users.

New competences

We will develop and attract competent and dedicated employees and managers to turn our vision into reality.

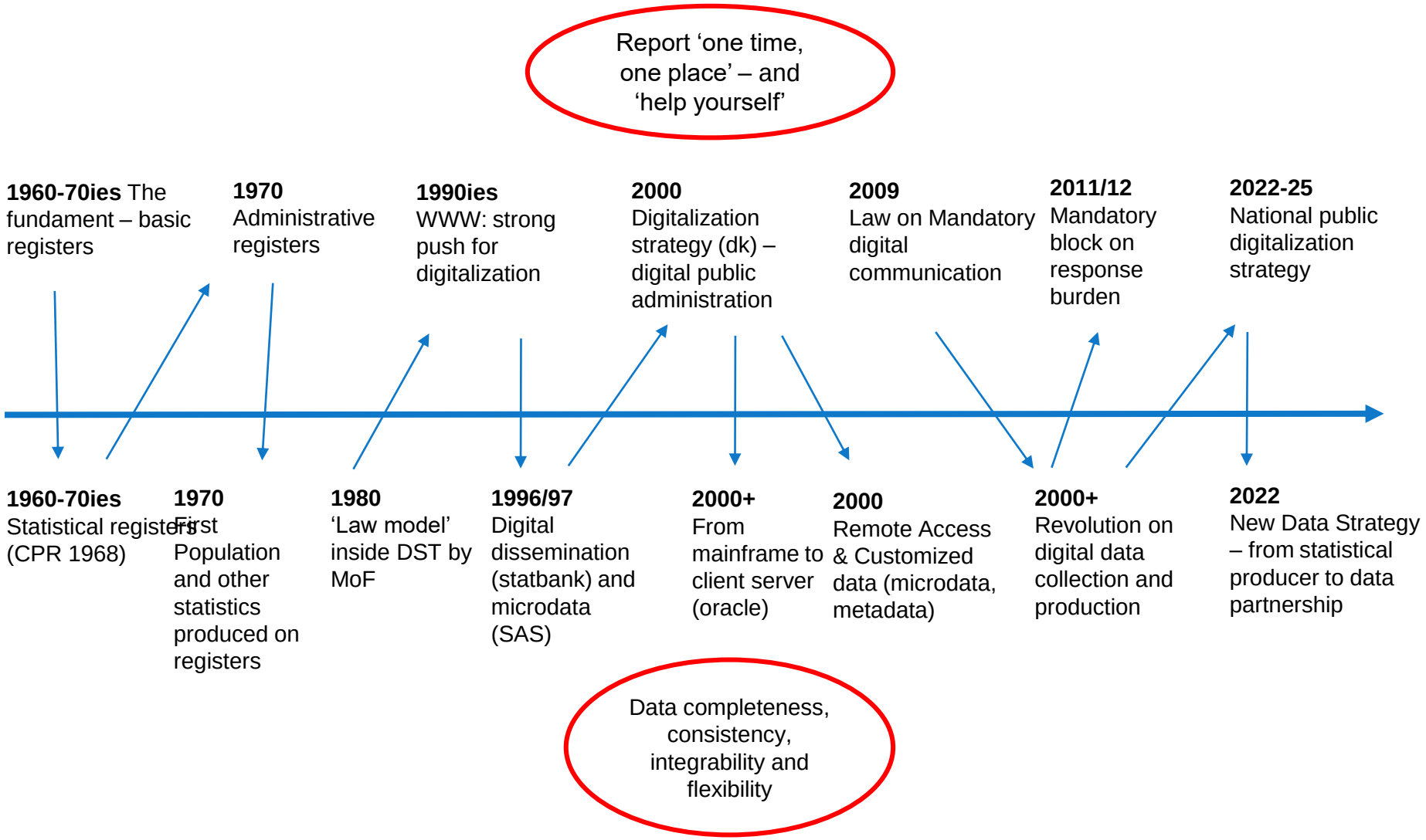
Increased efficiency

We will increase our efficiency using new technological options and work to ensure efficient use of our own resources as well as those of others.

STATISTICS DENMARK

The register based statistical system

The long and winding road

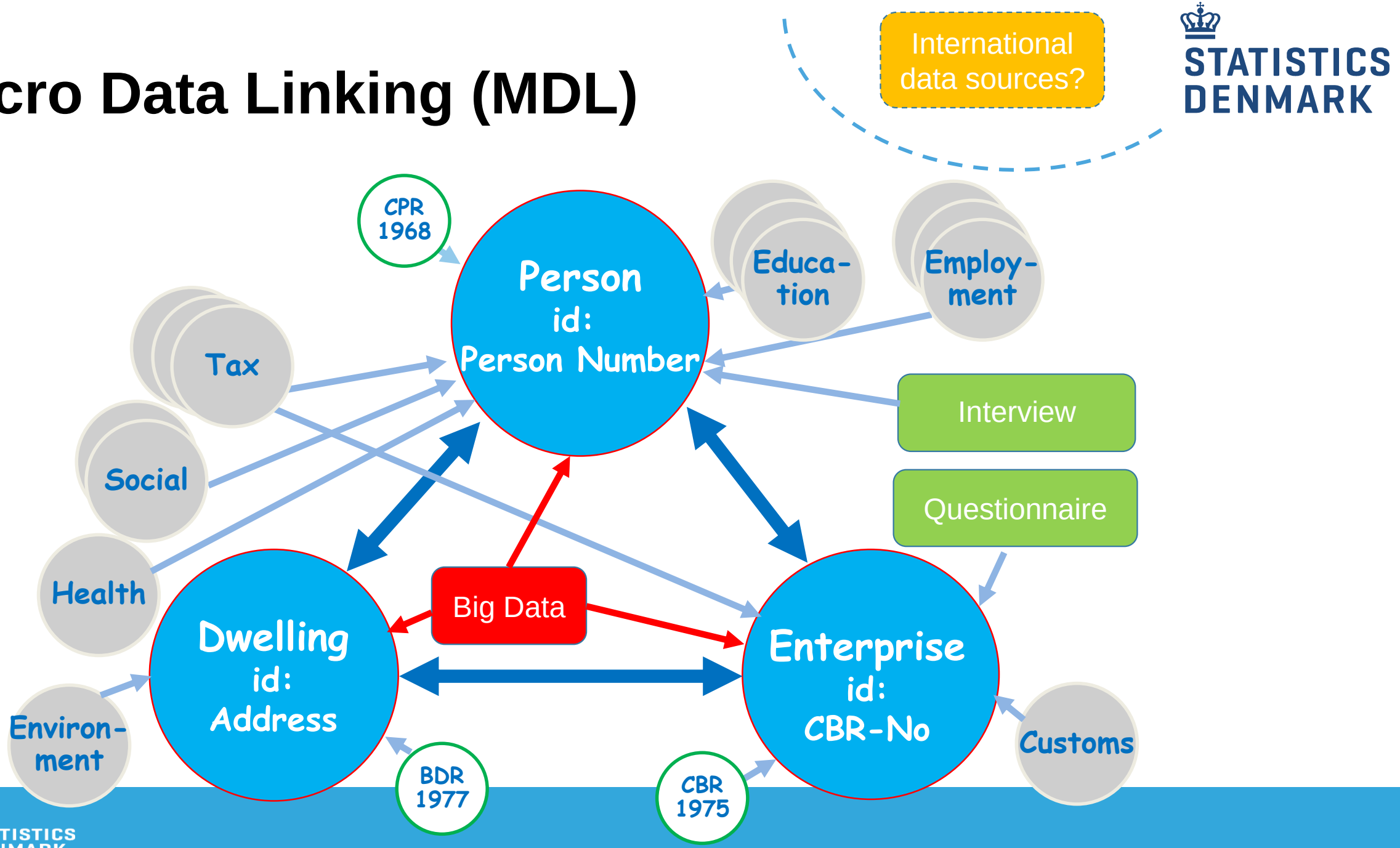


The register based system

STATISTICS DENMARK

- 3 basic registers: Persons (CPR), Business (CVR) and Buildings and Dwellings (BBR)
- 150 other admin. registers are received and linked via these basic registers
- For social statistics surveying is limited to LFS and EU-SILC variables, which are not available from admin. registers
- For business statistics fewer registers are available, but we use all relevant registers available, incl. VAT, Customs data, XBRL, automobiles, bankruptcies, farms etc.
- Other business data are collected system-to-system or via on-line questionnaires. (All data collection is digital by law sine 2010)
- The rest is received as Big Data, e.g. scanner data, Smart Meter, AIS, booking systems etc.

Micro Data Linking (MDL)



Examples outside social statistics

- Central Admin. Business Register (for SBR and business demography)
- Register of Buildings and Dwellings (for STS - among others)
- VAT (for STS, QNR and SBR)
- XBRL (for SBS and financial accounts)
- Customs data (for ITGS and BoP)
- Automobile register (for STS)
- Bankruptcies (incl. filings)
- Register of agricultural holdings (for SBR and FSS)
- Accounts for agricultural holdings and fishery / aquaculture
- Live stock, pesticides etc. (for agricultural statistics)
- Covid compensation schemes
- Booking and renting of holiday houses (tourism statistics)
- Scanner / barcode data (for CPI)
- Smart meter data (for STS and environmental statistics)
- AIS (gps data – sea transport / harbour statistics)
- Ticket sales, lending from libraries etc. (cultural statistics)
- Owners and Executive Board members

Danish statistics based on Tax data

- Structural Business Statistics
  - Enterprises' Tax Accounts
  - Covid compensation schemes for enterprises
- Short Term Statistics
  - VAT data (STS, foreign trade, SBR)
  - Customs data
  - Registration (sales) of new cars (register of vehicles)
  - Building and construction (register of buildings and dwellings)
- Social statistics
  - Employment and salaries ('e-Income' register)
  - Individuals' and households' income and wealth
  - Structural statistics on households' dwelling conditions
- Macro-economic
  - Tax payments from households and corporations
- Geospatial / environmental statistics
  - Land ownership

Data are delivered as identifiable micro data – confidentiality guaranteed

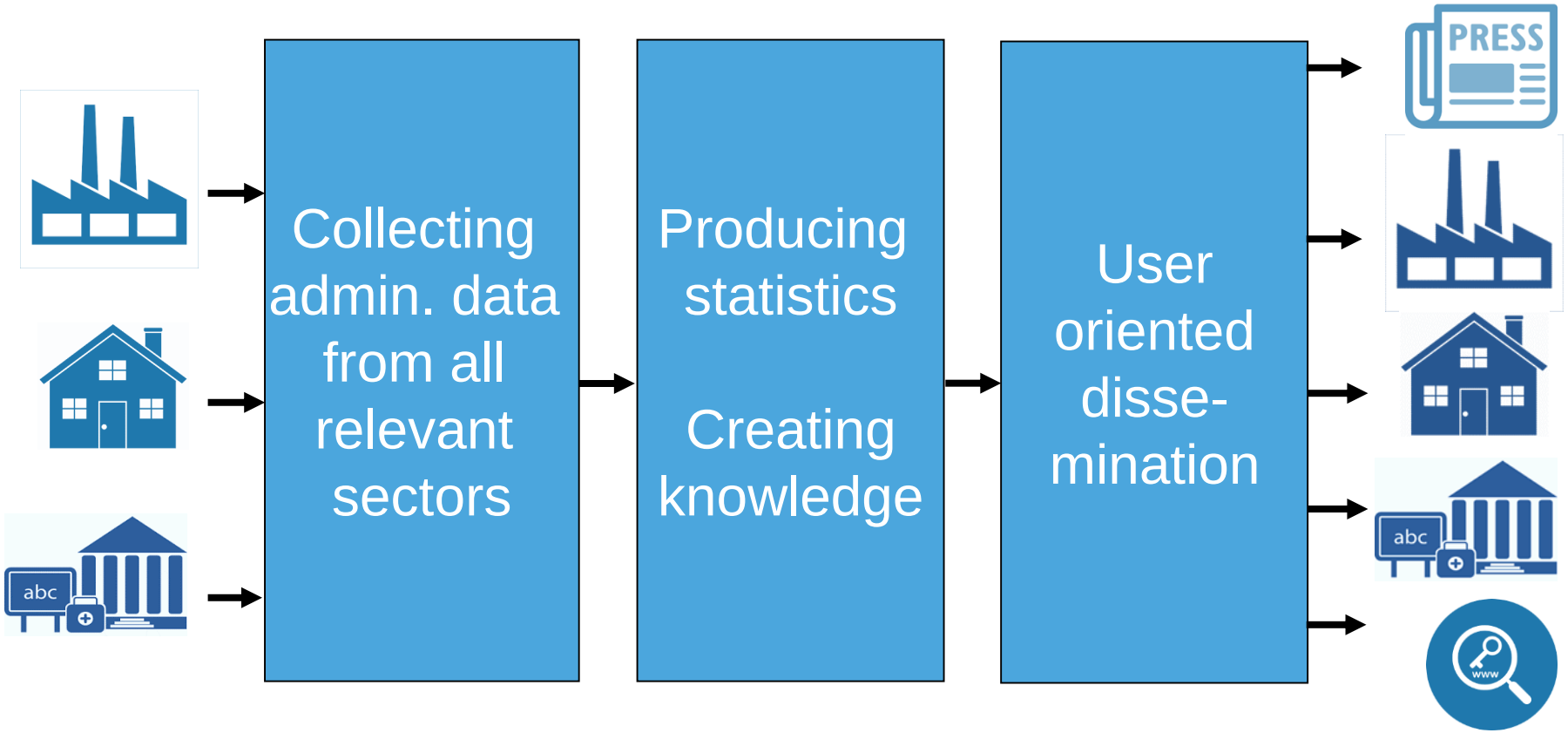
What is ‘admin data’ in this context?

- Structured information which is originally …
- collected by other authorities for their own purposes ...
- re-used by the NSI for official statistics.

What are the advantages?

- Coverage
- Structure
- Coherence
- Accuracy
- Combining & linking
- Longitudinal studies
- Cost-effectiveness (“average price per data point”)
- Flexibility (“n:m” combinations and MDL)

How does the process work?



How is data managed?

- Admin. registers are received, edited and organised in statistical registers
- Data is integrated across subject areas and time
- Statistical registers are combined to give comprehensive and new information
- Investment with high returns

How does data sharing materialize?

- Cooperation is key: Thorough discussions at strategic as well as practical level
  - Is it OK NOT to share to data?
- Focus on common interests / solutions – win-win situations and 'enabling incentives'
- Data sharing agreements (MoU) with data owners regarding content and conditions
- Raise awareness of the importance of high quality data registration
- Ensure adequate data protection

Standard Template for data delivery MoUs

1. Background and purpose
2. Delimitation of the data delivery in question
   - 2.1 Legal basis
   - 2.2 Definition / listing of data items covered
   - 2.3 Technical format for data delivery
   - 2.4 Time / frequency for data delivery
3. Data minimisation (cf. EU legislation)
4. Terms for amending and discontinuing the agreement
5. Signature by authorised managers

The template ensures that legal requirements are complied with

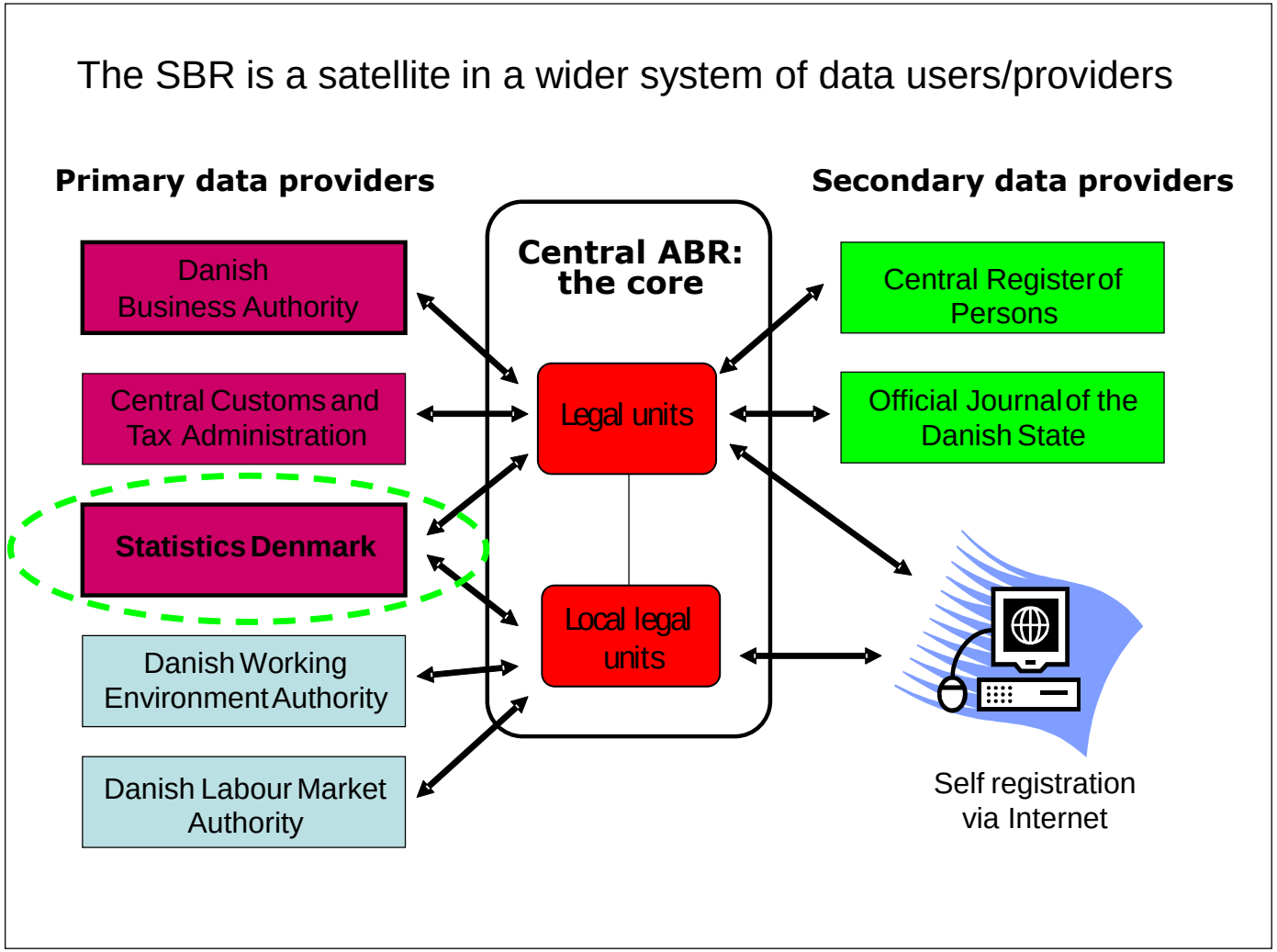
How is public trust maintained?

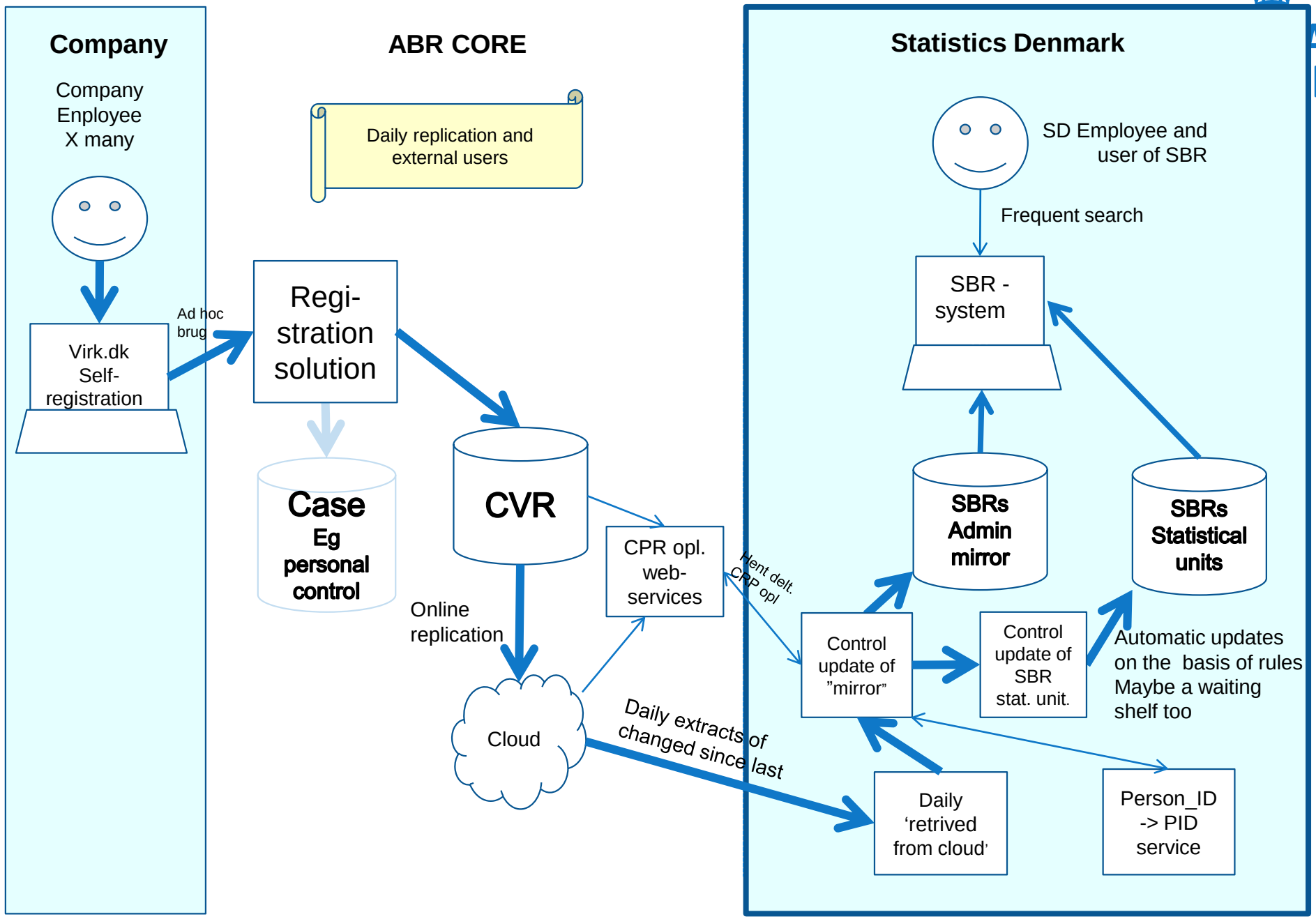
- Data protection legislation / agreements must be in place and complied with
- Data security rules and systems must be in place and used
- Individual data records/items (‘micro data’) are confidential
- Data on physical persons are given random identifiers
- No statistical info about individual persons or enterprises is ‘sent back’ (‘one way traffic’)

STATISTICS DENMARK

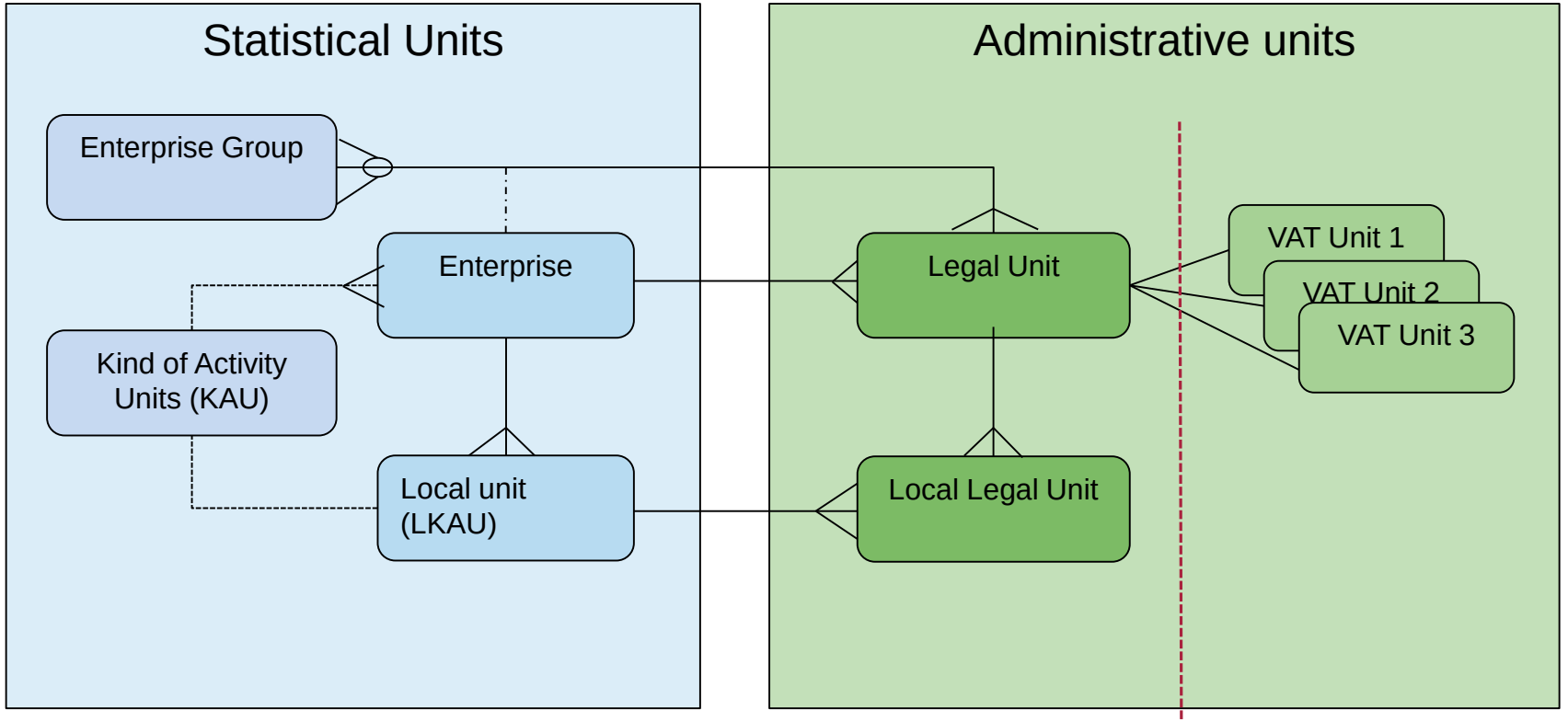
Zooming in on the statistical business register

Business Register – the national set-up




Company
Company
Enployee
X many
Virk.dk
Self-
registration
Ad hoc
brug
ABR CORE
Daily replication and
external users
Regi-
stration
solution
Case
Eg
personal
control
CVR
Online
replication
Cloud
CPR opl.
web-
services
Hent delt.
CPR opl
Daily extracts of
changed since last
Statistics Denmark
SD Employee and
user of SBR
Frequent search
SBR -
system
SBRs
Admin
mirror
SBRs
Statistical
units
Control
update of
"mirror"
Control
update of
SBR
stat. unit.
Automatic updates
on the basis of rules
Maybe a waiting
shelf too
Daily
'retrived
from cloud'
Person_ID
-> PID
service

Basic data model for the Danish SBR



Owner:
Statistics Denmark

Owner:
The Danish Business Authority
Responsibility:
All three players

Owner:
Customs and Tax Authorities

Examples of tables in the Danish SBR

Datamodel and temporal implementation

Enterprise (OK)

- OK_no
- Start_dato
- End_Dato
- Date_update
- Source_update
- ID_update
- Date_correction
- Source_correction
- ID_correction
- Type_correction

LKAU (ARB)

- ARB_no
- Start_dato
- End_Dato
- Date_update
- Source_update
- ID_update
- Date_correction
- Source_correction
- ID_correction
- Type_correction

OK_ARB_REL

- OK_no
- ARB_no
- From_Date
- To_Date
- Date_update
- Source_update
- ID_update
- Date_correction
- Source_correction
- ID_correction
- Type_correction

Activity Nace 2

- Unit_no
- Activity_code_Level
- Activity_code
- Unit_type
- Date_update
- Source_update
- ID_update
- Date_correction
- Source_correction
- ID_correction
- Type_correction

- Activity_code_Level = {0, 1, 2, 3}
  Level 0 = Principal activity
  Level 1 = Secondary activity no 1
  …
- Activity code = Danish Nace Rev. 2

SBR - context and content

The SBR has a up-to-date data base of all private and public enterprises in Denmark (low threshold for registration)

The SBR is maintained on the basis of the Law on Statistics DK

The SBR contains all basic data to construct/extract and maintain all survey populations

The most important characteristics/variables maintained are:

- ID_number, name and address
- Creation and cessation
- Ownership and institutional sector
- Activity (main and up to three secondary)
- Statistics (turnover and employment)
- Relations to other (types of) units
- Historical information
- Telephone no. and email-address

The Danish business structure cf. the SBR

STATISTICS DENMARK

- Approx. 800.000 enterprises ~ legal units
  - Activity Code
    - Main: all, Secondary 1: 10%, Secondary 2: 2 %, Secondary 3: 0,4 %
- Approx. 950.000 local units
  - Activity Code
    - Main: all, Secondary 1: 9%, Secondary 2: 1,7 %, Secondary 3: 0,4 %
- Non-active units (approx. 2%)
- Many very small (insignificant) units (approx 50 %)
  - 60 % has turnover; 25 % has employees
  - Below Danish threshold 60 %
  - 92 % not part of SDs data collection
- Very few very important units (approx. 10% of the units have 80% of the turnover; 0,2 % have 60%)
- Vice versa for the public limited companies
- Only few enterprises (less than 4%) have more than one LKAU
- Growing number of units in services, decline in manufacturing

SBR quality assurance in a nut shell

Automatic corrections

- Validity, consistency and coherence

Manual controls

- Workplace and employees
- NACE

Collaboration with the various statistics sections in SD

- Feedback from the respondents
- Data reported by the respondents

Use of SBR for surveys in Denmark

SBR is used in >100 different statistics (structural, short-term, register based, sample based) in 8 divisions / 3 directorates:

Either as the statistical business units, or by

Adding characteristics to other types of statistical units

Approx. 165 recurring extracts each year

Approx. 50 pct. of NSI staff are on-line users of SBR and extract system

Examples:

- Expected investments in the manufacturing sector
- Employment in construction sector
- Stocks in manufacturing and whole sale
- Purchases of goods and services in construction sector

Usage of SBR in Denmark – more examples

- Sales of goods in the in the manufacturing sector
- Consumer credits
- Financial sector companies
- Government finance statistics
- Statistics on utilities
- Retail trade index
- Companies' sales and purchases
- Inward FATS
- Business demography
- General accounts statistics (SBS)
- Structural employment statistics (obs)
- Tourism statistics

Usage of the SBR

- “Live” vs. “Frozen” register
- Survey samples are drawn using background (stratification) variables from SBR
- Enumeration
- All business statistics are distributed on the basis of either activities, size or location – all data drawn from the SBR
- Business Demography
- Response burden management – keeping track of the number of contacts to small/medium sized enterprises (SMEs)

Frozen frames from the SBR

- Quarterly and annual frames in a preliminary and a final version
- Only ”thawed” (= updated after freezing) in case of serious errors - and then by Steering Committee decision
- Is to be used as sample frame and for adjustments after data collection
- Name and address are validated against SBR before contacting the reporting entities

Frozen frames from the SBR (cont.)

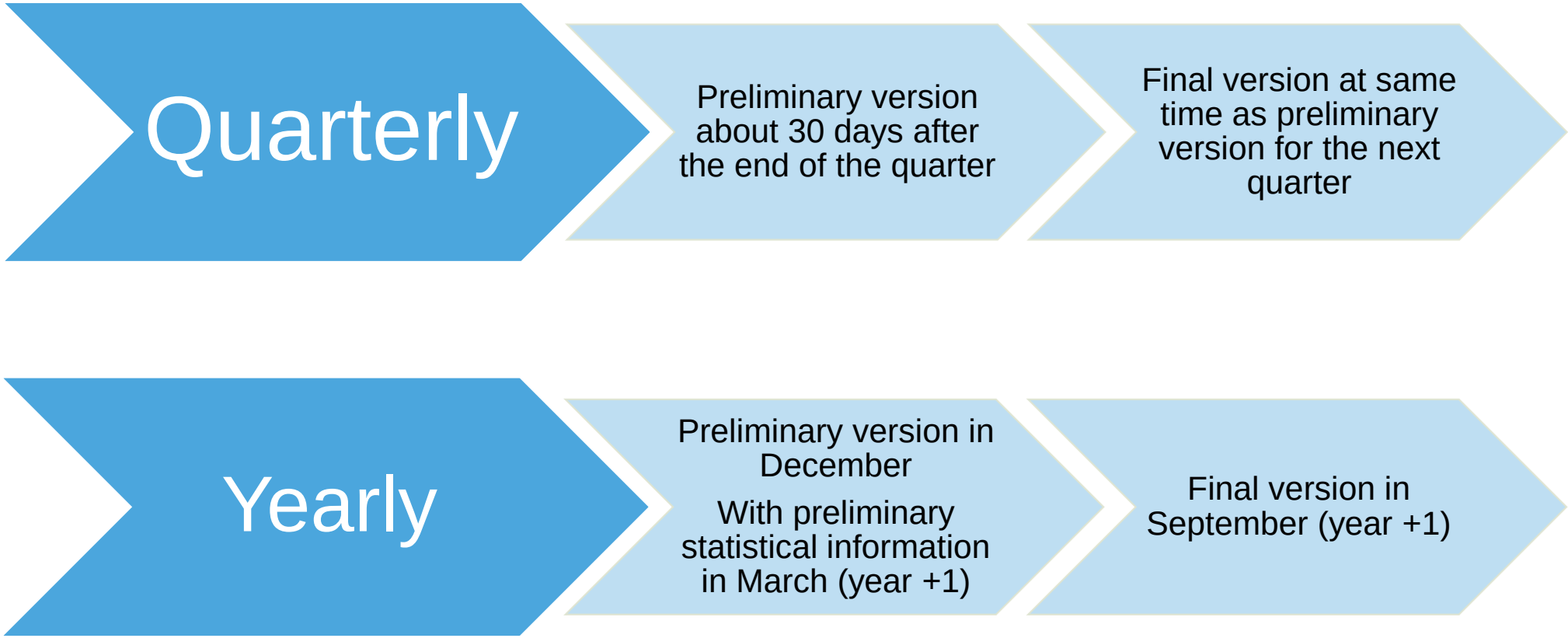
What to expect:

- The agreed quality assurance programs are done
- Consistency of relations between units
- Consistency of NACE-code between legal units and local units
- Observance of values (ex. NACE, legal form)
- Coverage of more than 99,9% - a few are deleted

What not to expect:

- Correct NACE-code for every unit
- Full address for every unit
- All demographic relations are found and implemented

Frozen frames from SBR - time schedule



STATISTICS DENMARK

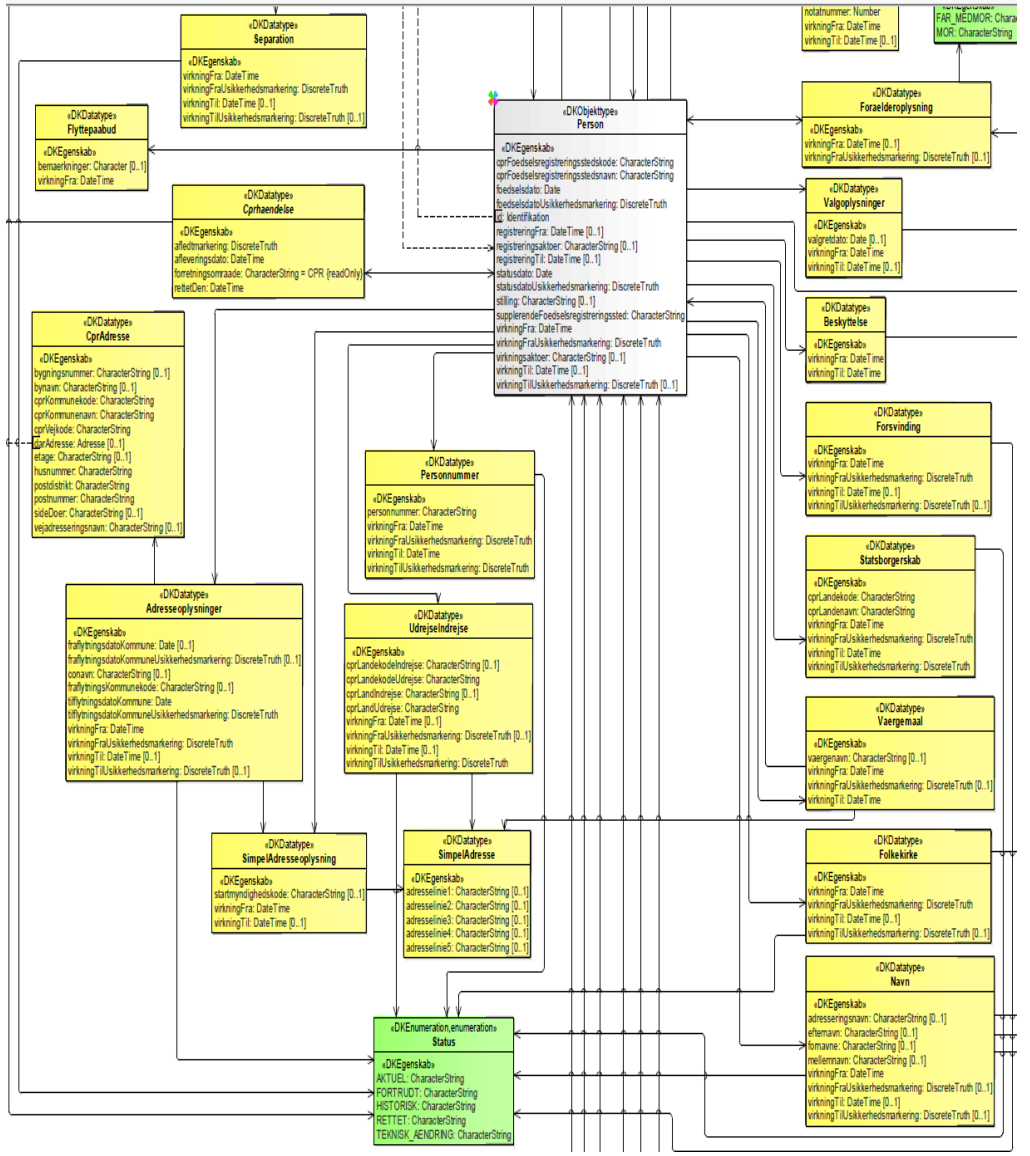
The national Basic Data Initiative

The Basic Data Initiative

- Initiated by the Danish Government, the municipalities and the Danish Regions
- Achieved through a series of sub-projects
- Fully implemented in 2020
- Aims at securing:
  - Free, fast and reliable access to Basic Data
  - High quality and coherent Basic Data updated at one place
- Aims at contributing to:
  - Increased effectiveness, modernization and high-quality public administration
  - Increased innovation, growth and job creation
  - Reduction of costs

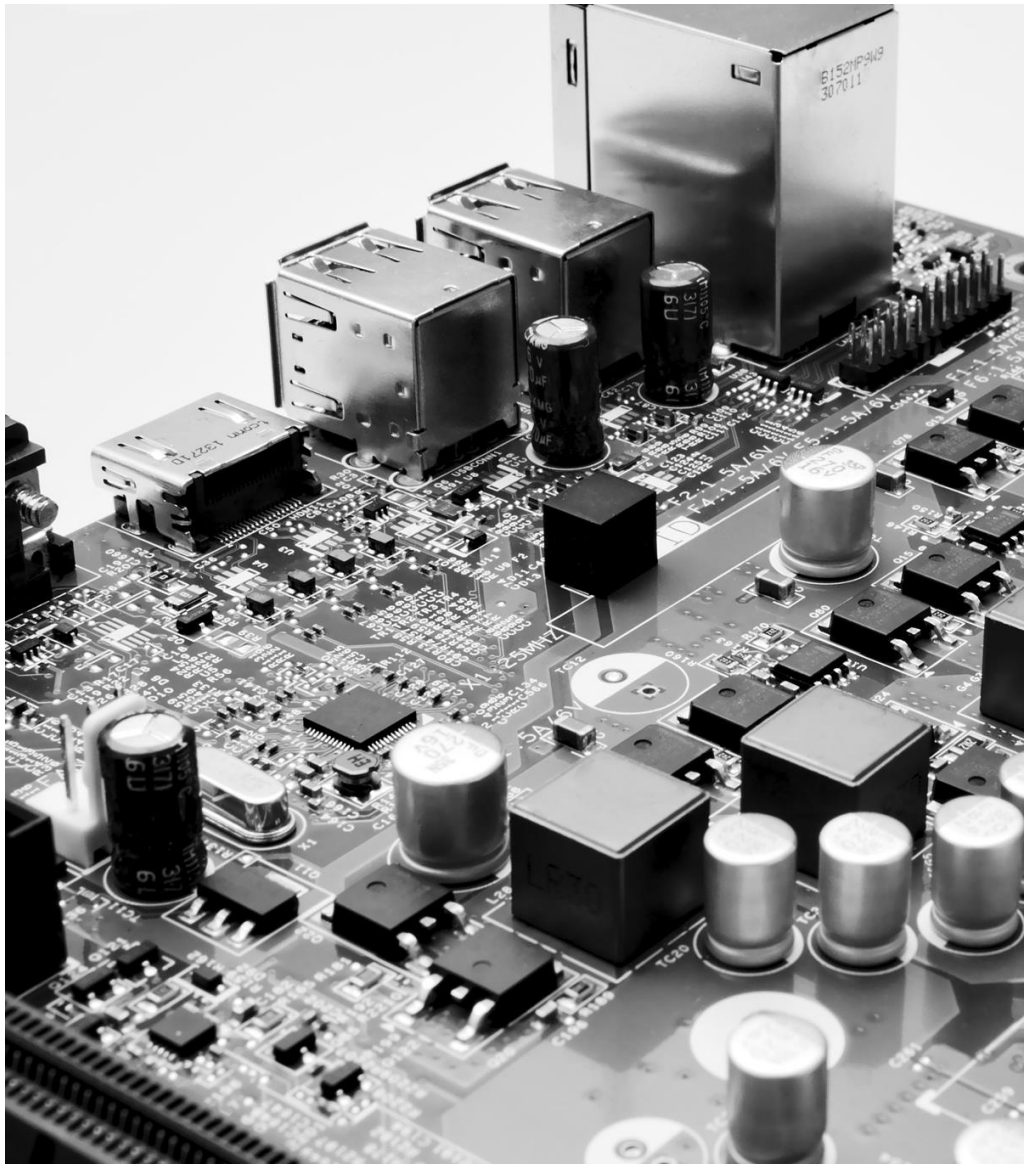
The Data Model

- Prior to the Basic Data Initiative, data was documented and modelled in different ways, making them difficult to combine
- Solution: Creation of a coherent common model for all Basic Data
- Characteristic feature of the model that objects and attributes are unique and can only be found in one Basic Data register



The Data Distributor

- Basic Data is distributed through The Data Distributor
- Secure, stable and scalable platform
- High availability (99,9%) and high performance
- Possible to compile data from different registers in a single service
- Much more than just a machine …



STATISTICS DENMARK

Usage of admin data in the production – overall view

Advantages of central data sources – the ‘ladder’

STATISTICS DENMARK

Level 5: Digital partnerships with central supplier of data – structured or unstructured

Level 5: Central provider supplies NSI with comprehensive data from “all” enterprises

Level 4: Automatic reporting

Level 4: ERP systems are adapted to automatically transfer data to NSI. Data approved by enterprise.

Level 2-3 Semi-automatic reporting

Level 3: ERP systems are adapted to facilitate transfer of data to NSI. Data approved by enterprise.

Level 2: APP exports data from internal business systems. Data validated and approved by enterprise.

Level 1: Manuel file upload

Level 1: Enterprise exports data from internal business system. Data file uploaded with/without validation at source.

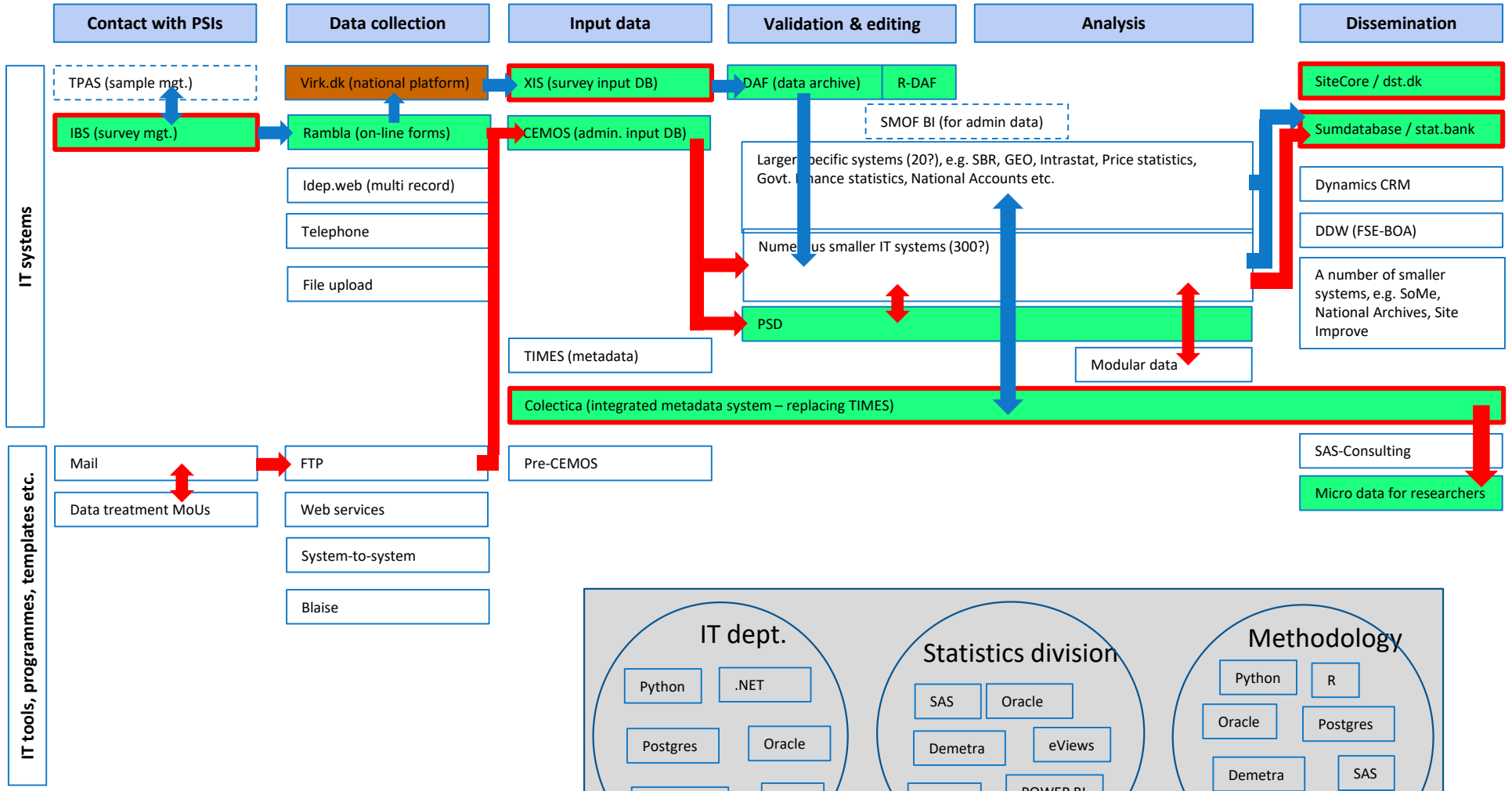
Baseline: Enterprise enters data into digital questionnaire or APP questionnaire. Data validated and approved by business.

Increasing value per observation

Decreasing cost per observation

Logical view of the standardized production

STATISTICS DENMARK



Admin data based

Survey data based

Highly standardised and mandatory

Highly standardised

Standard system under construction

External system