Steen Eiberg Jørgensen
Statistics Denmark
Session No. 5 – New data sources

Big Data – New data sources

Abstract

In this paper, Statistics Denmark presents some of the ideas that we are working on to change the data collection for the statistics from digitisation to automation.
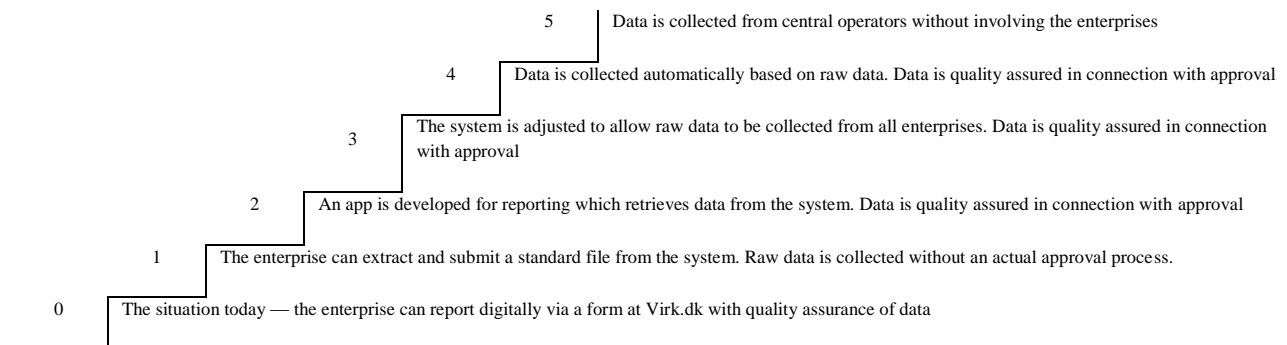
Over the last ten years, Statistics Denmark has carried through a number of initiatives for the purpose of digitising the data collection for business surveys, and the data transmission from the enterprises to Statistics Denmark is now essentially digitised.

In the course of the digitisation process, the technological development has resulted in even more work processes and systems becoming digital. In many fields, this means that data is born digital, and handled and stored digitally. This applies to e.g. ordering on an enterprise website which is invoiced digitally and subsequently stored automatically in a cloud-based accounting system. The digitisation implies that data is increasingly shared between systems and makes workflows smoother and faster. In addition, there is a tendency for enterprises to share data if it can contribute to reduce administrative tasks.

When systems become online-based, it further results in an alignment of the application of the systems by the enterprises. Customised application is not possible to the same extent and consequently, data is registered and stored in a consistent way. The consistent structure of data facilitates use of data for statistical data collection from many enterprises.

As a result, the technological development makes it is possible to extend the digitisation process into the enterprises so that data is collected directly from the enterprise systems. In this way, a change has happened by which digitisation transitions to automation. The digitisation into the enterprises and the automation of the work processes can be implemented at various levels.

Various levels of automation

5	Data is collected from central operators without involving the enterprises
4	Data is collected automatically based on raw data. Data is quality assured in connection with approval
3	The system is adjusted to allow raw data to be collected from all enterprises. Data is quality assured in connection with approval
2	An app is developed for reporting which retrieves data from the system. Data is quality assured in connection with approval
1	The enterprise can extract and submit a standard file from the system. Raw data is collected without an actual approval process.
0	The situation today — the enterprise can report digitally via a form at Virk.dk with quality assurance of data

Statistics Denmark wants to present our ideas at the Wiesbaden Group meeting.