

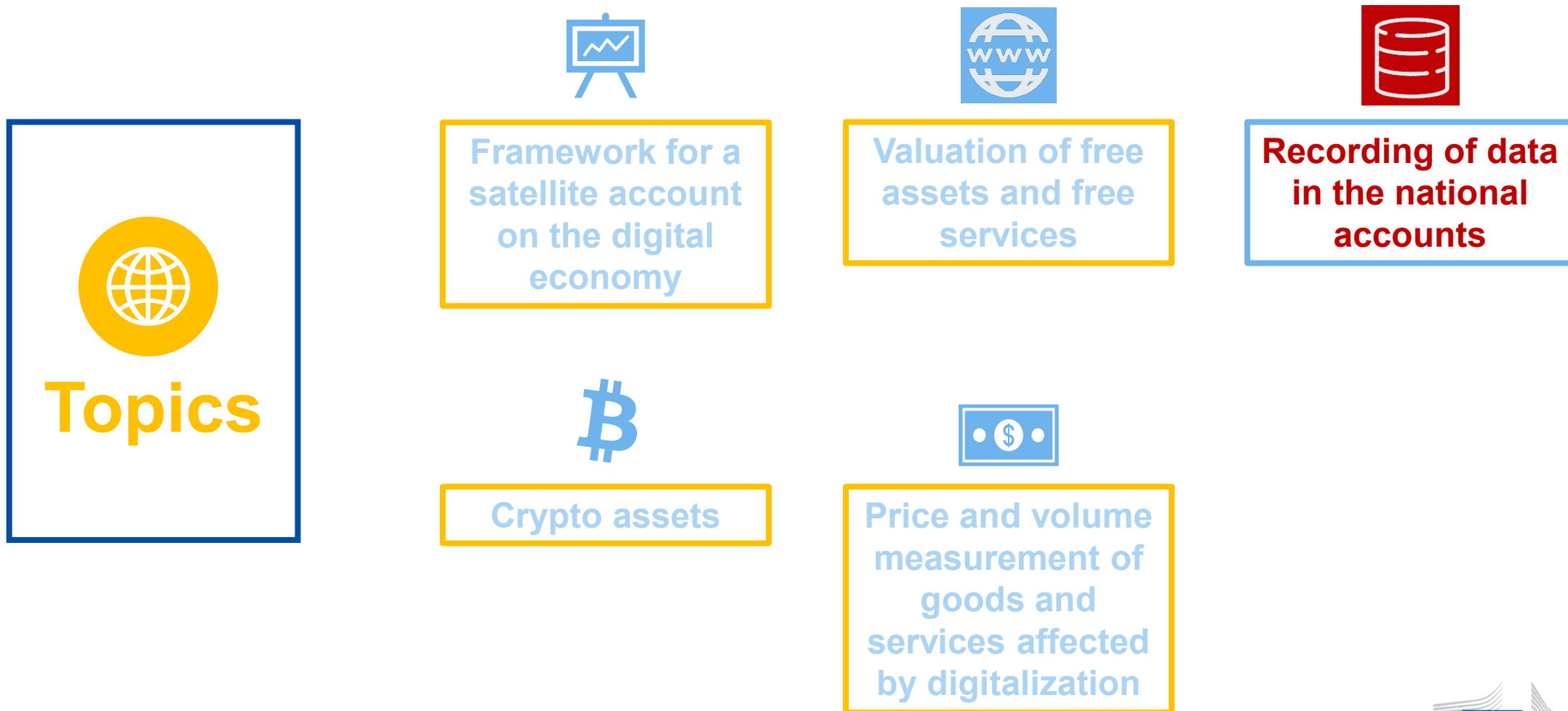
Recording and valuation of 'data' in National Accounts

Nicola Massarelli - Eurostat
on behalf of the ISWGNA Task Team on Digitalisation

Eurostat-OECD-UNECE Group of Experts on National Accounts

Webinar on Digitalisation, 14 September 2020

ISWGNA Task Team on Digitalisation



Background

Data is currently not covered in 2008 SNA. Data is valued together with a DB only if this is sold. Data in own-account DBs is not valued.

Perception: data are largely used in economic activities.

General questions:

- Are we missing an important part of GDP?
- What is the contribution of data to economic activity?

Background (cont'd)

Specific questions the SG is addressing:

- How should data be defined and classified for statistical purposes?
- Are data produced or non-produced?
- Do they qualify as assets?
- How should they be valued?
- How should they be recorded in a national accounts?
- Who is the economic owner of data?
- Should they be included in the core accounts or satellite accounts?

Databases and data in 2008 SNA

[2008 SNA (10.112)]

Databases consist of files of **data** organized in such a way as to permit resource-effective access and use of the **data**. Databases may be developed exclusively for own use or for sale as an entity or for sale by means of a license to access the information contained.

[2008 SNA (10.113)]

The creation of a database will generally have to be estimated by a **sum-of-costs** approach. [...]

The cost of preparing data in the appropriate format is included in the cost of the database **but not the cost of acquiring or producing the data**.

Task team's proposed definitions

Data

Data is information content that is **produced** by collecting, recording, organising and storing **observable phenomena** in a **digital** format, which can be accessed electronically for reference or processing and from which their owner(s) derive economic benefits by holding or using it.

Data as an asset

Data from which their owner(s) derive economic benefits by **using them in production** for **at least one year** (long-lived data) are fixed assets.

Task team's proposed definitions (cont'd)

Short-lived data:

Data from which their owner(s) derive economic benefits by using them in production for **less than one year** is to be considered as **intermediate consumption** when it is purchased from third parties, or as the product of an **ancillary activity**, i.e. an integral part of the primary activity, when the production is taking place within the same unit.

Observable phenomenon:

An observable phenomenon is the occurrence of a singular event or piece of information.

Treatment of 'observable phenomena'

Observable phenomena are ubiquitous. They can result from events, interactions and participation by actors in the economy. They can come into existence through the actions of, or information about, a single person; as the direct result of an interaction between two parties; or as a by-product of production.

In general they are **non-produced** and have **zero or negligible value**

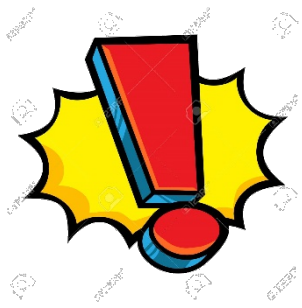
They are **outside the SNA production** and **asset boundaries**

Their possible treatment in a **satellite account** is being considered in combination with the treatment of free digital products

Data – main characteristics

According to definitions:

- Data is **produced** and included in the SNA production boundary
- Distinguished from underlying non-produced **observable phenomena**
- Focus on **digital** data (non-digital data exist but considered marginal)
- **Long-lived data** is an **asset** (therefore also included in the SNA asset boundary)
- Data is subject to **economic ownership, valuation** and **depreciation**
- **Short-lived data** are also produced but not an asset (not inventories or valuables)



Data as a produced asset: increases **output, value added** and **GDP**

Data producer

The **data producer** is the entity that collects, records, organises and stores observable phenomena in a digital format,

not the person or entity the underlying observable phenomena refer to

Digital platforms are a tool to ‘capture’ observable phenomena and turn them into ‘data’

Economic ownership

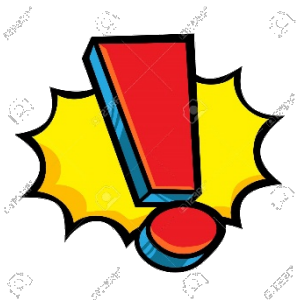
Own account data: economic owner is the producer (not the owner of the place where the data is stored)

Sold data (explicit transaction): economic owner is the data buyer

License to use data:

< 1 year: no change in economic ownership

≥ 1 year: the **licensee** assumes risks and rewards of ownership (2008 SNA 10.100)



*Multiple copies of data can be licensed to several licensees at the same time, which opens the possibility of **multiple owners** of the same data*

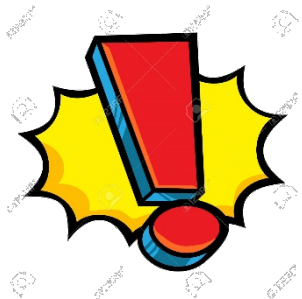
Recording data as an asset

Ideal recording: create a **new** specific fixed assets category under 'Computer software, **data** and databases (AN1173)'

Conceptual rationale: give relevance to data as a separate fixed asset, thus highlighting its importance in today's economy

Alternative recording: **expand** the scope of **databases** in capital formation, to include the costs to produce or acquire data

Practical rationale: it may not always be possible to separate the costs of database structure from that of the database content (i.e. data)



Practical tests of the two options will be crucial to make a final recommendation

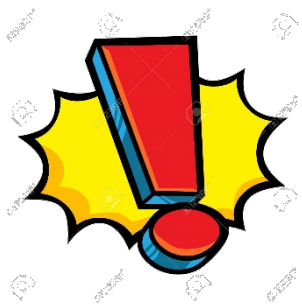
Valuation of data assets

Purchased data: valued at **market price**

(data that are part of pre-existing datasets are bought and sold in the commercial market as part of a database)

Own-account data: valued at **sum of costs**

Costs to include: **collecting** or **acquiring** observable phenomena, [**storing** and **processing** data]



*Storing and processing costs already included in databases' cost.
Practical tests will tell us if and how to disentangle data and database costs and avoid duplications*

Depreciation of data assets

PIM method with appropriate service lives and depreciation profiles

Data recorded separate from databases: ideally, data-specific service lives and depreciation patterns. Options (from best to worst), subject to **practical tests:**

- *Information from data-owner companies;*
 - *Empirically estimated;*
 - *Averaged across all data assets;*
 - *Geometric depreciation function with a common estimated service life (e.g. 10 years, as Eurostat and the OECD's approach to R&D) used by all countries, as a way to ensure consistency*

Data recorded with databases and software: same service lives can be assumed

Selected issues from summer consultations*

General:

- **Consistency** across guidance notes

Conceptual:

- Possible multiple **economic ownership** of licensed data
- **Ownership** of data derived from personal observations, in the light of **privacy legislation** and GDPR
- Treatment of **short-lived data**

Practical:

- **Disentangle data and database** costs not straightforward
- Need to also include non-digital data for **long time series**?
- Definition of **service lives and depreciation functions** for PIM model

* OECD's IAG Digital Economy, Eurostat's NAWG, EG NA, IARIW

Thank you

Contributors to this paper are: Andreas Dollt and Nicola Massarelli (Eurostat), Dylan Rassier and Rachel Soloveichik (BEA), John Mitchell (OECD), Marshall Reinsdorf, Jennifer Ribarsky, Jim Tebrake, Margarida Martins and Silvia Matei (IMF), Richard Heys (ONS), Ziad Ghanem (Statistics Canada), Sri Soelistyowati (BPS Indonesia), Stanimira Kosekova (ECB), Benson Sim (UNSD) and Kevin Fox (UNSW)

Contact: nicola.massarelli@ec.europa.eu