

# Modelling and harmonisation of units of measures

A preliminary study for SDMX Content-Oriented Guidelines

COSMOS 2024 presentation

Stéphane Crête, **György Gyomai**, Zlatina Hofmeister, Anastassia Samsonova

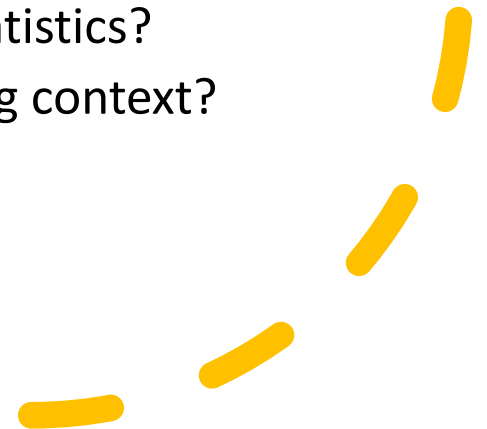
# Motivation

## Context

- The work initiated in the SDMX Statistical Working Group (SWG)
- It was preceded by several discussions in SDMX experts' meetings and global DSD exercises (Labour statistics, SDGs, etc.) over the last decade

## Rationale

- Alignment with established scientific best practice  
**reliance on dimensional analysis**
- What is special about socio-economic statistics?
- What is special about a data-warehousing context?



# Socio-economic statistics

- Economic value as a new dimension (straightforward)
  - although with complexities (time variance) -> or variations that are worth modelling on multiple dimensions
- Extrapolating from the 'amount of substance'
  - A special 'counting' dimension – or rather dimensions!
    - In socio-economic measurements it is hard to argue for a link with 'mass' and hence the postulation of the 'amount of substance' dimension for all counting quantities
  - Measurement context matters even in STEM (particles vs. atoms, protons, molecules)
  - A utilitarian approach
    - use UoMs to propose 'computation and comparability scope'



# Data warehousing

- SDMX and data-warehousing adds value by **revealing the structure and inter-connectedness** of data (as opposed to just a bunch of data)
- Two implications:
  - Fear of void: ratios of commensurable quantities -> favour the '**change of unit of measure**' representation of the data  
e.g. Debt to GDP ratio = 0.84 [? USD<sup>0</sup>] vs. Debt = 84 [% of GDP]
  - Units of measures might have a structure themselves, represented as a combination of multiple dimensions and attributes  
e.g. Currency: EUR, Price base: constant, Base year: 2010, Unit multiplier: Thousands

# Going forward

