SDMX Global Conference 2023

SDMX, an international standard for micro data
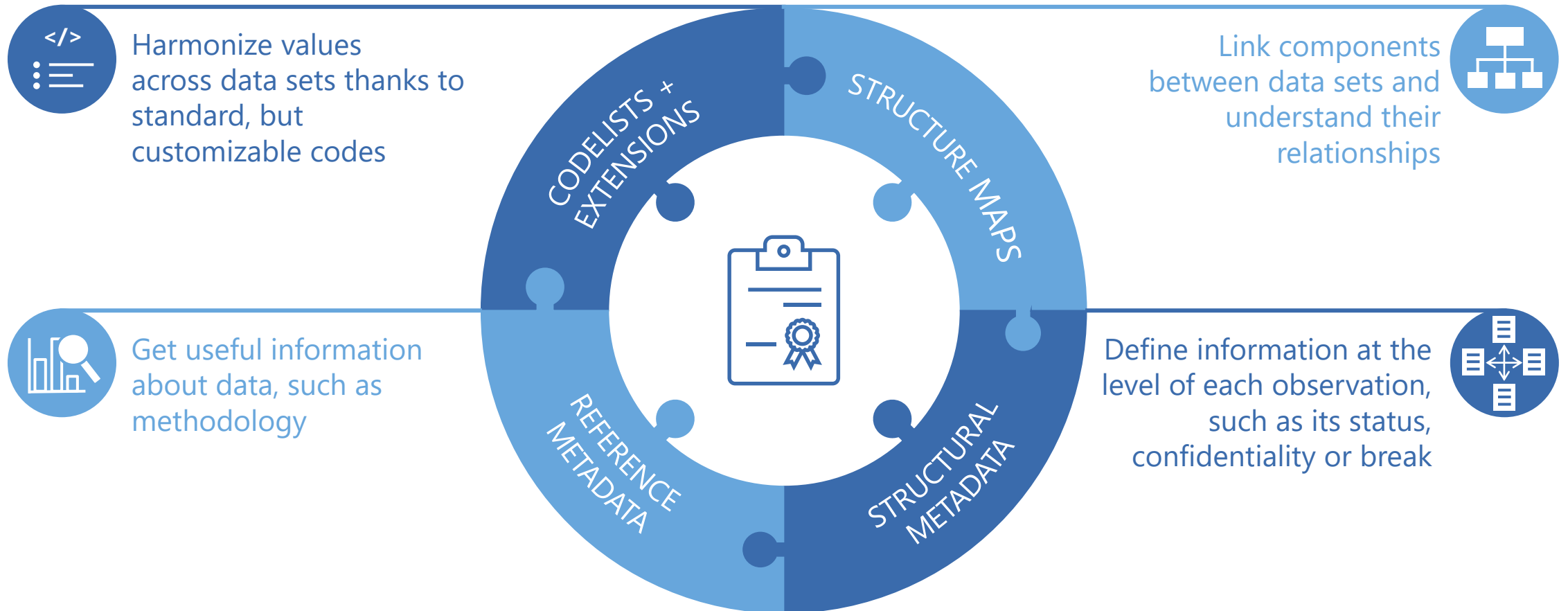
Stratos Nikoloutsos, Olivier Sirello (Bank for International Settlements)

A standard for micro data: trade-offs and challenges

1 **Standardisation but also customisation**
of data and metadata is key to facilitate interpretability, comparability and data lineage

2 **Reconciling micro and macro data: zooming in and out**
with the help of SDMX 3.0

3 **Proper modelling: from a top-down to a bottom-up approach**
to ensure consistency and standardisation across different data sets

4 **Data sharing with more performant and tailored queries**
made simpler and more efficient via SDMX

Standardisation

of data and metadata is key to facilitate interpretability, comparability and data lineage

Harmonize values across data sets thanks to standard, but customizable codes

CODELISTS + EXTENSIONS

STRUCTURE MAPS

Link components between data sets and understand their relationships

Get useful information about data, such as methodology

REFERENCE METADATA

STRUCTURAL METADATA

Define information at the level of each observation, such as its status, confidentiality or break

The right balance between standardisation and customisation

Harmonized values across data sets, but also customisable code lists

Customisation and extension of code lists

- Micro data often come with the **need for customised** and/or **extended code lists**
- This feature is key for **flexible maintenance of codes**, for instance during the collection and compilation phases

In a security-by-security data set, new ISIN codes can be appended after each data collection round
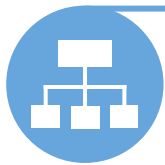
Structure and representations maps

- Data sets with micro data typically are split into **multiple tables**
- Structure maps are key to **describe the relationships** between each of them

Structure maps can be used to describe the relationships between the columns of multiple tables
(also allow to map custom internal codes to standard codes leveraging representation maps)

Structural metadata

- Structural metadata are key to describe statistical data, for each at data set, series, observation and measure level
- With **SDMX 3.0**, a **list of values for attributes is allowed**, increasing the flexibility of the data modelling notably required for micro data

In a security-by-security data set, SDMX 3.0 allows to set attributes for multiple measures,
such as face, nominal and market value per each security per period

Easy reconciliation of micro and macro data

with the help of SDMX

Hierarchies to quickly drill down from aggregates or conversely

Mappings to link representations and understand relationships between concepts

Track the validity of concepts and codes, their groupings, as well as their mappings, across time

Attributes at multiple levels with fine grained attachment

HIERARCHIES

MAPPINGS

VALIDITY

ATTRIBUTES + MULTIPLE MEASURES

From micro to macro and from macro to micro
## Quickly drill down from aggregates and conversely

Hierarchies to zoom in and zoom out

- Hierarchies are key to **drill down on the most granular level** from aggregates
- Groups and hierarchies share the same standardized codes, ensuring consistency

Share the codes across different groups, such as a country belonging to multiple economic groupings
Derive from the hierarchy the underlying entities that have been aggregated
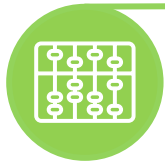
Mappings to better understand relationships between concepts

- Map representations, also **leveraging regular expressions**, to other representations and concepts
- Mapping also include **free text** and can be **one-to-many or many-to-many**

Map the initial two letters of the ISIN code to the country dimension
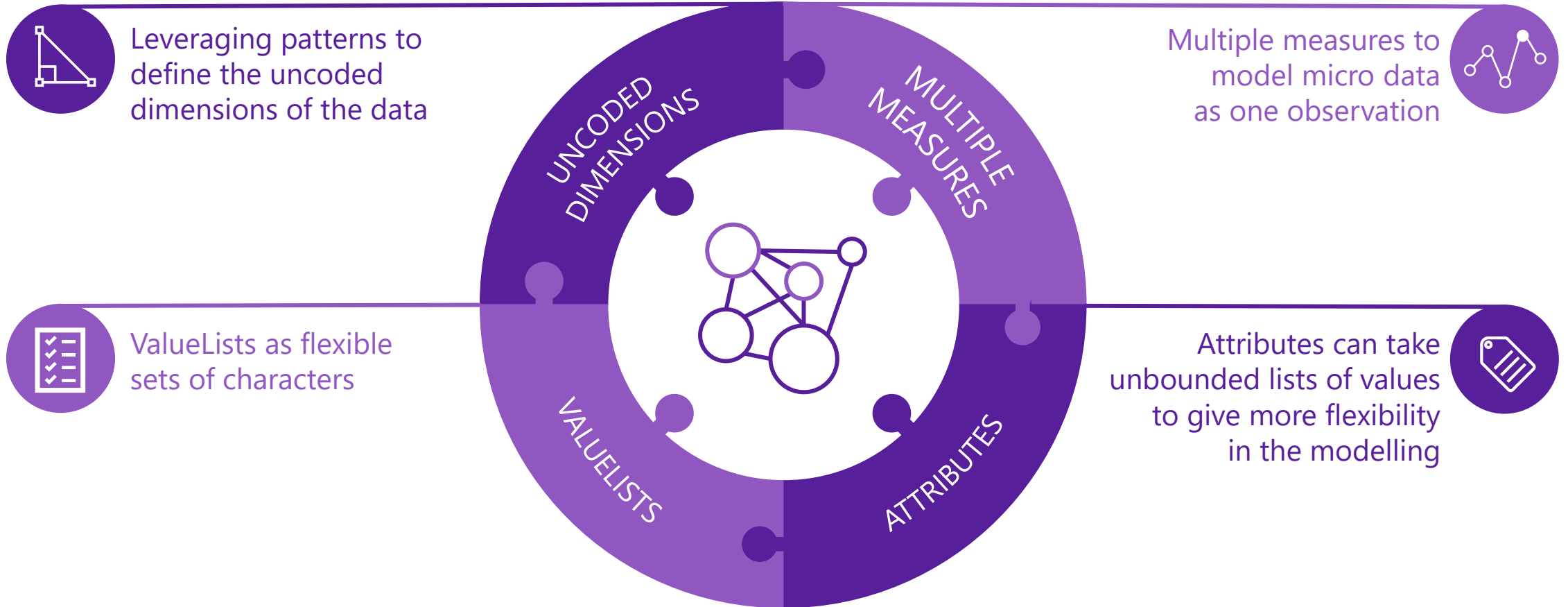
Attributes and multiple measures

- Measure-specific attributes: **an attribute** can be explicitly related **to one or more measures**

As an example, it might be possible define attribute A "1" for Gender and attribute A "2" for Occupation and filter according to their values
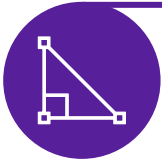
◆BIS

Proper modelling

of micro data ensures consistency and standardisation across different data sets

Leveraging patterns to define the uncoded dimensions of the data

Multiple measures to model micro data as one observation

ValueLists as flexible sets of characters

Attributes can take unbounded lists of values to give more flexibility in the modelling

UNCODED DIMENSIONS

MULTIPLE MEASURES

VALUELISTS

ATTRIBUTES

BIS

Achieve a flexible modeling, a bottom-up perspective

Leveraging patterns to define the uncoded dimensions of the data

A component based on a **value domain that follows a pattern**, without requiring the creation of a list of code

It is possible to derive the country ISO2 code from the first two letters of the column "ISIN code" from a security-by-security database
ISIN code –> Reference area, thus CH0000000000 -> CH

Multiple measures to model micro data as one observation

More than one measurement per record, allowing also to provide fine grained metadata per measure – rationalizing/simplifying data modelling

A **security-by-security table may contain three measures** for the amount outstanding, **face**, **nominal** and **market** value.
SDMX allows to **define attributes at the measure level**, for example to flag confidential only some specific values.
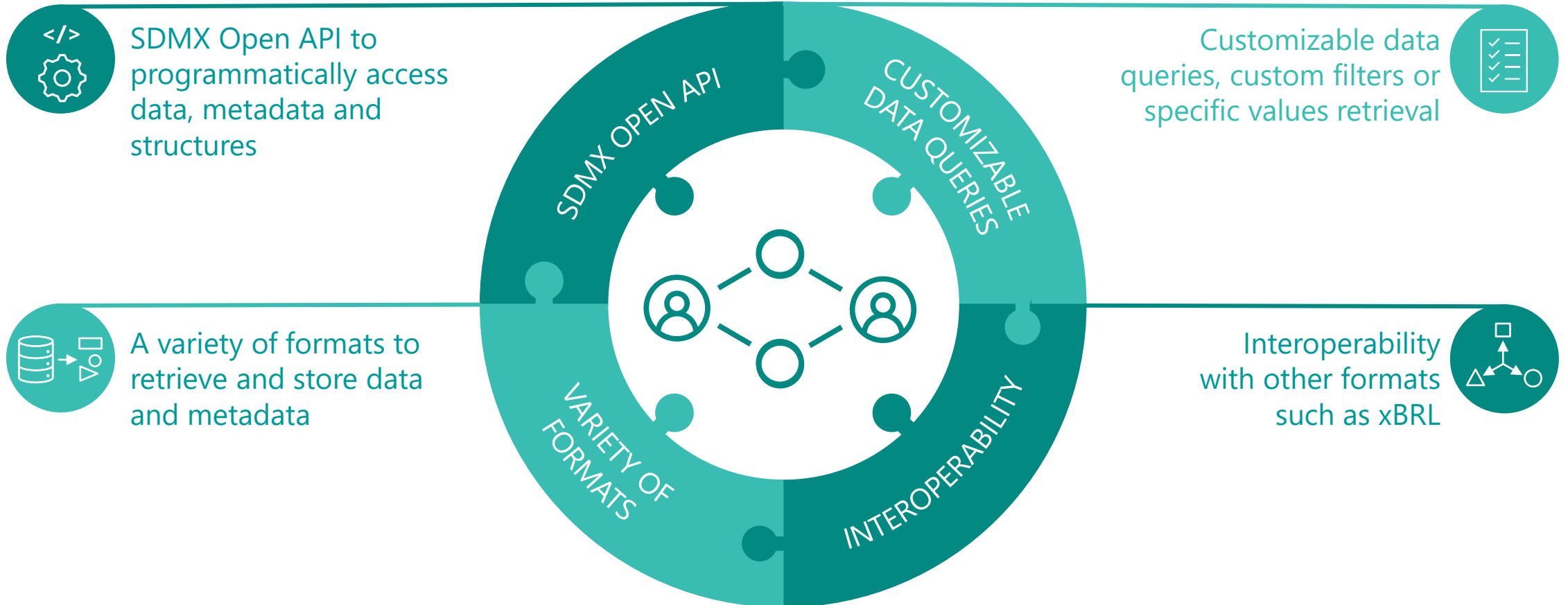It also allows to define **several statuses for a given value**, eg *provisional* and *unvalidated* value for market value on 2023-20

Time period	ISIN code	Face value	Nominal value	Market value
2023-10	CH0123456789	12	11.5	14[P, V]
2023-09	CH0123456789	12	11.6[CONF]	13
2023-08	CH0123456789	12	12	15

Data sharing
made simpler and more efficient via SDMX

SDMX Open API to programmatically access data, metadata and structures

A variety of formats to retrieve and store data and metadata

SDMX OPEN API

CUSTOMIZABLE DATA QUERIES

VARIETY OF FORMATS

INTEROPERABILITY

Customizable data queries, custom filters or specific values retrieval

Interoperability with other formats such as xBRL

Open API, more formats and interoperability

A powerful new Open API with increased flexibility

- Accessing data, metadata, structures within a **client application** also to ease their maintenance
- New parameters and operators for more flexible data and metadata querying

Querying for data for a range of values of a measure or attribute, including string matching
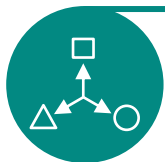
A variety of formats to retrieve and store data and metadata

- More data and metadata formats
- **Combining data with** reference **metadata**

New powerful XML messages, covering all standards
JSON targeting data visualization with combined data and structural metadata
CSV for human readable/processable datasets

Interoperability with other formats (xBRL)

- Ongoing work to **link the two standards**

xBRL-SDMX converter to facilitate the interoperability

BIS

Thank you!

Olivier.Sirello@bis.org

Stratos.Nikoloutsos@bis.org