



**United
Nations**

DESA

Statistics Division

Importance of Metadata

**Training Workshop on the Use of
Administrative Data for Migration
Statistics in Maldives**

10-12 December 2024

Elena de Jesús, UNSD

Data without metadata

1985

78 %

1 540 000

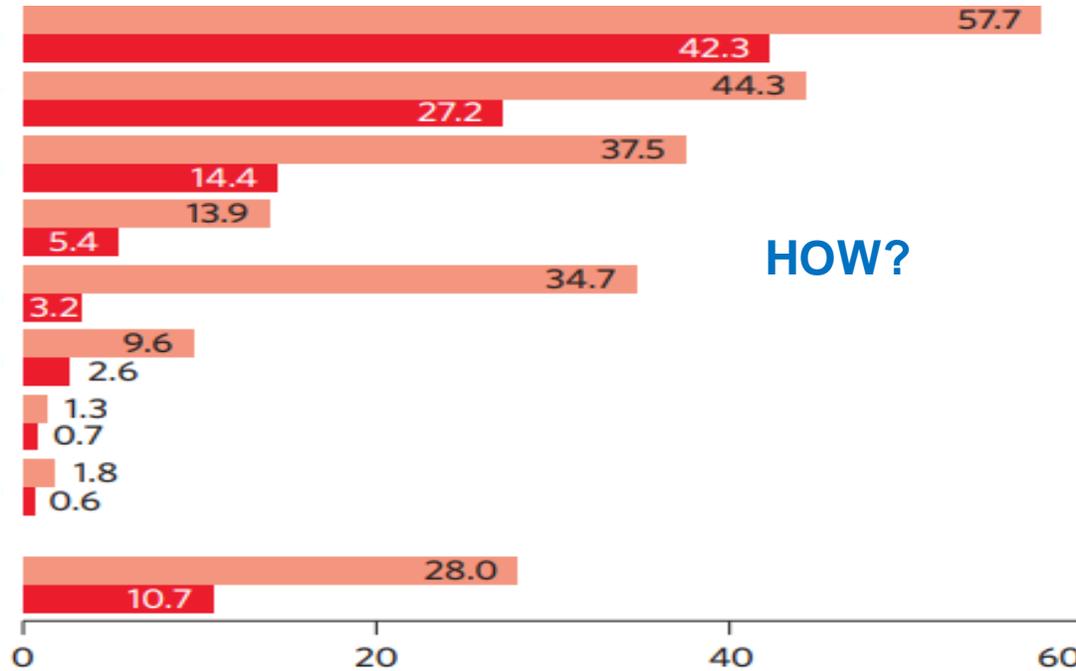
$\frac{3}{4}$

17,9



Data without metadata

WHAT?



HOW?

WHEN?

WHERE?

What is metadata?

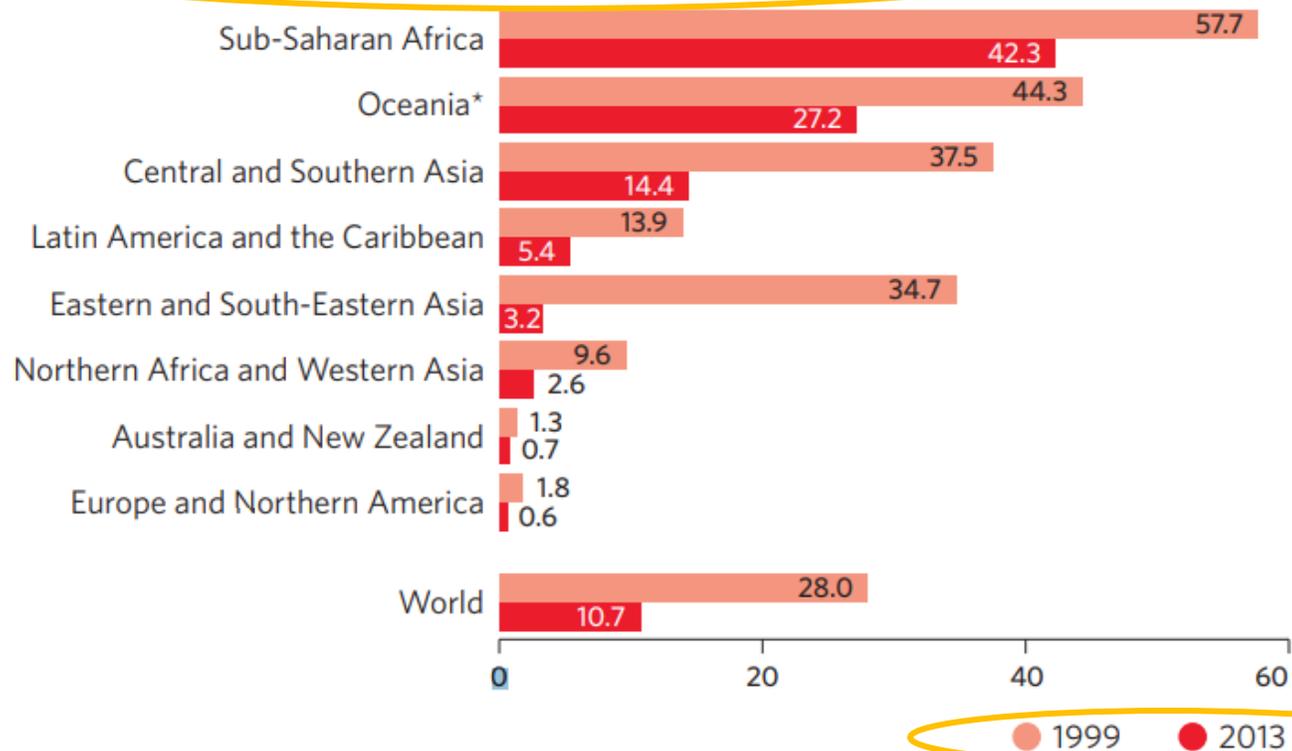
- Quite simply, metadata is **data that defines or describes other data**
- Metadata **helps explain and understand the data** or values being presented.

The African Charter for Statistics:

“the range of information, generally textual, that fosters understanding of the context in which statistical data have been collected, processed and analyzed with the objective of creating statistical information ...”

Why does metadata make a difference?

Proportion of the population living below 1.90 US dollars a day, 1999 and 2013 (percentage)



Note: Oceania* refers to Oceania excluding Australia and New Zealand throughout the publication.

Metadata in migration statistics

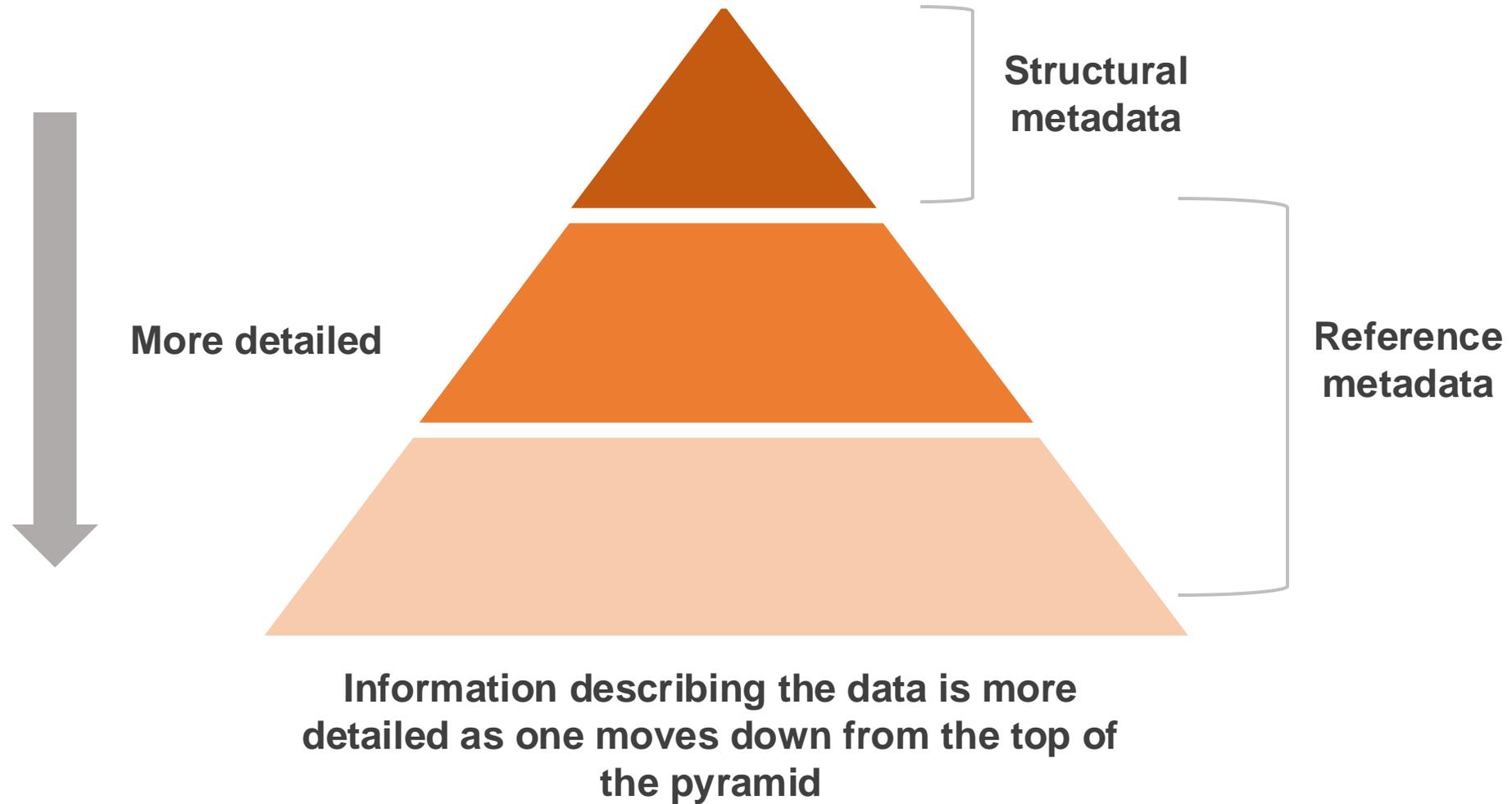
“There are 132,493 foreign residents in Maldives.”

With metadata:

- **Definition:** Foreigners residing in Maldives for one year or more (or intending to reside for a year or more)
- **Time period:** As on September 13, 2022
- **Exclusions:** Diplomatic staff, tourists
- **Data source:** Census 2022
- **Geographic coverage:** 463 islands
- **Coverage:** ? %



Hierarchical way of viewing metadata

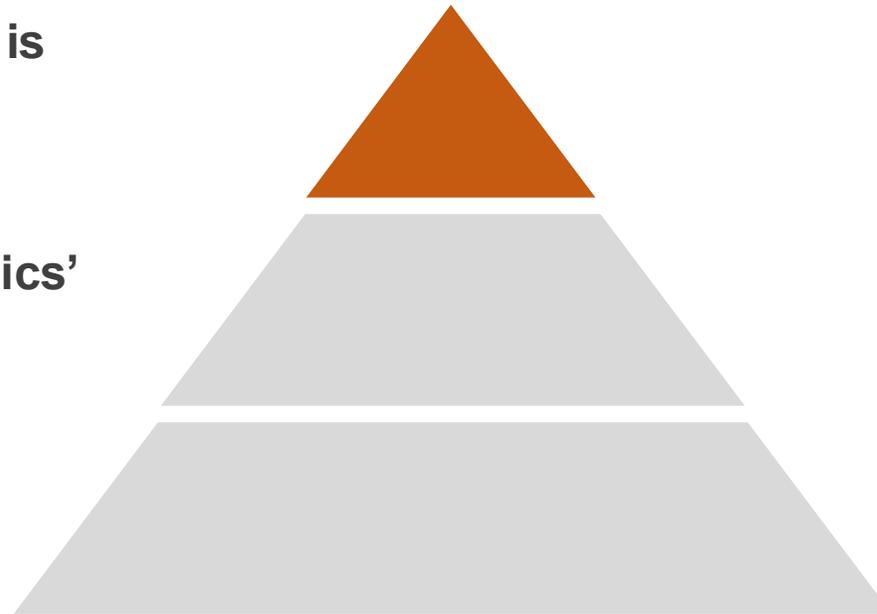


Structural metadata

Metadata that acts as identifiers and descriptors of data

At the top of the pyramid is information essential for understanding the data

Needs to explain the 'basics' of when, where, who and what?



In practical terms this means at the least information about: the series or indicator, the reference period, source, geographic scope, and the unit

Structural metadata

Metadata that acts as identifiers and descriptors of data



Indicator 8.5.2

Unemployment rate, by sex, age and persons with disabilities^{a,b}

Indicator

Unit

(a) Unemployment rate, both sexes^c
(Percentage)

Geographic
scope

Regions	2000	2005	2010
World	6.4	6.2	6.1
Sub-Saharan Africa	8.2	7.8	7.6
Northern Africa and Western Asia	11.8	11.7	10.2
Northern Africa	15.1	12.7	10.3
Western Asia	9.0	10.8	10.1
Central and Southern Asia	4.9	5.2	4.4
Central Asia	10.5	9.1	8.4
Southern Asia	4.6	5.0	4.3
Eastern and South-Eastern Asia	4.6	4.7	4.4
Eastern Asia	4.5	4.2	4.3
South-Eastern Asia	4.8	6.2	4.6
Latin America and the Caribbean	10.9	9.2	7.7
Oceania	6.2	4.9	5.4
Australia and New Zealand	6.3	4.8	5.4
Oceania (excluding Australia and New Zealand)	6.0	5.2	5.3
Europe and Northern America	7.9	7.5	9.2
Europe	9.7	8.5	9.0
Northern America	4.3	5.3	9.4
Landlocked developing countries	6.2	6.5	6.0
Least developed countries	5.1	5.6	5.5
Small island developing States	8.6	9.3	8.5

Reference period

Source: ILO Database of Labour Statistics (ILOSTAT), International Labour Organisation (ILO), ILO modelle 2016.

Source



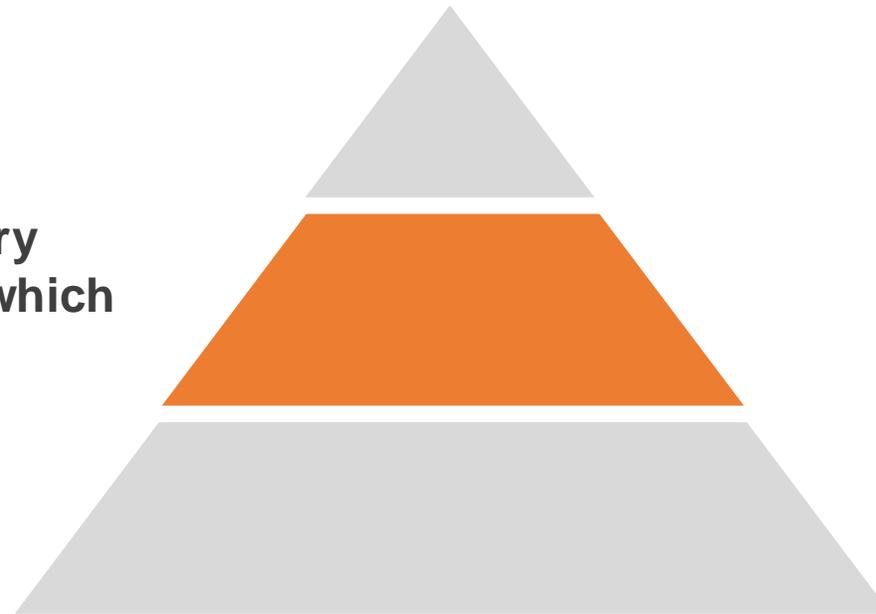
United Nations

DESA
Statistics Division

Conceptual reference metadata

Reference metadata that describes the key concepts and definitions

In the middle are explanatory notes and descriptive text which describe the statistics



In practical terms this means description about: definition, key concepts, standards and classifications used

Conceptual reference metadata

Reference metadata can include information such as key concepts and definitions



Definitions

Definition:

Percentage of a cohort of children or young people aged 3-5 years above the intended age for the last grade of each level of education who have completed that grade.

Key concepts

Concepts:

The intended age for the last grade of each level of education is the age at which pupils would enter the grade if they had started school at the official primary entrance age, had studied full-time and had progressed without repeating or skipping a grade.

For example, if the official age of entry into primary education is 6 years, and if primary education has 6 grades, the intended age for the last grade of primary education is 11 years. In this case, 14-16 years ($11 + 3 = 14$ and $11 + 5 = 16$) would be the reference age group for calculation of the primary completion rate.

Classifications

Standards and classification

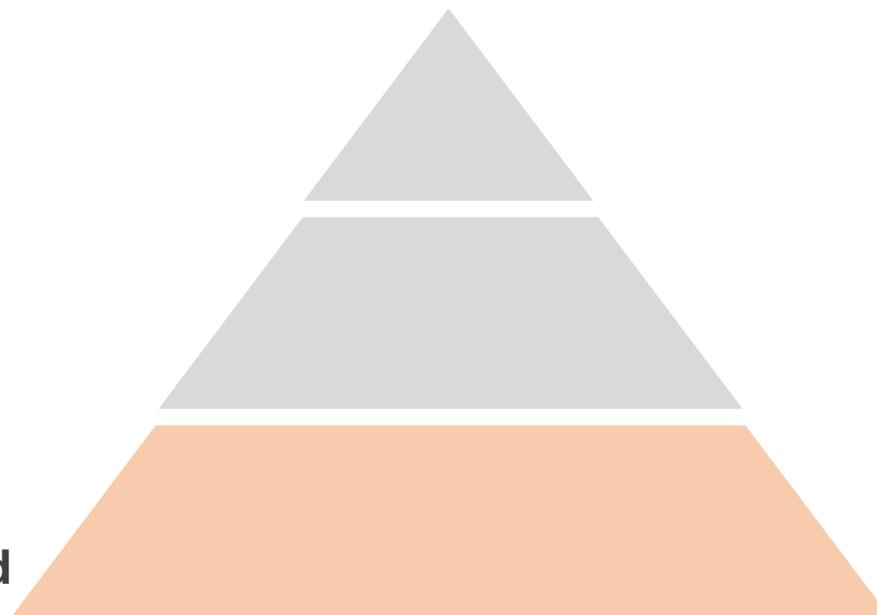
The International Standard Classification of Education (ISCED) is used to define primary, lower secondary and upper secondary education.



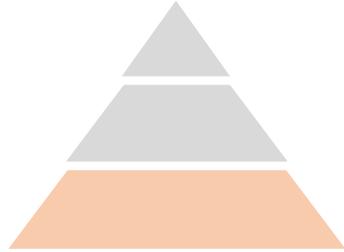
Detailed reference metadata on methods and quality

Further down the hierarchy of metadata, there is more detailed information

In practical terms this means descriptions of: the process of collecting data, calculations, quality aspects, limitations, and other detailed information.



Detailed reference metadata



How are data collected?

Data Sources and Collection Method

The data for this indicator is collected through household-based labour force surveys, population census, and any other nationally representative household surveys with an appropriate employment module. Such surveys are generally conducted by the ministries or bureaus of labour or national statistical offices.

Unemployment registers, under social insurance administrative systems, can also serve as instruments to collect data on unemployment levels, and used to supplement the information obtained by household surveys.

Method of Computation and Other Methodological Considerations

Computation Method:

The unemployment rate (U) is calculated using the following formula:

$$U = \frac{\text{Number of unemployed persons}}{\text{Persons employed} + \text{Persons unemployed}} \times 100$$

How is the indicator calculated?

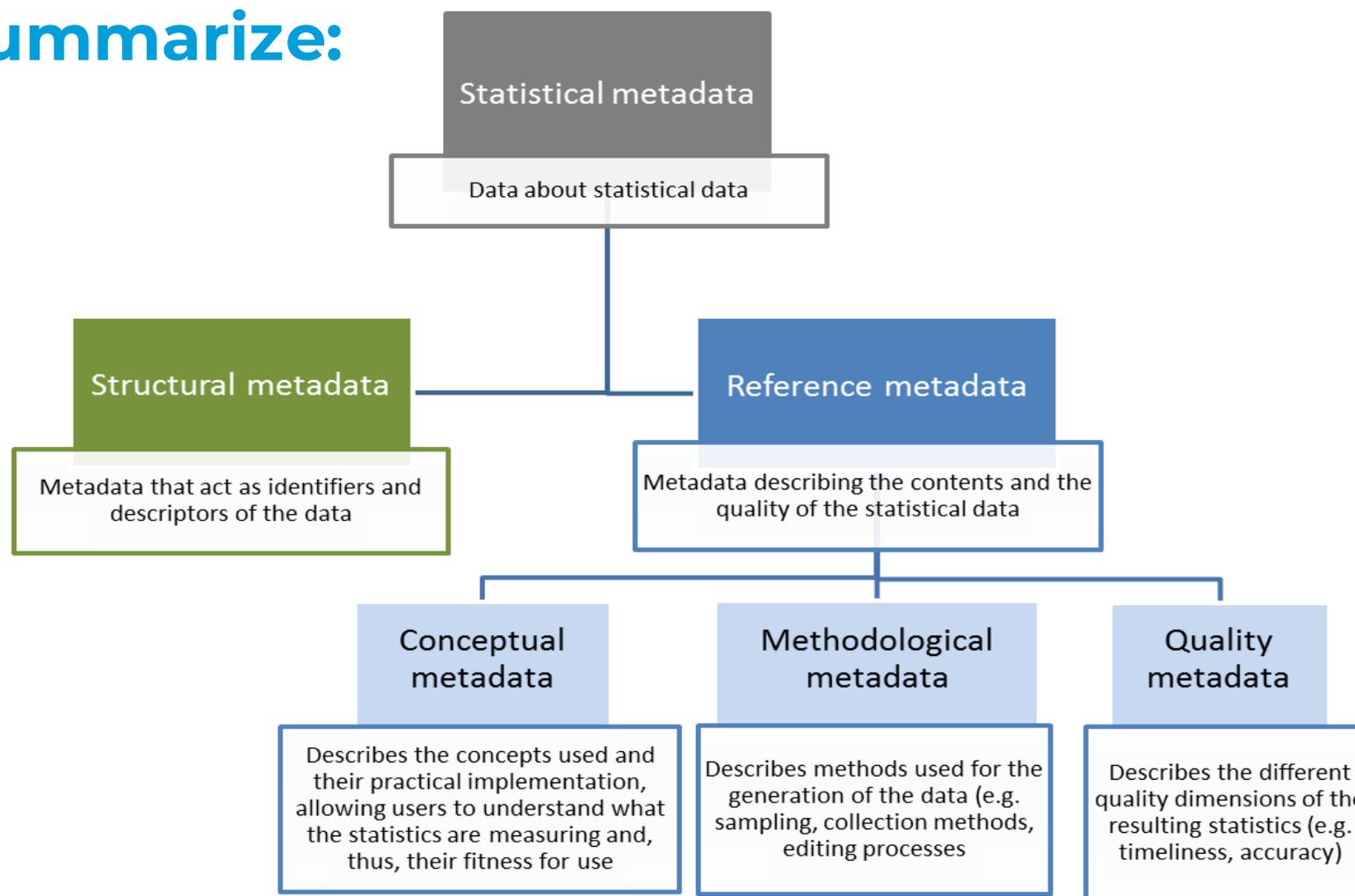
Comments and limitations:

The significance of the unemployment rate depends on context. It is not to be interpreted the same way universally. In the absence of unemployment insurance systems or social safety nets, persons of working age must avoid unemployment, resorting to engaging in some form of economic activity, however insignificant or inadequate. Thus, in this context, other measures should supplement the unemployment rate to comprehensively assess labour underutilization, such as the time-related underemployment rate or measures of the potential labour force. In this regard, the 2013 *Resolution*

What are the limitations?



To summarize:



Data dictionaries

Structured references that define the data in a dataset:

- **Variables:** Names, descriptions
- **Metadata:** Information about the variables (structural and reference)
- **Relationships:** How variables connect



Key components of a data dictionary

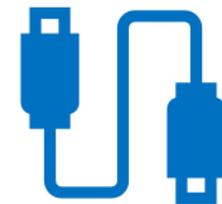
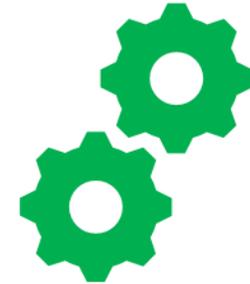
- **Variable name:** Unique identifier for each data element
- **Definition:** Description of what the variable measures
- **Data type:** Format of the variable (e.g., text, number, date)
- **Units:** Measurement units (e.g., age in years)
- **Source:** Origin of the data (e.g., administrative records, census, survey)
- **Coverage:** Population or time period the variable pertains to
- **Relationships:** Specifies how variables are linked across datasets or records
- **Allowed values/Validation rules:** Defines the permissible values or ranges for each variable; codelists

Data dictionary example

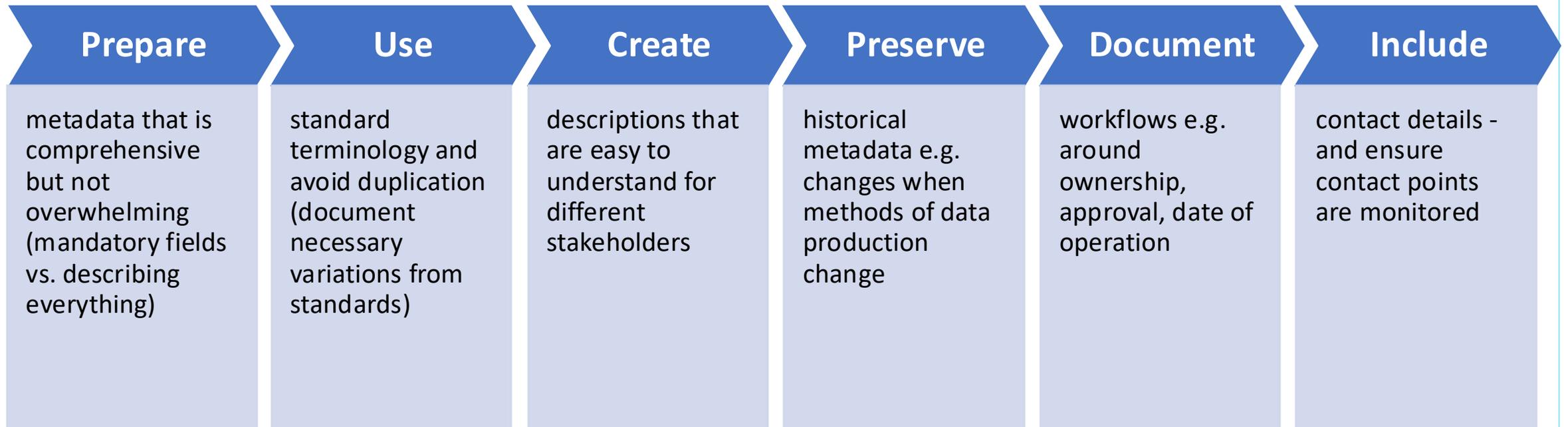
Variables				
Name	Occupation	OccupationCode	WorkVisaArrivalDate	LocationIsland
Description	Current Occupation Title	OccupationCode	Date of Arrival On Work Visa	
Definition	The title or nature of the main job or profession of the individual.	ISCO coding adopted to Maldivian context	The recorded date when an individual with a work visa first entered the country, as per immigration and work permit administrative records.	The Island in the application of the employer applied for
Limitations	Titles may be too specific and not directly aligned with standardized occupation classifications.		May not reflect subsequent entries or exits from the country. Does not account for changes in visa status post-arrival.	Exact location they work may not be captured
Reference period	Current at the time of the administrative data extraction or update.		The specific date when the individual entered the country under the current work visa.	
Categories	<i>(As per the administrative records. <u>Aligning with major groups of the International Standard Classification of Occupations (ISCO) ?</u>)</i>		N/A	
Data manipulations	Standardization of occupation titles into ISCO codes for consistency in reporting and analysis, which may involve mapping job titles to the closest ISCO category.			
Source	Xpat (MED)		Xpat work permit database	
Type	Categorical	Discrete	Date	
Format	Text/String	Integer	Date (commonly YYYY-MM-DD)	?
Size	64 characters		10 characters (including separators for year, month, and day)	
Range	All recognized occupations in the administrative source.		1990 - 2023	
Data availability	Available for 95% of records			

Metadata (data dictionaries) essential for data integration

- Ensures consistency in variable definitions across datasets
- Helps standardize formats, units, and classifications
- Reduces mismatches and misinterpretations during integration
- Highlights discrepancies in variable definitions or units
- Speeds up the process of linking and harmonizing data sources
- Facilitates inter-agency collaboration
- Provides clear documentation for data users and stakeholders



Recommendations for drafting and compiling metadata





**United
Nations**

DESA
Statistics Division

Thank you!

Questions?

Elena de Jesus, UNSD

DATA DICTIONARY PRESENTATIONS



Maldives Bureau of Statistics

Ministry of National Planning Housing & Infrastructure



Ministry of Higher Education,
Labour and Skills Development
Republic of Maldives



Ministry of Homeland Security and Technology

REPUBLIC OF MALDIVES



Maldives Immigration

Republic of Maldives



United Nations

DESA
Statistics Division