

Paris, 11 of April, 2024

COSMOS 2024 EDITION

The role of the term in defining centralised standard metadata

Claudia Brunini, Emanuela Recchini



Toward a metadata management: a complex process of transition



- Siloed repositories
- Ad-hoc, proprietary, in-house developed tools
- Absence of integration and harmonization
- Manual processes with large overheads



- Partially centralized siloed repositories
- Ad-hoc, proprietary, in-house developed tools
- Limited, time-consuming integration and harmonization

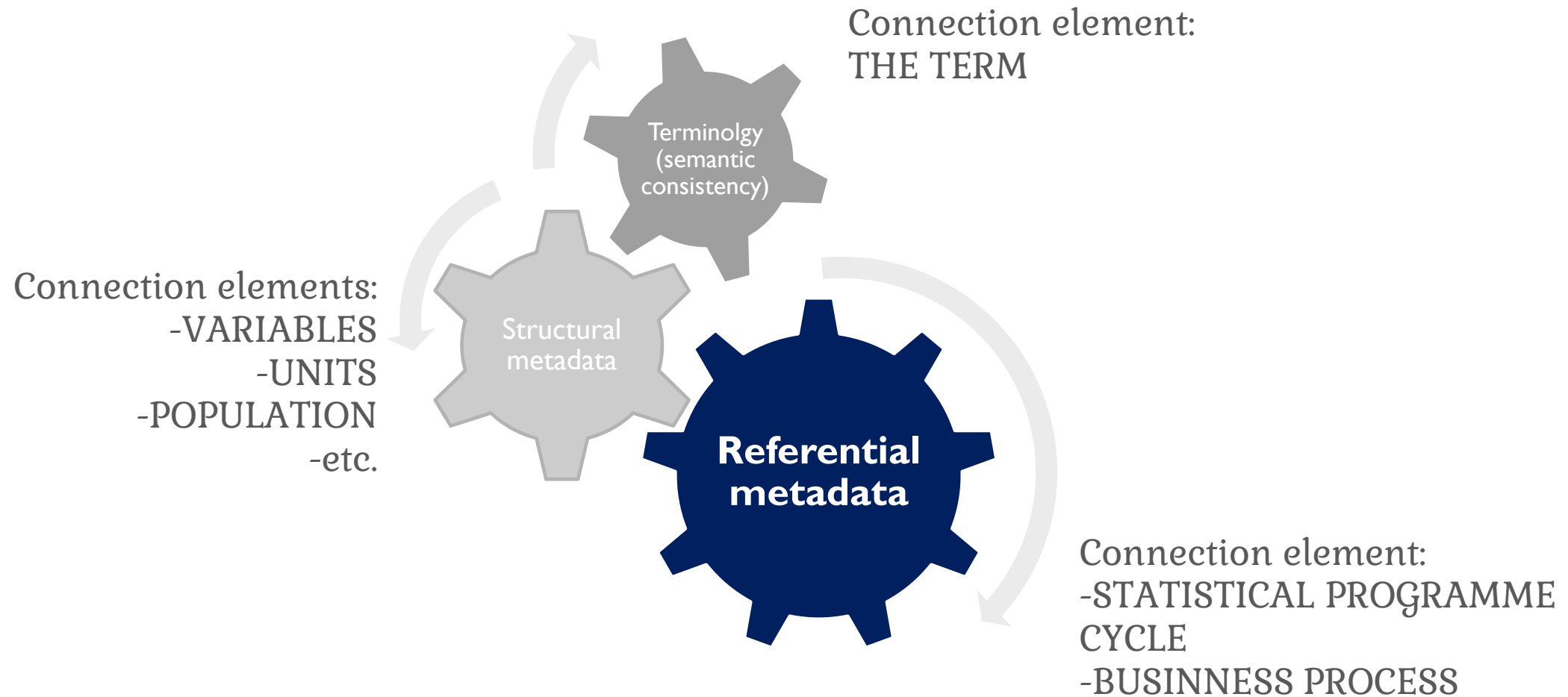


- Standards-based repositories
- Standards-based (off-the-shelf and open source) tools
- Standards-enabled integration and harmonization
- Standards-driven automation

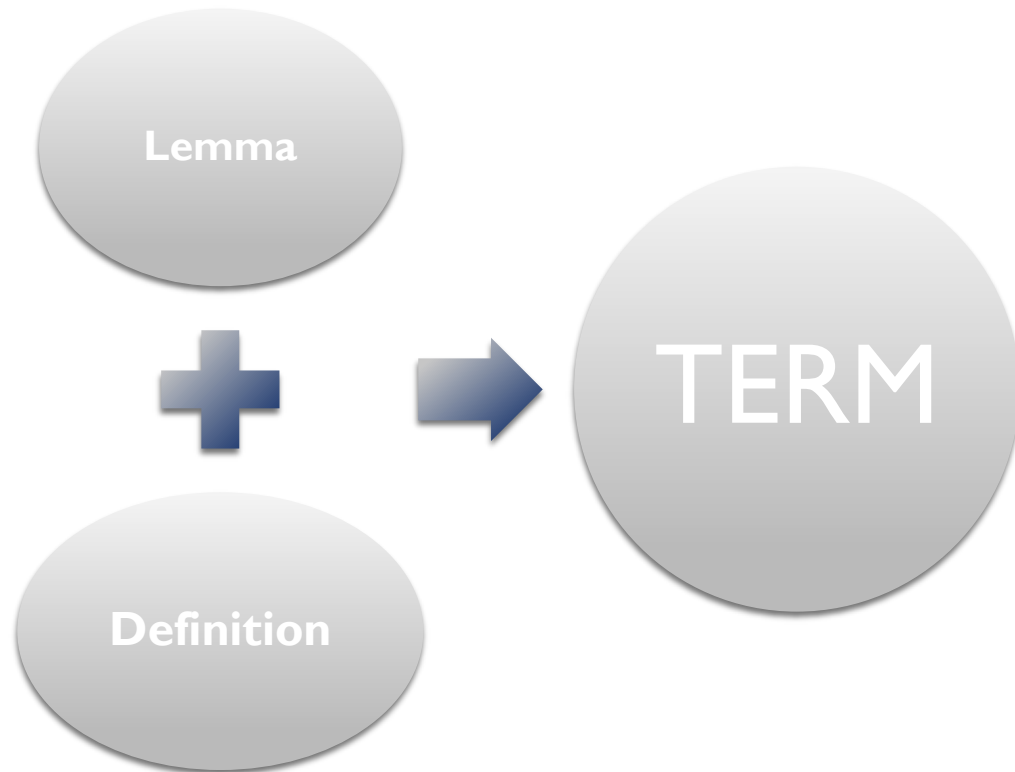
METAstat: a new governance on metadata

- Metadata form an **ecosystem**
- Metadata have **semantic consistency**
- Metadata are **compliant with standards**
- Metadata are **centralized** and have a **strong management**
- Metadata are **actionable**
- Metadata are **managed throughout their lifecycle**
- Metadata are **used throughout the organization and all processes**

METAstat: a system with 3 core module strictly connected



THE TERM: the key element of the semantic infrastructure

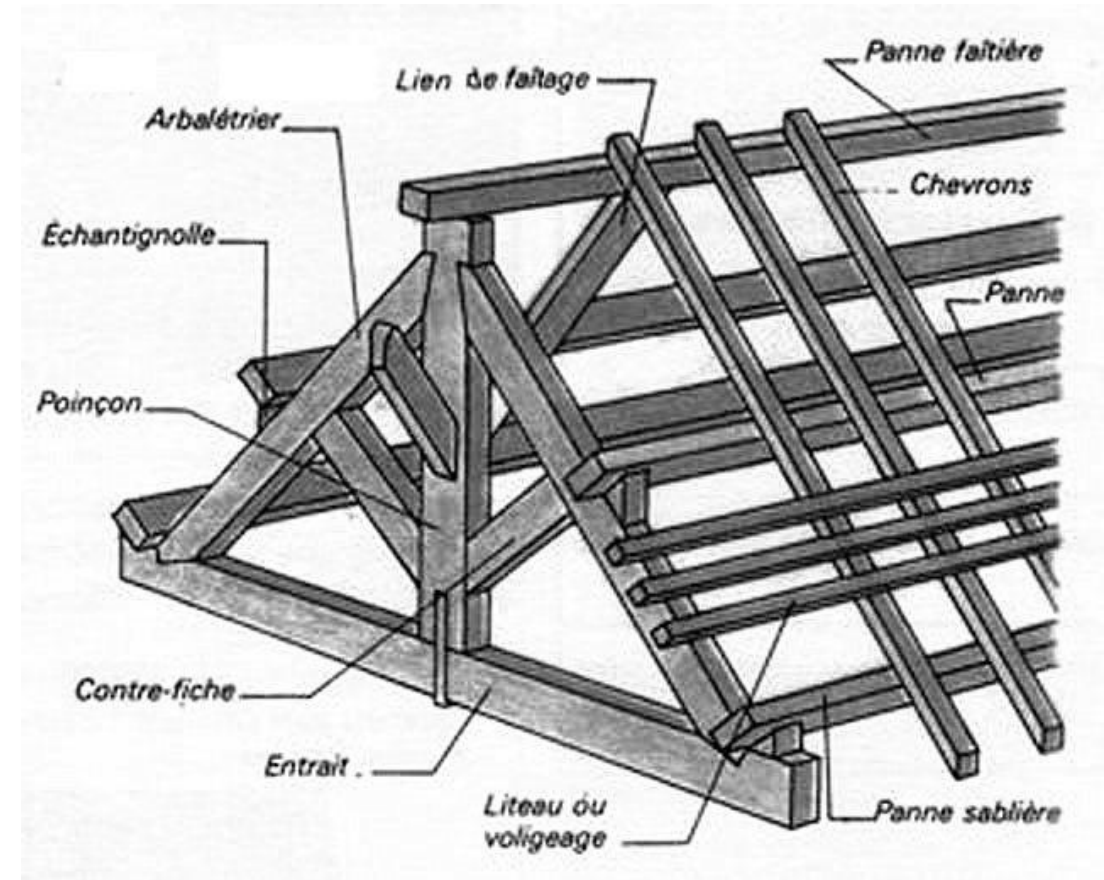


- The term, **MINIMAL UNIT OF LANGUAGE**, is defined as the association of two elements: the lemma and its definition
- It is necessary to refer to the lemma-definition pair as a uniquely determined element because, in the language of NSOs, terms with the same lemma but a different definition (**polysemy**) are currently used and vice versa, terms with the same definition but different lemma (**synonymy**).

THE TERM: terminological standard

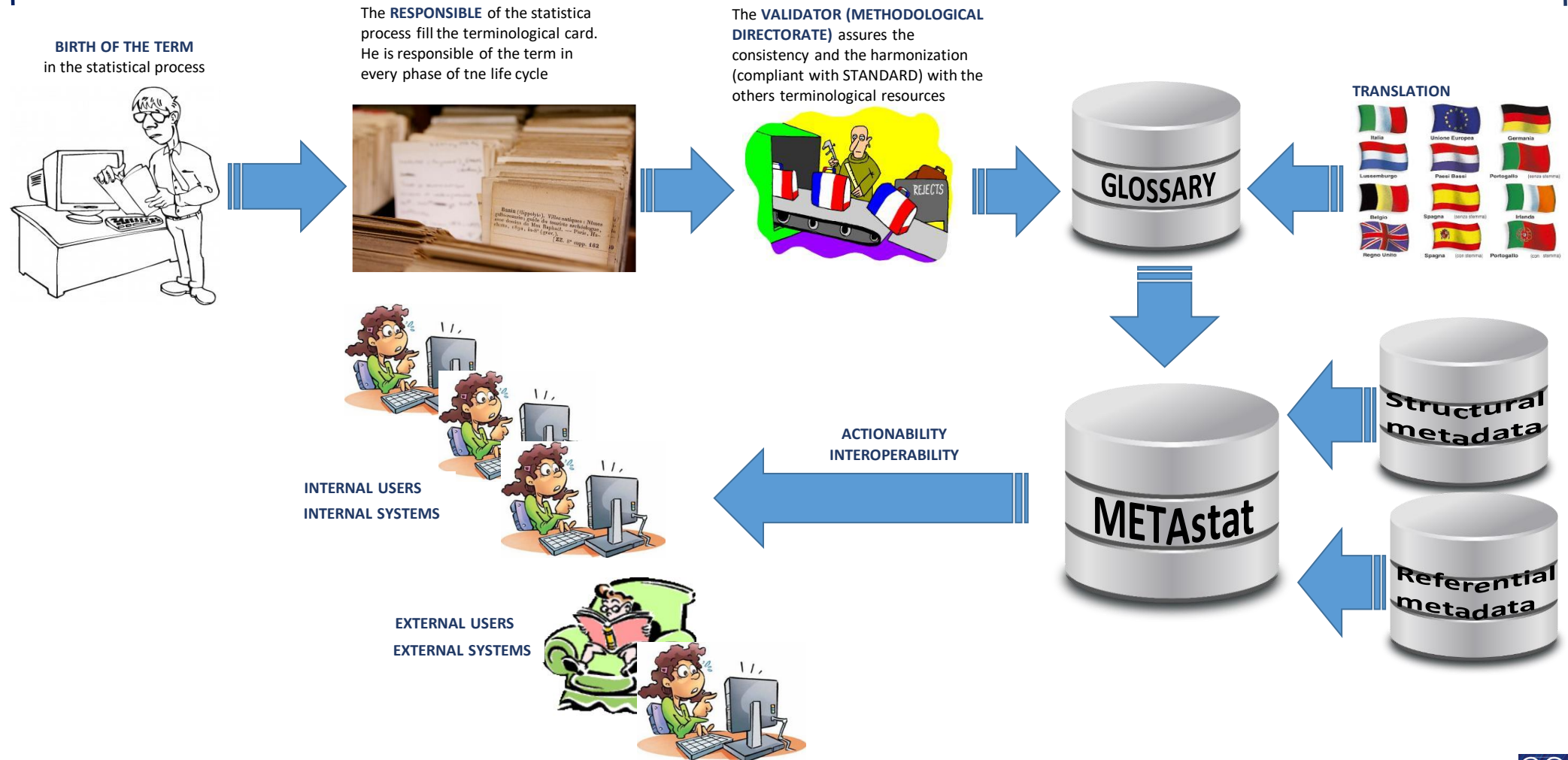
○The **LEMMA** performs the onomasiological function of appointing the concept. It may consist of a single word, a group of words or a more or less complete sentence. It is therefore a medium-length text field.

○The **DEFINITION** has the function of specifying the role assigned to the lemma in the reference domain. It requires a text field of considerable length. Semantically, it is formulated by first placing the concept in the class of objects to which it belongs, thus making the term's domain of reference immediately clear. It then expresses any further explanations, including examples. The specialised definition, from the language of official statistics, can be followed by a generic one.



THE TERM: its governance

THE TERM ACROSS THE GSBPM PHASES



From a Centralized Glossary to a Controlled Vocabulary

CENTRALIZED GLOSSARY

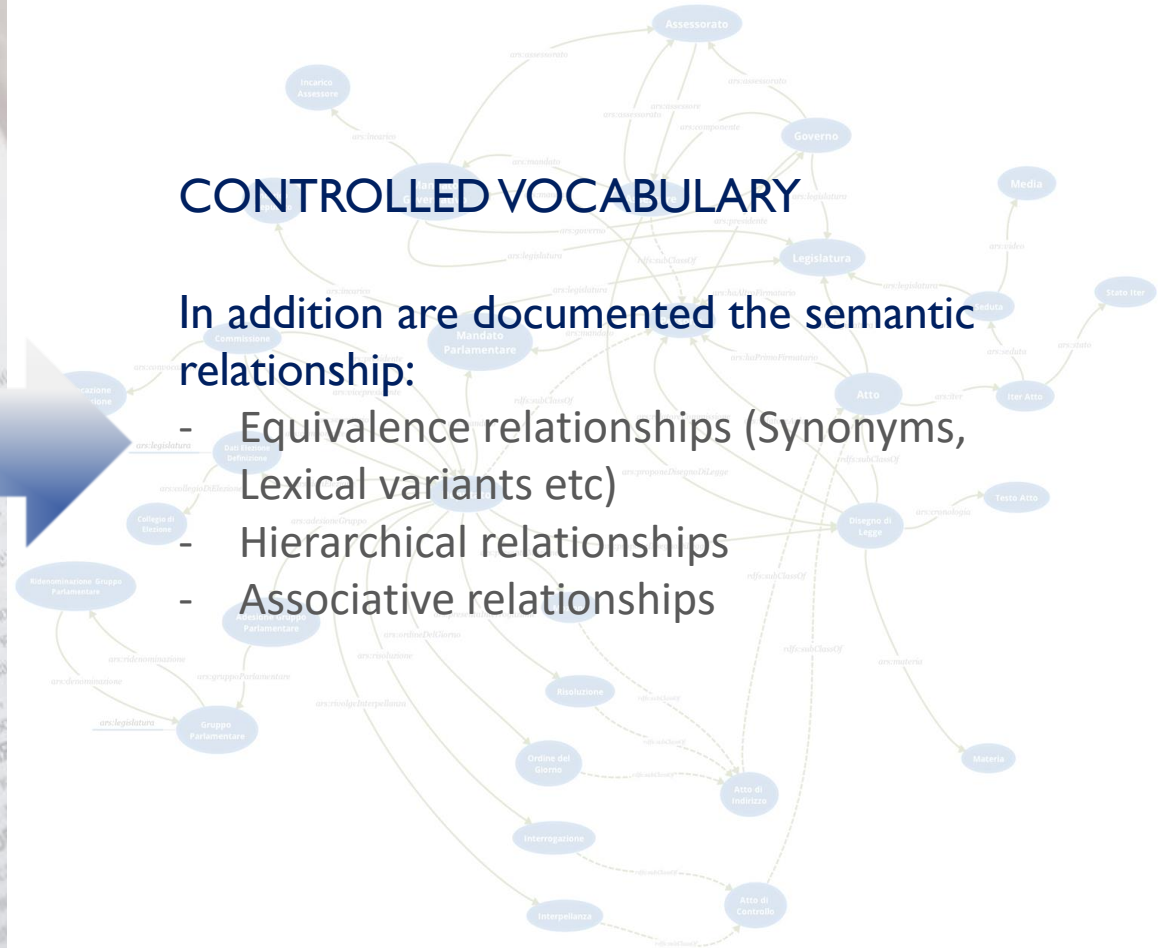
The terms are documented with:

- Start date of activity
- Validity dates
- End of validity date
- Relationship with structural metadata
- Responsible for managing content and modifications
- Supervision on reuse
- Uses

CONTROLLED VOCABULARY

In addition are documented the semantic relationship:

- Equivalence relationships (Synonyms, Lexical variants etc)
- Hierarchical relationships
- Associative relationships



Thanks for the attention

brunini@istat.it

