

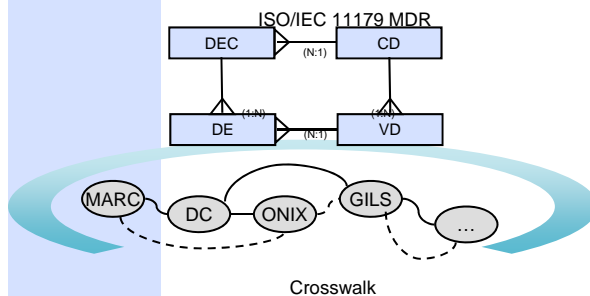
Update for ISO/IEC PDTR 20943-5 Semantic Metadata Mapping Procedure (SMMP)

May, 2010

Tae-Sul Seo and Sung-Joon Lim

tsseo@kisti.re.kr

joon@kdb.or.kr



Contents

- ◆ Brief History
- ◆ Scope
- ◆ Normative References
- ◆ Related Standards
- ◆ Semantic Metadata Mapping Procedure
 - ◆ Data element
 - ◆ **Value domain**
- ◆ Types of Heterogeneity
- ◆ Relationship to ISO/IEC 11179 MDR
- ◆ Future Work

0. Brief History

- ◆ 2007-07-19 ~ 2009-06-22/Study Period
 - ◆ Semantic Harmonization of Metadata

- ◆ 2008-11-20/WG2 Resolution
 - ◆ SC32N1812/WG2N1215 **Project sub-division was approved.**
 - ◆ Semantic Metadata Mapping Procedure (SMMP)
 - ◆ Project Number 1.32.16.01.05.00

- ◆ 2009-04-12/Summary of Voting
 - ◆ SC32N1856 **Proposed SC32 resolution for sub-division was adopted.**

- ◆ 2009-06-26/SC32 Resolution Draft
 - ◆ SC32N1887 **Registration as PDTR was instructed.**

- ◆ 2009-11-19/Update WD
 - ◆ WG2N1317 **ISO/IEC PDTR 20943-5**

- ◆ 2010-05-17/Update WD
 - ◆ WG2N1392 **ISO/IEC PDTR 20943-5**

1. Scope

◆ Background

- ◆ Semantic metadata mapping is required to mediate among data elements to be interoperably used.
- ◆ Metadata crosswalk needs to be elaborated to have semantics.

◆ Purpose

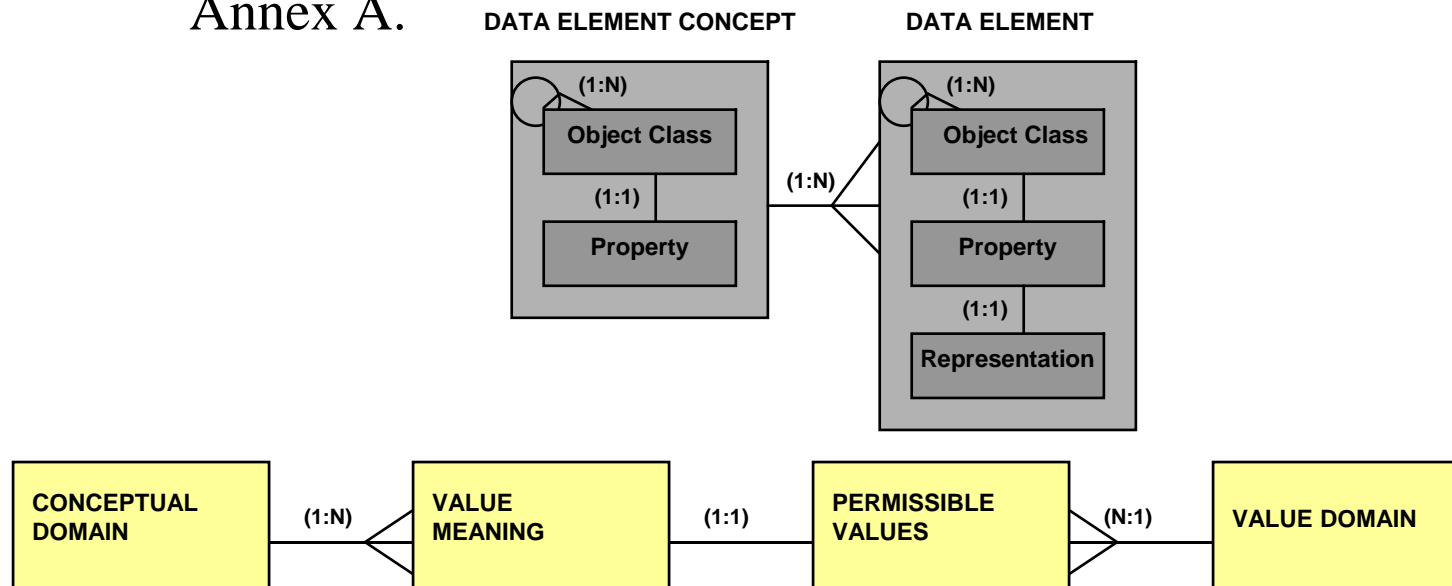
- ◆ To set up a procedure for improving the semantic harmonization of metadata.

◆ Scope

- ◆ The procedure is recommended for use in a specific domain.
- ◆ **The procedure consider the interoperability of data element and value domain.**

2. Normative references

- ◆ ISO/IEC 11179-1:2004
 - ◆ Borrowing the fundamental model of data element.
- ◆ ISO/IEC 11179-5:2005
 - ◆ Borrowing the principles of naming convention and example in Annex A.



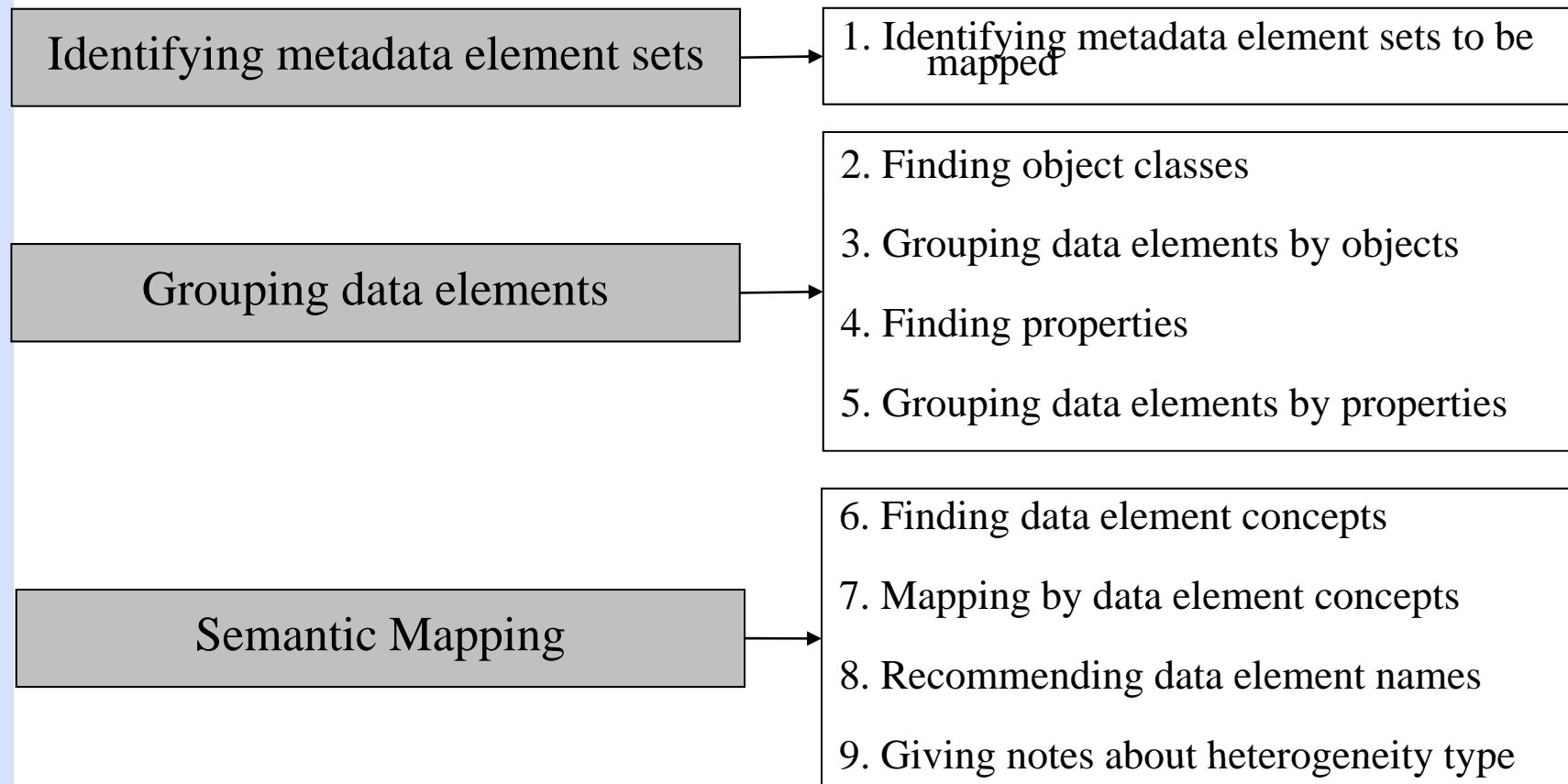
3. Related standards

- * ISO 25964-2 WD
 - * Interoperability with other vocabularies.
 - * ISO 2788:1986 + ISO 5964:1985 = ISO 2596
 - * Types of heterogeneity (Hierarchical and lexical differences) are related.
- * ISO/IEC 15944-7 WD
 - * 5.2 Principles and rules for the development of a definition for a concept and assignment of an appropriate term in a controlled vocabulary
 - * A kind of general rule for the development of term definition
- * ISO/IEC 24751-8
 - * Language Accessibility and Human Interface Equivalencies (HIEs) in e-Learning Applications: Principles, Rules and Metadata Elements
 - * A type of heterogeneity (domain difference) is related.

4. Semantic Metadata Mapping Procedure : Data Element

Main Processes

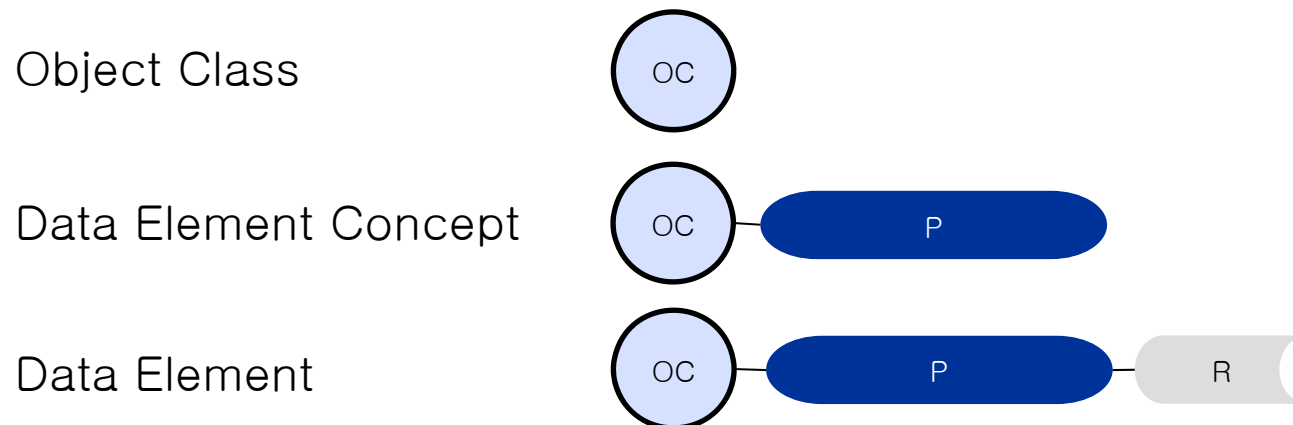
Sub-Processes



4. Semantic Metadata Mapping Procedure : Data Element

Premises of the SMMP

- * Object class should be same for data element concept mapping.
- * Data elements are semantically identical if data element concepts are same.
- **If data element concept (= object class + property) is same, the data element is semantically same.**



4. Semantic Metadata Mapping Procedure : Value Domain

The value of each data element may be converted or rearranged to solve the inconsistency of data type, when the data types are not equal.

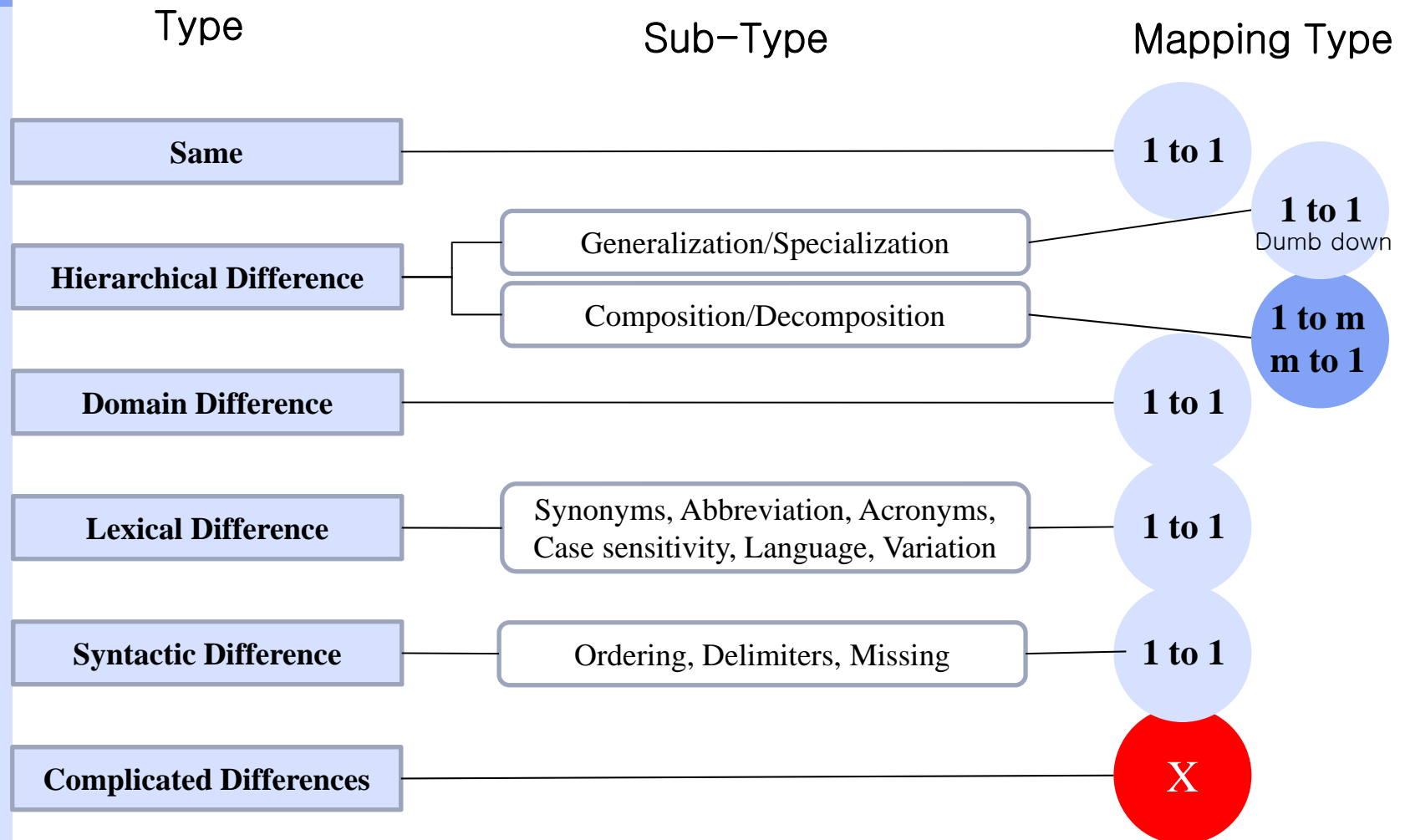
* **Conversion**

- **Code conversion:** In case of using different code sets sharing same conceptual domain for the value domain, a code set is converted to the other code set.
- **Letter conversion:** In case of using characters such as abbreviation, acronyms, different case, lexical variation, or different delimiters for the value domain, a part of the value is converted to a preferred representation.
- **Unit conversion:** In case of using numeric value for the value domain, the numeric value is converted using an appropriate conversion rule.

* **Structural rearrangement**

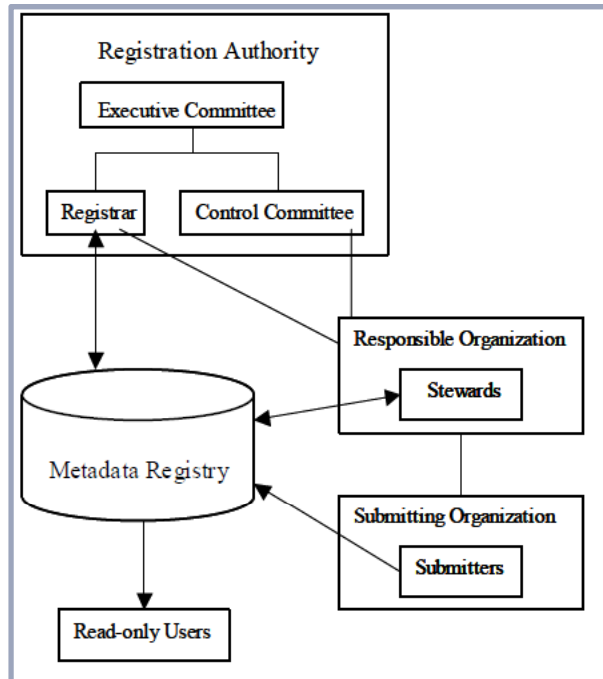
- **Composition:** If a value is separated into two or more data elements, those values are composed into one value of an appropriate data element.
- **Decomposition:** If a value should be separated into two or more data elements, the value is decomposed into two or more values of appropriate data elements.
- **Order change:** If a value is composed of two or more words which are ordered differently, the order of the words are changed in an appropriate order.

A. Types of Semantic Heterogeneities



B. Relationship to ISO 11179 MDR

ISO/IEC 11179 MDR



Read only user

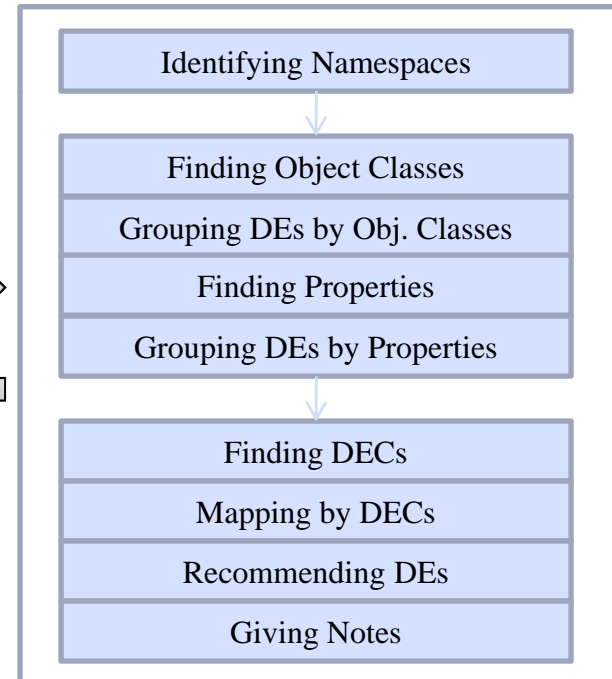
Context
Obj. Class
Property
DEC
DE



Submitter

Obj. Class
Property
DEC
DE

Metadata Mapping Agent



* Future Work

- ◆ ISO/IEC TR 20943-5 is being reviewed and improved to go into the next stage.
 - ◆ The WD is being updated.
- ◆ It is expected to be PDTR stage by July 15 2010.
 - ◆ We update the WD and submit it in Kunming meeting on May 2010.