

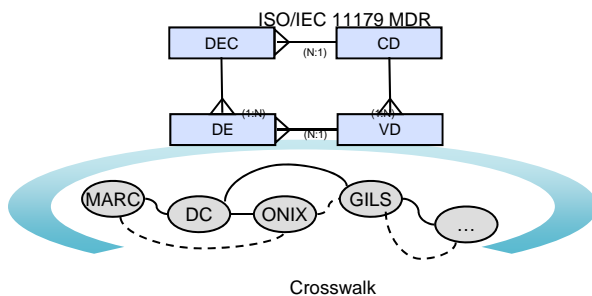
Update for ISO/IEC PDTR 20943-5 Semantic Metadata Mapping Procedure (SMMP)

November 19, 2009

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
- * Schematic Explanation
- * Example
- * 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.**

- * 2009-10-15/WD submission
 - * WG2N1317 **ISO/IEC WD 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.

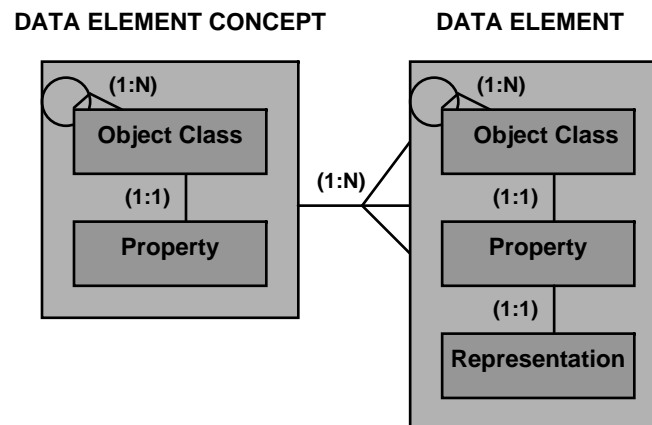
* Scope

- * The procedure is recommended for use in a specific domain.
- * The procedure doesn't consider the interoperability of values.



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

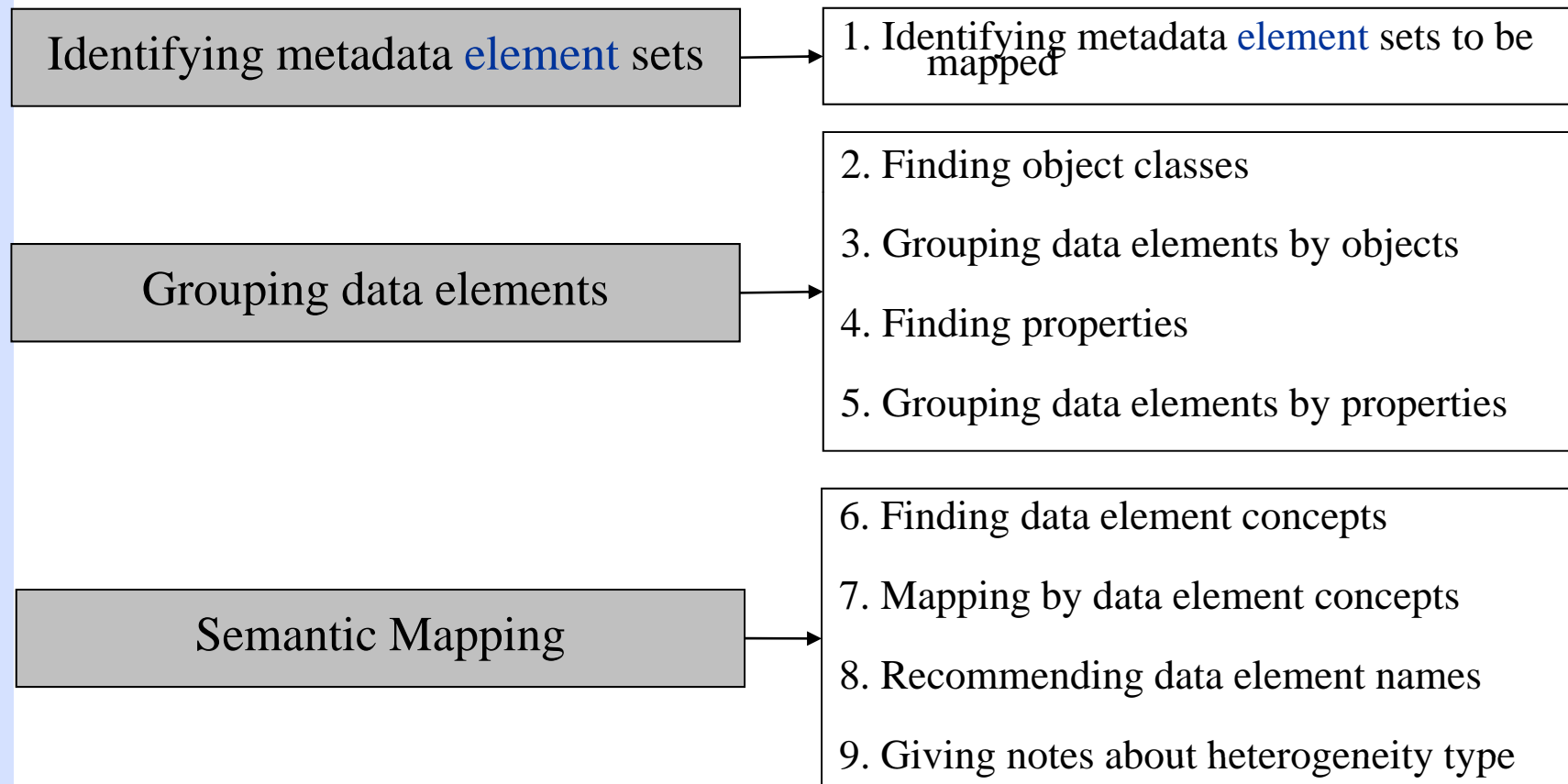
- * ISO 25964-2 WD
 - * Interoperability with other vocabularies.
 - * ISO 2788:1986 + ISO 5964:1985 = ISO 2596
 - * Types of heterogeneity (Hierarchical and lexical differences) are related.
- * ISO/IEC 15944-7 WD
 - * 5.2 Principles and rules for the development of a definition for a concept and assignment 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

Main Processes

Sub-Processes

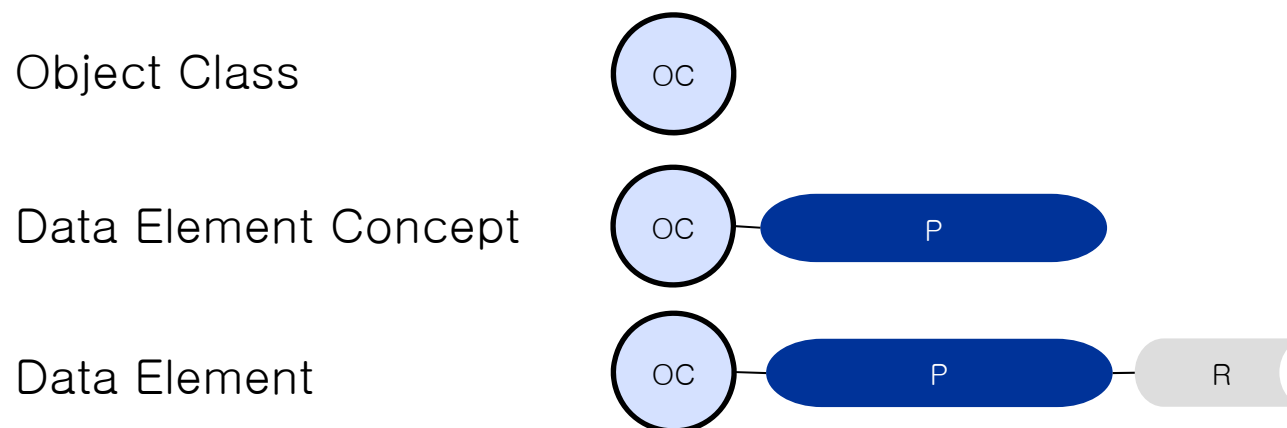




4. Semantic Metadata Mapping Procedure

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 and property) is same, the data element is semantically same.**

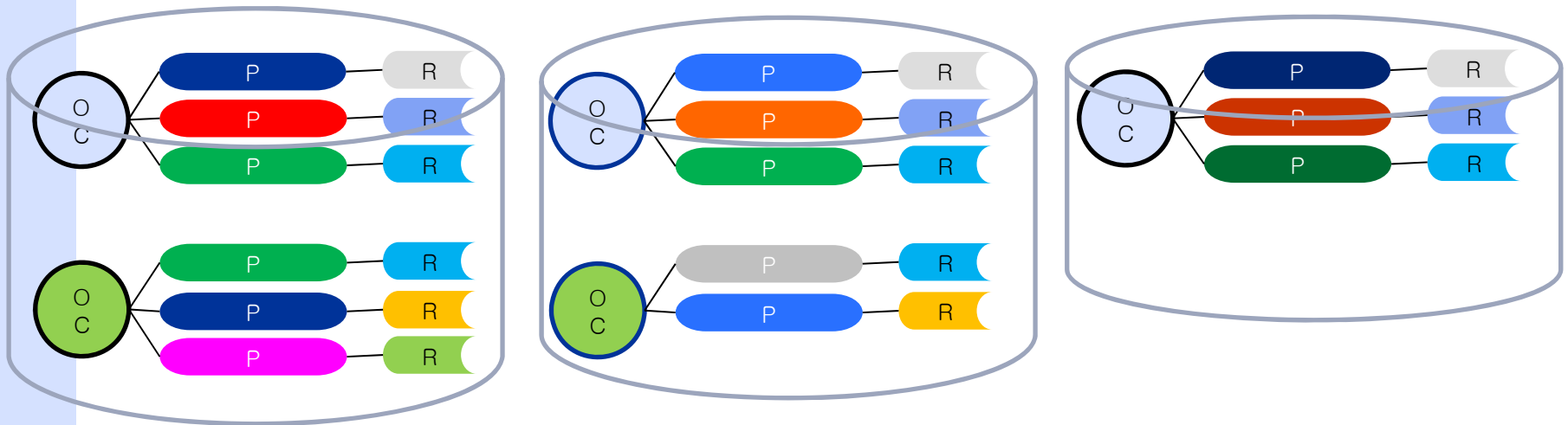




5. Schematic Explanation

Identifying metadata element sets

1. Metadata element sets to be mapped: 3



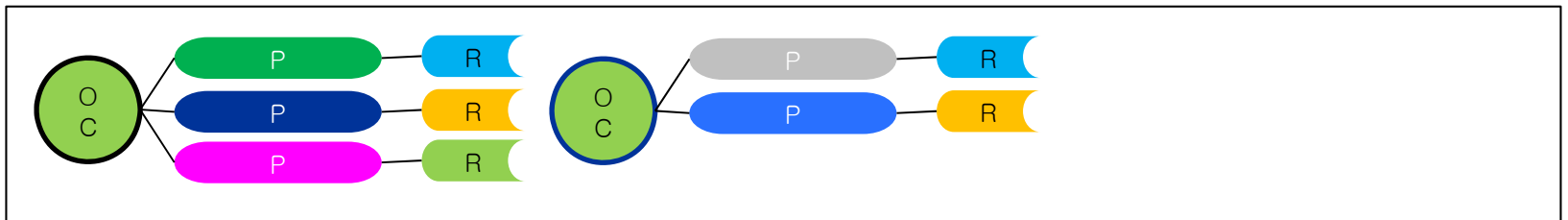
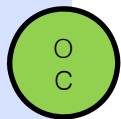
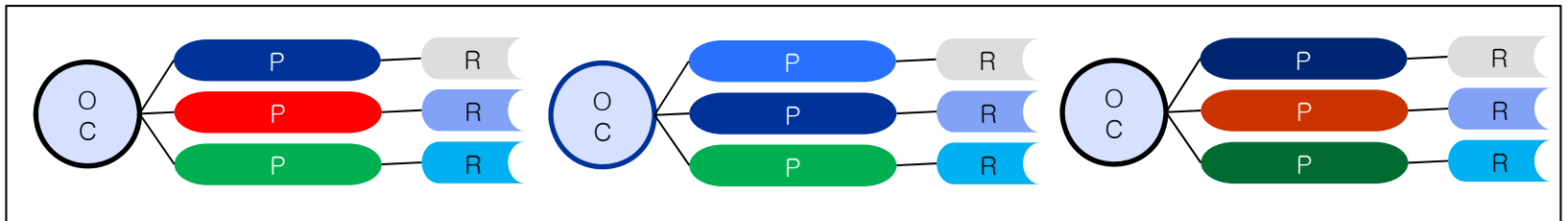
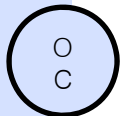
* Same color group means same meaning.



5. Schematic Explanation

Grouping data elements

- 2. Finding object classes
- 3. Grouping data elements by object classes

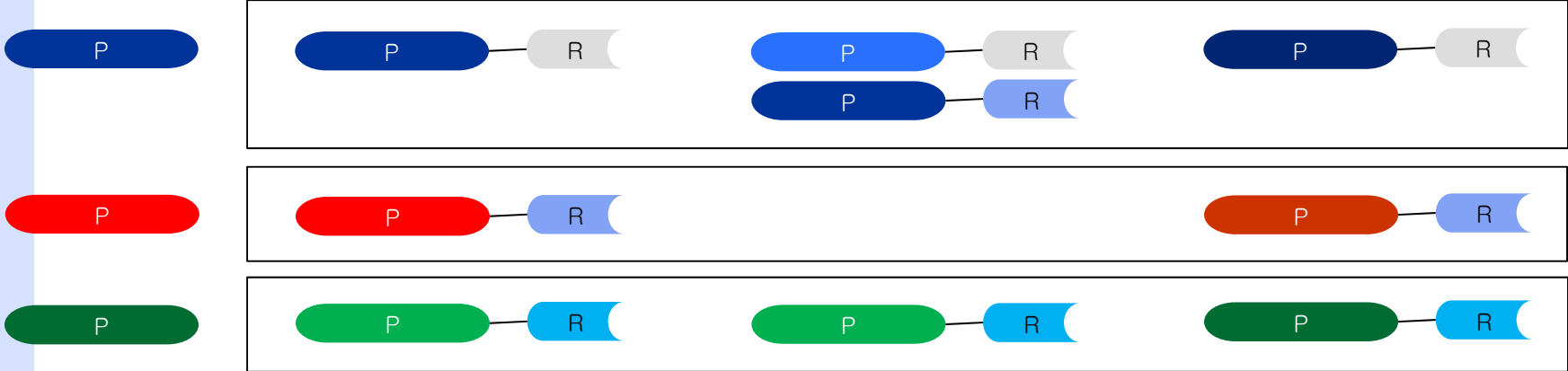




5. Schematic Explanation

Grouping data elements

- 4. Finding properties
- 5. Grouping data elements by properties

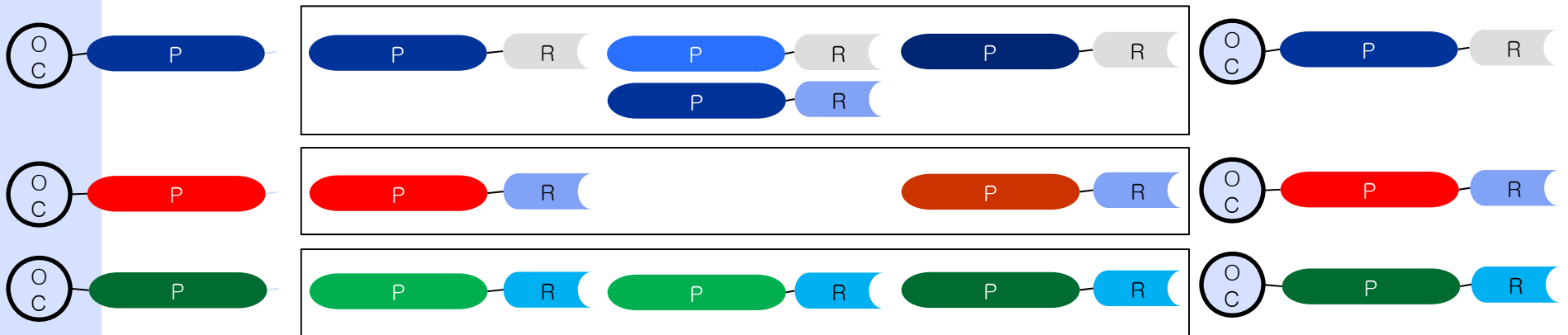




5. Schematic Explanation

Semantic Mapping

- 6. Finding data element concepts
- 7. Mapping by data element concepts
- 8. Recommending data element concepts



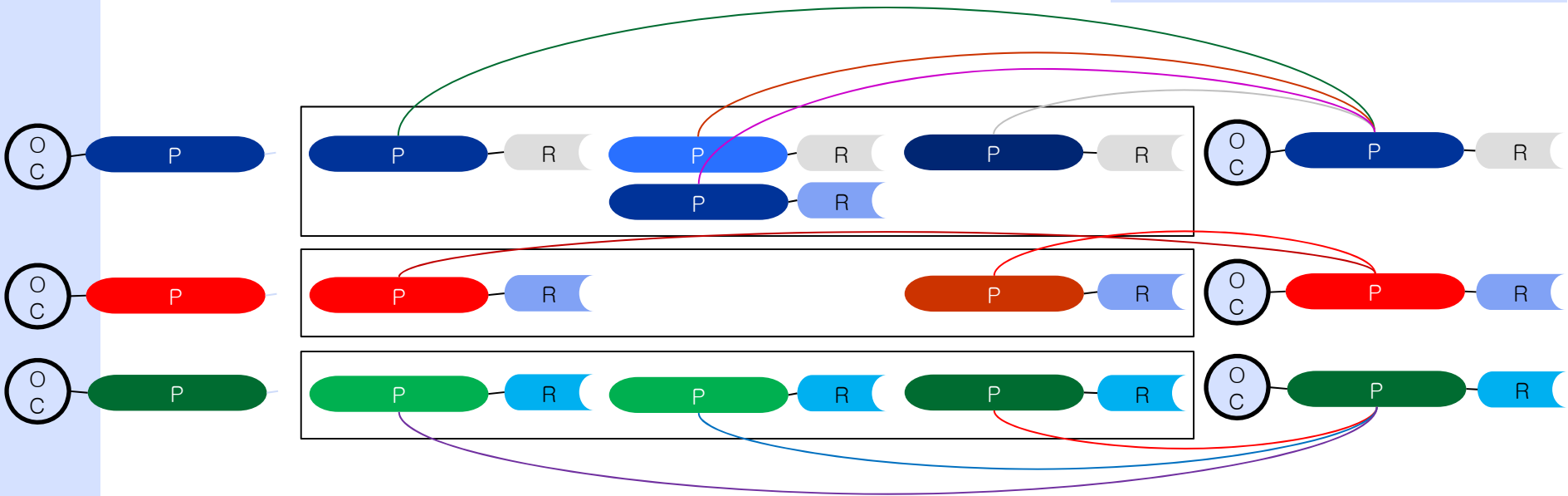


5. Schematic Explanation

Semantic Mapping

9. Giving notes according to types of heterogeneity

- H – Hierarchical Differences
- D – Domain Differences
- L – Lexical Differences
- S – Syntactic Differences
- C – Complicated Differences





6. Example

Identifying metadata element sets

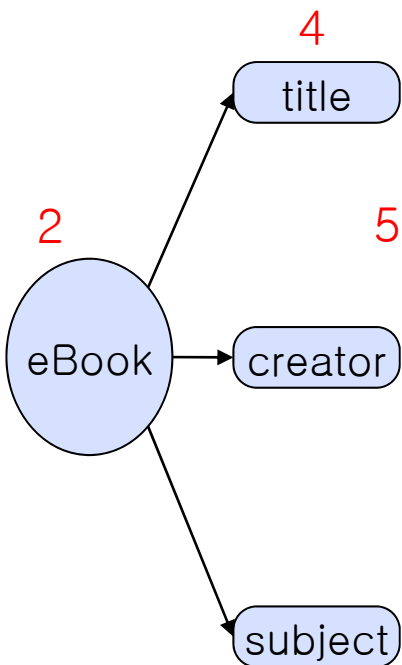
1. Metadata sets to be mapped: 3

	OpenEBPS	MODS	TEI header
Context	Description of Electronic Book	Description of Library resources	Encoding methods for machine-readable texts
Number of fields	15	About 60 (top level: 20)	Over 20
Authority	Open eBook Forum	Library of Congress	TEI Consortium



6. Example

Grouping data elements



OpenEBPS	MODS	TEI
title	title subTitle partNumber partName nonSort	title seriesStmt:title seriesStmt:idno
creator(role) creator(file-as)	name:role name:namePart name:displayForm name:affiliation name:discription	author
subject	topic classification catographics occupation	keyword class catRef

3



6. Example

Semantic Mapping

6

7

8

DEC	OpenEBPS		MODS		TEI		Recommaned DE
ebookTitle	title		title	H/dec	title	H/dec	ebookTitle
			subTitle	H/dec	seriesStmt:title	H/dec, S/del	
ebookAuthor	creator(role)	C	name:role	C	author	S/mis	ebookAuthorName
	creator(file-as)	D	name:namePart	D			
ebookSubject	subject		topic	L/syn	keyword	L/syn	ebookSubject
			classification	L/syn	class	L/syn	

9

H – Hierarchical Differences

D – Domain Differences

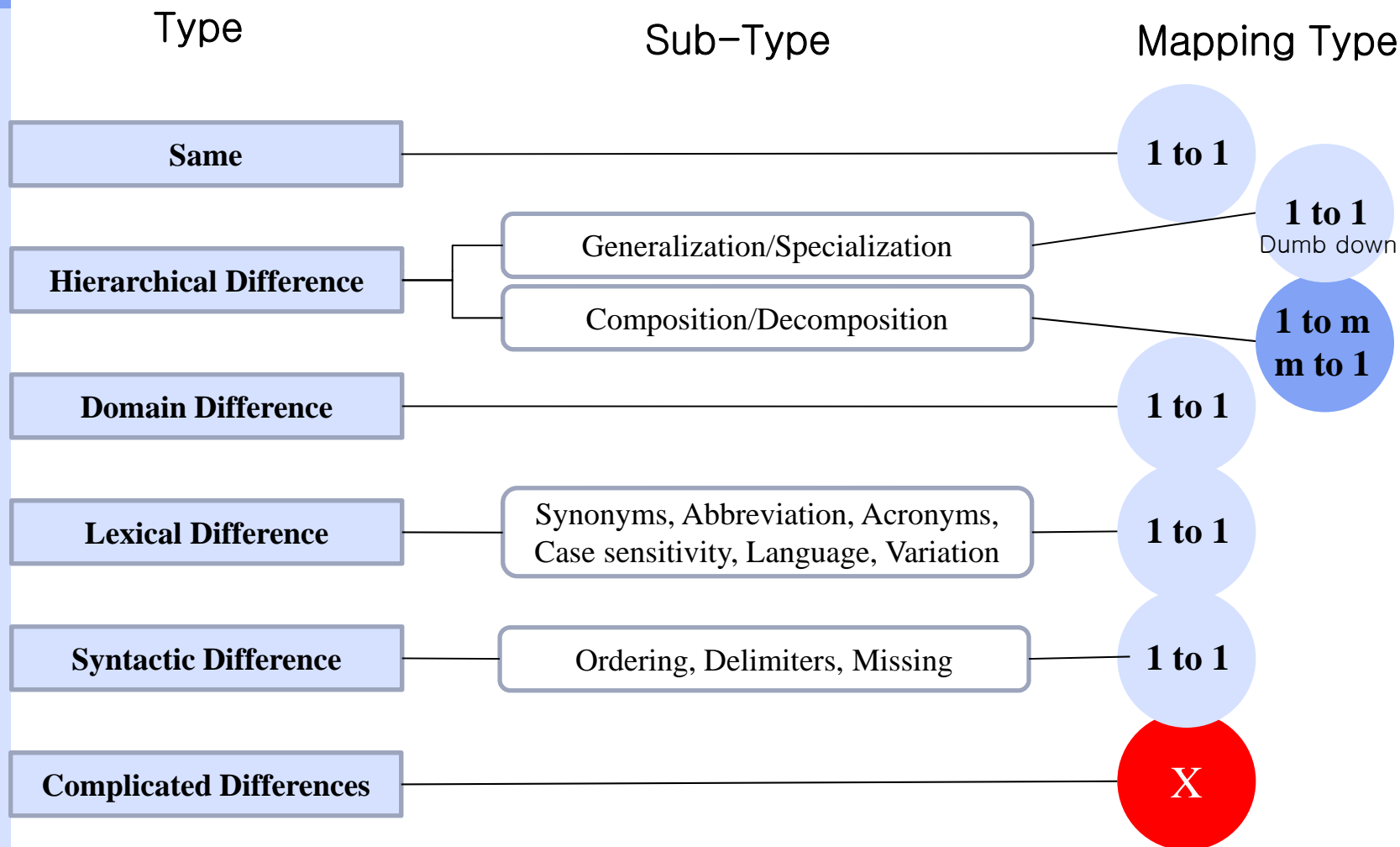
L – Lexical Differences

S – Syntactic Differences

C – Complicated Differences



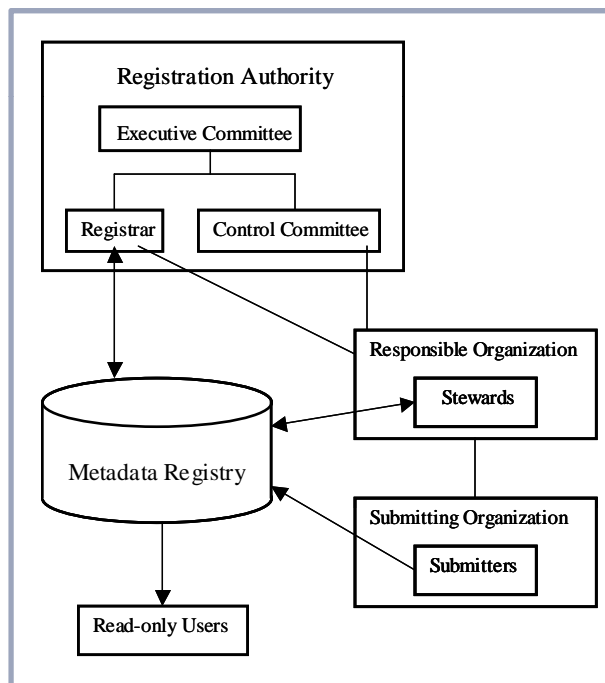
A. Types of Semantic Heterogeneities



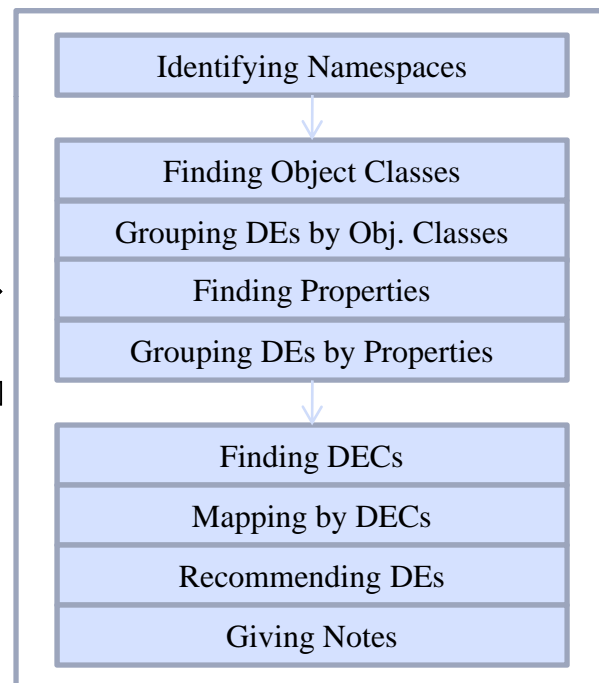


B. Relationship to ISO 11179 MDR

ISO/IEC 11179 MDR



Metadata Mapping Agent



Read only user

Context
Obj. Class
Property
DEC
DE



Submitter

Obj. Class
Property
DEC
DE





* 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 until February 1 2010.
 - * We will update the WD and submit it within December 31 2009.