



# **Study Report and NP Proposal on Semantic Metadata Mapping Procedure**

2008. 11. 19

Tae-Hoon Lim and Tae-Sul Seo

[taehoon@dpc.or.kr](mailto:taehoon@dpc.or.kr)

[tsseo@kisti.re.kr](mailto:tsseo@kisti.re.kr)



# Brief History

- \* 2007-07-19/New York, USA
- \* Semantic harmonization of metadata
- \* WG2N1042(SC32N1658)
- \* Request for a new study period
  
- \* 2007-12-06/Seoul, Korea
- \* Semantic harmonization of metadata
- \* WG2-N1077, WG2-N1078
- \* Study report
  
- \* 2008-05-xx/Sydney, Australia
- \* Semantic Metadata Mapping (SMM) Processes
- \* WG2N1122
- \* Request for extending the study period



# Motivation

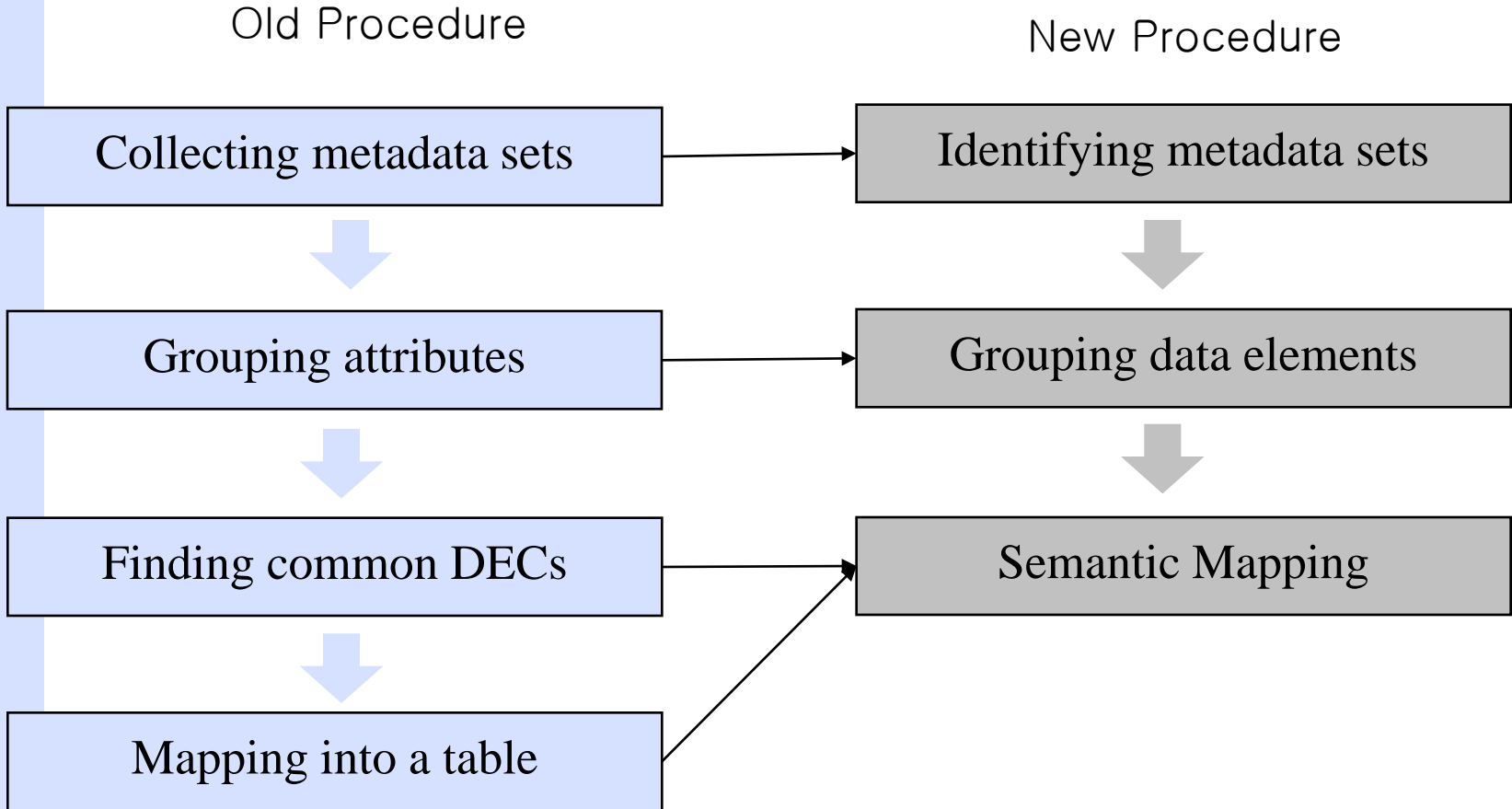
- \* A measure is needed to mediate between metadata sets already used.
- \* A metadata crosswalk has poor semantics because it is usually based on a simple one-to-one mapping.
- \* Therefore, a metadata crosswalk is required to be elaborated in order to have semantics.
- \* The idea of ISO/IEC 11179 can be helpful to improve metadata crosswalk semantically.
- \* This candidate standard includes a procedure called semantic metadata mapping procedure (SMMP), which deals how to identify metadata to be mapped and how to categorize the relationship of each pair of metadata.



# Purpose

- \* The purpose of this standard is to set up a procedure for making a metadata mapping table that conform to ISO/IEC 11179 standard, and thus, to improve semantic harmonization of metadata.

# Semantic Metadata Mapping Procedure

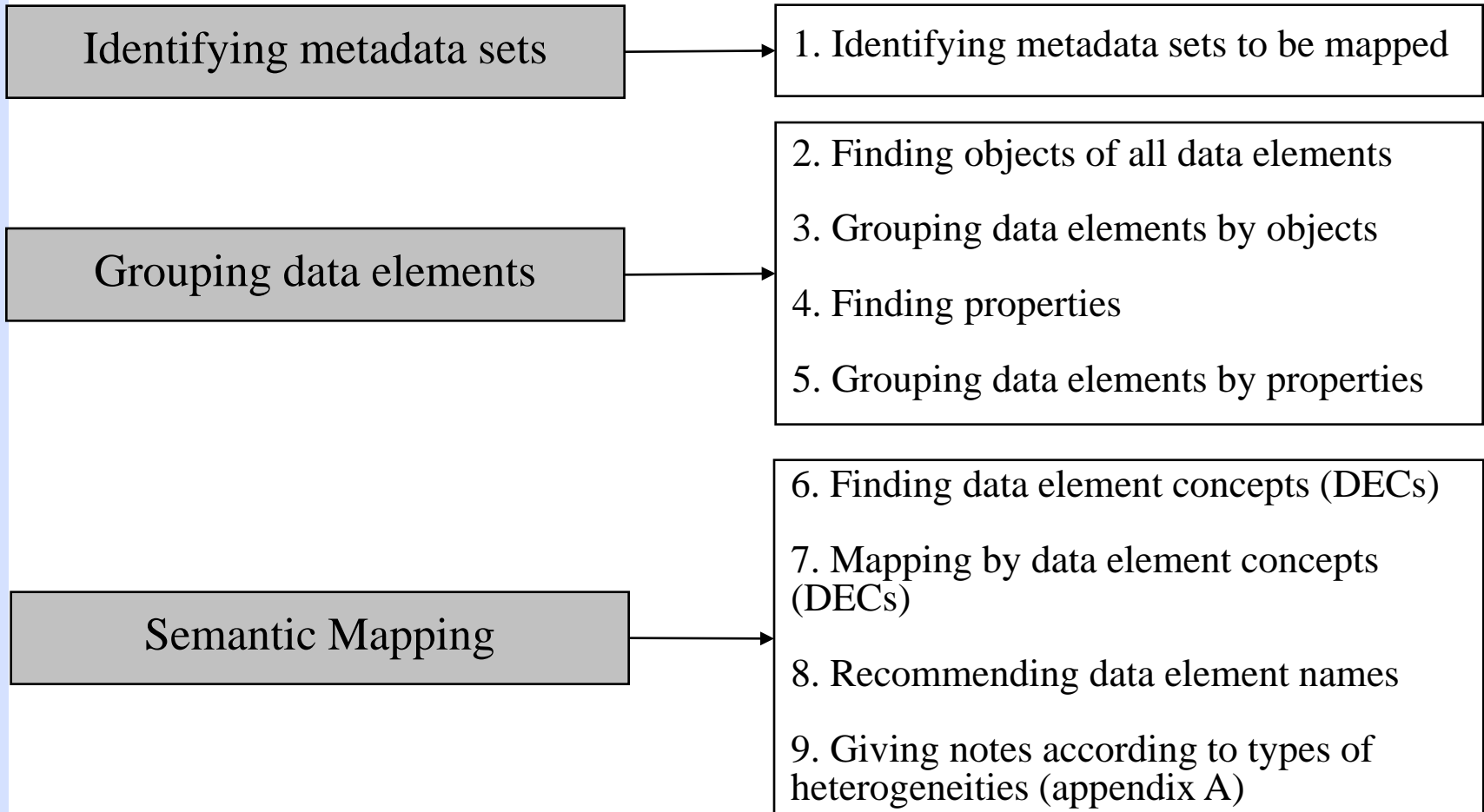




# Semantic Metadata Mapping Procedure

## Main Processes

## Sub-Processes





# Example of eBook

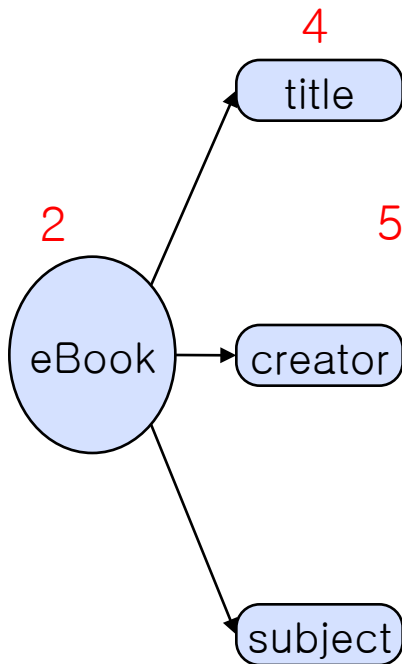
## Identifying metadata sets

1. Metadata sets to be mapped: three metadata sets

	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
Sample data	yes	no	yes
Authority	Open eBook Forum	Library of Congress	TEI Consortium

# Example of eBook

## 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



# Example of eBook

## Semantic Mapping

6

7

8

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

H – Hierarchical Differences

D – Domain Differences

L – Lexical Differences

S – Syntactic Differences

C – Complicated Differences

9



# Types of Semantic Heterogeneities

- \* **Hierarchical Difference;** due to different level of detail
  - \* Generalization : Specialization, Composition : Decomposition
- \* **Domain Difference;** due to different context and culture
- \* **Lexical Difference;** is different appearance
  - \* Synonyms, Abbreviation, Acronyms, Case sensitivity, Language, Variation
- \* **Syntactic Difference;** due to different arrangement of parts
  - \* Ordering, Delimiters, Missing
- \* **Complicated Differences;** due to different policies



# Summary

- \* The working draft has been made during the study period.
- \* The description system for mapping, the types of semantic heterogeneity, was elaborated.
- \* It is desirable to be proposed as an IS in the future.



Thank you!