
**Information technology — Procedures
for achieving metadata registry
content consistency —**

Part 5:
Metadata mapping procedure

iTeh STANDARD PREVIEW
*Technologies de l'information — Procédures pour réaliser la
consistance du contenu de l'enregistrement des métadonnées —
Partie 5: Procédure de mappage des métadonnées*
(standards.iteh.ai)

[ISO/IEC TR 20943-5:2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

[https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-
adb44c1c0da0/iso-iec-tr-20943-5-2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC TR 20943-5:2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

[https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-
adb44c1c0da0/iso-iec-tr-20943-5-2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2013

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

	Page
Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	2
5 Metadata mapping procedure	3
5.1 Overview	3
5.2 Data element concept mapping	3
5.3 Conceptual domain mapping	7
Annex A (informative) Relationship between metadata mapping and registry	9
Annex B (informative) Factoring and types of heterogeneity	10
Bibliography	13

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC TR 20943-5:2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

[https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-
adb44c1c0da0/iso-iec-tr-20943-5-2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard (“state of the art”, for example), it may decide to publish a Technical Report. A Technical Report is entirely informative in nature and shall be subject to review every five years in the same manner as an International Standard.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 20943-5 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 32, *Data management and interchange*.

ISO/IEC TR 20943 consists of the following parts, under the general title *Information technology — Procedures for achieving metadata registry content consistency*:

- *Part 1: Data elements*
- *Part 3: Value domains*
- *Part 5: Metadata mapping procedure*
- *Part 6: Framework for generating ontologies*

Introduction

Metadata differs in many ways, yet for the purposes of data and semantic interoperability, those differences must be bridged or harmonized. For example, synonyms may cause discrepancies. If two data elements mean the same thing but have different names, that difference may cause data under each to be flagged as incompatible when they aren't.

Usually, two systems do not share the same model, and that is because the categories represented in the models were not factored in the same way. This situation inhibits interoperability. There may be two or more metadata element sets applicable to an information object. For example, metadata schemas, such as DC (Dublin Core),^[10] MARC (MACHine Readable Cataloguing),^[11] and MODS (Metadata Object Description Schema)^[12] can be used to describe books. So, the result might be significantly different data elements in different schemas, even though the schemas describe the same objects.

ISO/IEC 11179 provides a framework for achieving interoperability of metadata between systems. A metadata registry based on ISO/IEC 11179 offers a good way to secure interoperability among databases. However, there are many metadata sets which are not following ISO/IEC 11179. In order to mediate among plural data elements already developed or used, other measures are necessary. In general, interoperability may be achieved through conformity to some set of provisions. For example, metadata crosswalk is the most commonly used way to map a data element to another data element. However, the metadata crosswalk is meaningful only for exact matching; it provides a simple one-to-one mapping table between data elements without any explanation about the relationship. Therefore, the metadata crosswalk needs to be elaborated in order to cover cases other than exact matching.

Some other approaches have been tried to provide guidelines or a model for harmonization of metadata and data in especially the transport industry. A Model for Semantic Equivalence Discovery for Harmonizing Master Data,^[9] presented at OTM 2009 Workshops, suggests a model for semantic mapping of master data. ISO/TR 25100^[8] reviews four harmonization processes applicable to an ITS/TICS central data registry and ITS/TICS data dictionaries and recommends practical guidelines. Both literatures are examples restricted within the transport industry but including a rationale for the need of metadata mapping in general.

This part of ISO/IEC TR 20943 describes a metadata mapping procedure (MMP), which can maximize the interoperability among metadata.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC TR 20943-5:2013](#)

[https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-
adb44c1c0da0/iso-iec-tr-20943-5-2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

Information technology — Procedures for achieving metadata registry content consistency —

Part 5: Metadata mapping procedure

1 Scope

The purpose of this part of ISO/IEC TR 20943 is to describe a procedure for establishing metadata crosswalks based on the ISO/IEC 11179 series, subsequently improving mapping quality between metadata.

Therefore, this part of ISO/IEC TR 20943 describes a metadata mapping procedure (MMP), which can maximize the interoperability among ISO/IEC 11179-based registries through achieving metadata registry content consistency.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 11179-1:2004, *Information technology — Metadata registries (MDR) — Part 1: Framework*

ISO/IEC 11179-3:2013, *Information technology — Metadata registries (MDR) — Part 3: Registry metamodel and basic attributes*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 11179 and the following apply.

3.1

crosswalk

mapping (3.8) of the elements, semantics, and syntax from one metadata scheme to those of another

Note 1 to entry: Definition taken from *Understanding Metadata*, 2004, NISO Press.[15]

3.2

complicated difference

type of *heterogeneity* (3.4) that cannot be harmonized

3.3

domain difference

type of *heterogeneity* (3.4) arising from different kinds of contexts or cultures

3.4

heterogeneity

difference arising from different descriptions of the same concept

3.5

hierarchical difference

type of *heterogeneity* (3.4) arising from the different levels of details

3.6 interoperability

ability of multiple systems with different hardware and software platforms, data structures, and interfaces to exchange data with minimal loss of content and functionality

Note 1 to entry: Definition taken from *Understanding Metadata*, 2004, NISO Press.^[15]

3.7 lexical difference

type of *heterogeneity* (3.4) arising from different designations

[SOURCE: ISO 1087-1:2000, 3.4.1]

3.8 mapping

correspondence between instances of one model and instances of another model that represent the same meaning

[SOURCE: ISO/TS 18876-1:2003, 3.1.14]

3.9 metadata element set

small and fundamental group of data elements or data element concepts through which resources can be described and catalogued in a domain

EXAMPLE DC (Dublin Core), MARC (MACHine Readable Cataloguing), MODS (Metadata Object Description Schema), etc.

ITeH STANDARD PREVIEW
(standards.iteh.ai)

3.10 primary metadata element set

metadata element set (3.9) to which other *metadata element sets* are grouped

[https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-
adb44c1c0da0/iso-iec-tr-20943-5-2013](https://standards.iteh.ai/catalog/standards/sist/159725d3-a74f-4dc0-ac48-adb44c1c0da0/iso-iec-tr-20943-5-2013)

3.11 secondary metadata element set

metadata element set (3.9) which are grouped according to the *primary metadata element set* (3.10)

3.12 syntactic difference

type of *heterogeneity* (3.4) arising from varying arrangement of parts

4 Abbreviated terms

DC	Dublin Core
DCMI	Dublin Core Metadata Initiative
DE	Data Element
DEC	Data Element Concept
MARC	MACHine Readable Cataloguing
MDR	MetaData Registry
MMP	Metadata Mapping Procedure
MODS	Metadata Object Description Schema
NISO	National Information Standards Organization (NISO is a USA organization: www.niso.org)

OEBPS Open EBook Publication Structure

TEI Text Encoding Initiative

5 Metadata mapping procedure

5.1 Overview

This part of ISO/IEC TR 20943 covers data element concept mapping and conceptual domain mapping. The mapping procedure for data element concept, described in this part of ISO/IEC TR 20943, is applicable to data element mapping, while the mapping procedure for conceptual domain includes value domain mapping.

5.2 Data element concept mapping

5.2.1 Overview

The procedure for data element concept mapping consists of three main processes as shown in [Figure 1](#).

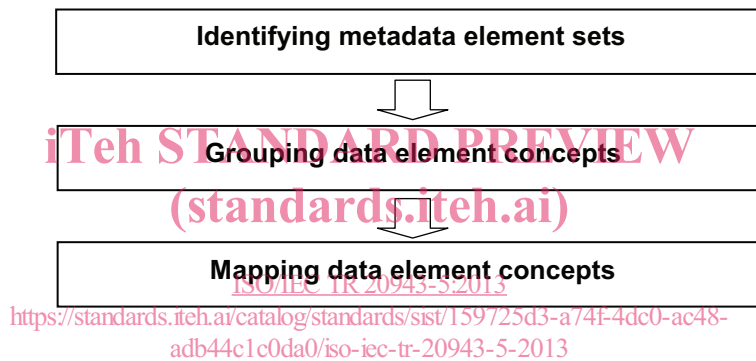


Figure 1 — Procedure for data element concept mapping

The first process is to identify metadata element sets required to be mapped. It is necessary to survey available metadata element sets (in a specific domain).

The second process is to group data element concepts obtained from the identified metadata element sets. This includes four consecutive sub-processes: finding object classes, grouping all data element concepts by object class, finding their properties, and grouping all data element concepts by property.

The last process involves mapping data element concepts. In this process, it is necessary to arrange all data element concepts into a table. Notes on the type of mapping are included in every slot of the table.

[Figure 2](#) shows all sub-processes related to corresponding main processes.

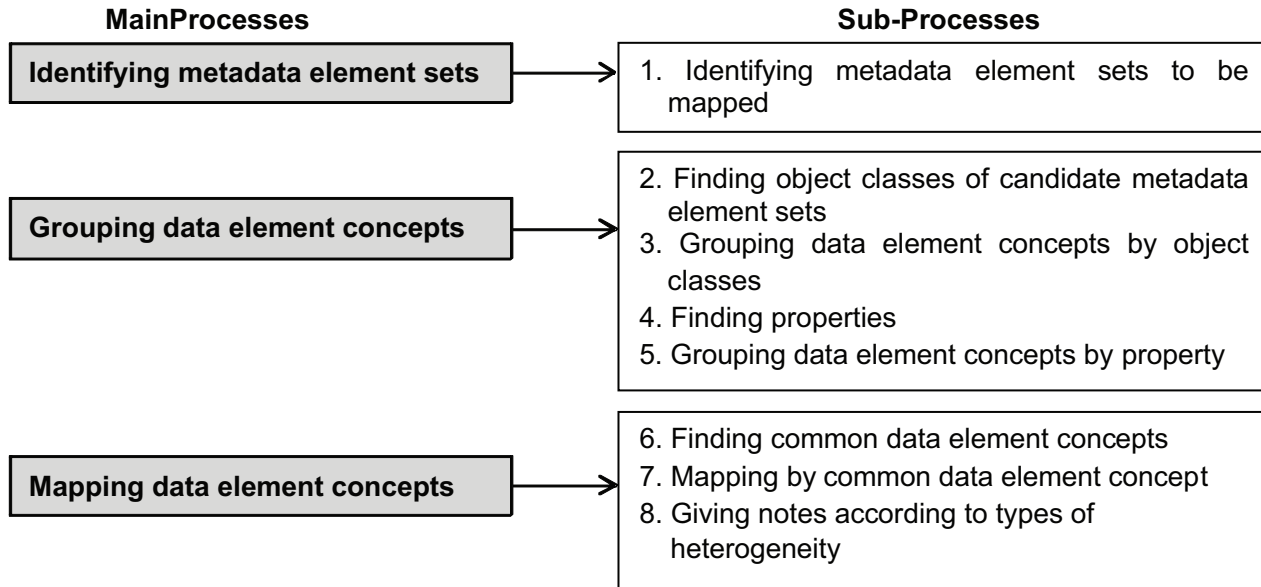


Figure 2 — Main and sub-processes for data element concept mapping

5.2.2 Identifying metadata element sets (First process)

iTeh STANDARD PREVIEW
(standards.iteh.ai)

5.2.2.1 Method

First, it is necessary to collect available metadata element sets and to identify candidate metadata element sets to be mapped. Then, what the purpose of the metadata element set is should be checked, how many numbers of fields should be counted, and whether sample data exists or not should be checked. Who or which organization has the authority over each metadata element set should also be checked.

5.2.2.2 Examples

- Domain: e-book cataloguing
- Available metadata element sets: OEBPS,[13] DC,[10] MODS[12] and TEI[14]

Table 1 — Analysing available metadata element sets

Metadata element set name	DC	OEBPS	MODS	TEI header
Purpose	Description of internet resource	Description of electronic book	Description of library resources	Encoding methods for machine-readable texts
Number of fields	15	15	About 60	Over 20
Sample data	yes	yes	No	Yes
Authority	DCMI	Open eBook Forum	Library of Congress	TEI Consortium

5.2.3 Grouping data element concepts (Second process)

5.2.3.1 Method

The next process is to group data element concepts by object class, and then to find properties associated with the objects and sub-group the data element concepts by property.

For convenience, it is helpful to select a primary metadata element set among the candidate metadata element sets and aggregate them by the primary metadata element set. The metadata element set having the least number of elements is recommended to be the primary metadata element set.

All data element concepts included in the candidate metadata element sets should be aggregated by property. Some data elements, which cannot be grouped, are supposed to be set aside.

In this process, metadata experts should perform the work along with domain experts.

5.2.3.2 Examples

In the sample object class, the e-book has plural properties as shown below.

- Object class: e-book
- Properties: title, author, subject, ..., edition

Table 2 — Example of grouping data element concepts by property

Primary metadata element set	Secondary metadata element set		
DC	OEBPS	MODS	TEI
Title	Title	Title subTitle partNumber partName nonSort	title seriesStmt:title seriesStmt:idno
Creator	Creator(role) Creator(file-as)	name:role name:namePart name:displayForm name:affiliation name:description	author
Subject	Subject	Topic classification catographics occupation	keyword classCode catRef
...
(none)	(none)	Edition	fileDesc_editionStmt_date fileDesc_editionStmt_edition fileDesc_editionStmt_respStmt fileDesc_editionStmt_respStmt_name fileDesc_editionStmt_respStmt_resp

Similar properties of OEBPS, MODS and TEI are grouped according to those of the primary metadata element set.

5.2.4 Mapping data element concept (Third process)

5.2.4.1 Method

After identifying object classes and properties related to the primary metadata element set, common Data Element Concepts (DECs) can be created according to ISO/IEC 11179-1 and ISO/IEC 11179-3.

The third process starts from finding common data element concepts in each group of data element concepts based on object classes and properties found in the second process. The existence of a domain taxonomy or ontology will make it easier to construct the common DECs.