
Data quality —

Part 120:

**Master data: Exchange of
characteristic data: Provenance**

Qualité des données —

*Partie 120: Données permanentes: Échange des données
caractéristiques: Provenance*

*iTeh STANDARD PREVIEW
(standard number)
Full standard available on <https://standards.iteh.ai/catalog/standards/sist/658a153d-7d8f-4b4d-92f6-e6b44cef0756/iso-8000-120-2016>*

PROOF / ÉPREUVE



iTeh STANDARD PREVIEW
(standards.iteh.ai)

Full standard:
<https://standards.iteh.ai/catalog/standards/sist/658a15d1-7d8f-4b4d-92f6-e6b44cef0756/iso-8000-120-2016>



COPYRIGHT PROTECTED DOCUMENT

© ISO 2016, Published in Switzerland

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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

	Page
Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	2
4 Abbreviated terms and notation	2
4.1 Abbreviated terms	2
4.2 Notation	2
5 Fundamental concepts and assumptions	2
6 Provenance data model	3
6.1 Diagram	3
6.2 date_and_time	3
6.3 ISO_6523_identifier	3
6.4 property_value_assignment	3
6.5 provenance_event	4
7 Data provenance record	4
8 Conformance requirements	6
Annex A (normative) Document identification	7
Annex B (informative) Information to support implementations	8
Annex C (informative) Scenarios	9
Annex D (informative) Use cases	15
Bibliography	23

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is Technical Committee ISO/TC 184, *Automation systems and integration*, Subcommittee SC 4, *Industrial data*.

This first edition of ISO 8000-120 cancels and replaces ISO/TS 8000-120:2009, which has been technically revised.

ISO 8000 is organized as a series of parts, each published separately. The structure of ISO 8000 is described in ISO/TS 8000-1.

Each part of ISO 8000 is a member of one of the following series: general data quality, master data quality, transactional data quality, and product data quality. This part of ISO 8000 is a member of the master data quality series.

A list of all parts in the ISO 8000- series can be found on the ISO website.

Introduction

The ability to create, collect, store, maintain, transfer, process and present data to support business processes in a timely and cost effective manner requires both an understanding of the characteristics of the data that determine its quality, and an ability to measure, manage and report on data quality.

ISO 8000 defines characteristics that can be tested by any organization in the data supply chain to objectively determine conformance of the data to ISO 8000.

ISO 8000 provides frameworks for improving data quality for specific kinds of data. The frameworks can be used independently or in conjunction with quality management systems.

ISO 8000 covers industrial data quality characteristics throughout the product life cycle from conception to disposal. ISO 8000 addresses specific kinds of data including, but not limited to, master data, transaction data, and product data.

ISO 8000-110 specifies requirements that can be checked by computer for the exchange, between organizations and systems, of master data that consists of characteristic data. It provides requirements for data quality, independent of syntax. This part of ISO 8000 specifies requirements for capture and exchange of data provenance information and supplements the requirements of ISO 8000-110. This part of ISO 8000 includes a conceptual data model for data provenance.

Data provenance information can be used to detect data echoes, and can be used to determine the credibility, currency or value of data. Data provenance information provides a necessary capability to support claims of data accuracy.

NOTE Requirements regarding claims of data accuracy are covered in ISO 8000-130.

Any claim of conformance to this part of ISO 8000 implies a claim of conformance to ISO 8000-110.

This part of ISO 8000 can be used with any other standard that specifies a formal syntax for a data set.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Full standard:
<https://standards.iteh.ai/catalog/standards/sist/658a15d1-7d8f-4b4d-92f6-e6b44cef0756/iso-8000-120-2016>

Data quality —

Part 120:

Master data: Exchange of characteristic data: Provenance

1 Scope

This part of ISO 8000 specifies requirements for the representation and exchange of information about the provenance of master data that consists of characteristic data, and supplements the requirements of ISO 8000-110.

NOTE 1 ISO 8000-110 specifies that such data be represented as property values. This part of ISO 8000 provides additional requirements for property values when provenance information needs to be captured. Provenance is the history or pedigree of a property value.

This part of ISO 8000 does not specify a complete model for characteristic data, nor does it specify an exchange format for characteristic data with provenance information.

NOTE 2 This is done in other standards that reference this part of ISO 8000, e.g. ISO 22745-40.

The following are within the scope of this part of ISO 8000:

- scenarios for data provenance;
- requirements for capture and exchange of data provenance information;
- conceptual data model for data provenance information.

The following are outside the scope of this part of ISO 8000:

- exchange format for data provenance information;
- scheme for registering and resolving organization identifiers and person identifiers;
- provenance of data that are not characteristic data represented as property values;
- configuration management;
- change control;
- syntax of identifiers;
- resolution of identifiers.

Some of the requirements in this part of ISO 8000 can apply to exchange of data that is not master data which consists of characteristic data represented as property values.

NOTE 3 ISO 8000-100 provides an overview of the master data quality series of parts of ISO 8000 and a description of its overall structure.

2 Normative references

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

ISO 8000-2, *Data quality — Part 2: Vocabulary*

ISO 8000-110, *Data quality — Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 8000-2 apply.

4 Abbreviated terms and notation

4.1 Abbreviated terms

id	Identifier
UML	Unified Modeling Language
XML	Extensible Markup Language

4.2 Notation

p	Property
v	Value
e	event1)
pv(p, v)	property value asserting that property p has value v
pv(p, v, (e1, e2,...))	property value asserting that property p has value v, with provenance information given by events e1, e2, etc.
query(I, p)	query for current value of property p of item I
create(orgID, personID, t)	event that is the creation of a property value by the organization and person at time t
extract(orgID, personID, t)	event that is the extraction of a property value by the organization and person at time t
MDM(I, (pv1, pv2,...))	master data message containing property values pv1, pv2, etc.
null	no value given

5 Fundamental concepts and assumptions

The Oxford English Dictionary defines provenance as:

- the fact of coming from some particular source or quarter; origin, or derivation;
- the history or pedigree of a work of art, manuscript, rare book, etc.; concretely a record of the ultimate derivation and passage of an item through its various owners.

In this part of ISO 8000, the term “data provenance” corresponds to the first part of the second definition above: “history or pedigree of a property value”; the term “provenance record” corresponds to the second half of second definition above: “record of the ultimate derivation and passage of a property value through its various custodians”.

6 Provenance data model

6.1 Diagram

The UML class diagram for the provenance model is given in [Figure 1](#).

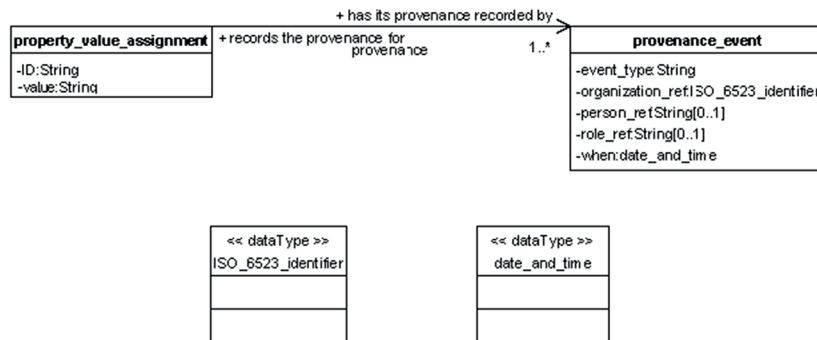


Figure 1 — UML class diagram for provenance

NOTE The entity `property_value_assignment` is the intersection point between this data model and the target data model: the model of data for which provenance information is to be recorded. When the data model in [Clause 6](#) is integrated with the target data model, this entity needs to be replaced with the appropriate entity from the target data model.

6.2 date_and_time

A `date_and_time` is a point in time, expressed in Universal Coordinated Time (UTC).

6.3 ISO_6523_identifier

An `ISO_6523_identifier` is an identifier of an organization and possibly a subdivision of an organization, conforming to the structure specified in ISO/IEC 6523-1.

NOTE The syntax of ISO/IEC 6523-1 identifiers is not specified in this part of ISO 8000.

6.4 property_value_assignment

A `property_value_assignment` is a `data_object` that is a pair of a value and an identifier to a property defined in a data dictionary.

Attribute definitions:

ID: the string that unambiguously identifies the `property_value_assignment` within the organization that created it.

NOTE 1 The identifier need only be unique or meaningful within the organization that created the `property_value_assignment`.

NOTE 2 The format of the identifier is not specified in this part of ISO 8000.

provenance: the `provenance_event` that provides information on the provenance of the `property_value_assignment`.

Assertions:

Each property_value_assignment has its provenance recorded by one or many provenance_event objects. Each provenance_event records the provenance for exactly one property_value_assignment.

6.5 provenance_event

A provenance_event is an event for which data provenance information is recorded.

Attribute definitions:

- event_type: the event for which data provenance information is recorded.
- organization_ref: the unambiguous identifier of the organization and possibly the subdivision of the organization that performed the event, conforming to the structure defined in ISO/IEC 6523-1 and assigned in accordance with ISO/IEC 6523-2.
- person_ref: the identifier assigned by the organization to the person who performed the event.
NOTE 1 The identifier need only be unique within the organization.
NOTE 2 The format of the identifier is not specified in this part of ISO 8000.
- role_ref: the identifier assigned by the organization to the role within the organization played by the person who performed the event.
NOTE 3 The identifier need only be unique within the organization.
NOTE 4 The format of the identifier is not specified in this part of ISO 8000.
- when: the point in time at which the event took place.

Assertions:

Each provenance event records the provenance for exactly one property_value_assignment. Each property_value_assignment has its provenance recorded by one or many provenance_event objects.

7 Data provenance record

The data provenance record for a property value is the record of the ultimate derivation and passage of the property value through its various custodians.

The data provenance record for a property value shall be either:

- included in the structure that represents property value;
EXAMPLE 1 The following is an XML code fragment in which the data provenance record is included in the XML structure that represents property value.

Coded:

```
<property-value property-ref="0161-1#02-015007#1">  
  <controlled-value value-ref="0161-1#07-000435#1"/>  
  <provenance-event event-type="create" organization-ref="0161-ABCDE" person-ref="ROLLINS1"  
date="2008-10-27T15:40:31.287"/>  
</property-value>
```

Decoded:

```
<property-value property-ref="inclosure material">
```

```

    <controlled-value value-ref="ceramic"/>
    <provenance-event event-type="create" organization-ref="ABC Company" person-ref="William
    F. Rollins" date="2008-10-27T15:40:31.287"/>
  </property-value>

```

— stored separately and referenced from the structure that represents property value.

EXAMPLE 2 The following is an XML code fragment in which the data provenance record is referenced from the XML structure that represents property value.

Coded:

```

<property-value property-ref="0161-1#02-015007#1" provenance-ref="p4153">
  <controlled-value value-ref="0161-1#07-000435#1"/>
</property-value>
...
<provenance-record id="p4153">
  <provenance-event event-type="create" organization-ref="0161-ABCDE" person-ref="ROLLINS1"
  date="2008-10-27T15:40:31.287"/>
</provenance-record>

```

Decoded:

```

<property-value property-ref="inclosure material" provenance-ref="p4153">
  <controlled-value value-ref="ceramic"/>
</property-value>
...
<provenance-record id="p4153">
  <provenance-event event-type="create" organization-ref="ABC Company" person-ref="William F.
  Rollins" date="2008-10-27T15:40:31.287"/>
</provenance-record>

```

NOTE 1 See [D.1](#) for an explication of the codes in the examples above.

NOTE 2 This part of ISO 8000 does not require that a master data message use the specific structures shown in the examples above, or that it use XML syntax at all.

NOTE 3 In the examples above, data are given in coded form followed by decoded form for clarity. An actual master data message is required to be in coded form (see ISO 8000-110).

The data provenance record for a property value shall include:

- identification of the data creator (the organization, and possibly the person and role within the organization, that created the data);
- identification of each data extractor (the organization, and possibly the person and role within the organization, that extracted the data).

NOTE 4 As specified in [6.5](#), the following information is required for a provenance event:

- event type (e.g. "create", "extract");