



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 the International Standard for Data Quality

Peter Benson  
ISO 8000 Project Leader  
Executive director and chief technical officer  
Electronic Commerce Code Management Association (ECCMA)



The MIT 2008 Information Quality Industry Symposium



### ISO 8000 - Data Quality

- ISO 8000 addresses data quality. ISO 8000 is concerned with:
- the principles of data quality;
- the characteristics of data that determine its quality;
- the processes to ensure data quality.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – ISO Definitions

### **information**

knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning [ISO/IEC 2382-1:1993]

meaningful data [ISO 9000:2005]

### **data**

re-interpretable representation of information in a formalized manner suitable for communication, interpretation, or processing [ISO/IEC 2382-1:1993]

### **quality**

degree to which a set of inherent characteristics fulfils requirements [ISO 9000:2005]

### **characteristic**

distinguishing feature [ISO 9000:2005]

### **requirement**

need or expectation that is stated, generally implied or obligatory [ISO 9000:2005]



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Parts under development

Part 1: Overview, principles and general requirements

Part 2: Terminology

Part 100: Master data: Overview

**Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification**

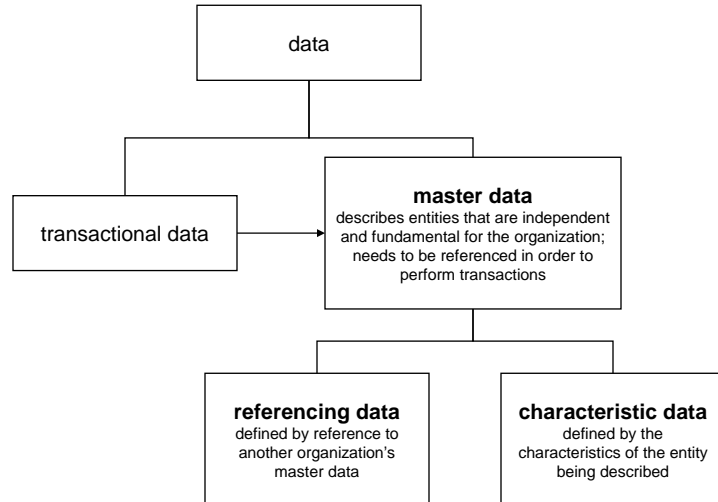
Part 120: Master data: Provenance

Part 130: Master data: Accuracy

Part 140: Master data: Completeness



## ISO 8000-100: – Master Data Quality



## ISO 8000-100: – Master Data Quality

### master data

data held by an organization that describes the entities that are both independent and fundamental for an enterprise, that it needs to reference in order to perform its transactions

A master data record is a collection of data element values. Data element values are the fundamental building blocks of electronically stored information, the quality of the data element values is a major determinant of the quality of the information and consequently the accuracy and reliability of the knowledge that can be derived from the information.

The following are considered to be intrinsic components of establishing data quality:

1. **Syntax**: The syntax or arrangement of data element values determines the ease with which data can be integrated within and across organizations.
2. **Semantics (metadata)**: The ability to retrieve the definition of metadata (data label) and the quality of the definition in terms of clarity of meaning to all stakeholders determines the portability of the data within and across organizations.
3. **Source of data (provenance)**: The ability to track the organization that owns the process that created, validated or transferred the data and the time when the process was performed determines the traceability of the data.
4. **Fitness**: The ability to assess if data meets the requirements of a specific function determines its fitness for the purpose.
5. **Accuracy**: The method through which accuracy is asserted determines the ability to validate accuracy.
6. **Completeness**: The method through which completeness is asserted determines the ability to validate completeness.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Part 110

This part of ISO 8000 specifies requirements that can be checked by computer for the exchange, between organizations and systems, of master data that consists of characteristic data.

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

- requirements regarding conformance to a formal syntax for master data messages;
- semantic encoding requirements for master data messages;
- requirements regarding conformance to data specifications for master data messages;

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

- requirements regarding data not in messages;
- requirements regarding exchange of data that are not master data;
- requirements regarding master data that are not characteristic data;
- Records of the history of the origination, modification, and transfer of custody or ownership of data are commonly referred to as the data provenance ( these are in part 120)
- requirements regarding recording the history of master data; (these are in part 120)
- requirements regarding accuracy of master data; (these are in part 130)
- requirements regarding the management of master data internally within an organization;
- *Data within an organization's enterprise resource planning (ERP) or product data management (PDM) system is out of scope.*



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Part 110

### Syntax

Each data set shall contain a reference to the syntax to which the data set complies....The reference shall be resolvable to the specification of the syntax through a mechanism that is publicly available.

### Semantic encoding

Each data element value shall reference all concepts necessary to unambiguously define its meaning. Each reference shall be to a concept dictionary entry contained in a concept dictionary that supports an interface for resolution of a concept identifier.

**Syntax and semantic resolution shall be available at no charge unless the data carries a "fee based encoding" warning label.**

### Conformance to requirements

Each data set shall contain a reference to the data requirements statement to which the data set complies. The reference shall be a globally unambiguous identifier that was used to encode the data set. The reference shall be resolvable to the data requirements statement. The data requirements statement shall be publicly available.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Part 120

### Provenance

This part of ISO 8000 describes requirements for representation and exchange of information about provenance of property value pairs and data sets.

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

- scenarios for data provenance;
- data provenance roles;
- requirements for capture and exchange of 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 property value pairs or data sets.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Part 130

### Accuracy

This part of ISO 8000 describes requirements for representation and exchange of information about accuracy of property value pairs, records, and data sets.

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

- scenarios for data accuracy;
- data accuracy roles;
- requirements for capture and exchange of data accuracy information;

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

- exchange format for data accuracy information;
- accuracy of data that are not master data;
- accuracy of data that are not property value pairs, records, or data sets.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000 - Data Quality – Part 140

### Completeness

This part of ISO 8000 describes requirements for completeness of master data.

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

- scenarios for completeness;
- completeness roles;
- requirements for completeness of master data;
- requirements for exchange of information about master data completeness.

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

- exchange format for completeness information;
- completeness of data that are not master data;
- completeness of data that are not property value pairs, records, or data sets.



The MIT 2008 Information Quality Industry Symposium



## ISO 8000-110:2008 Certification

- **Certified Software applications and data cleaning services**
  - Know how to use an open technical dictionary for metadata encoding and mapping
  - Know how to read and write master data requirement specifications
  - Know how to generate requests for master data
  - Know how to read and write encoded master data messages
- **Certified Master Data Quality Managers (data requestors)**
  - Know how to specify their master data requirements
  - Know how to ask for they data they need to validate or complete their master data
- **Certified Quality Master Data Providers**
  - Know how to respond to a request for master data

