Measuring Data Quality Using ISO 8000

Peter R. Benson
Executive Director
ECCMA

Project leader for ISO 8000 and ISO 22745
When you order seafood from Quality Fresh Seafood, you can be confident that you are receiving the very best quality of seafood and delivery.
ISO 9000:2005(E)

3.1.1 quality
degree to which a set of inherent characteristics fulfils requirements

3.1.2 requirement
need or expectation that is stated, generally implied or obligatory

Quality data is data that meets stated requirements
nothing less and nothing more!
The International Standard for Data Quality

ISO 8000

General Principles

Part 1 Introduction
Part 2 Terminology

Part 100 Introduction

Part 110

Part 120 Provenance
Part 130 Accuracy
Part 140 Completeness

Syntax
Semantic encoding
Meets requirements

Transaction data
ISO 8000 Quality Data Summarized

Data that meets requirements

Data that is Portable
A Form is a Specification of a Requirement for Data
Publicly Visible Terminology in a Standard Model

- Public domain concept identifiers
- Free identifier resolution to underlying terminology (web services)
- Hyperlink to source standards
- Multilingual
- Multiple terms, definitions and images linked to single concept identifier
The eOTD (ECCMA Open Technical Dictionary) is an ISO 22745-20 compliant central registry of terminology. Each concept and terminological component in the eOTD is assigned a unique and permanent public domain identifier.

Users create corporate preferred subsets of the eOTD and use the eOTD concept identifiers to manage concept equivalence mapping with the concepts used by their trading partners.
Data requirements support a business function, granting access to a computer, a website or a software program, simply asking for the data needed to deliver the right product or service or to comply with a regulation, these are all data requirements.

*Be careful what you ask for – data quality starts with the quality of the request for data*
Clear and unambiguous requests for:

1. Reference data (identifiers)
2. Characteristic data (descriptions)
3. The validation of reference and characteristic data,

ISO 22745-35 is a standard format for the generation and distribution of requests for data in a simple XML format that can be automated by the sender and recipient to create an integrated data exchange system.
Reply to a Request for Data

ISO 22745-40 is a standard format for the **exchange of reference and characteristic data** in a simple XML format that can be automated by the sender and recipient to create an integrated data exchange system.
A data provider may not have all the data requested so they in turn send a request through their data supply chain using the same ISO 22745 standard exchanges.
Quality data is **portable data**; it is independent of the software application and accessible by any application.
Motivation for ISO 22745 and ISO 8000

Controlling costs requires better asset, product, component and process visibility. This is achieved through faster, better and lower cost access to authoritative sourced data.
Motivation for ISO 22745 and ISO 8000

Controlling costs requires better asset, product, component and process visibility. This is achieved through faster, better and lower cost access to authoritative sourced data.
The contractor, sub-contractor or supplier shall, as and when requested to do so, supply technical data in electronic format on any of the items covered in this contract as follows:

**The data shall be ISO 8000-110:2009 compliant.**

- The data shall comply with registered ISO 22745-30 compliant data requirements
- The data shall be encoded using concept identifiers from an ISO 22745 compliant open technical dictionary that supports free resolution to concept definitions.
- The data shall be provided in ISO 22745-40 compliant Extensible Markup Language (xml).

- Providing the data necessary for the safe and efficient operation of plant and equipment is an established legal requirement in most countries.
- Creating ISO 8000-110:2009 compliant data does not require the payment of any license fees or the use of specialized software, it is within the technical ability of all suppliers regardless of their size.
1. Document your business language
2. Document your requirements for data
3. Measure the quality of your data
   (by comparing your data to your requirements for data)

*There are low cost software tools that make this process very straightforward*