Meeting the ISO 8000 Requirements for Quality Data

ABSTRACT----------

Leaders of the NATO committee on codification coined the phrase "data is the DNA of logistics" but we know that it is much more than that. We are a data dependent society, every event, every individual, every organization, all locations, goods and services are represented by data. Beyond the accuracy of the data and its ability to accurately represent reality, data has its own history and intrinsic qualities that make it portable from one application to another and allow it to be safely preserved as it moves over time from one electronic media to another. As we migrate towards an ever more dynamic Event Driven Architecture (EDA) and an ever more interdependent Service Oriented Architecture (SOA), the accuracy and portability of the master data, data that describes individuals, organizations, locations, goods and services or the who, what and where, become ever more critical. ISO 8000 is the international standard that defines the requirements for quality data, understanding this important standard and how it can be used to measure data quality is an important first step in developing any information quality strategy.

BIOGRAPHY----------

Peter R. Benson
Executive Director and Chief Technical Officer
Electronic Commerce Code Management Association

Peter R. Benson is the Executive Director and Chief Technical Officer of ECCMA; he is an expert in distributed data systems and master data management. Peter developed data collection and reporting systems for large agricultural businesses as well as for the public relations, advertising, healthcare and financial industries. Peter was granted a British patent in 1992 covering the maintenance and usage tracking of distributed data. Peter developed the UNSPSC, an internationally recognized commodity classification for spend analysis and went on to develop the eOTD a leading open technical dictionary used to create unambiguous language independent descriptions of individuals, organizations, locations, goods, services, processes, rules and regulations. Peter served as the elected chair of ASC X12E the US Standards Committee responsible for the development and maintenance of EDI standard for product data. Peter is part of the US expert delegation to ISO TC 184/SC 4 and the project leader for the international standard ISO 8000 (data quality) and ISO 22745 (open technical dictionaries). Peter has a baccalaureate in mathematics and physics from the Academy of Bordeaux, France, a bachelor of science in agriculture from London University, England and a master of science in agricultural marketing from London University, England.

Meeting the requirements of ISO 8000

ECCMA
electronic commerce code management association

July 15, 2009

Massachusetts Institute of Technology (MIT)

Information Quality Industry Symposium (IQIS 2009)

Cambridge, Massachusetts

Peter R. Benson

Peter.Benson@eccma.org

Slide 1

International Organization for Standardization (ISO)

- 156 National standard organization members (one per country)
- (AFNOR, **ANSI**, BSI CNIS, DIN, GOST, Standards Australia …….)
- 192 Technical Committees
  - 3 000 Technical bodies
  - 50 000 domain experts
- Central Secretariat in Geneva
  - 150 staff

ISO International Organization for Standardization

- ISO TC 184 Industrial automation systems and integration
  - ISO TC184 SC4 Industrial data (STEP)
    - **ISO 22745 (open technical dictionaries and their application to master data)**
    - **ISO 8000 Data Quality**
    - **ISO 8000-100 Master Data Quality**

Slide 2

Vision for the Future

What is impossible to do right now, but, if you *could* do it, would fundamentally change your business?

1990 Joel Arthur Barker

- Automated data mapping,
- reliable data mapping,
- an end to incomplete data,
- an end to inaccurate information

Slide 3

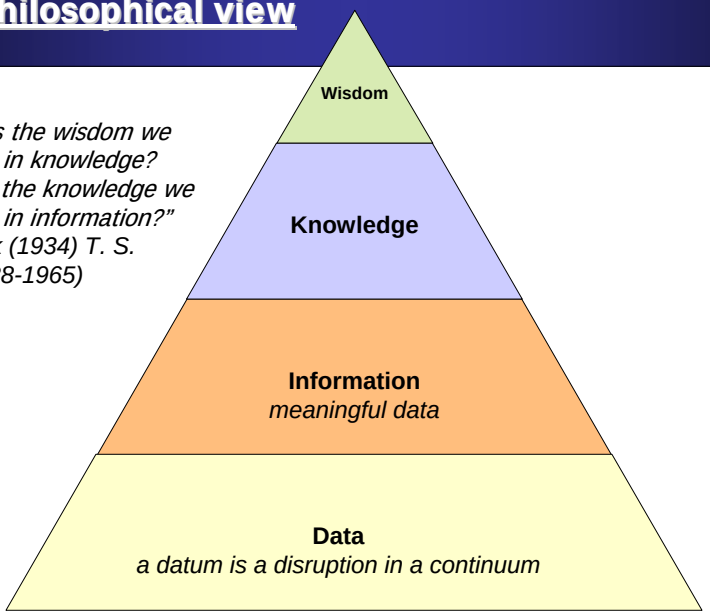
ISO 8000 - Data Quality

Understanding the difference between <u>data</u> and <u>information</u> is the key to solving data quality

Slide 4

The philosophical view

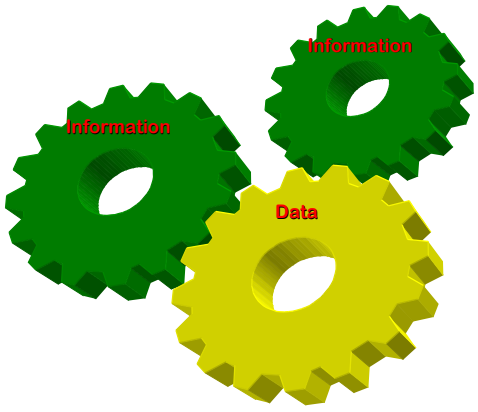
"Where is the wisdom we have lost in knowledge? Where is the knowledge we have lost in information?" The Rock (1934) T. S. Eliot (1888-1965)



Slide 5

The theoretical view

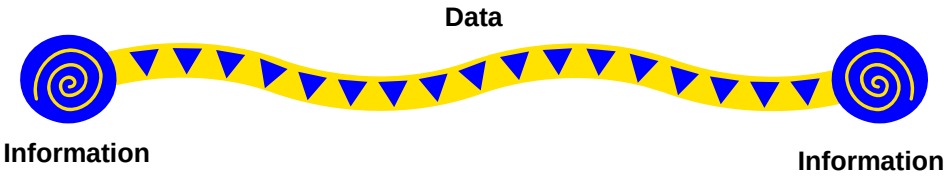
Data transfers information - *perfectly*



Slide 6

The practical view

Data transfers information - *imperfectly*



Slide 7

Tackling the problem - ISO 8000 Data Quality

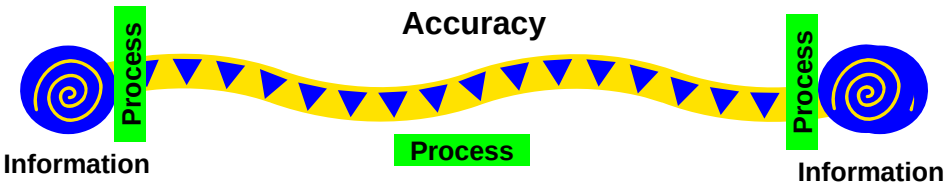
Syntax

Semantic encoding

Conformance to requirements

Provenance

Accuracy



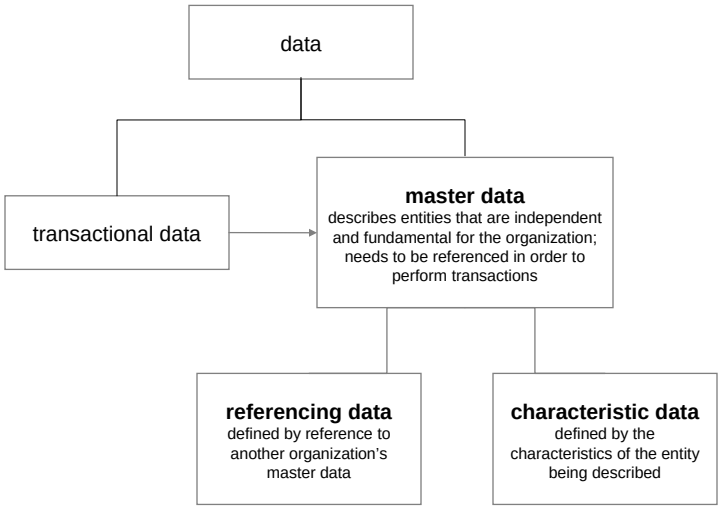
Slide 8

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**

Slide 9

ISO 8000-100 Master Data Quality



Slide 10

ISO 8000-110:2008 MASTER DATA QUALITY

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.

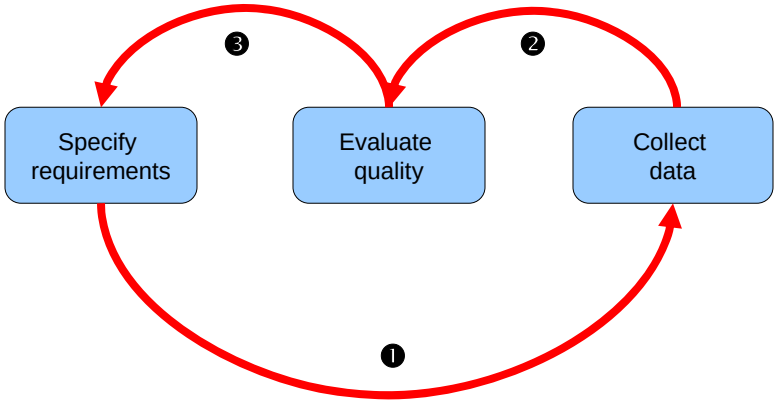
Conformance to requirements

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

Slide 11

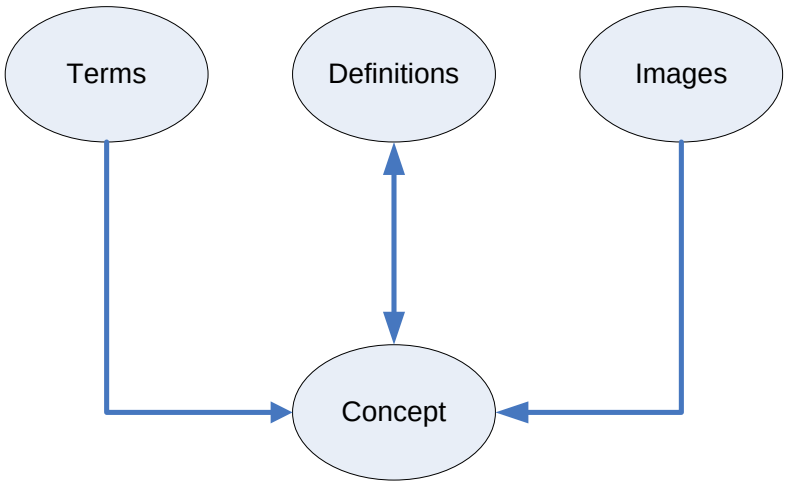
A practical solution to data quality

As simple as 1-2-3



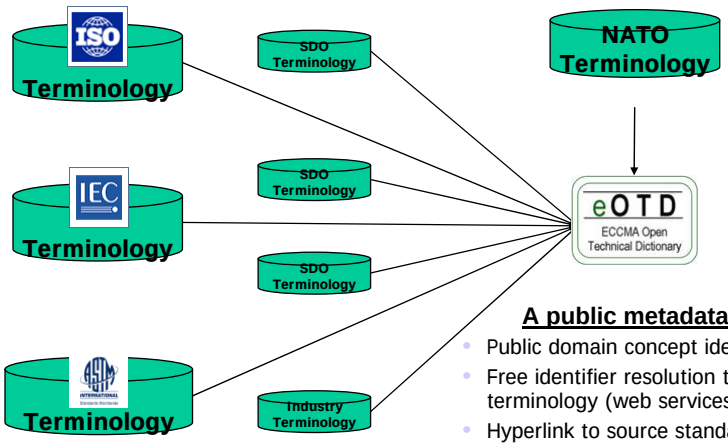
Slide 12

Open Technical Dictionary ISO 22745



Slide 13

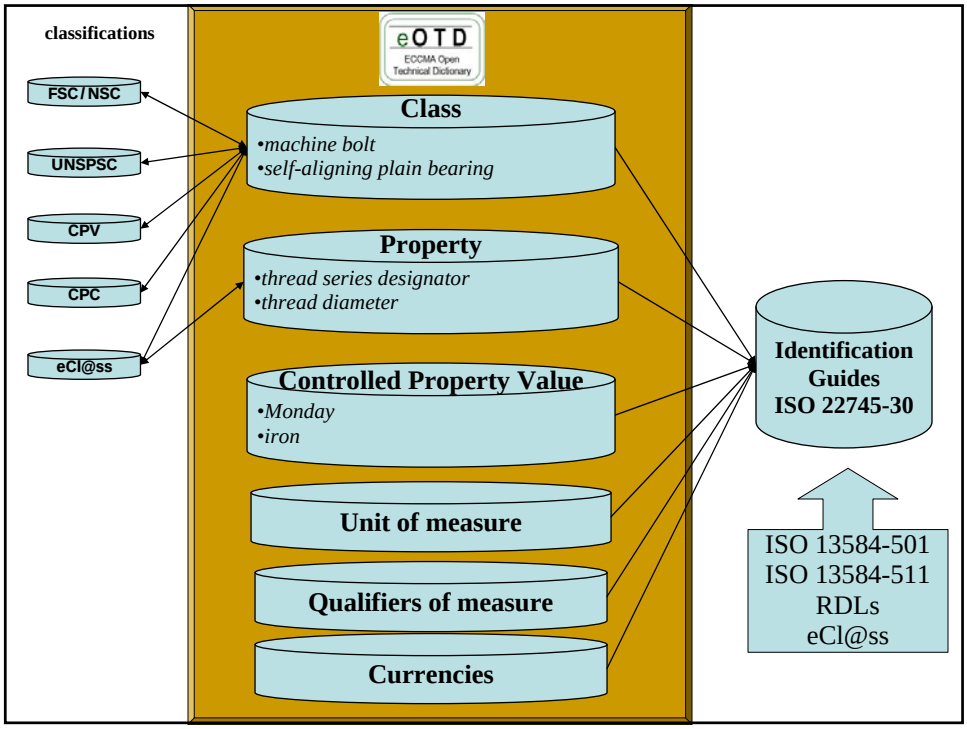
ECCMA Open Technical Dictionary (eOTD) ISO 22745



A public metadata registry

- Public domain concept identifiers
- Free identifier resolution to underlying terminology (web services)
- Hyperlink to source standards
- Multiple terms, definitions and images linked to single concept identifier
- Multilingual

Slide 14



Examples of eOTD Concept Types

- **01 - Class**
  - machine bolt
  - self-aligning plain bearing
- **02 - Property**
  - thread series designator
  - thread diameter
- **03 - Feature**
  - flange
  - inner liner
  - outer ring
  - second hole
- **04 - Representation**
  - string
  - decimal measure
  - rational measure
- **05 - Unit of Measure**
  - degree
  - radian
  - kilogram
  - newton per square millimeter
  - bolt
- **06 - Qualifier of Measure**
  - nominal
  - minimum
  - maximum
- **07 - Controlled Property Value**
  - Monday
  - Tuesday
  - iron
- **08 - Currency**
  - US Dollar
  - Euro

Slide 16

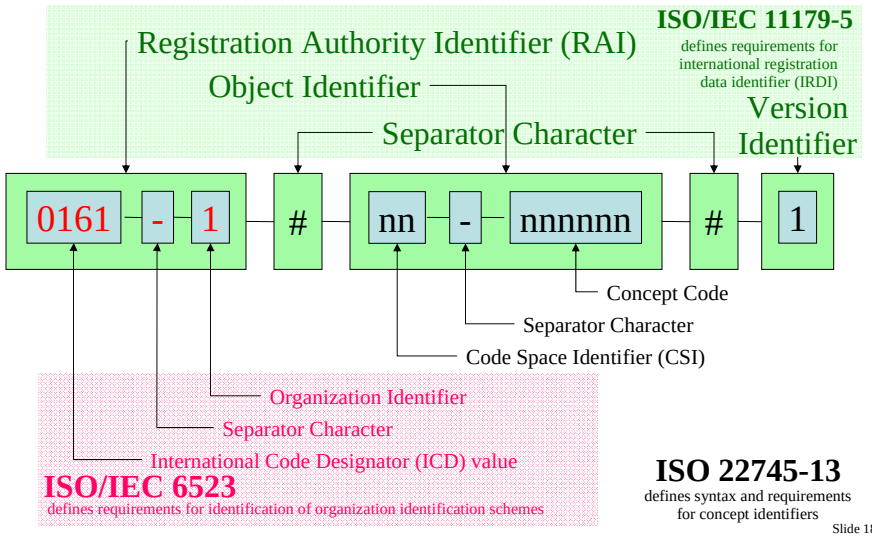
eOTD Dictionary

- Contains
  - Concepts with identifiers
  - Terminology to specify meaning of concepts
- Does not contain*
  - Relationships between concepts
  - Constraints on property values
  - Data types
  - Reply instructions

***These are all contained in identification guides**

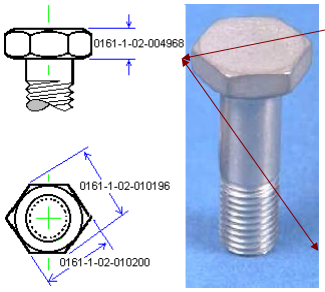
Slide 17

eOTD Concept Identifier



Slide 18

Semantic encoding using the eOTD



Property ID	Value	Measure ID
0161-1#02-046898#1	0161-1#07-014684#1	
0161-1#02-027375#1	3225020037	
0161-1#02-023822#1	1.0	0161-1#05-000798#1
0161-1#02-010200#1	1.450	0161-1#05-000798#1
0161-1#02-010196#1	1.653	0161-1#05-000798#1
0161-1#02-004968#1	0.591	0161-1#05-000798#1
0161-1#02-027376#1	10	
0161-1#02-027378#1	0.80	0161-1#08-000168#1

eOTD Identifier Coded

Property term	Value	Measure term
eOTD CLASS NAME	BOLT:MECHANICAL	
PRODUCT NUMBER	3225020037	
NOMINAL THREAD DIAMETER	1.0	INCHES
WIDTH ACROSS FLATS	1.450	INCHES
WIDTH ACROSS CORNERS	1.653	INCHES
HEAD HEIGHT	0.591	INCHES
COUNT PER PACK	10	
PACK PRICE	0.80	US DOLLAR

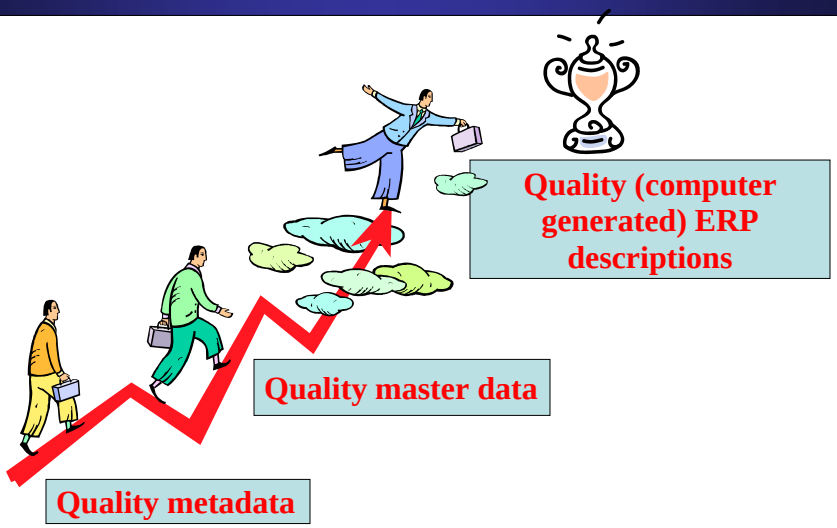
eOTD Identifiers Resolved

Machine Bolt; Product Number: 3225020037; Nominal thread diameter: 1.0 inches; Width across flats: 1.450 inches; Width across corners: 1.653 inches; Head height: 0.591 inches; Count per pack: 10; Pack price: $0.80 (M-Bolt;NTD1.0";WAF1.45";CPP10)

Rendered

Slide 19

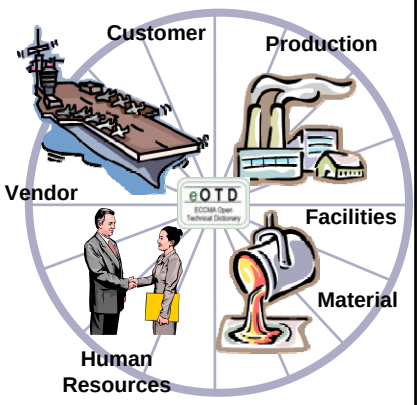
The steps to quality ERP descriptions



Slide 20

Common Concept Encoding

- **Across the supply chains**
- **ERP masters:**
  - **vendor / customer / material**
- **Manufacturing / production**
  - **CAx / PDM**
- **Facilities / raw materials**
- **Human Resources**
- **Data life cycle management:**
  - **design through disposal**



Common metadata mapping across applications!

Slide 21

ISO 8000-110:2008 MASTER DATA QUALITY

Providing the data necessary for the safe and efficient operation of plant, and equipment is a legal requirement in most countries

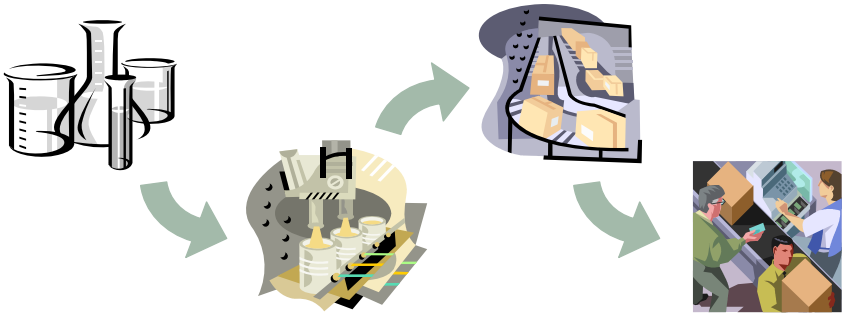
The contractor, sub-contractor or supplier shall, as and when requested to do so, supply technical data in electronic format on any of the items covered in this contract as follows:

- The data shall be ISO 8000-110:2008 compliant.
- The data shall comply with registered ISO 22745-30 compliant Identification Guides.
- 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 an ISO 22745-40 compliant Extensible Markup Language (xml) format.

Slide 22

✓ A Vision Realized

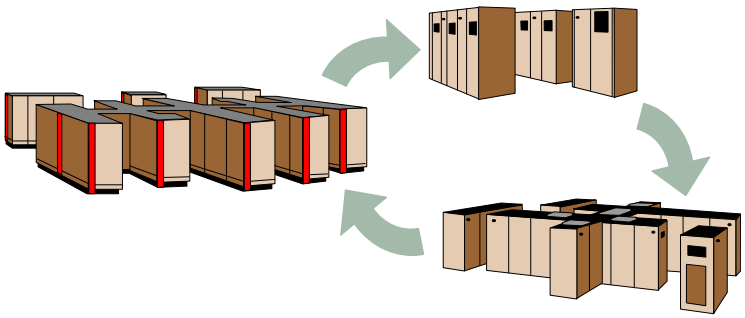
A single process for identifying and describing individuals, organizations, locations, goods, services, processes rules and regulations



Slide 23

✓ A Vision Realized

Portable data – data that is independent of hardware, operating system *and application software*.



Slide 24

ISO 8000-100 Master Data Quality
A data provider's perspective

Data providers recognize that:

- data integration is one of the keys to a long term relationship
- the ability to provide their customers with quality data is a significant differentiating factor.
- ***There is growing resistance to "data lock-in"***

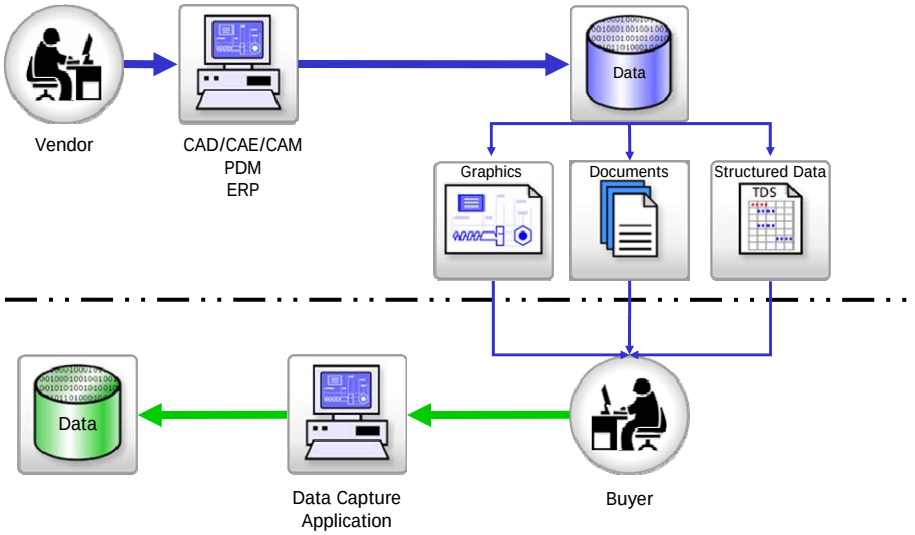
Data providers are:

- looking to increase their visibility and understand that the best way to do this is to improve the quality of their data.
- looking for a Standard that they can use to identify the quality of their data.

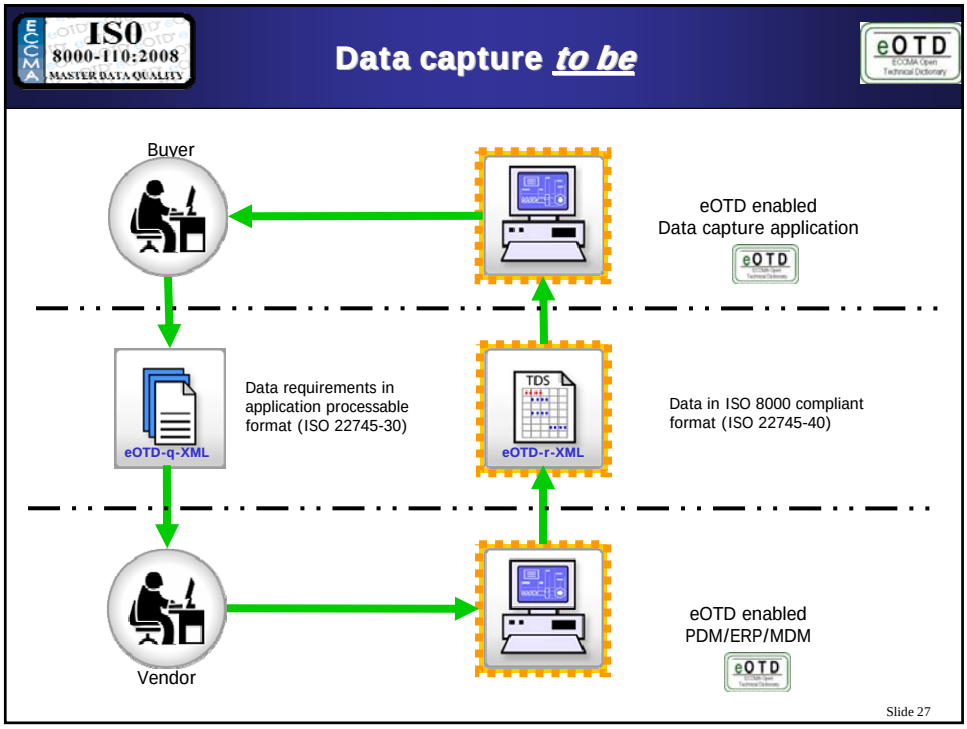
ISO 8000-110:2008 MASTER DATA QUALITY

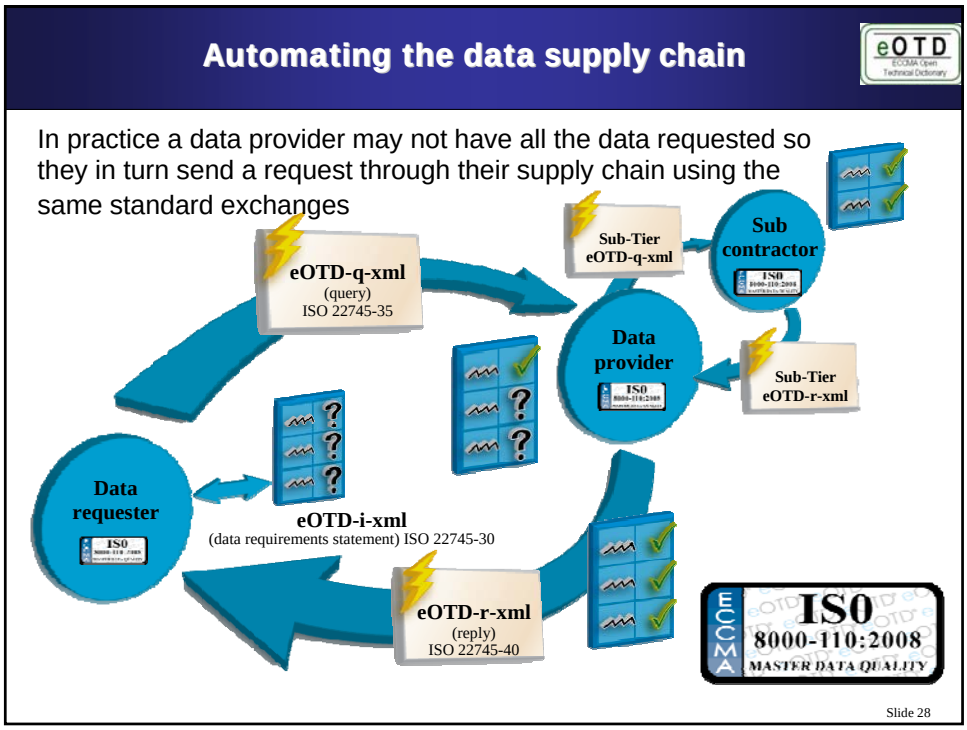
Slide 25

Data capture *as is*

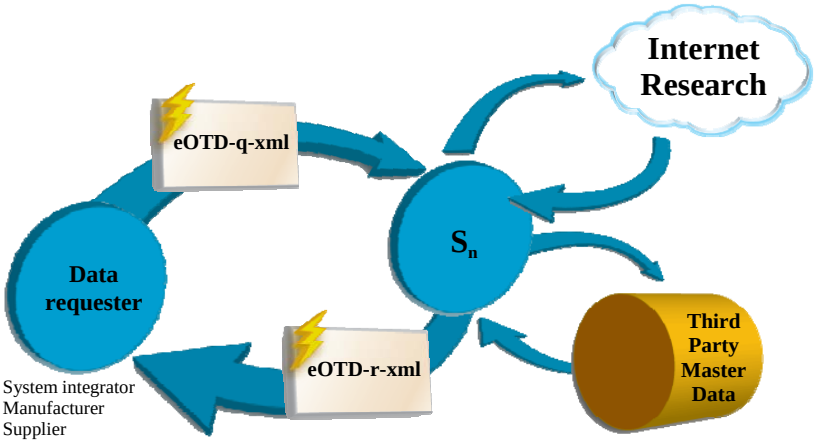


Slide 26


ISO 8000-110:2008 MASTER DATA QUALITY
Data capture to be
eOTD
Buyer
eOTD enabled
Data capture application
eOTD
Data requirements in
application processable
format (ISO 22745-30)
eOTD-q-XML
TDS
eOTD-r-XML
Data in ISO 8000 compliant
format (ISO 22745-40)
Vendor
eOTD enabled
PDM/ERP/MDM
eOTD
Slide 27


Automating the data supply chain
eOTD
In practice a data provider may not have all the data requested so
they in turn send a request through their supply chain using the
same standard exchanges
eOTD-q-xml
(query)
ISO 22745-35
Sub-Tier
eOTD-q-xml
Sub
contractor
Data
provider
Sub-Tier
eOTD-r-xml
Data
requester
eOTD-i-xml
(data requirements statement) ISO 22745-30
eOTD-r-xml
(reply)
ISO 22745-40
ISO 8000-110:2008 MASTER DATA QUALITY
Slide 28

Distributed master data gathering scenarios



S1 System integrator
S2 Manufacturer
S3 Supplier
S4 Wholesaler
S5 Company that does manual internet research
S6 Company that does automated internet research
S7 Company that sells parts data
S_n… Company that provides data services

Slide 29

ISO 8000-110:2008 MASTER DATA QUALITY

The three steps to ECCMA Master Data Quality Manager (MDQM™) Certification

- Create an Identification Guide (eOTD-i-xml)
- Create a request for data (eOTD-q-xml)
- Create a reply to a request for data (eOTD-r-xml)

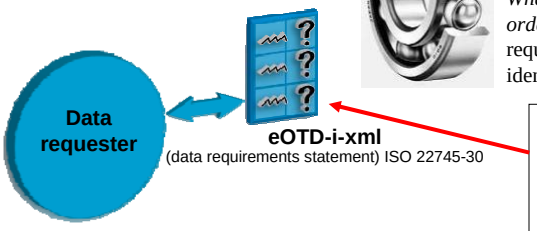
Slide 30

Step 1- Creating an Identification Guide

The buyer, item manager or user as the <u>data requester</u> defines their requirements for data:

What is the Class (item name) and what characteristic data do I need?
A data requirement statement is created as an eOTD-i-xml identification guide, an XML file that conforms to ISO 22745 part 30 in which the item name is the class and the characteristic data is represented by property-value pairs

1.A
What is the common name in the eOTD ?
Bearing: Ball; Annular
1.B
What do I need to know about this bearing in order to buy or manage it? This is a data requirement statement also known as an identification guide or a cataloguing template.



BEARING,BALL,ANNULAR
- Inner Diameter
- Outer Diameter
- Width
- # Rows
- Sealing Type
- Load Rating
- Speed Rating

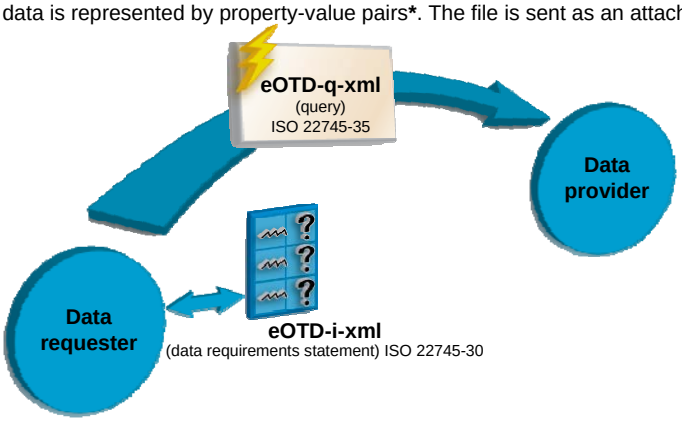
Properties selected from the eOTD

Slide 31

Step 2 – Creating a request for data

The buyer, item manager or user, as the <u>data requester</u> creates a query and sends it to the supplier or manufacturer as the <u>data provider</u>

"*Please provide (validate) the specified characteristic data for the item identified by the reference (part number) or Please provide a reference (part number) for an item(s) that matches the specified characteristic data."* A query is created as an eOTD-q-xml message, an XML file that conforms to ISO 22745 part 35 in which the item name is the class and the characteristic data is represented by property-value pairs*****. The file is sent as an attachment to an email.



*****It is a simple task to convert an eOTD-q-xml file to plain text in any format such as a spreadsheet, a word processed document, or even to a web page that can be easily filled in by the data provider

Slide 32

Examples of data requirements for Material and Service Master Data Quality

Item data can be categorized in accordance with one of the following quality levels:

Level I: The item is identified in terms its source of supply and has a reference number sufficient to successfully place an order for the item.

Level II: The item is identified and has been assigned a class sufficient to reference an Identification Guide and analyse spend (NATO Type 2) .

Level III: The item is identified and partially described, a class has been assigned and some of the properties specified in the Identification Guide for the class have been provided (NATO Type 4).

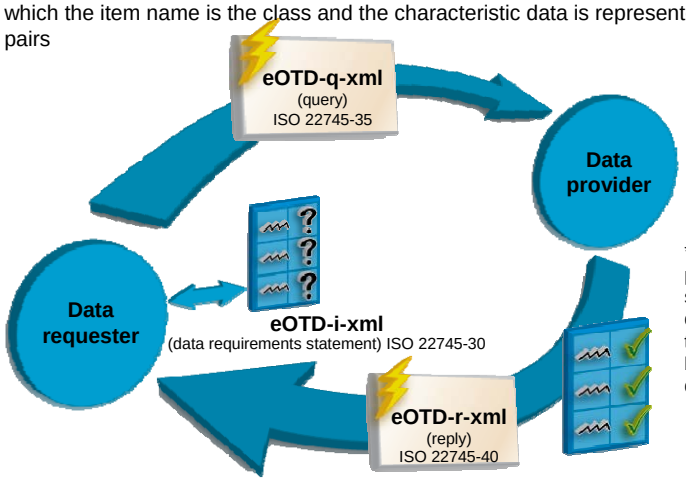
Level IV: The item is identified and completely described in accordance with the Identification Guide (NATO Type 1).

Slide 33

Step 3 – Creating a response to a request for data

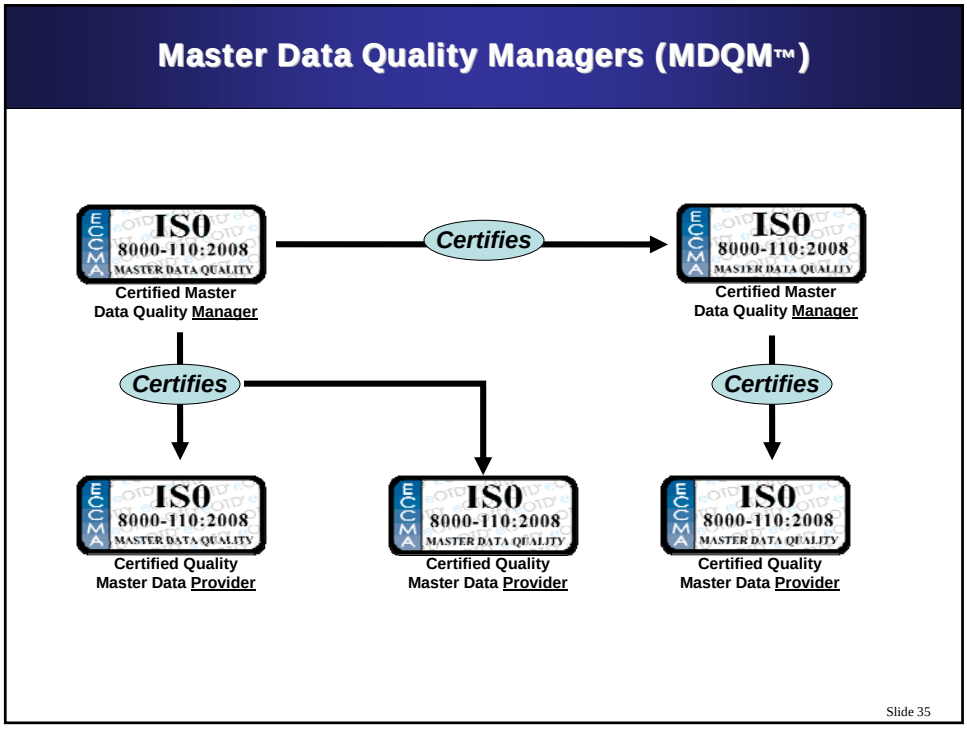
The supplier or manufacturer as the data provider , sends the reply to the buyer, item manager or user, as the data requester

The reply is created as an eOTD-r-xml*, an XML file that conforms to ISO 22745 part 40 in which the item name is the class and the characteristic data is represented by property-value pairs



*As an alternative the data provider can return a filled in spreadsheet, word processed document or web page as this can easily be converted back to eOTD-r-xml by the data requestor

Slide 34


Master Data Quality Managers (MDQM™)
ISO 8000-110:2008 MASTER DATA QUALITY
Certifies
Certified Master Data Quality Manager
Certified Master Data Quality Manager
Certifies
Certifies
ISO 8000-110:2008 MASTER DATA QUALITY
Certified Quality Master Data Provider
ISO 8000-110:2008 MASTER DATA QUALITY
Certified Quality Master Data Provider
ISO 8000-110:2008 MASTER DATA QUALITY
Certified Quality Master Data Provider
Slide 35