FEDERAL DATA ARCHITECTURE
Definitions and Structure

Michael F. Reed
Managing Partner, EWSolutions
mreed@ewsolutions.com

Executive Summary/Abstract: ISO/IEC 11179 is one of the few mature standards for storing enterprise metadata in a controlled environment. When data elements are well documented according to ISO/IEC 11179 and the documentation is managed in a Data Element Registry, finding and retrieving them from disparate databases as well as sending and receiving them via electronic communications are made easier.
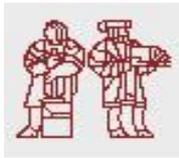This brief describes the elements and usage of ISO 11179.

The MIT Information Quality Industry Symposium, 2007
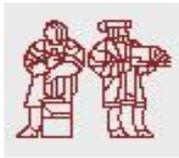
ISO 11179

- ISO/IEC JTC1 SC32 WG2 11179 1-6

- In our experience, the most commonly used metadata (data) standard after Dublin Core

- ISO/IEC 11179 is one of the few mature standards for storing enterprise metadata in a controlled environment

The MIT Information Quality Industry Symposium, 2007
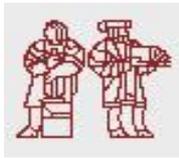
ISO 11179

Part 1: Framework, introduces and discusses fundamental ideas of data elements, value domains, data element concepts, conceptual domains, and classification schemes essential to the understanding of this set of standards and provides the context for associating the individual parts of ISO/IEC 11179.

The MIT Information Quality Industry Symposium, 2007

ISO 11179

- **Part 2: Classification**, provides a conceptual model for managing classification schemes. There are many structures used to organize classification schemes and there are many subject matter areas that classification schemes describe.
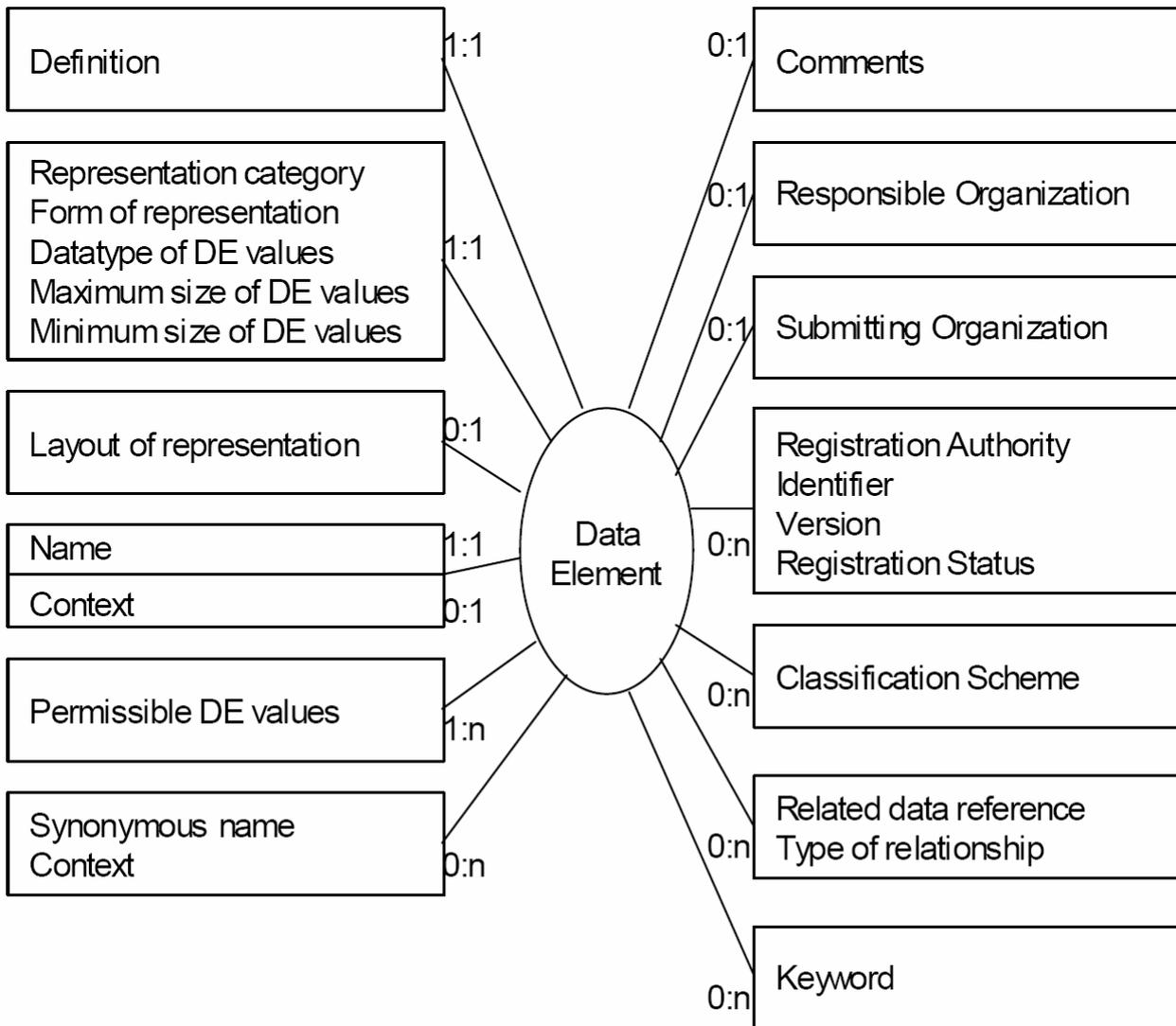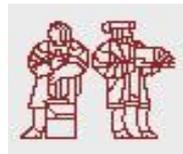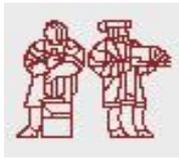
The MIT Information Quality Industry Symposium, 2007

ISO 11179

- **Part 3: Registry Metamodel and Basic Attributes**, specifies a conceptual model for a metadata registry. It is limited to a set of basic attributes for data elements, data element concepts, value domains, conceptual domains, classification schemes, and other related classes, called administered items.

# The MIT Information Quality Industry Symposium, 2007

Definition	1:1

Representation category
Form of representation
Datatype of DE values
Maximum size of DE values
Minimum size of DE values

1:1

Layout of representation — 0:1

Name — 1:1

Context — 0:1

Permissible DE values — 1:n

Synonymous name
Context — 0:n

Data Element

Comments — 0:1

Responsible Organization — 0:1

Submitting Organization — 0:1

Registration Authority
Identifier
Version
Registration Status — 0:n

Classification Scheme — 0:n

Related data reference
Type of relationship — 0:n

Keyword — 0:n

The MIT Information Quality Industry Symposium, 2007

ISO 11179

- **Part 4: Formulation of Data Definitions**, provides guidance on how to develop unambiguous data definitions. A precise, well-formed definition is one of the most critical requirements for shared understanding of an administered item; well-formed definitions are imperative for the exchange of information.
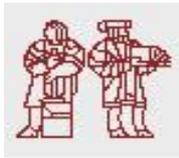
The MIT Information Quality Industry Symposium, 2007
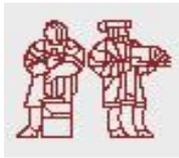
ISO 11179

- Examples of unambiguous guidelines

- Requirements:

  A data definition shall:

  - a) be stated in the singular

  - b) state what the concept is, not only what it is not

  - c) be stated as a descriptive phrase or sentence(s)

  - d) contain only commonly understood abbreviations

  - e) be expressed without embedding definitions of other data or underlying concepts

The MIT Information Quality Industry Symposium, 2007

ISO 11179

## Examples of unambiguous guidelines

- Requirements:

  A data definition shall:

  – be stated in the singular

  – state what the concept is, not only what it is not

  – be stated as a descriptive phrase or sentence(s)

  – contain only commonly understood abbreviations

  – be expressed without embedding definitions of other data or underlying concepts

The MIT Information Quality Industry Symposium, 2007
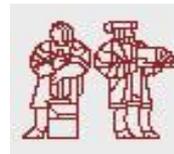
ISO 11179

Examples of unambiguous guidelines

- Recommendations:

A data definition should:

 – state the essential meaning of the concept

 – be precise and unambiguous

 – be concise

 – be able to stand alone

 – be expressed without embedding rationale, functional usage, or procedural information

 – avoid circular reasoning

 – use the same terminology and consistent logical structure for related definitions

 – be appropriate for the type of metadata item being defined

The MIT Information Quality Industry Symposium, 2007
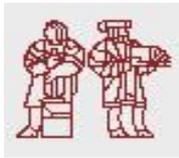
ISO 11179

- Example of an unambiguous data definition:
    - EXAMPLE - "Freight Cost Amount"

  good definition: Cost amount incurred by a shipper in moving goods from one place to another

  poor definition: Costs which are not related to packing, documentation, loading, unloading, and insurance

    - REASON - The poor definition does not specify what is included in the meaning of the data (do not define a concept in terms of what it is *not)*

The MIT Information Quality Industry Symposium, 2007
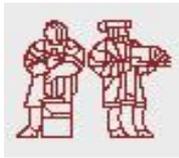
ISO 11179

- Example of a poor data definition:
  - EXAMPLE - two data elements with poor definitions:

  Employee ID Number - Number assigned to an employee.

  Employee - Person corresponding to the employee ID number.
  - REASON - Each definition refers to the other for its meaning. The meaning is not given in either definition. (Do not use circular arguments!)
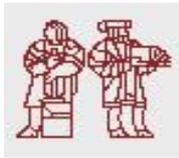
ISO 11179

- **Part 5: Naming and Identification Principles**. Identification includes the assignment of numerical identifiers that have no inherent meanings to humans; icons; and names with embedded meaning, usually for human understanding, that are associated with the data item's definition and value domain.
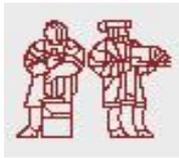
ISO 11179

- **Part 6: Registration**, provides instruction on how a registration applicant may register a data item with a central Registration Authority and the allocation of unique identifiers for each data item. Maintenance of administered items already registered is also specified in this document.

The MIT Information Quality Industry Symposium, 2007

- ❖ ISO 11179 is a very useful standard
- ❖ It is an easy way for your organization to agree on naming conventions
- ❖ Avoids creating a committee!