

WG2 N1337

First Principles for Data Semantics Standards

Frank Farance

Dan Gillman

US Experts

2009-11-13

Terminology Theory

- Object
 - Anything perceivable or conceivable
- Property
 - Determinant (of an object)
 - Differentiate objects
 - Result of a determination
- Characteristic
 - Determinable (capable of being determined)

Terminology Theory

- Concept
 - Unit of thought differentiated by characteristics
- Property and Characteristic
 - Concepts in roles
- Characteristic
 - Concept – feature common to set of objects
- Property
 - Concept – how the feature is determined

Terminology Theory

- Definitions
 - Intensional – based on superordinate concept and differentiating characteristics
 - Extensional – based on a set of subordinate concepts
- Set of determinants for a determinable
 - Extensional definition of a characteristic
- Follows ISO 704:2000
- Extension of concept
 - Set of corresponding objects

Terminology Example

- Concept – people living in the UK
- Characteristic – eye color
- Properties – brown, hazel, green, blue, grey
- Definition
 - Intensional – Humans living in the UK
 - Extensional – UK Children, UK Adults
- Designation = association of a concept with a signifier which denotes it

Data

- Datum = the designation of a value, where a value is a concept with a notion of equality defined
- Additional semantics
 - Concept whose extension is set of objects
 - Allowed values
 - Concept the allowed values define extensionally
 - Set of allowed values = partition of extension

Data

- Follows ISO/IEC 11179-3:2003
- Described in draft ISO/IEC 11179-4 Ed3
- Object Class = Concept – people in the UK
- Characteristic = Characteristic – eye color
- Property = Value meaning – brown, hazel, etc
- Each is a concept
- Concepts convey semantics

Data Example

- Object Class = adults in the UK
- Characteristic = marital status
- Value meanings = {single, married, etc}
- Datatype = state

Data and Metadata

- Framework describes data
- Any data used to describe some object = metadata
- Therefore, metadata are data
- Description of metadata is terminological

Datatype

- Follows ISO/IEC 11404:2007
- Datatype = computational description of data
 - Value space
 - Assertions
 - Characterizing operations
- Metadata has a datatype, since metadata are data
- Then, metadata have computational description

Attributes

- Purpose – semantics for descriptors
- Follows FDIS ISO/IEC 19773
- IKV tuples
- I (Identifier) – name of characteristic of concept
- K (Kind) – datatype
- V (Value) – selected value in value space of datatype

Groupings

- More complex datatypes generated from simpler ones and a rule, called the generator
 - Again, follows ISO/IEC 11404:2007
- Sets of attributes can be built in the same way
- Using the same generation rules
- Groupings
 - Arbitrary, user defined
 - Defined in advance

Groupings Example

- Describe people
 - Attribute 1: sex, state, male
 - Attribute 2: marital status, state, married
- Grouping
 - Person { marital status, sex }

Ontology

- Concept system = a set of concepts structured according the relations among them
- Ontology = a concept system with a computational description
 - Follows draft ISO/IEC 11179-4 Ed3
- Examples
 - Datatypes
 - UML models

Services and Processes

- Operations on data (or metadata)
- Part of computational description for data
- In general, the steps in a service or process contain semantics

Conclusion

- Framework
 - Terminology
 - Metadata
 - Attributes
 - Generators
- Result
 - Self-describing version of any description framework