

LMI

GOVERNMENT CONSULTING

THE OPPORTUNITY TO MAKE A DIFFERENCE HAS NEVER BEEN GREATER



Federal XML Naming and Design Rules and Guidelines

Mark Crawford

Agenda

- Purpose
- Scope
- Audience
- Sources
- Terminology
- Modularity
- Namespaces
- Versioning
- Content
- Next Steps



The purpose of this document is to provide a set of rules and guidelines that will enable development of:

- a flexible federal modularity model that defines the structure for creating interoperable schema and schema modules
- a clearly defined namespace scheme that ensures consistency across Agencies
- a versioning scheme that will support consistency in versioning of government schema
- a Federal canonical schema for base Data Types
- specific NDR's by government agencies or communities of practice that build on this document
- a reference to use for a mapping of different agency NDR's to each other



The purpose of this document is to provide a set of rules and guidelines that will enable development of: (cont'd)

- consistent, reusable XML components that may be made available for reuse such as:
 - Schema
 - Schema Modules such as reusable code lists and identifier lists
 - Simple and Complex Types
 - Elements
 - Attributes
- a set of tools to facilitate ease of development, validation, and interoperability



Scope

This Federal XML Naming and Design Rules and Guidelines document is intended for use by all Executive Branch Departments and Agencies (hereinafter referred to as Agency) that employ XML, including all commercial and government off-the-shelf XML related product implementations. It should be used by all contractors and vendors doing XML development work on behalf of Departments and Agencies. Agencies developing specific XML Naming and Design Rules and Guidelines should use this document as the baseline for those efforts.



Audience (not yet vetted)

- Developers of Federal Enterprise Schema
- Government organizations looking for guidance
- Agency level developers interested in fostering interoperability
- Private sector organizations who wish to track government efforts



Sources

- Voluntary Consensus Standards Bodies
 - OASIS Universal Business Language Technical Committee
 - UN/CEFACT
- Government NDRs
 - Department of the Navy
 - Environmental Protection Agency
 - Global Justice



Terminology

- The key words **MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY,** and **OPTIONAL** in this document are to be interpreted as described in Internet Engineering Task Force (IETF) Request for Comments (RFC) 2119. Non-capitalized forms of these words are used in the regular English sense.



Modularity is key to reuse

- Modularity Model Must Be:
 - Structured
 - Flexible
 - Consistent

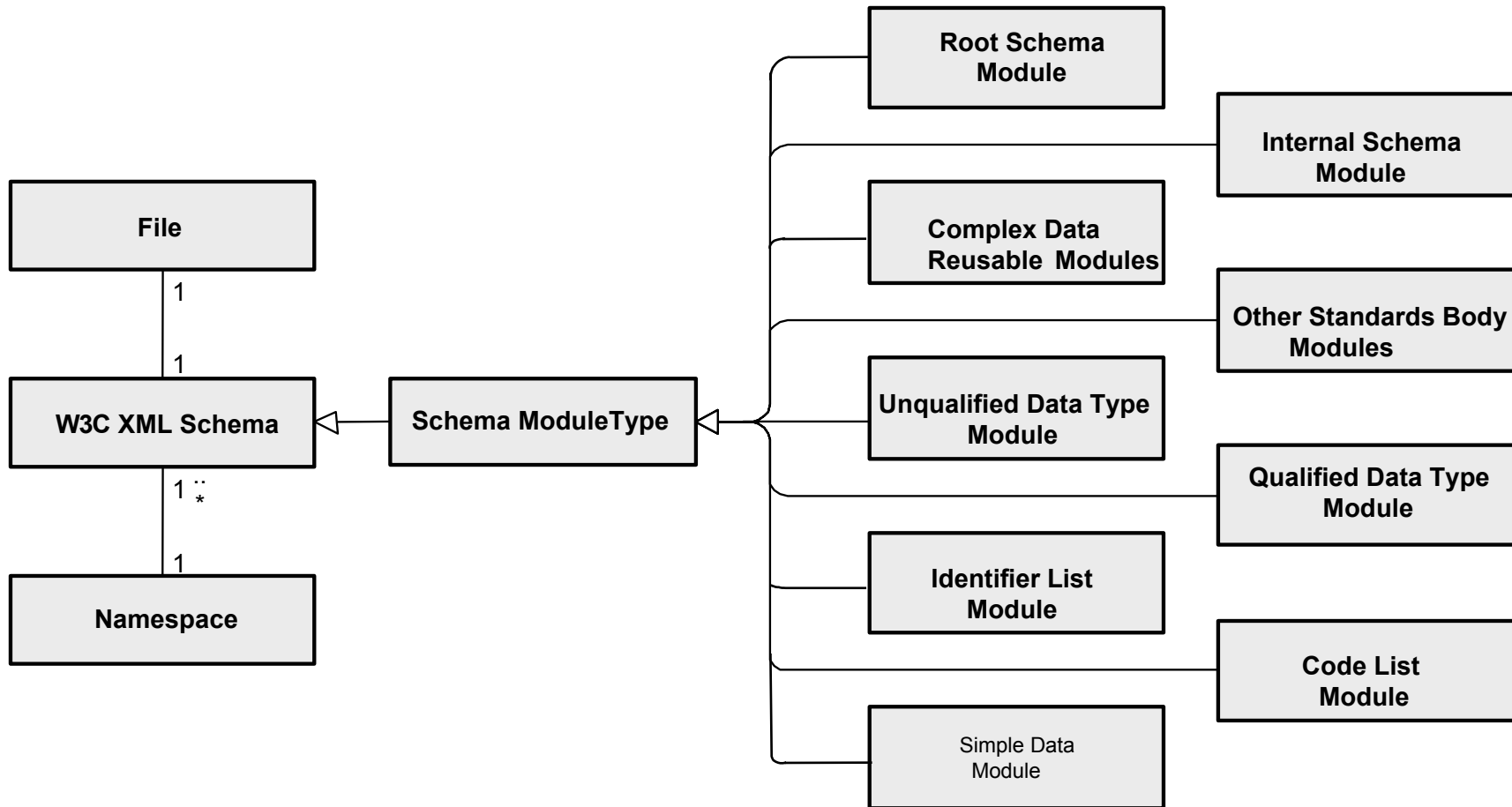


Modularity

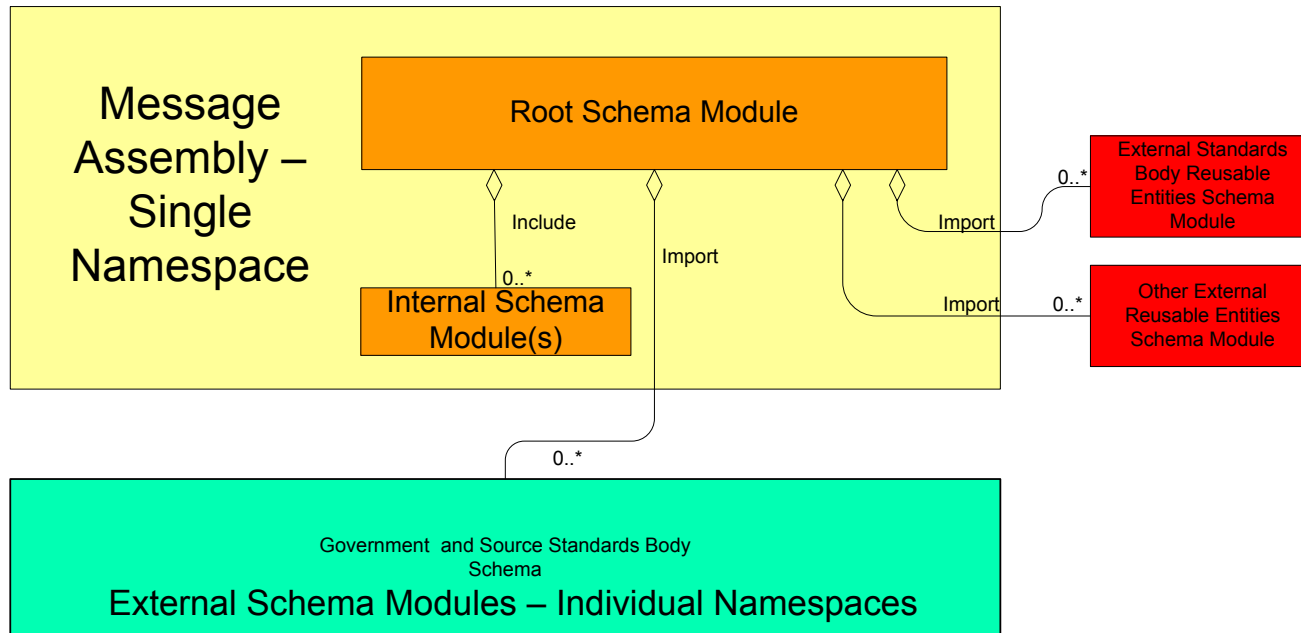
- Three Approaches under consideration
 - Monolithic
 - Single Namespace
 - One schema per process, idea, or information requirement
 - No imports
 - Reuse
 - Root Schema with all content imported
 - Unique namespace for each schema
 - Common modules for reusables (data elements and generic data elements), standard data types, code lists, identifier lists
 - Unique
 - Root Schema with all content imported
 - Unique namespace for each schema
 - Unique schema modules for each data element i.e. an address would be a stand alone schema in a unique namespace



Modularity Model – Reuse Approach



Modularity Model – Root Schema

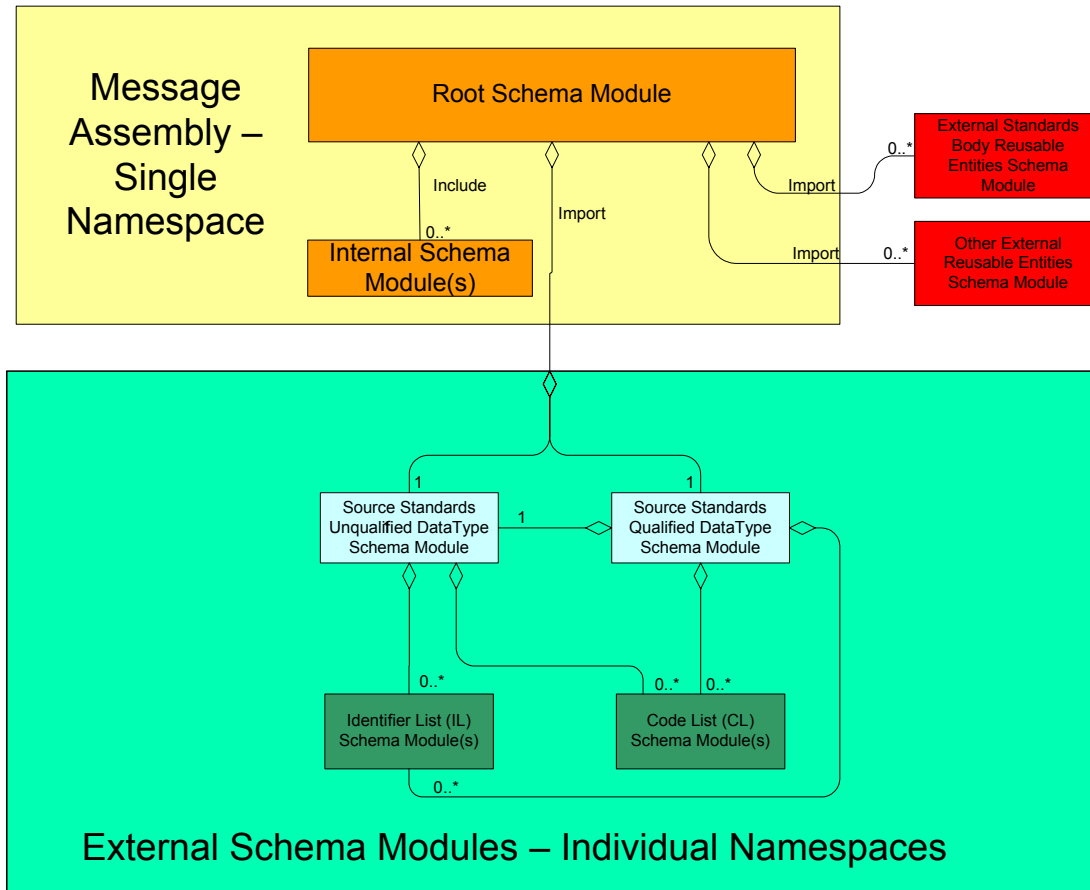


- A Root Schema is a single concept like an Order, Invoice, OMB 300, SF12, Document, Book, or some other logical grouping of information.



Modularity Model – Importing Data Types From Standards

- UBL and other standards bodies are converging on UN/CEFACT Data Type schema modules
- UBL and other standards bodies are converging on single approach to code lists



What Data Types

- Amount (xsd:decimal)
- Binary Object (xsd:base64Binary)
 - Graphic
 - Sound
 - Video
- Code (xsd:normalizedString)
- Date Time (xsd:dateTime)
 - Date (xsd:date)
 - Time (xsd:time)
- Identifier (xsd:normalizedString)
- Indicator (xsd:boolean)
- Measure (xsd:decimal)
 - Value
 - Rate
 - Percent
- Numeric (xsd:decimal)
- Quantity xsd:decimal)
- Text (xsd:string)
 - Name (xsd:string)



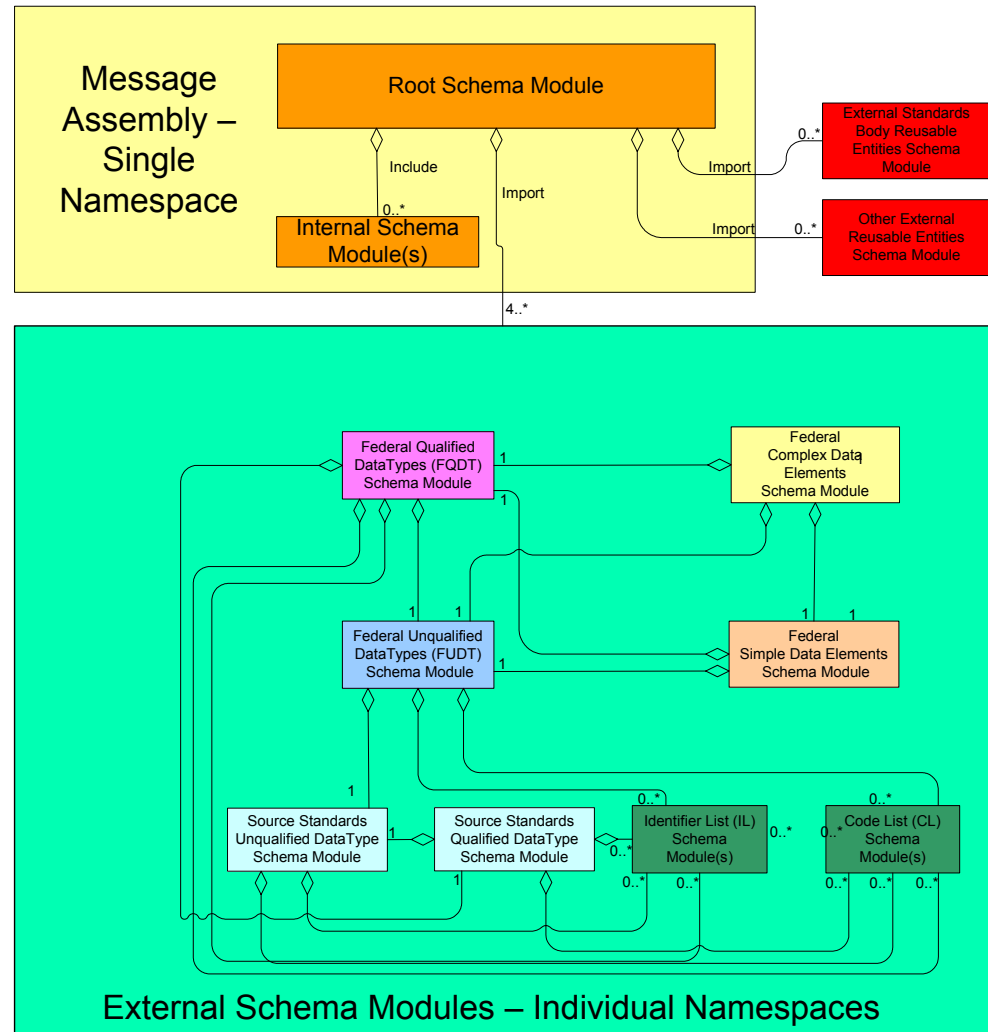
A Data Type Example

```
<xsd:complexType name="AmountType">
  <xsd:annotation>
    <xsd:documentation>
      <Component>
        <ComponentType>DT</ComponentType>
        <DictionaryEntryName>Amount. Type<:DictionaryEntryName>
        <Definition>A number of monetary units specified in a currency where the unit of the currency
is explicit or implied.</Definition>
        <RepresentationTerm>Amount</RepresentationTerm>
      </Component>
    </xsd:documentation>
  </xsd:annotation>
  <xsd:simpleContent>
    <xsd:restriction base="cct:AmountType">
      <xsd:attribute name="amountCurrencyID" type="xsd:normalizedString" use="required"/>
      <xsd:attribute name="amountCurrencyCodeListVersionID" type="xsd:normalizedString"
use="optional"/>
    </xsd:restriction>
  </xsd:simpleContent>
</xsd:complexType>
```



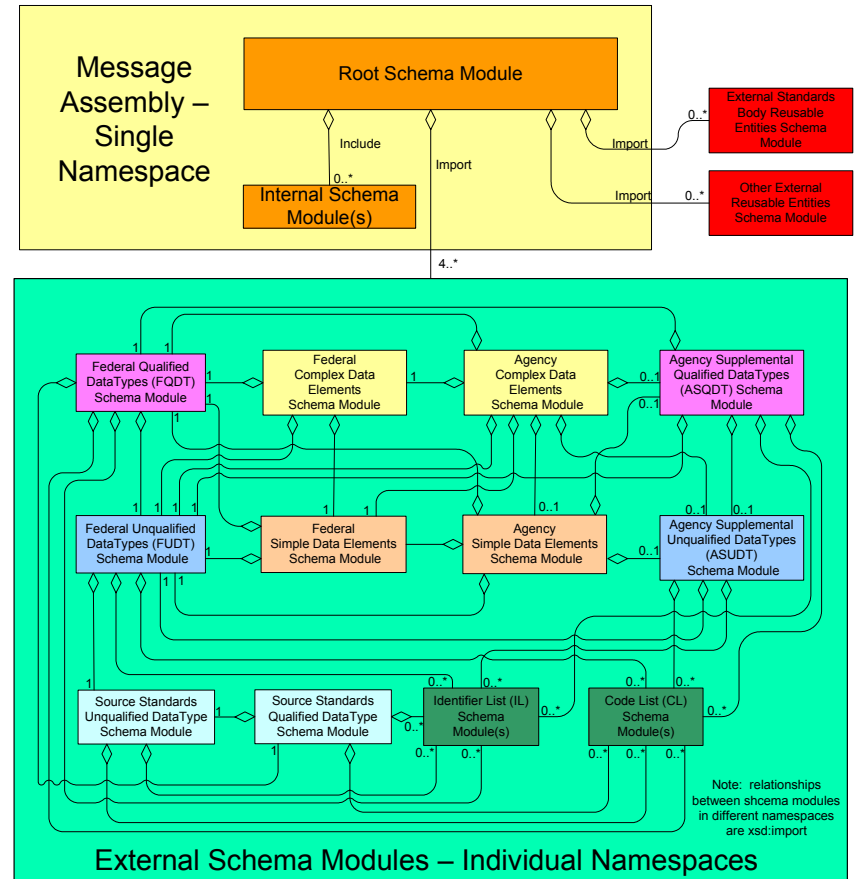
Modularity Model – Introducing the Federal Common Schema

- Reusing the Standards Body Schema
- Creating 4 Federal Enterprise Schema
 - Unqualified DataTypes
 - Qualified DataTypes
 - Complex Data Element Reusables
 - Address
 - Person
 - Organization
 - Simple Data Element Reusables
 - First Name
 - Last Name
 - Birthdate
- Code and Identifier Lists



Flexibility for Everyone

- Government organizations and initiatives can leverage standards body and federal schema
- Organizations can create organizationally unique
 - Qualified and unqualified Data Types
 - Complex Data Element reusables
 - Simple data element reusables
 - Code Lists
 - Identifier Lists



Namespaces

- Initial Recommendation –Mandate hierarchical URN scheme
- Consensus – SHOULD use URN and hierarchy scheme. MAY use URL hierarchy scheme



Namespaces

- 1st Level Domain – NID of US
- Second Level Domain – Organization Hierarchy. In this case GOV
- Third Level Domain – Specific Government Hierarchy – EPA, OMB, DoD, Treasury
- Fourth Level Domain – Agency Level Hierarchy – USN, USAF, IRS, FMS
- Fifth Level Domain – Resource Type – Schema or Other as Identified
- Sixth Level Domain – Resource Status
- Seventh Level Domain – Resource Name



Versioning

- Consensus on:
 - <name>-<major >.<non-zero>[.<revision>]
- To be determined:
 - Minor versioning of namespaces
 - Use of namespaces for schema location
 - URN or URL for schema location



Schema Content

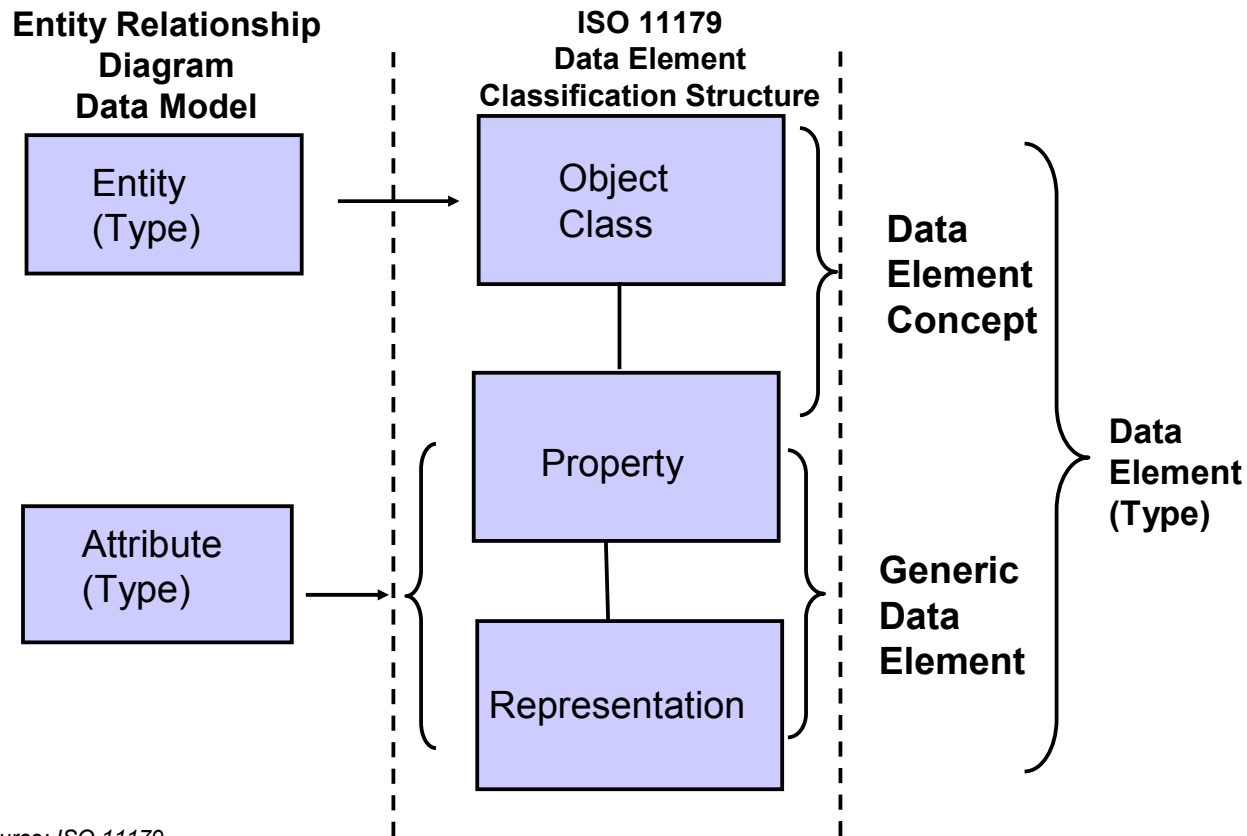
- Modularity, Namespaces and Version is important, But:

What about the Schema Content?



Schema is All About Data

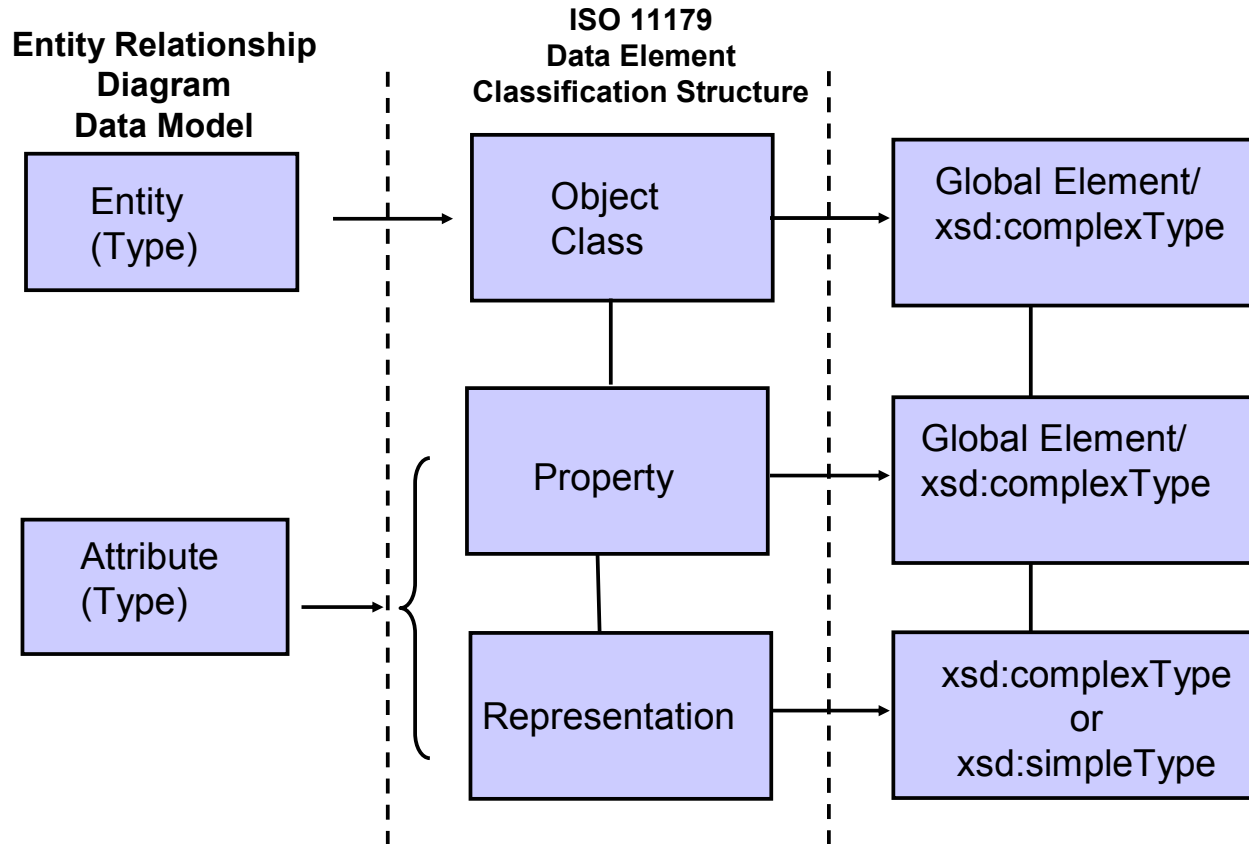
Data Concept



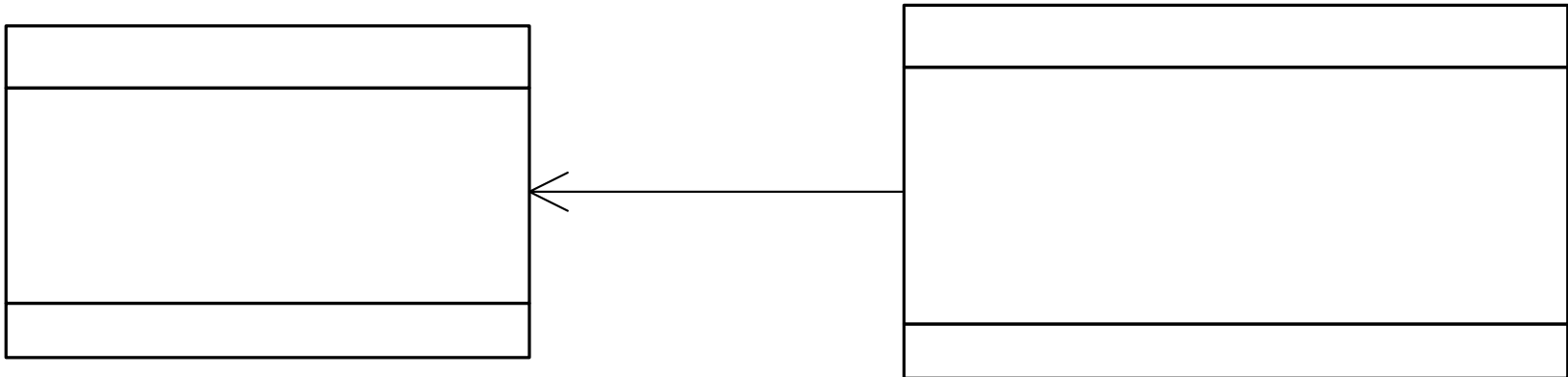
Source: ISO 11179

Transforming Data to XML

Simple XSD Transformation



Handling Class Associations



- An associated class is treated as a simple data element with a global element declaration and complex type definition



Next Steps

- Continue to work through comments to first draft
 - Finalize modularity
 - Finalize versioning
 - Finalize namespace
- Flesh out content

