Harmonizing XML Tag Semantics via UDEF Intelligent Identifiers

XML ‘99

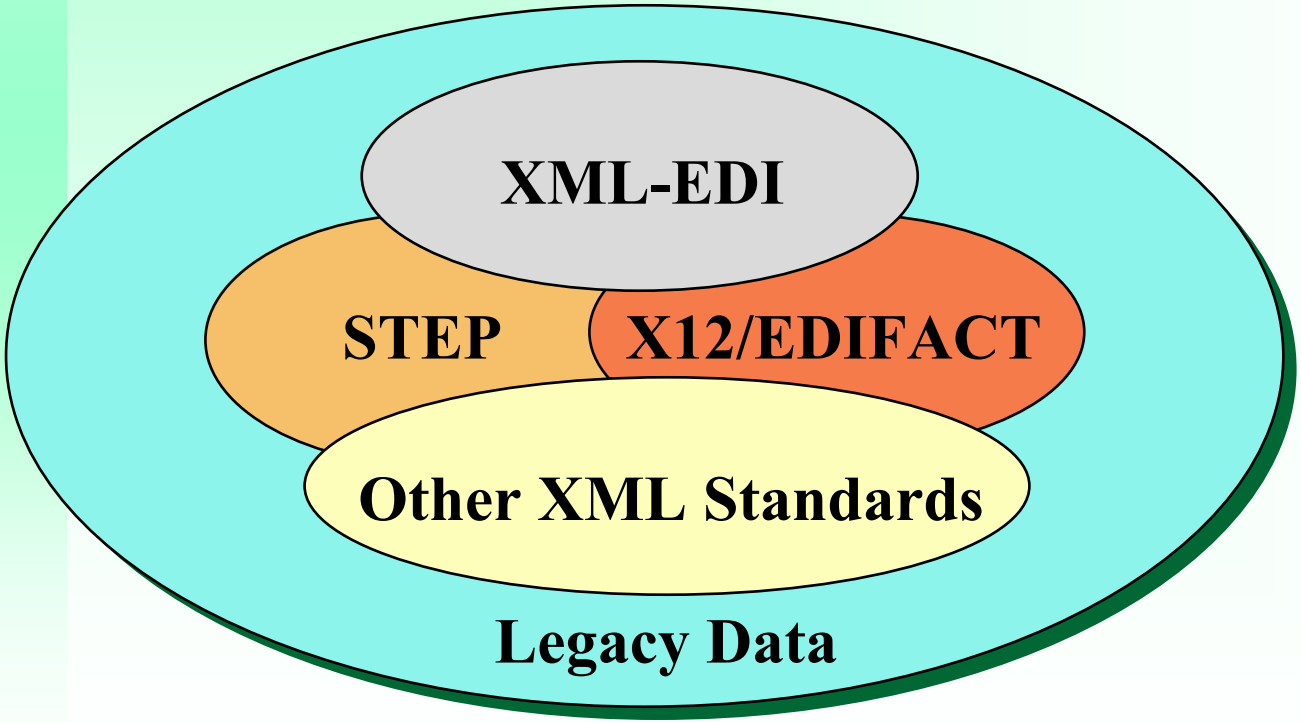
Ron Schuldt, December 7, 1999

XML Interest Domains

- Accounting
- Education
- Electronic Data Interchange
- Enterprise Resource Planning
- Finance
- Healthcare
- Human Resources
- Insurance
- News
- Real Estate
- Science
- Travel
- Web Applications
- … Others

The Problem

Overlapping Standards



Why Harmonize Tags?

- To take advantage of XML's potential for finding the information or data diamond in the rough
  - Within the enterprise
  - Across enterprises
- Semantically equal tags are necessary across multiple interest domains
  - Enterprise address
  - Enterprise name
  - Product attributes
  - ... Hundreds of other key tags

Proposed Approach

- Global intelligent data index - like barcode
  - **OASIS and XML.ORG managed**
  - **Optional mapping of DTD tags when submitted to XML.ORG registry and repository**
  - **Index added as option to each tagged data element**
- Built on rules based Universal Data Element Framework (UDEF)
  - **Developed within CALS**
  - **Tested within Electronic Industries Association**
  - **Conforms to ISO/IEC 11179 data element naming principles**

Finding the Diamond

XML '99

UDEF Participants

CALS ISG - Developed UDEF

Industry

Ron Schuldt (Chair) - Lockheed Martin
Barbara Barman (Vice Chair) - Raytheon
Rob Bryant - DynCorp
Ruey Chen - David Taylor Research Center
Bob Hodges - Texas Instruments
Neal McNamara - Analysis & Technology Inc.
Bud Orlando - TRW
Madelyn van der Bokke - ASEC
George Walther - Lockheed Martin

Government

Norma Kornwebel - PM JCALS
Dinah Beres - NAWC
Steve Waterbury - NASA

EIA - Applied UDEF

Industry

Ron Schuldt (Chair) - Lockheed Martin
Rick Lang - Texas Instruments
Pam Stanfield - Lockheed Martin
Gary O'Hara - Hughes Space and Comm
Tony DiPerna - Ericsson Communications
Ken McTee - Texas Instruments
Cindy Hauer - Mevatec Corp
Fred Bahrs - CMstat Corporation
Doug Drury - ITT Federal Services
Lee LeClair - Texas Instruments

Government

Deborah Cornelius - US Army Missile Cmd
C. H. VanLandingham - NOAA Nat'l Wea Svc

What is the UDEF?

- Rules based approach for naming data elements - for example:

  "part number", "delivery date" and "country code" are transformed to more rigorous names when the rules of UDEF are applied

- Uses tree structure taxonomies that easily transform into an intelligent universal identifier

UDEF Objects and Properties

Entity = 0
Asset = 1
Document = 2
Enterprise = 3
Environment = 4
Human = 5
Law/Rule = 6
Place = 7
Process = 8
Product = 9
Program = 10
Condition = 11

+

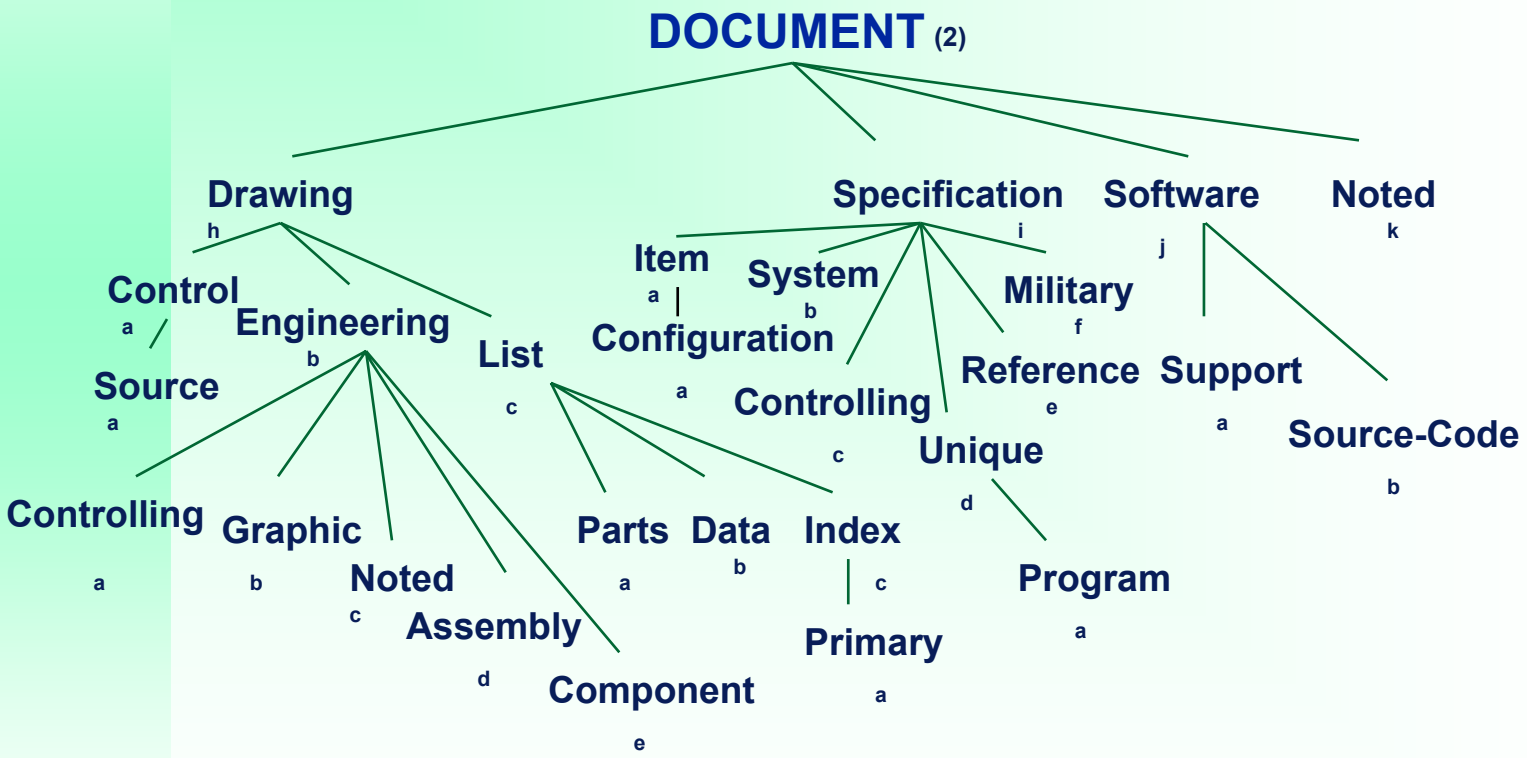
From DoD 8320.1-M-1

Amount = 1
Angle = 2
Area = 3
Code = 4
Coordinate = 5
Date = 6
Dimension = 7
Identifier = 8
Mass = 9
Name = 10
Quantity = 11
Rate = 12
Temperature = 13
Text = 14
Time = 15
Volume = 16
Weight = 17

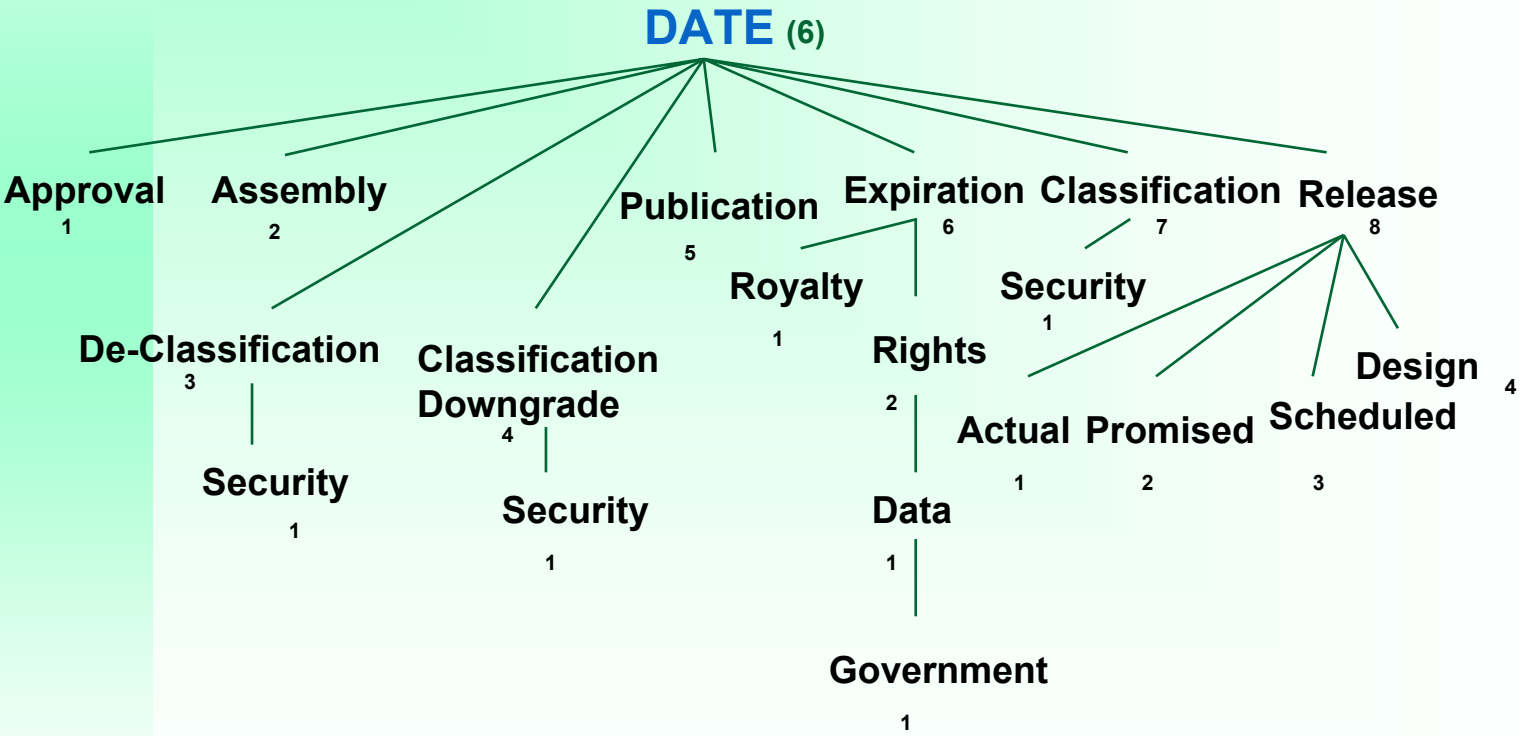
Finding the Diamond

XML '99

"Document" Object Example



"Date" Property Example



Finding the Diamond

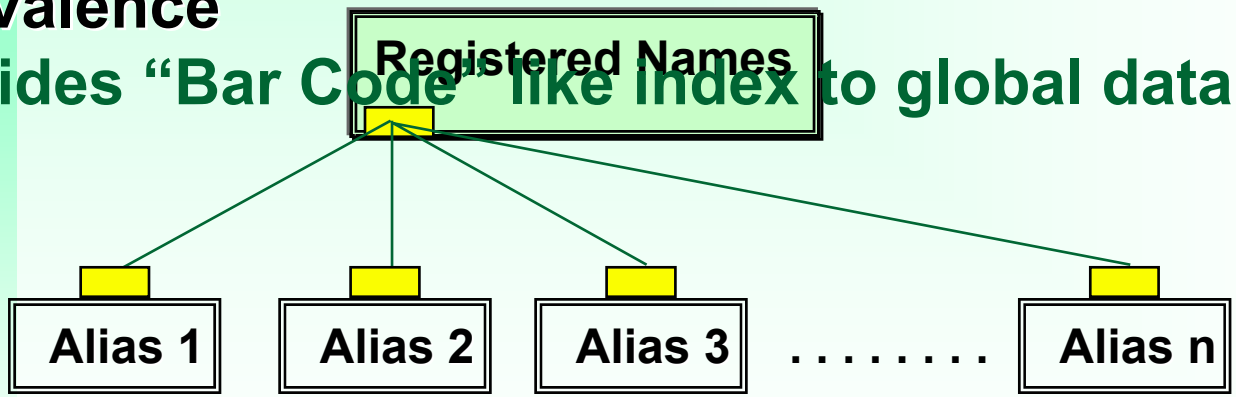
XML '99

Universal ID is Key

Universal ID is the "KEY"

- **allows for an unlimited number of aliases**
- **no system needs to change its data element names**
- **mapping to UDEF is based on semantic equivalence**

Provides "Bar Code" like index to global data



Universal ID =

Mapping Examples

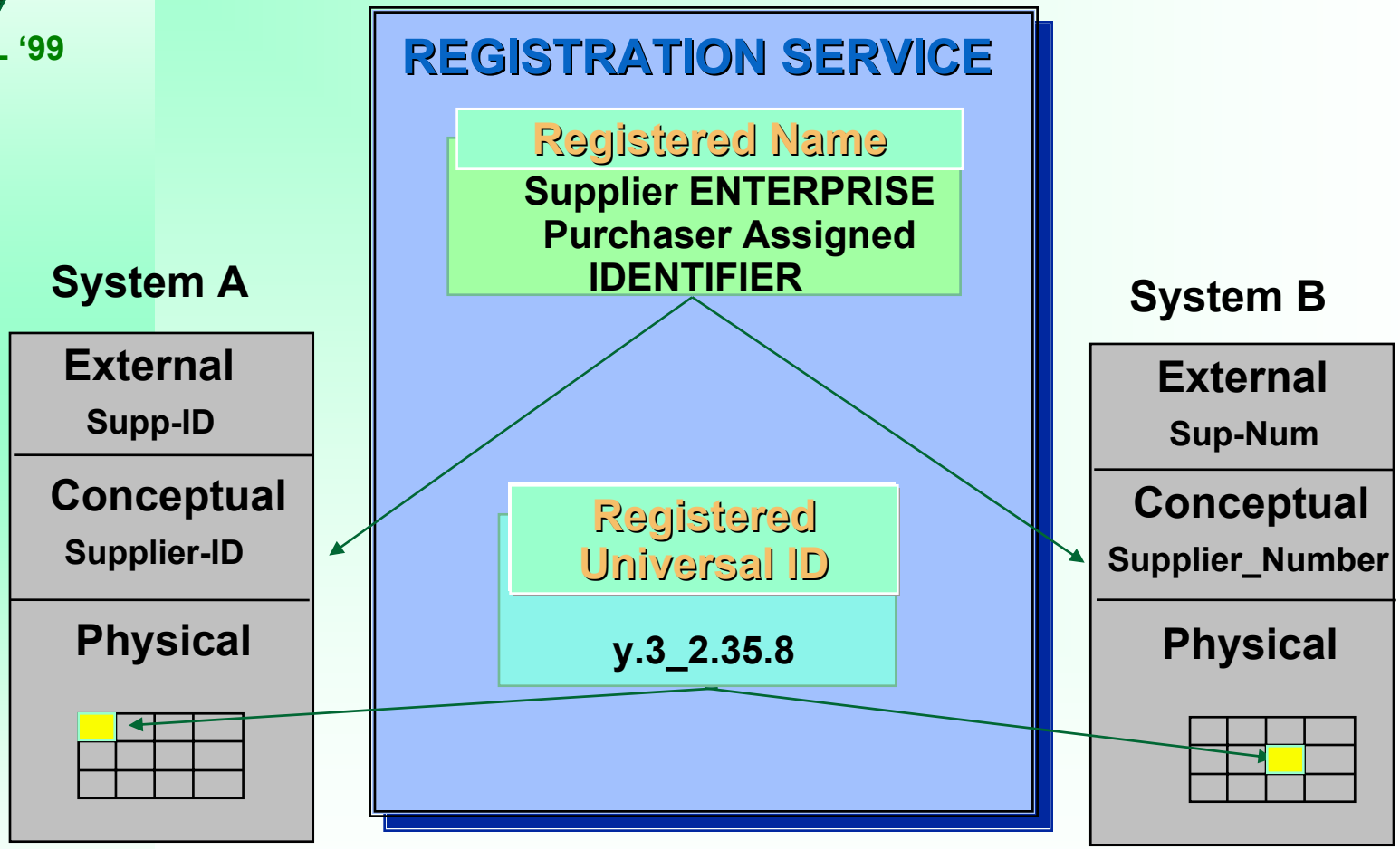
X12 & EDIFACT Data Elements	UDEF ID
country code	e.7_4
invoice number- assigned by issuer	bd.2_1.35.8
purchase order type code	d.t.2_33.4
postal code	7_1.10.4
location qualifier	7_20.33.4
location identifier	7_8.4
contract effective date	e.2_13.6
expiry date of import license	a.be.2_6.6
item number - product	9_8
item number - service	f.9_8
price	9_2.1

UDEF ID	MIL-STD-2549	X12 (EDI)	System A
g.9_8	Part Product Identifier	Product/Service ID	Part No
g.9_9	Part Product Name	Product/Service Name	
y.3_9		Entity (Supplier) Name	Supplier
e.2_8	Contract Document Identifier	Buyer's Contract Number	Contract No
f.g.9_11	Component Product Quantity		
2_33.4	Document Type Code	Report Type Code	Doc Type

Finding the Diamond

XML '99

Require Registration Service



Propose that OASIS and XML.ORG provide this service

Some of the Benefits

- UDEF Indexed Data Elements (XML Tags)
  - **Rapidly aligned**
  - **Language independent**
  - **Provide ID required by ISO/IEC 11179**
  - **Allow rapid comparison of XML schema content**
  - **Support harmonization across domains**
- UDEF Top Level Objects & Properties
  - **Intuitive**
  - **Tested**
  - **Provide a common frame of reference to establish context (key to semantics)**