

An Introduction to the Universal Data Element Framework (UDEF)

**Ron Schuldt, Chair
The Open Group UDEF Forum
October 20, 2005**

Agenda

- **Brief background – where the UDEF originated**
- **The problem and associated UDEF business value**
- **What the UDEF is and what it is not**
- **A brief description of the UDEF**
- **The six basic steps to map enterprise data to the UDEF**
- **A metadata managed architecture that UDEF supports**
- **UDEF status**
- **A brief description of a UDEF pilot to support disaster response teams**

UDEF Background

CALS ISG - Developed UDEF

in the late 80s – early 90s

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

in the mid 90s and 2002

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

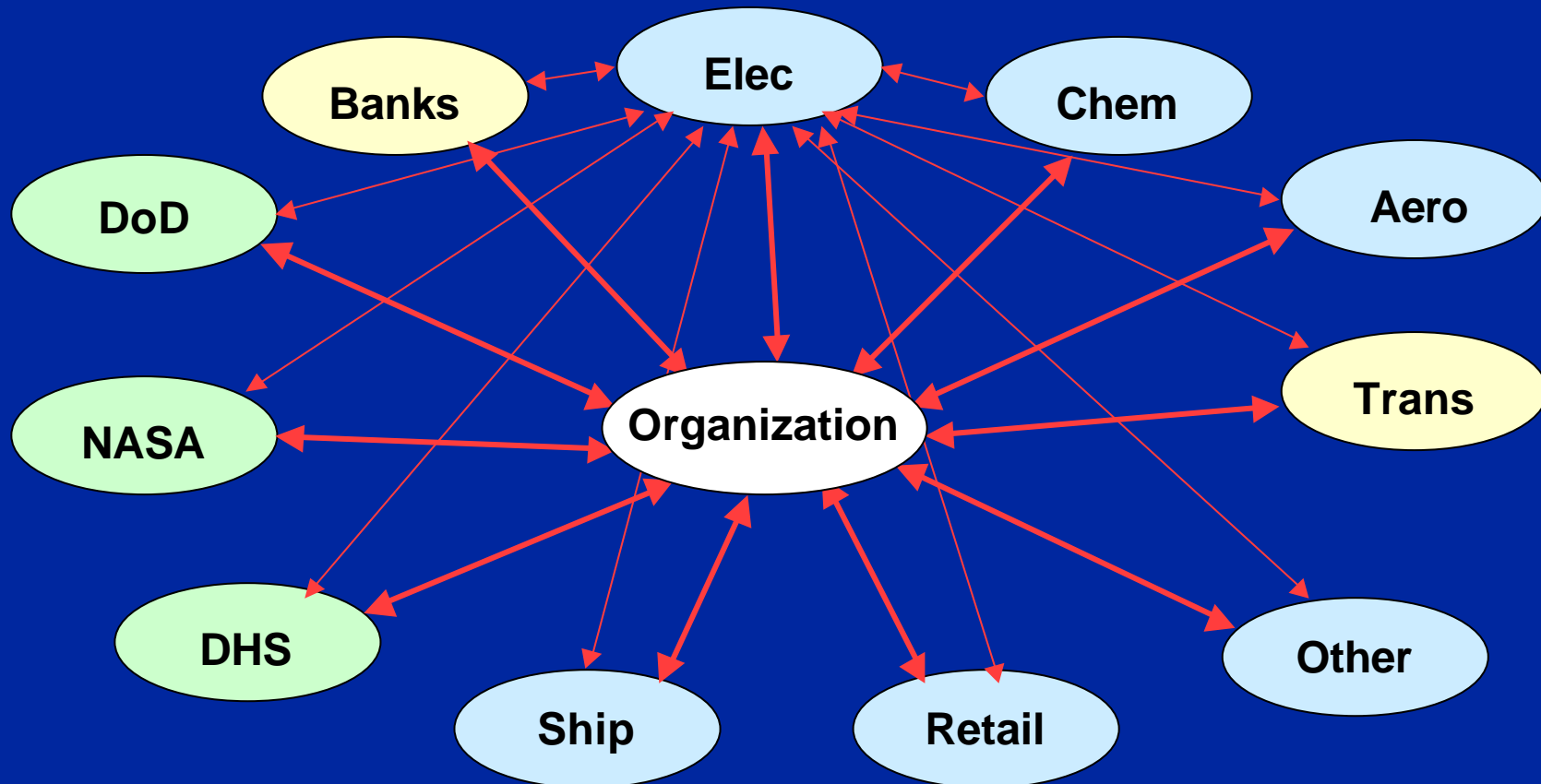
C. H. VanLandingham - NOAA Nat'l Wea
Svc



Today AFEI (formerly CALS ISG) holds the Intellectual Property Rights to the UDEF

Problem – Global Perspective

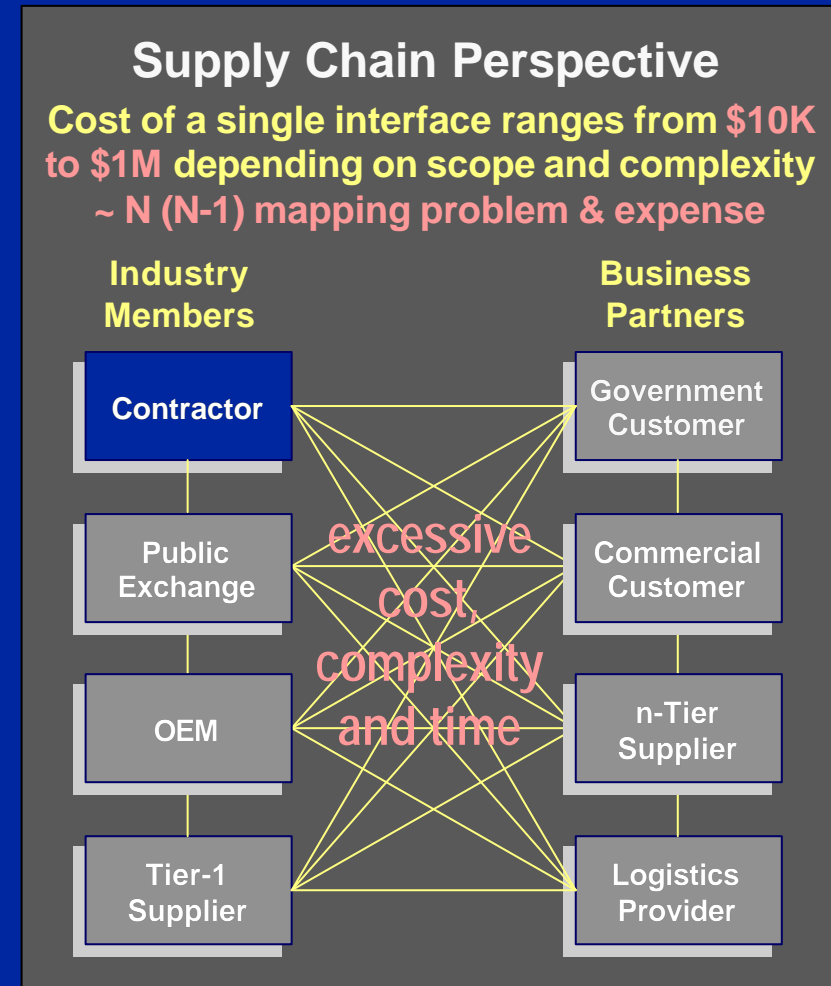
Each organization is attempting to set its own semantics standard
Each must interface with organizations they do not control



The problem is the lack of common semantics and schema between organizations

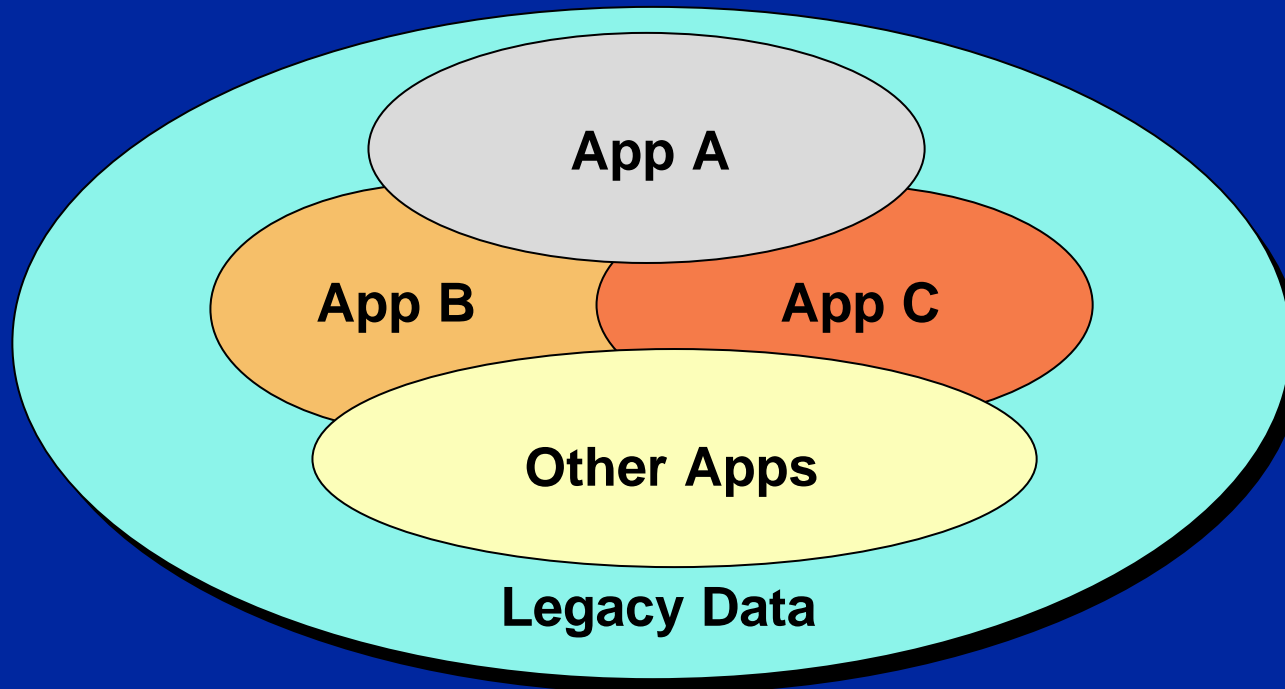
Problem – Supply Chain Perspective

- Today's reality... point-to-point solutions introduce excessive cost, complexity and time
- Examples of gaps and/or inefficiencies:
 - Increasing number of unique portals adding cost to suppliers
 - Multiple, redundant, incompatible “portal systems”
 - Incompatibilities in information exchange contribute to delay, rework, and error
- Excessive cost, complexity and time impeding supply chain agility



Problem – Internal Perspective

Conflicting semantic overlaps between back-office systems



Though semantically equal, the following are 4 different XML tag names

```
<PARTNUMBER>111-222-333</PARTNUMBER>
```

```
<partNumber>111-222-333</partNumber>
```

```
<PartNumber>111-222-333</PartNumber>
```

```
<partnumber>111-222-333</partnumber>
```

The Impact on Integration

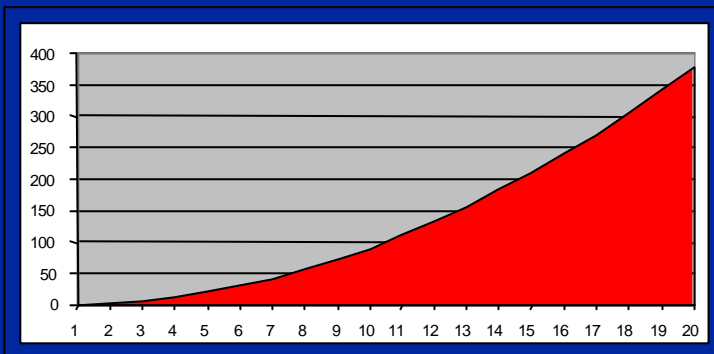
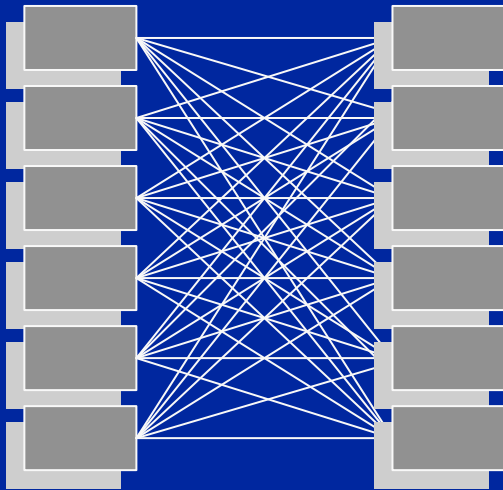
To interface two or more applications either within the enterprise or between enterprises ----

- **Middleware vendors build and sell “adaptors” for interfaces between major commercial applications (i.e., SAP - PeopleSoft)**
 - As the version of a given application changes it typically drives the need for changes to the adaptors
- **Application interface developers (data architects) spend considerable time and effort analyzing the semantics of data that need to be exchanged between systems**
 - Semantics analysis **requires up to 35-40%** of the requirements and design-time phase labor cost of building a given interface
 - To minimize the semantics analysis effort typically **requires support (availability) from the “experts”** of each application since most applications are poorly documented

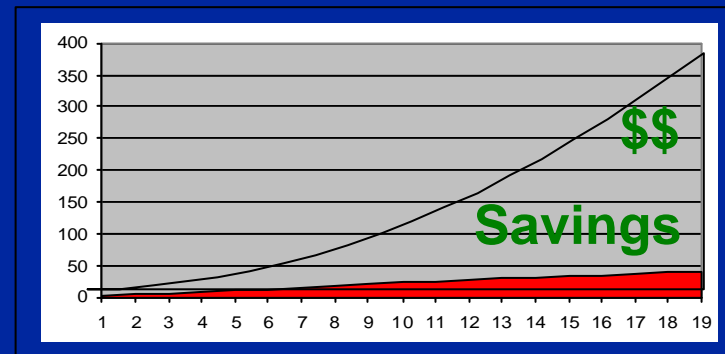
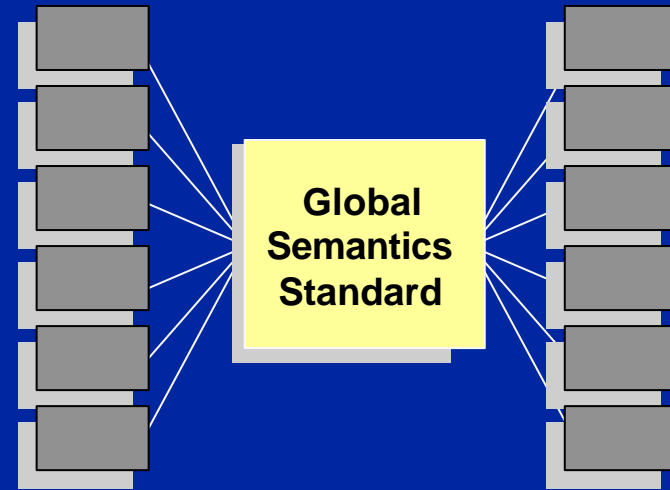
The Goal

Reduce Requirements and Design-Time Phase Semantics Analysis Time and Cost

**Current Point-to-Point
Approach --- $n(n-1)$**



**Global Semantics Standard
Approach --- $2n$**



UDEF Value Proposition

Typical Interface Build Tasks

- Analyze and document the business requirements.
- Analyze and document the data interfaces (design time)
 - ◆ Compare data dictionaries
 - ◆ Identify gaps
 - ◆ Identify disparate forms of representation
- Perform data transformations as required at run time
 - ◆ Transform those data that require it

UDEF Value

- ➡ Reduces dependency on system expert
- ➡ Allows automated compare
- ➡ Reduce design time labor
- ➡ Step toward automated transform



UDEF – Is and Is Not

UDEF Is

- **A “loose coupling” method for integrating semantics across disparate systems**
 - Allows automation to help align disparate systems – thereby reducing the design-time phase that is typically manual
- **A standards compliant rules-based naming convention**
 - Conforms to ISO/IEC 11179 and to ebXML Core Components
- **Evolving to become a global open standard**

UDEF Is Not

- **A software application**
 - However, software applications need to adopt UDEF within their APIs
- **Another data standard**
 - It acts as the glue to tie data standards together
- **Another ontology**
 - It could help build the bridge between ontologies

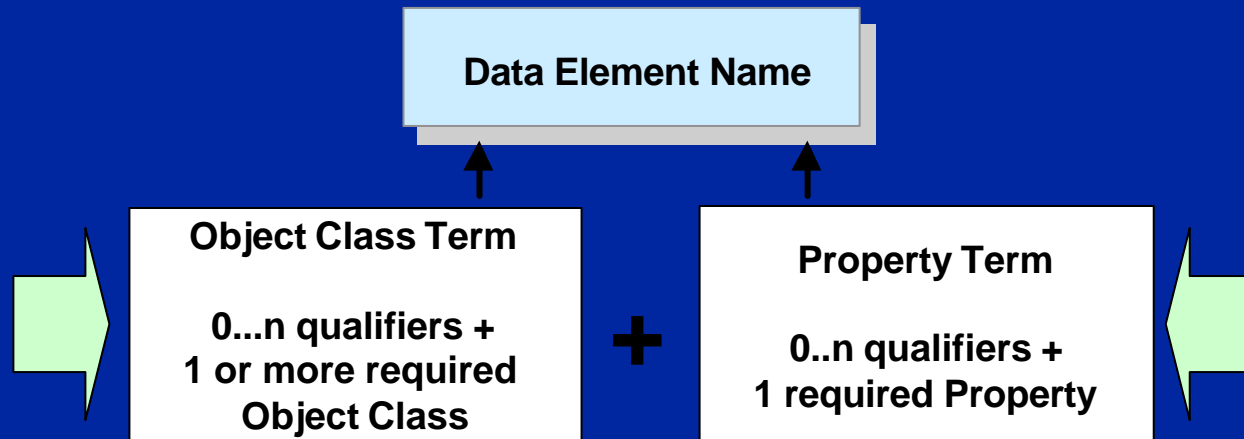
UDEF Built on Standards

An Instantiation of ISO 11179-5 Naming Convention and Supports ebXML

UDEF Object Class List

- Entity
- Document
- Enterprise
- Place
- Program
- Product
- Process
- Person
- Asset
- Law-Rule
- Environment
- Condition
- Liability
- Animal
- Plant
- Mineral
- Event

ISO 11179-5 Naming Convention



ebXML Property List

- Amount
- Code
- Date
- Date Time
- Graphic
- Identifier
- Indicator
- Measure
- Name
- Percent
- Picture
- Quantity
- Rate
- Text
- Time
- Value
- Sound
- Video

Example UDEF-Based Data Element Names

Document Abstract Text

Enterprise Name

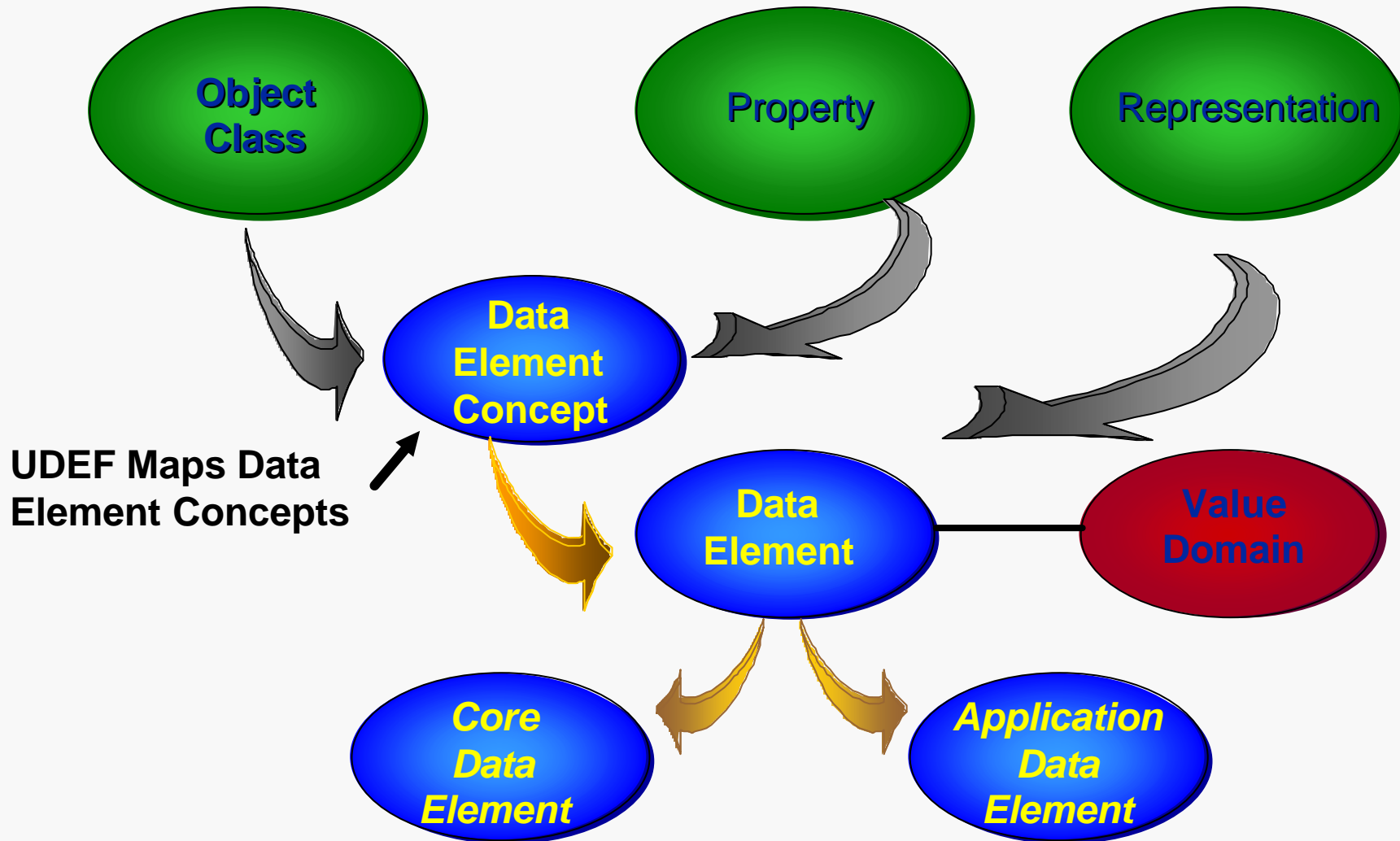
Product Price Amount

Product Scheduled Delivery Date

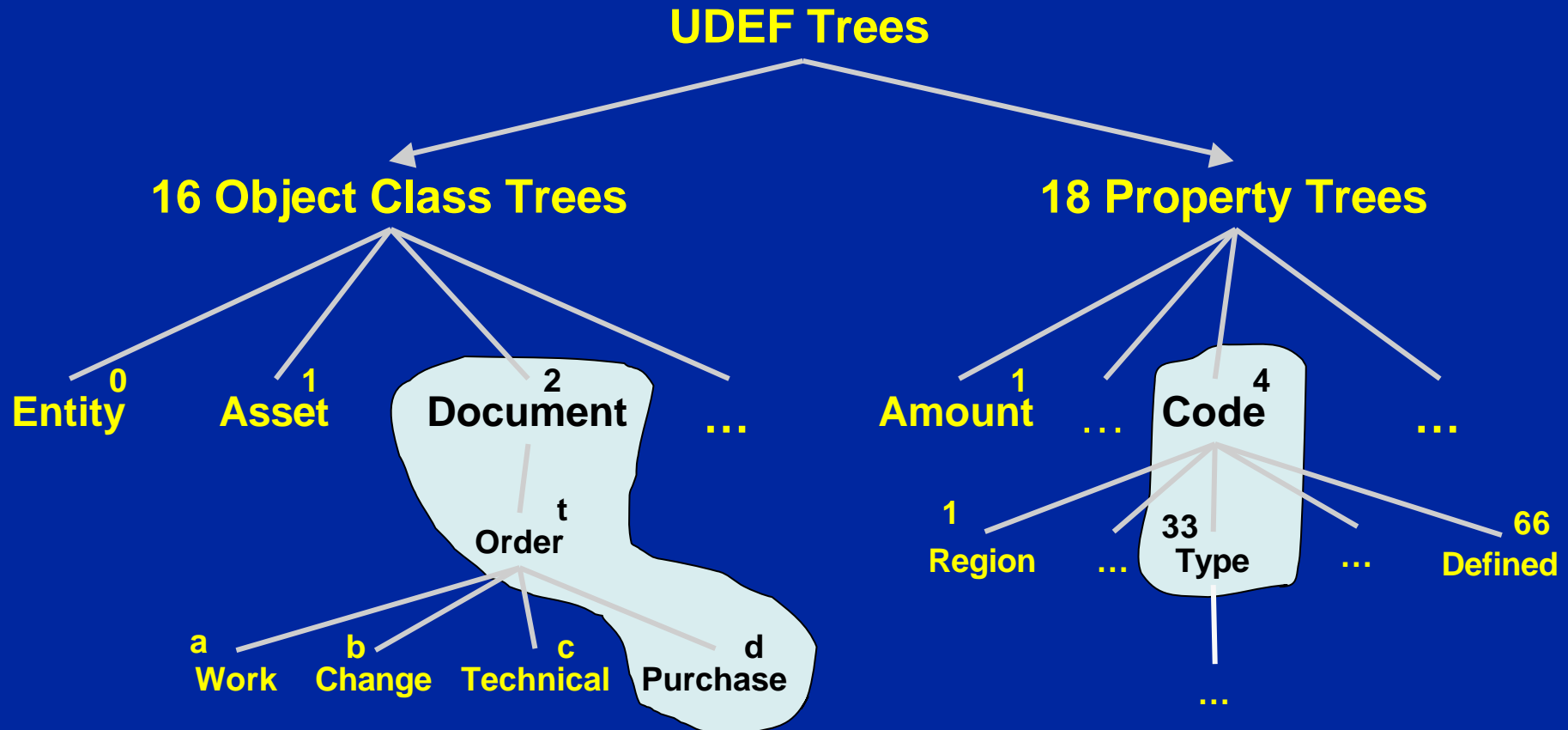
Engineering Design Process Cost Amount

UDEF names follow the rules of English – qualifiers precede the word they modify

ISO/IEC 11179 – Data Element Concept



Creating UDEF IDs



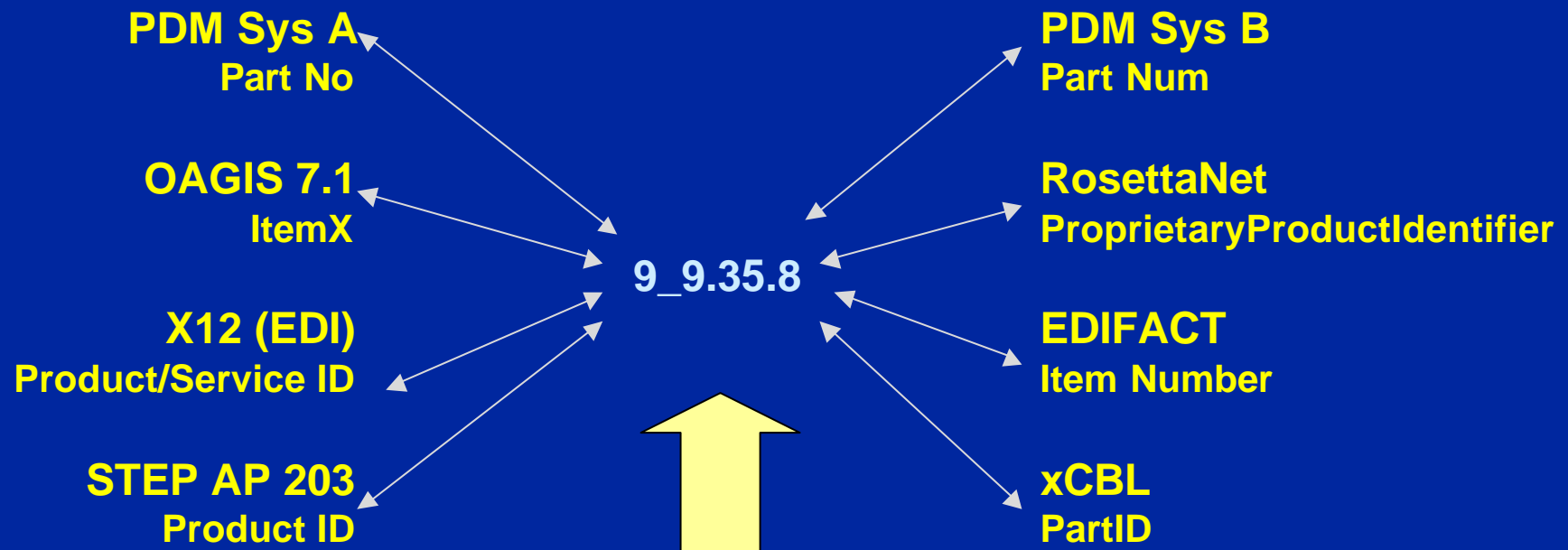
Purchase Order Document_Type Code has UDEF ID = d.t.2_33.4

See <http://www.udef.org/>

Example Mappings

Organizations cannot avoid multiple data standards

**** Need global semantics standard ****



UDEF Universal Identifier

Product(9)_Manufacturer(9).Assigned (35).Identifier(8)

N (N-1) mapping effort instead becomes a 2N mapping effort

Mapping Concepts to UDEF

1. Identify the applicable UDEF property word that characterizes the dominant attribute (property) of the data element concept. For example, Name, Identifier, Date, etc.
2. Identify the dominant UDEF object word that the dominant property (selected in step 1) is describing. For example, Person_Name, Product_Identifier, Document_Date, etc.
3. By reviewing the UDEF tree for the selected property identified in step 1, identify applicable qualifiers that are necessary to unambiguously describe the property word term. For example, Last Name
4. By reviewing the UDEF tree for the selected object identified in step 2, identify applicable qualifiers that are necessary to unambiguously describe the object word term. For example, Customer Person
5. Concatenate the object term and the property term to create a UDEF naming convention compliant name where it is recognized that the name may seem artificially long. For example, Customer Person_Last Name
6. Derive a structured ID based on the UDEF taxonomy that carries the UDEF inherited indexing scheme. For example <CustomerPersonLastName GUID="as.5_5.10">

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property

1. **Identify the applicable UDEF property word that characterizes the dominant attribute (property) of the data element concept.** For example, Name, Identifier, Date, etc.

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property
				Identifier

2. Identify the dominant UDEF object word that the dominant property (selected in step 1) is describing. For example, Person_Name, Product_Identifier,

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property
		Document		Identifier

3. By reviewing the UDEF tree for the selected property identified in step 1, identify applicable qualifiers that are necessary to unambiguously describe the property word term. For example, Last Name

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property
		Document	State Department Government Assigned	Identifier

4. By reviewing the UDEF tree for the selected object identified in step 2, identify applicable qualifiers that are necessary to unambiguously describe the object word term. For example, Customer Person

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property
	United States Passport	Document	State Department Government Assigned	Identifier

5. **Concatenate the object term and the property term to create a UDEF naming convention compliant name where it is recognized that the name may seem artificially long.** For example, Customer Person_Last Name

Homeland Security Typical Example

Data Element Concept Name

United States Passport Number

UDEF ID	Object Type or Role	Object Class	Property Type	Property
a.ck.2_2.13.35.8	United States Passport	Document	State Department Government Assigned	Identifier

6. **Derive a structured ID based on the UDEF taxonomy that carries the UDEF inherited indexing scheme.** For example <CustomerPersonLastName GUID="as.5_5.10">

Possible DHS-DoJ UDEF Example

UDEF-Based Common Vocabulary

Person Identifier	5_8
Person Last Name	5_5.10
Person First Name	5_4.10
Person Birth Date	5_51.6
Person Birth City Name	5_2.10.10

Local Police – Suspect Role

Suspect Id	aw.5_3.13.35.8
Suspect Name	aw.5_5.10 & aw.5_4.10
Suspect Date of Birth	aw.5_51.6
Suspect City of Birth	aw.5_2.10.10

Local Police – Criminal Role

Criminal Id	av.5_3.13.35.8
Criminal Name	av.5_5.10 & av.5_4.10
Criminal Date of Birth	av.5_51.6
Criminal City of Birth	av.5_2.10.10

UDEF Use with XML

- UDEF derived intelligent ID is the “key” to systems integration
- Mapping to UDEF is a rules-based approach to determine semantic equivalence

UDEF Name
Purchase Order Document_Identifier

UDEF ID = d.t.2_8

- **Optional UDEF IDs applied only when aligning ANY two systems for the first time (design time) – stored as an alias within each system**
 - Non-intrusive – does not require change to data element names
- **Leverage the UDEF Web Service to generate a gap analysis report**

<http://www.maphin.net:8080/UDEFServlet/UdefReport1>

Two different systems with need to exchange purchase order data

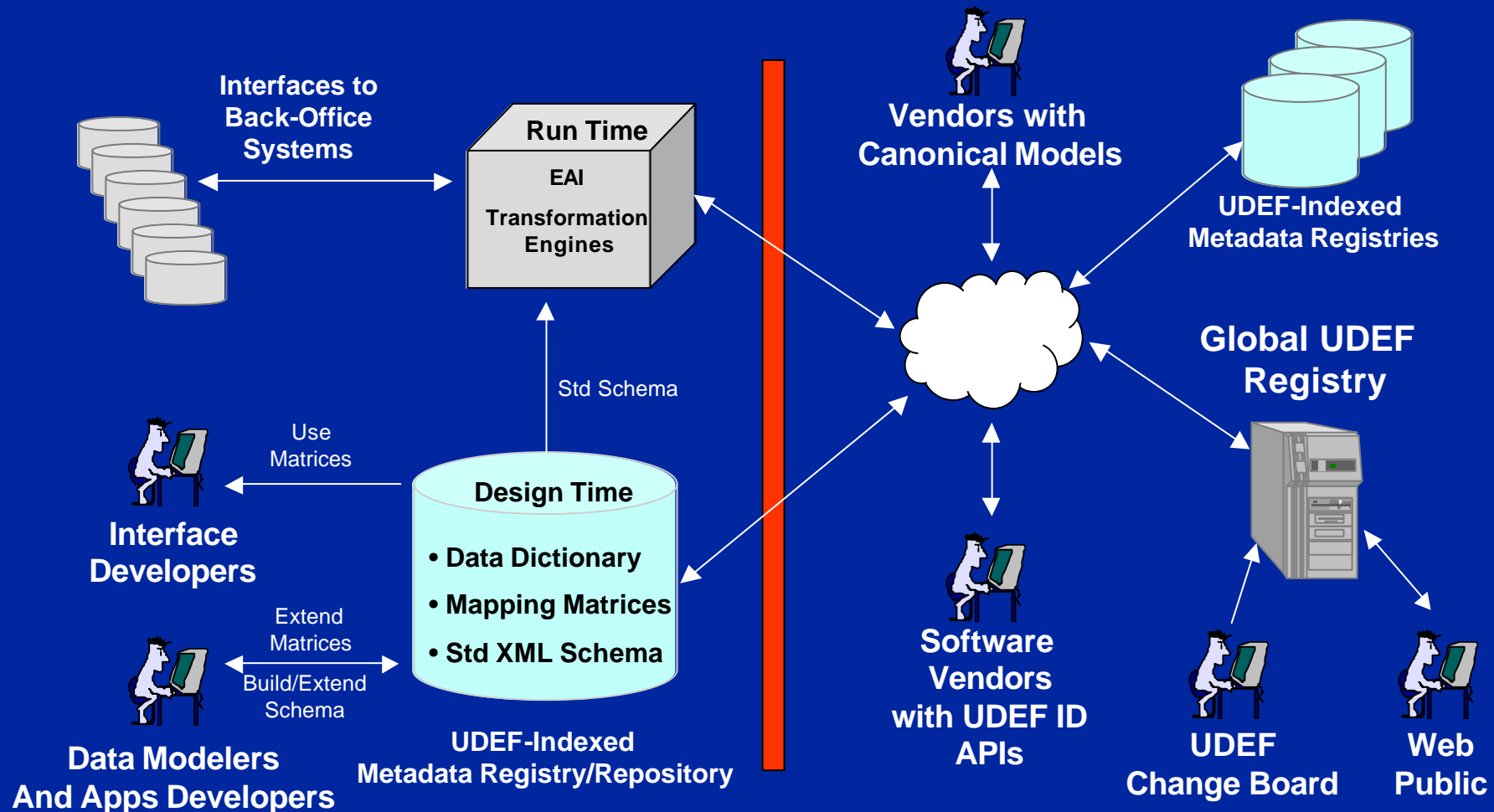
System A (xCBL 3.x Format)

```
<Order>
- <OrderHeader>
  - <OrderNumber>
    <BuyerOrderNumber GUID="d.t.2_8">ZT8923</BuyerOrderNumber>
  </OrderNumber>
  <OrderIssueDate GUID="d.t.2_5">19991108T23:59:59</OrderIssueDate>
```

System B (OAGIS 7.x Format)

```
<POID GUID="d.t.2_8">ZT8923</POID>
<POTYPE GUID="d.t.2_33.4">224</POTYPE>
<ACKREQUEST>1</ACKREQUEST>
<CONTRACTB GUID="e.2_8">94-A17</CONTRACTB>
<TAXEXMPT GUID="q.3_1.4.9.4">FOR RESALE</TAXEXMPT>
- <PARTNER>
  <NAME GUID="q.3_10" index="1">RUNNER UNLIMITED</NAME>
  <ONETIME>0</ONETIME>
  <PARTNRID GUID="q.3_8">A123</PARTNRID>
  <PARTNRTYPE GUID="a.at.3_1.37.4">SoldTo</PARTNRTYPE>
  <TAXID GUID="e.2_8">NT132765</TAXID>
```

A Metadata Managed Architecture



Centralized metadata registry/repository

- Enables reuse to reduce costs
- Encourages standardization

UDEF Current Status

As of May 2005, UDEF Officially Adopted by The Open Group

- Goal - establish the UDEF as the universally-used categorization system for data
 - Develop and maintain the UDEF as an open standard
 - Advocate and promote it
 - Put in place a technical infrastructure to support it
 - Implement a Registry for it, and
 - Set up education programs to train information professionals in its use.

Actual UDEF Trees are Viewable Online

<http://www.opengroup.org/udedefinfo/defs.htm>

UDEF Pilot Submitted to Federal Semantic Interoperability Community of Practice (SICoP)

- To enable manufacturing industry to support disaster response teams
- Gain visibility to the highest levels of government and possibly (??) obtain funding

Pilot Summary

Natural disaster response team shows up lacking batteries to operate GPS system and walkie-talkie for 200 search and rescue workers – need four hundred 9-volt batteries to even begin the search and rescue effort

- Goal – determine if resources might be available nearby within a manufacturer's inventory
- Use two UDEF tags (IDs) to locate available resources in a battery manufacturer's inventory about 5 miles from the response team command center – an ad hoc query since formal interface not previously defined
- Use UDEF tags to support semantic integration of disparate procurement applications that use different purchase order semantics

Leverage OASIS Emergency Management TC defined message wrappers

- Open Group UDEF Forum defines message content
- Two vendors currently involved – Unicorn and Safyre Solutions

For Additional Information

ISO/IEC 11179 – Specification and standardization of data elements

http://isotc.iso.ch/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm

OPENGROUP UDEF Forum Web Site

<http://www.opengroup.org/udf/>

UDEF.ORG Web Site

<http://www.undef.org/>

Dr. Chris Harding – c.harding@opengroup.org

Ron Schuldt – ron.l.schuldt@lmco.com