

*Guide*

## **Extending the Universal Data Element Framework (UDEF)**



Copyright © 2006, The Open Group

All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of the copyright owner.

It is fair use of this document for implementers to use the names, labels, etc. contained within the document. The intent of publication of the document is to encourage implementations of the document.

This document has not been verified for avoidance of possible third-party proprietary rights. In implementing this document, usual procedures to ensure the respect of possible third-party intellectual property rights should be followed.

Guide

**Extending the Universal Data Element Framework (UDEF)**

ISBN: 1-931624-65-8

Document Number: G064

Published by The Open Group, October 2006.

Comments relating to the material contained in this document may be submitted to:

The Open Group  
Thames Tower  
37-45 Station Road  
Reading  
Berkshire, RG1 1LX  
United Kingdom

or by electronic mail to:

[ogedit@opengroup.org](mailto:ogedit@opengroup.org)

# Contents

Contents.....	iii
Preface.....	iv
Trademarks.....	vi
Acknowledgements.....	vii
Referenced Documents.....	viii
1 Introduction.....	1
1.1 What is the UDEF?.....	1
1.2 The UDEF is Enterprise-Centric.....	3
1.3 Intended Audience.....	4
1.4 Guide to Using this Document.....	4
1.5 Terminology.....	4
2 Key Terminology.....	5
2.1 Data Element Concept.....	5
2.2 Object Class.....	6
2.3 Property.....	6
2.4 Qualifier Term.....	6
3 Key Principles for Extending the UDEF.....	7
3.1 Develop the Data Element Concept.....	7
3.2 Use Existing Qualifier Terms Where Possible.....	7
3.3 Select Qualifier Terms Words that Specialize the Word Above.....	7
3.4 Avoid Proprietary References.....	8
3.5 Verify the Syntax of the Submitted Data Element Concept.....	8
3.6 Qualifier Term Conventions.....	8
4 UDEF Extension Submittal and Approval Process.....	9
4.1 Create the Data Element Concept.....	9
4.2 Map Each Data Element Concept to the UDEF.....	10
4.3 If the UDEF is Adequate, Record the UDEF ID.....	10
4.4 If the UDEF is Inadequate, Use Chapter 3 of this Document.....	10
4.5 Submit Proposed UDEF Extensions Plus Data Element Concepts to The Open Group.....	11

# Preface

## The Open Group

The Open Group is a vendor-neutral and technology-neutral consortium, whose vision of Boundaryless Information Flow™ will enable access to integrated information within and between enterprises based on open standards and global interoperability. The Open Group works with customers, suppliers, consortia, and other standards bodies. Its role is to capture, understand, and address current and emerging requirements, establish policies, and share best practices; to facilitate interoperability, develop consensus, and evolve and integrate specifications and Open Source technologies; to offer a comprehensive set of services to enhance the operational efficiency of consortia; and to operate the industry's premier certification service, including UNIX® certification.

Further information on The Open Group is available at [www.opengroup.org](http://www.opengroup.org).

The Open Group has over 15 years' experience in developing and operating certification programs and has extensive experience developing and facilitating industry adoption of test suites used to validate conformance to an open standard or specification.

More information is available at [www.opengroup.org/certification](http://www.opengroup.org/certification).

The Open Group publishes a wide range of technical documentation, the main part of which is focused on development of Technical and Product Standards and Guides, but which also includes white papers, technical studies, branding and testing documentation, and business titles. Full details and a catalog are available at [www.opengroup.org/bookstore](http://www.opengroup.org/bookstore).

As with all *live* documents, Technical Standards and Specifications require revision to align with new developments and associated international standards. To distinguish between revised specifications which are fully backwards-compatible and those which are not:

- A new *Version* indicates there is no change to the definitive information contained in the previous publication of that title, but additions/extensions are included. As such, it *replaces* the previous publication.
- A new *Issue* indicates there is substantive change to the definitive information contained in the previous publication of that title, and there may also be additions/extensions. As such, both previous and new documents are maintained as current publications.

Readers should note that updates – in the form of Corrigenda – may apply to any publication. This information is published at [www.opengroup.org/corrigenda](http://www.opengroup.org/corrigenda).

## **This Document**

This Guide specifies key terminology and the process to extend the Universal Data Element Framework (UDEF) beyond its existing coverage of data element concepts. The UDEF is a global instantiation of ISO/IEC 11179-5, Information Technology, Metadata Registries (MDR) – Naming and Identification Principles.

Any organization or individual that wishes the UDEF to be extended must review and understand this document before attempting to submit proposed extensions to The Open Group.

## Trademarks

Boundaryless Information Flow<sup>™</sup> and TOGAF<sup>™</sup> are trademarks and Making Standards Work<sup>®</sup>, The Open Group<sup>®</sup>, and UNIX<sup>®</sup> are registered trademarks of The Open Group in the United States and other countries.

The Open Group acknowledges that there may be other brand, company, and product names used in this document that may be covered by trademark protection and advises the reader to verify them independently.

## Acknowledgements

The Open Group gratefully acknowledges the contribution of the following members of the committee in the development of this document:

- Arnold Van Overeem, Capgemini
- Ronald L. Schuldt, Lockheed Martin
- Dr. Chris Harding, The Open Group

## Referenced Documents

The following documents are referenced in this Guide:

### Normative

- ISO/IEC 11179-1, Information Technology – Metadata Registries (MDR) – Framework, 15 September 2004, available at:  
[isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf\\_Home/PubliclyAvailableStandards.htm](http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm)
- ISO/IEC 11179-5, Information Technology – Metadata Registries (MDR) – Naming and Identification Principles, 1 September 2005, available at:  
[isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf\\_Home/PubliclyAvailableStandards.htm](http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm)
- IETF RFC 2119, Key Words for use in RFCs to Indicate Requirement Levels, S. Bradner, March 1997, available at:  
[www.ietf.org/rfc/rfc2119.txt?number=2119](http://www.ietf.org/rfc/rfc2119.txt?number=2119)

### Non-Normative

- UDEF (current version, continually expanding), The Open Group UDEF Forum, go to:  
[www.opengroup.org/udefinfo/defs.htm](http://www.opengroup.org/udefinfo/defs.htm)
- ANSI X12, American National Standards Institute Committee X12: Standards for Electronic Data Interchange, available from the Data Interchange Standards Association at: [www.disa.org/bookstore/public/index.cfm](http://www.disa.org/bookstore/public/index.cfm)
- ISO/IEC 9735:2002, Electronic Data Interchange for Administration, Commerce, and Transport (EDIFACT) – Application-level Syntax Rules
- EIA 836, Configuration Management Data Exchange and Interoperability, Electronic Industries Alliance Standard 836, available from the Defense Collaboration Network/International Collaboration Network at: [www3.dcnicn.com](http://www3.dcnicn.com)



# 1 Introduction

---

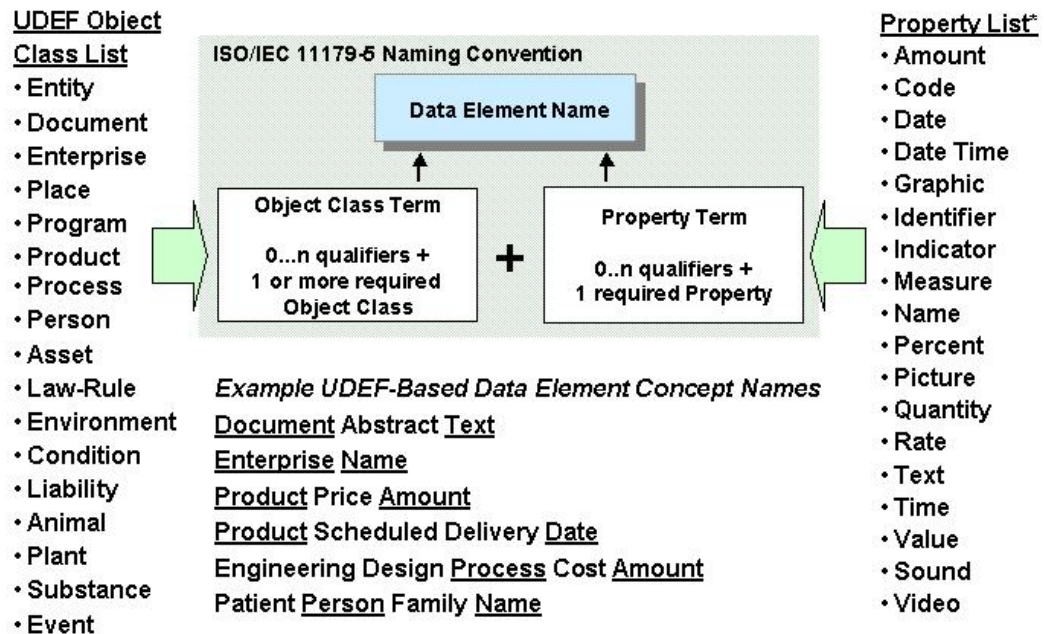
The Universal Data Element Framework (UDEF), which was developed by the US CALS ISG in the late 1980s and early 1990s, represents a small but potentially crucial aspect of information management: specifically, simplification of information management through consistent classification and assignment of a structured indexing identifier to the names (metadata) of data. The UDEF provides a conceptual naming convention framework with associated structured unique identifiers for indexing and aligning semantically equivalent concepts. It provides a means to associate different data element names (i.e., vocabulary terms) that semantically refer to the same concept, to a standard data element concept name provided by the framework that conforms to the relevant international standard on naming conventions, ISO/IEC 11179-5. A key definition provided in paragraph 3.3.9 of ISO/IEC 11179-1 is:

**“data element concept – concept that can be represented in the form of a data element, described independently of any particular representation”**

Interpreting the meaning of data element concepts is the essential first step of enabling semantic interoperability between disparate applications.

## 1.1 What is the UDEF?

The UDEF is a proposed global instantiation of the naming convention and unique identification requirements specified by ISO/IEC 11179-5. It also supports key portions of the ebXML Core Components Technical Specification (CCTS), published as ISO 15000-5. The UDEF controlled vocabulary includes 17 object class terms that are intrinsically intuitive to nearly everyone and 18 property terms that are the same as the allowable core components representation types specified by ISO 15000-5 in Tables 8-1 and 8-3. This is illustrated in Figure 1-1.



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

\* Based on Tables 8-1 and 8-3 in ISO 15000-5

**Figure 1-1: UDEF is based on International Standards**

The UDEF trees (one for each UDEF object class and property) provide taxonomies of sub-types and roles. Each UDEF object class and property is a top-level abstract concept super type. The words selected in each tree are abstract terms commonly used within domains that are commonly used by enterprises and are based on existing standards such as EDI ANSI X12 and UN/EDIFACT and EIA-836 or normalized data models from existing applications. By design, the UDEF trees have unlimited extensibility and can theoretically accommodate any data element concept within any domain relevant to any enterprise. **This document provides the foundation for extending the UDEF into domains that the existing UDEF does not address.**

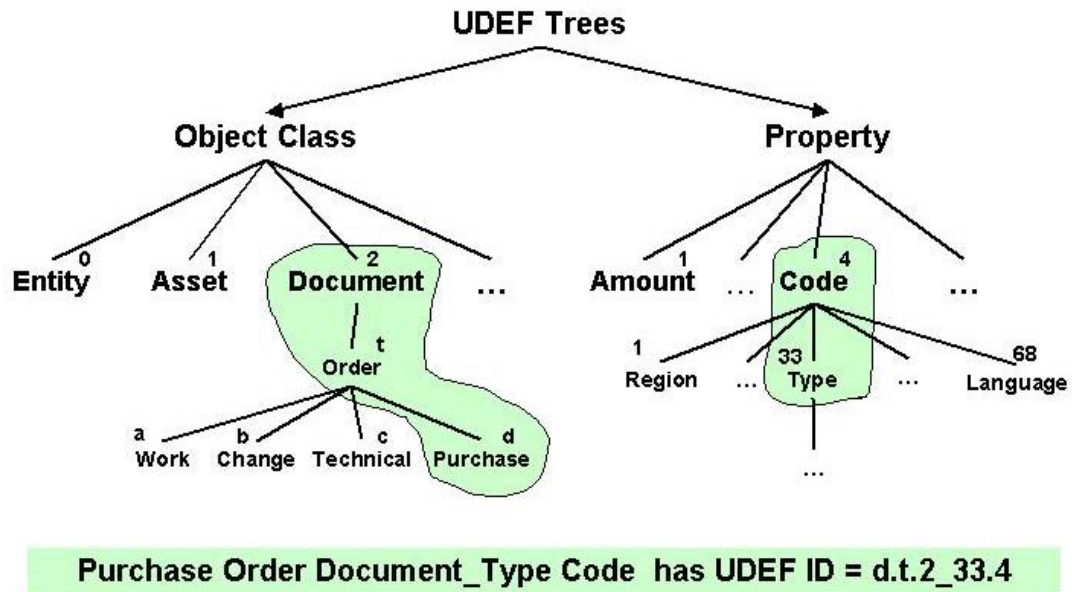


Figure 1-2: An Example Data Element Concept and its UDEF Identifier

## 1.2 The UDEF is Enterprise-Centric

The UDEF is enterprise-centric since the 17 top-level object classes are all defined in the context of an enterprise. Within the UDEF, an enterprise can be any organization – whether commercial or government or other. The UDEF is built around the basic premise that the enterprise establishes the critical cornerstone context for the semantic meaning of concepts that need to be exchanged between applications or systems either within the enterprise or between enterprises. The underlying assumption is that the enterprise manages data that is relevant to that enterprise. When a particular enterprise needs to share data with another enterprise, each needs to interpret the meaning of the data in the context of the other enterprise. Once an organization establishes the type of enterprise it is and the role it plays, the meaning (interpretation) of all of the other UDEF objects will fall into place in the context of that enterprise.

For example, the widely used and widely understood concepts of “product” and “asset” are tightly inter-connected and very dependent on the “enterprise” role context. The term “product” conveys an implied context that associates the term with an enterprise that performs a product producer or product seller role. The term “asset” conveys an implied asset possessor or asset buyer role by the enterprise. The product of one enterprise could be the asset of another enterprise. Tax laws and standard accounting practices re-enforce this distinction even though the two terms are addressing the same object. Each enterprise (product producer and asset owner) typically manages data that is common to the other as well as data that is unique to each. For example, an asset purchased by the owner enterprise may have a depreciation schedule that would be irrelevant to the product producer enterprise. Yet the two enterprises would likely share some common data such as the name, identifier, dimensions, etc. of the product/asset.

## 1.3 Intended Audience

The intended audience of this document includes non-exhaustively:

- Architects and developers designing, identifying, or developing a system who have decided to implement the UDEF and need to extend the UDEF to accommodate data element concepts that are not currently within the UDEF
- Standards architects and analysts developing specifications for data standards within a given domain who have decided to cross-reference their standard with the UDEF and need to extend the UDEF to accommodate data element concepts that are not currently within the UDEF

## 1.4 Guide to Using this Document

Readers are encouraged to read this document in its entirety. Concepts are presented in an order that the authors hope will promote rapid understanding:

- This chapter introduces the UDEF concept and the basic UDEF architecture.
- Chapter 2 introduces the reader to key terminology that must be understood if the reader intends to submit proposed extensions to the UDEF.
- Chapter 3 introduces the reader to the key principles that must be applied when the reader submits proposed extensions to the UDEF.
- Chapter 4 introduces the reader to the process for submitting proposed extensions to the UDEF.

## 1.5 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 IETF RFC 2119.

## 2 Key Terminology

---

This chapter introduces the reader to key terminology that must be understood if the reader intends to submit proposed extensions to the UDEF.

### 2.1 Data Element Concept

A **Data Element Concept** as defined by ISO/IEC 11179 is a “**concept** that can be represented as a **data element**, described independently of any particular representation”.

An example data element concept used by many enterprises is “product scheduled delivery date”. There are many possible representation forms for “date” such as July 5, 2006 or 20060705 or 07/05/2006, etc. The semantic meaning of the data element concept remains the same regardless of the representation form. The UDEF provides a global framework for aligning data element concepts. When interfacing disparate applications, one of the initial design-time steps is to determine whether the semantics are the same prior to determining a suitable transformation for those data element concepts that have different representation forms.

ISO/IEC 11179 uses a diagram (see Figure 2-1) to further describe a data element concept and to compare it to a data element. The diagram highlights the fact that a data element concept includes one or more object classes and only one property. For each data element concept there are one or more possible data elements since there are potentially multiple ways of representing a given data element concept.

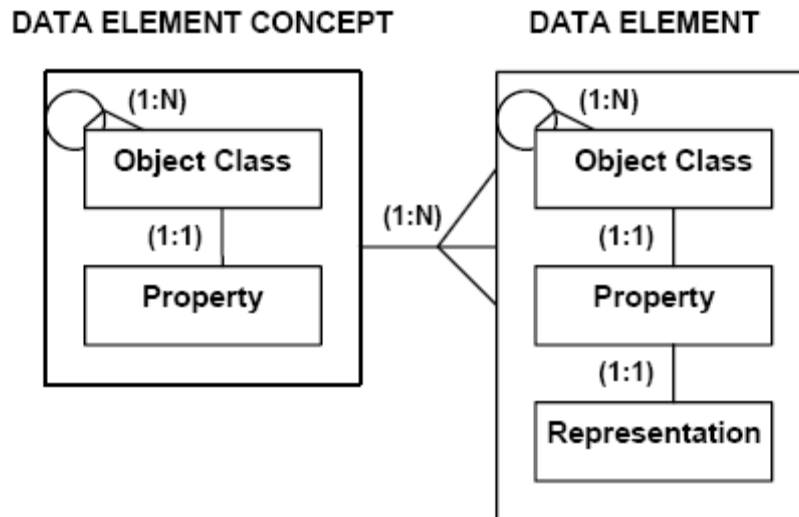


Figure 2-1: Metamodel of Data Element Concepts and Data Elements

## 2.2 Object Class

An **Object Class** as defined by ISO/IEC 11179 as a “set of ideas, abstractions, or things in the real world that are identified with explicit boundaries and meaning and whose properties and behavior follow the same rules”.

Within the context of any enterprise and the data that the enterprise needs to manage, a universal set of object classes includes the 17 identified in the UDEF Object Class List in Figure 1-1. If the list of 17 is inadequate to accommodate all domains of all enterprises, then it is possible to extend the UDEF with one or more additional object classes.

## 2.3 Property

A **Property** as defined by ISO/IEC 11179 as a “**characteristic** common to all members of an **object class**”.

Within the context of the data that the enterprise needs to manage, a universal set of properties includes the 18 identified in the UDEF Property List in Figure 1-1. The list and their definitions are derived from Tables 8-1 and 8-3 in ISO 15000-5 (same as the ebXML Core Components Technical Specification).

## 2.4 Qualifier Term

A **Qualifier Term** as defined by ISO/IEC 11179 as “word or words that differentiate a concept”.

Within the context of the data that the enterprise needs to manage, qualifier terms enable sub-type and role specialization of the universal set of UDEF object classes and properties. The primary approach for extending the UDEF is by identifying suitable qualifier terms for the UDEF object class and/or the UDEF property for a given data element concept – assuming that the data element concept does not already exist within the UDEF.

## 3 Key Principles for Extending the UDEF

---

This chapter identifies the key principles that shall be applied before submitting proposed extensions to the UDEF.

### 3.1 Develop the Data Element Concept

The submitter shall develop the applicable *data element concept* based on the existing UDEF *object classes* and *properties* and *qualifier terms* (as applicable) in the context of the affected enterprise. If the existing UDEF object classes are inadequate for the data element concept, then the submitter should propose a new UDEF object class as a part of the submittal.

Note: Although not impossible, it is unlikely that new UDEF object classes will be approved. In nearly all cases, the existing UDEF object classes will be adequate – simply requiring new qualifier terms. In the unlikely event of a new object class or property being needed, its name and description must avoid the use of terms that are specific to particular products or that might give commercial advantage to any individual or organization.

In nearly all cases, the proposed extension to the UDEF will identify new qualifier term extensions – either to the UDEF object class and possibly portions of existing qualifier terms and/or to the UDEF property and possibly portions of existing qualifier terms.

The complete data element concept with proposed new UDEF extensions shall be included in the submittal. If built properly (following the UDEF naming convention), the submitted data element concept should be self-explanatory and will not require a separate definition. The submitted data element concept shall also include the enterprise context.

### 3.2 Use Existing Qualifier Terms Where Possible

The submitter shall use existing UDEF object class qualifier terms and UDEF property qualifier terms where possible to minimize the chances of semantic redundancy within the existing UDEF.

### 3.3 Select Qualifier Terms Words that Specialize the Word Above

The submitter shall select qualifier term words that are a specialization (sub-type or role) of the word above.

### **3.4 Avoid Proprietary References**

The submitter shall select qualifier term words that do not include trademarks, are not product-specific, and do not in any way give any commercial advantage to any individual or organization.

### **3.5 Verify the Syntax of the Submitted Data Element Concept**

Prior to submitting a data element concept, the submitter shall verify that the UDEF property word is last and any qualifier terms modify the preceding word. Similarly, the submitter shall verify that any qualifier terms to the UDEF object class modify the preceding word. When the object class string (object class plus any qualifiers) is combined with the property string (property plus any qualifiers), the submitter shall verify that the resultant data element concept accurately describes the intended meaning within the context of the affected enterprise.

### **3.6 Qualifier Term Conventions**

The following conventions shall be applied to all proposed qualifier term extensions to the UDEF.

#### **Capitalize first letter of each word**

The first letter of each word in a proposed qualifier term extension to the UDEF shall be capitalized.

#### **Hyphenate multi-word extensions that are at the same level**

In many cases it is not practical for the sake of re-usability to extend the UDEF by placing each qualifier word at a different level. If a single word (qualifier term) can stand alone and conveys meaning when combined with the words above, then that is the preferred approach. However, in some cases it is not practical and in those cases the multi-word extension shall be hyphenated between each word at the same level.

#### **Avoid conjunctions**

The submitter shall avoid the use of conjunctions such as “of, and, or”. For example, rather than “Color-Of-Hair” the proposed UDEF extension submittal could be “Hair-Color” if the two words are at the same level or “Hair.Color” if the two words within the qualifier term are at different levels.

#### **Possessive apostrophe**

Since the UDEF source is in XML format and since XML does not accept an apostrophe, the convention for expressing an apostrophe within the UDEF shall be “Word-s”. For example, to express “Seller’s” within UDEF, the submitter shall use “Seller-s” instead.



## 4 UDEF Extension Submittal and Approval Process

This non-normative chapter identifies the process for submitting a proposed UDEF extension and the general approval procedure. The process is summarized in Figure 4-1.

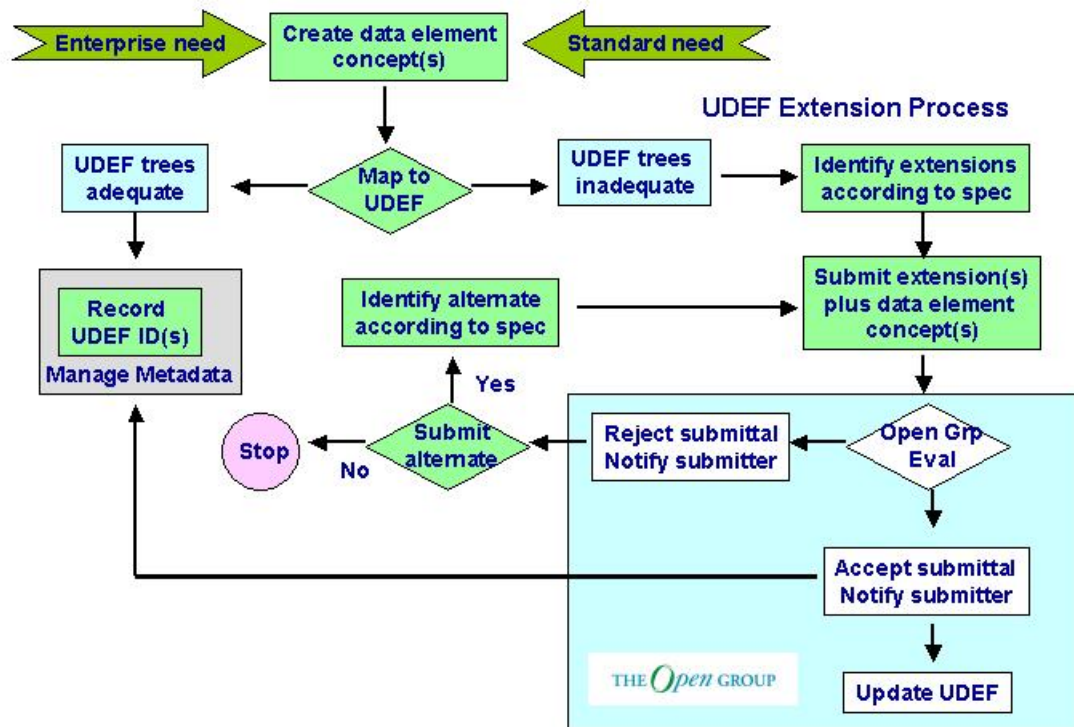


Figure 4-1: UDEF Extension Submittal Process

The following sub-sections provide additional details and examples for each major step in the process.

### 4.1 Create the Data Element Concept

Each enterprise needs to manage data relevant to the enterprise. The challenge is to establish a common vocabulary that conveys shared meaning across all users (people and system applications) of the data. Recognized experts within a domain typically define the terminology used within a given enterprise. For example, those in the engineering discipline commonly use terminology that is taught in engineering schools. However, most applications used within an enterprise do not have a common vocabulary foundation.

If a given application is expected to share data with one or more other applications, then a key first step is to identify the shared data element concepts between the applications. The lack of a shared vocabulary between the applications can be overcome by applying the UDEF to each data element concept that needs to be shared between the applications.

## 4.2 Map Each Data Element Concept to the UDEF

There are six basic steps that should be followed when mapping data element concepts to the 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, *Family 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\_Family Name*.
6. Derive a structured ID based on the UDEF taxonomy that carries the UDEF inherited indexing scheme; for example, `<CustomerPersonFamilyNameUDEFID="as.5_11.10">`.

## 4.3 If the UDEF is Adequate, Record the UDEF ID

To leverage the advantages that the UDEF can provide, the UDEF ID should as a minimum be recorded as an optional alias for the data element concept. If the enterprise has multiple applications, then a metadata registry should capture the UDEF name and ID for each data element concept that is shared between applications. The UDEF name and ID pair provides a simple but effective indexing mechanism for discovering re-usable data element concepts across the enterprise.

## 4.4 If the UDEF is Inadequate, Use Chapter 3 of this Document

To extend the UDEF, develop a data element concept with proposed UDEF extensions that adheres to the requirements specified in Chapter 3 of this document. Be certain that the enterprise context for the data element concept is also established.

## **4.5 Submit Proposed UDEF Extensions Plus Data Element Concepts to The Open Group**

Submit proposed UDEF extensions and associated data element concepts and the enterprise context to the email address identified on The Open Group UDEF Forum website ([www.opengroup.org/udef](http://www.opengroup.org/udef)). By making a submission, the submitter agrees to the terms of the UDEF Contribution Agreement that is published on The Open Group website.

The submittal procedure is subject to change in future versions of this document.