



Internal Revenue Service
Information Technology

A large, semi-transparent American flag is positioned in the background, waving from left to right. The stars and stripes are clearly visible, though slightly faded due to the transparency.

Canonical Model Management Forum
IRS Transition to a Common XML Vocabulary

John Triplett
March 8, 2010

Teaming for Results

Who we are:

IRS Enterprise Data Management Mission

Define an enterprise-wide data environment to more easily and efficiently organize, identify, share, reuse and correlate data that enables the business to consume information and maximize the value to the agency.

Activities under Enterprise Data Management can be bundled into four major areas:

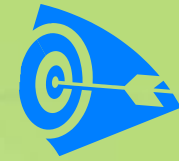
- Strategic Planning and Alignment
- Quality Management
- Architecture Management
- **Transparency Management**

Transparency Management

- Implement and maintain a registry and repository of XML schema components and related component assets.
- Develop and maintain IRS XML Naming and Design Rules (NDR).
- Develop and maintain a controlled vocabulary of IRS terms for common data interchange.
- Implement a metadata strategy which establishes a framework of content, definition and standards that document data resources and artifacts for business and technical use.
- Expose the data structures to business users to help them understand what is available.

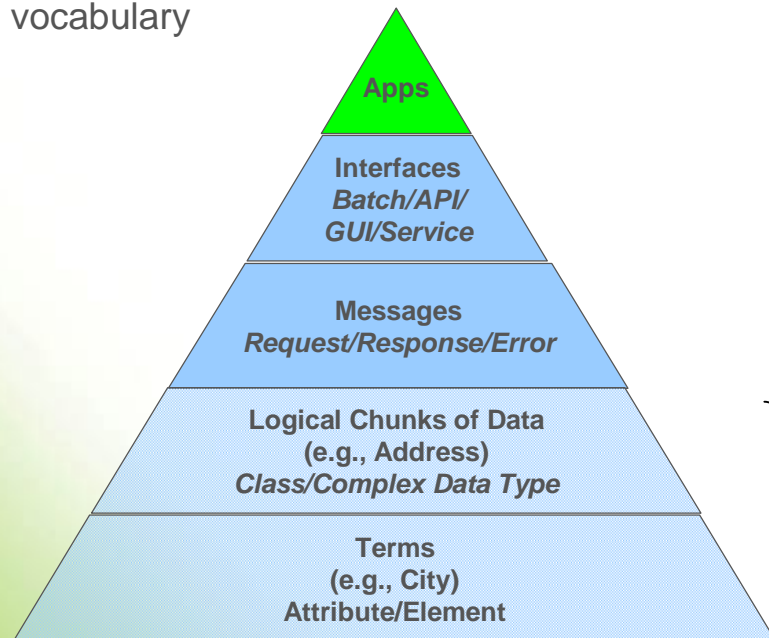


What we hope to accomplish



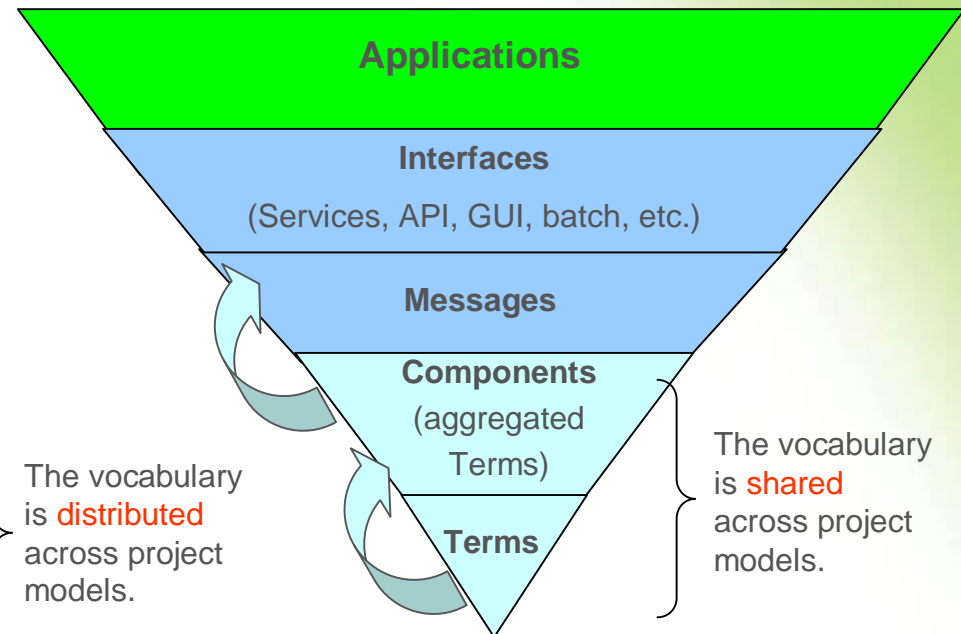
TODAY

Large vocabulary with numerous **redundant** terms, message components, messages, and interfaces. A small set of business applications results in a large distributed vocabulary



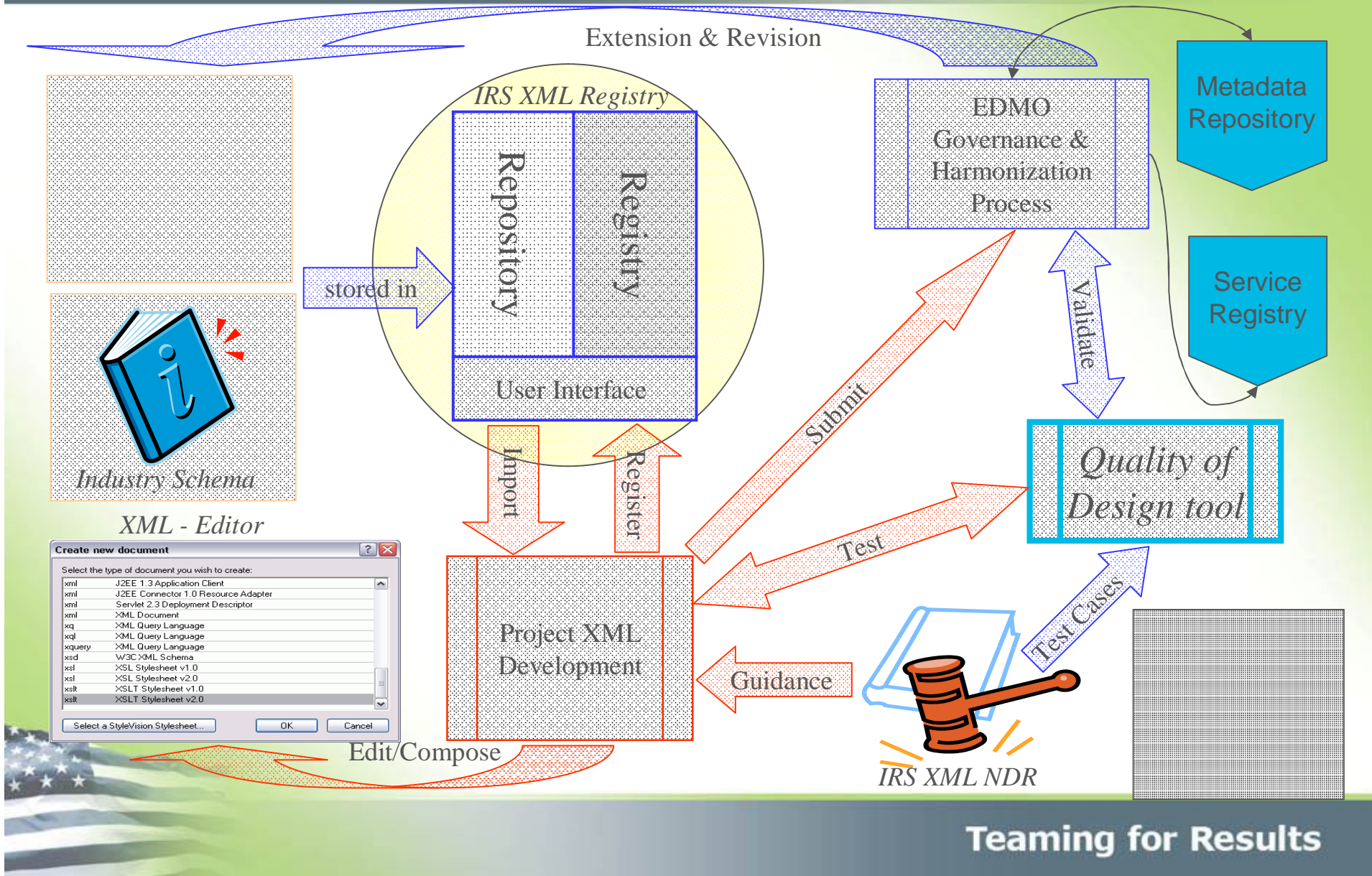
TOMORROW

Each layer is assembled from components at the lower levels by following the Naming and Design Rules and discovering re-usable components.



WHY: A controlled vocabulary supports a broad portfolio of business applications with a shorter time to market

How this will be accomplished



Where the Common Vocabulary is used

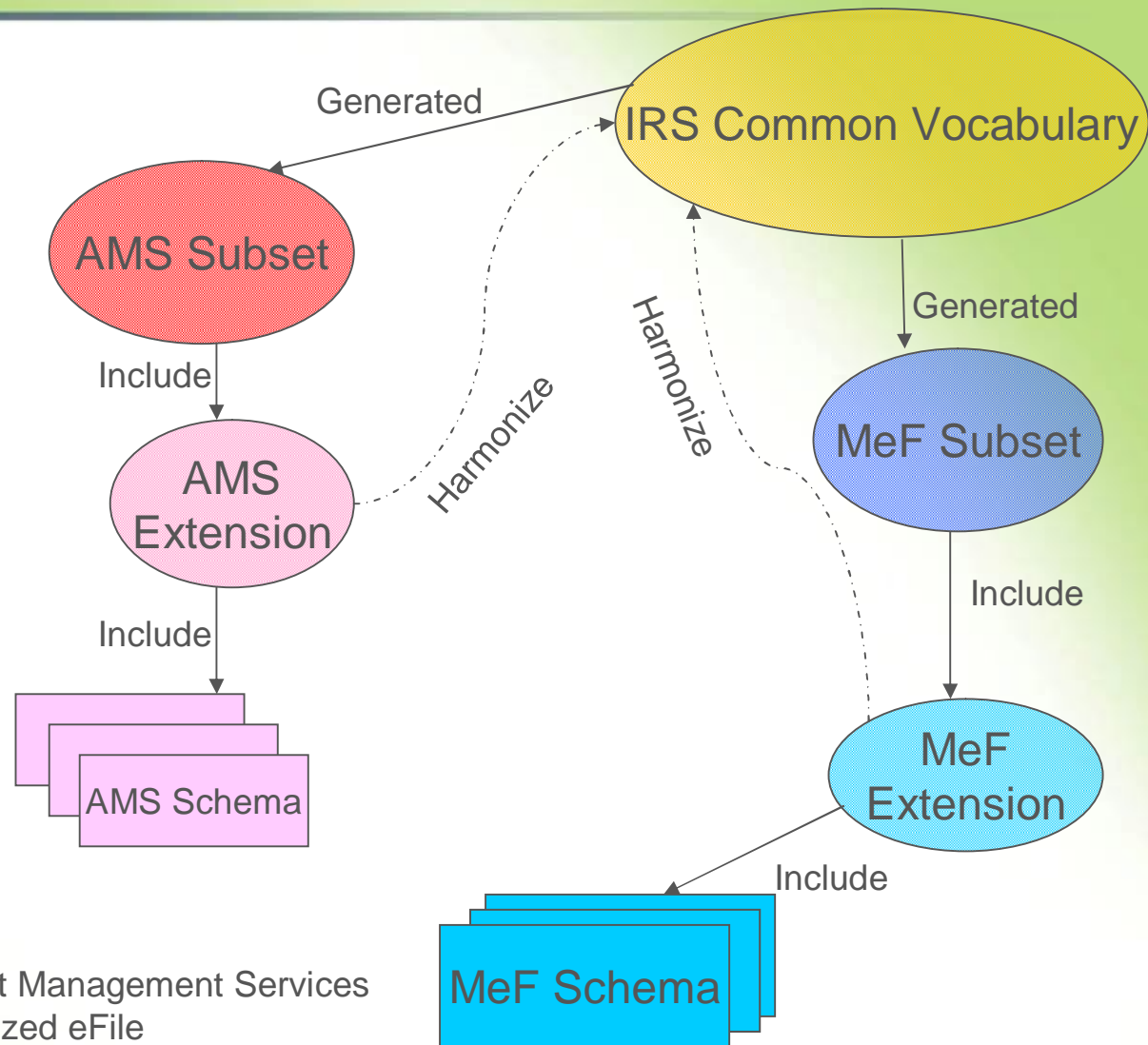
The IRS XML Common Vocabulary is the approved set of components that will be used by projects for their XML implementations.

The Common Vocabulary encompasses the entire IRS XML vocabulary and is quite large; therefore, projects will only need to use a subset of this vocabulary.

Implementation specific requirements not incorporated into the Common Vocabulary can be managed as project extensions to the library.

This approach attempts to strike a balance between the conflicting needs for project autonomy and enterprise data management (governance)

AMS: Account Management Services
MeF: Modernized eFile

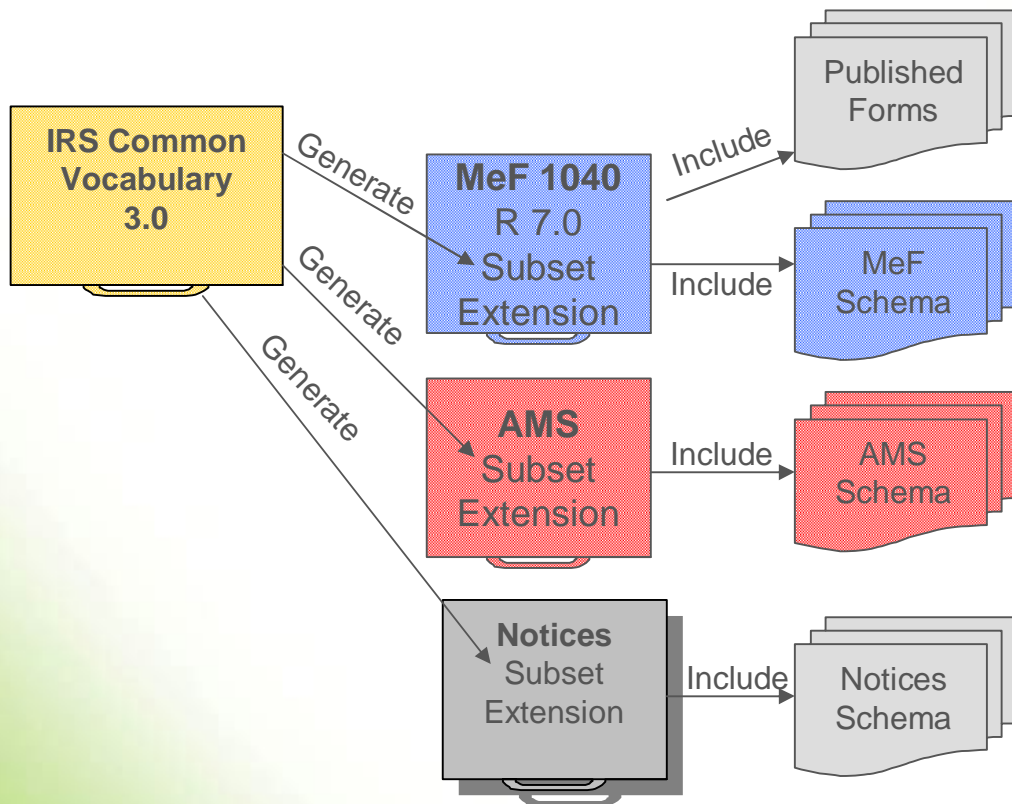


igniteXML: Where does it fit?

- IRS XML Strategy
 - XML Standards
 - XML Vocabulary
 - XML Tools
- igniteXML is a content management system optimized for the management of XML schema content. That is, igniteXML provides for the publication, subscription, association, search, discovery, versioning, and analysis of XML schema components.
- igniteXML is an XML **tool** that will make it easier for projects to deploy XML components built from the IRS XML Common **Vocabulary** in accordance with IRS XML **standards**.



igniteXML Content Organization

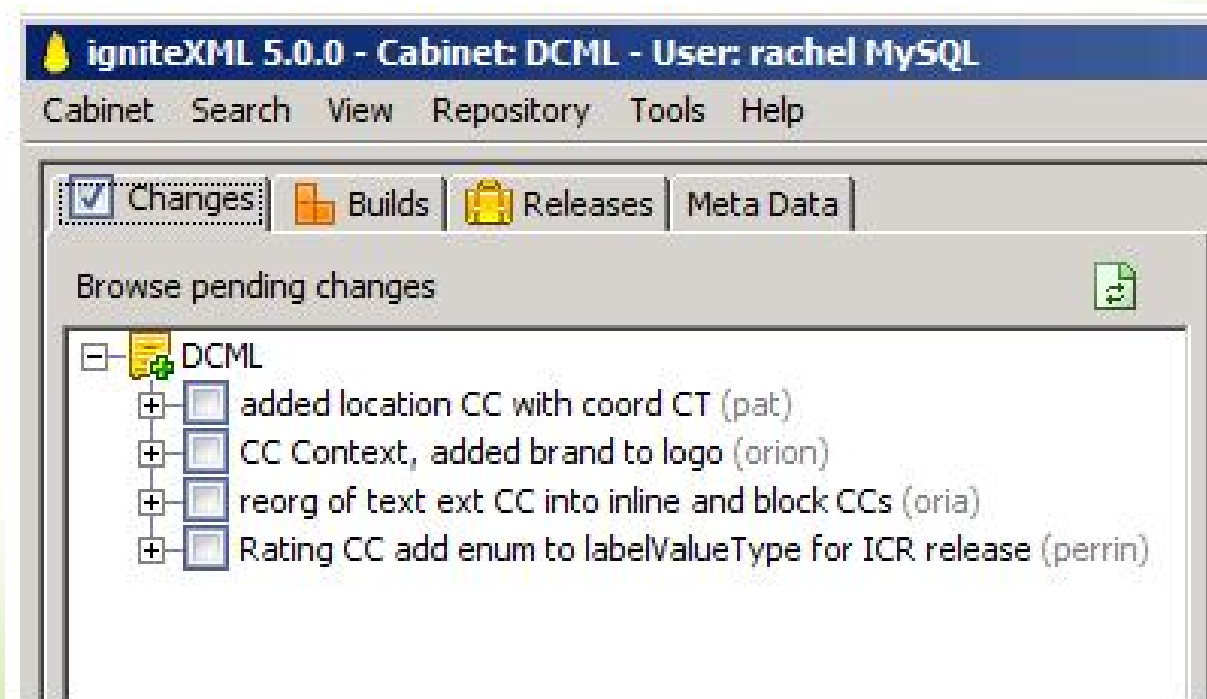


- Schema objects are constructed from existing objects in igniteXML, maintaining often complex relationships between objects so that schema can be created from high-level types, and the tool can automatically bring in the subset of data required to ensure schema is complete, e.g. from a reference model.
- igniteXML provides version control, impact analysis and build mechanism as part of a governance piece which can be extended to integrate with external Registry and Repository and Policy validation suites. igniteXML provides mechanisms to segregate data into different Cabinets through which access is controlled through Users, Groups and Roles.

igniteXML uniquely manages schema by decomposing schema into smaller re-usable business concepts (Core-Components) that can be used as building blocks to produce new consistent schema typically required to support XML messages in SOA.

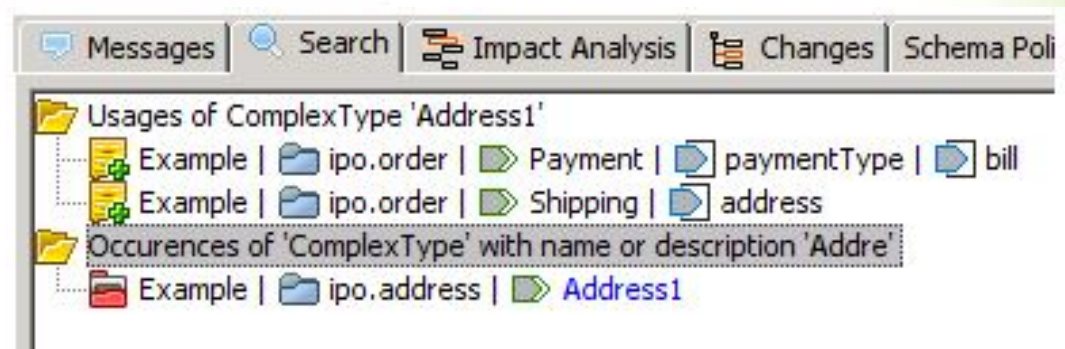
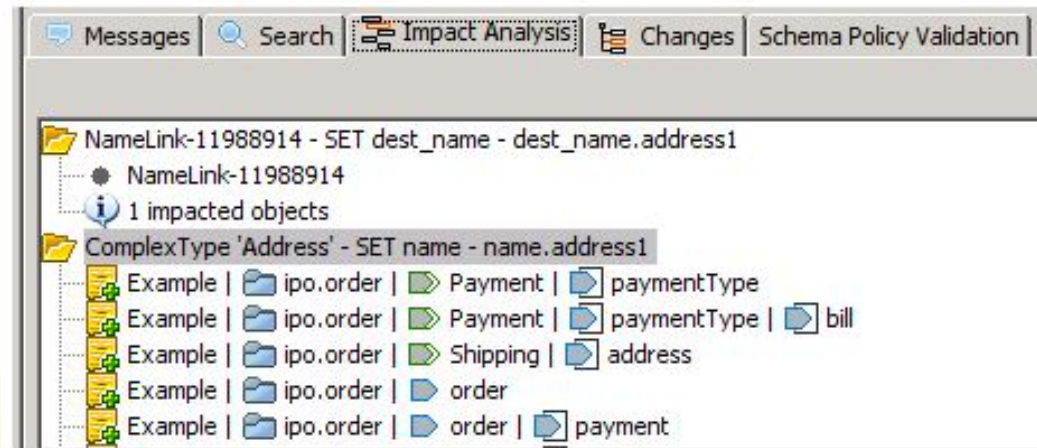
Collaboration

Multiple developers make changes through their personal Workspaces: Changes are checked into the Cabinet and merged into a Build by the Cabinet Administrator, where any conflicts are resolved.



Impact Analysis

As every object relationship is modeled, igniteXML can determine the impact of a change:



Subset Schema Generation

Schema that use only parts of a Reference schema only require a subset of the Reference schema to be consistent. IgniteXML can create a subset Reference schema so that schema size and complexity are substantially reduced.



Schema generated in ICS are composed of associated subset reference schema containing only the contents of the selected elements.

Schema Documentation

igniteXML and ICS produce schema documentation with details of Build versions and included Changes. Schema may also include attachments and instance documents that are stored in igniteXML's database.

XML Schema Documentation

Table of Contents

- [Schema Document Properties](#)
- [Global Declarations](#)
 - [Element: AdjustedGrossIncomeAmt](#)
 - [Element: DocumentTypeCd](#)
 - [Element: EducationCreditAmt](#)
 - [Element: IndividualReturnIRADistributionAmt](#)
 - [Element: IRADeductionAmt](#)
 - [Element: PartyBirthDt](#)
 - [Element: PensionsAnnuitiesAmt](#)
 - [Element: PersonFirstName](#)
 - [Element: PersonLastName](#)
 - [Element: SSN](#)
 - [Element: TaxExemptInterestAmt](#)
 - [Element: TotalExemptionsCnt](#)
 - [Element: TotalTaxesPaidAmt](#)