

ISO/IEC JTC 1/SC 32 N 0444

Date: 2000-02-02

REPLACES: --

<p style="text-align: center;">ISO/IEC JTC 1/SC 32</p> <p style="text-align: center;">Data Management and Interchange</p> <p style="text-align: center;">Secretariat: United States of America (ANSI) Administered by Pacific Northwest National Laboratory on behalf of ANSI</p>

DOCUMENT TYPE	Other Document (Open)
TITLE	Tutorial on Metadata Registries
SOURCE	Bruce Bargmeyer
PROJECT NUMBER	
STATUS	Presented at the SC 32 Tutorial 2000-01-24
REFERENCES	
ACTION ID.	FYI
REQUESTED ACTION	
DUE DATE	
Number of Pages	26
LANGUAGE USED	English
DISTRIBUTION	P & L Members SC Chair WG Conveners and Secretaries

Douglas Mann, Secretariat, ISO/IEC JTC 1/SC 32

Pacific Northwest National Laboratory *, 901 D Street, SW., Suite 900, Washington, DC, 20024-2115, United States of America

Telephone: +1 703 575 2114; Facsimile: +1 703 681 9180; E-mail: MannD@battelle.org

available from the JTC 1/SC 32 WebSite <http://bwonotes5.wdc.pnl.gov/SC32/JTC1SC32.nsf>

*Pacific Northwest National Laboratory (PNL) administers the ISO/IEC JTC 1/SC 32 Secretariat on behalf of ANSI

Semantics Management - Terminology Extensions for 11179



SC 32

Tutorial on Metadata Registries

Bruce Bargmeyer

Chair, SC 32 – Data Management and Interchange

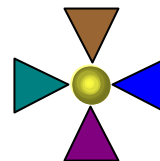


U.S. Environmental Protection Agency

Telephone: (202) 260-5306

bargmeyer.bruce@epa.gov

WWW URL: <http://www.nist.gov/itl/div897/staff/bargmeyer>



January 19, 2000

SDC-0002-021-JE-2030

ISO 11179, Parts 1-6

- Part 1: *Framework for the Specification and Standardization of Data Elements*
- Part 2: *Classification for Data Elements*
- Part 3: *Basic Attributes of Data Elements*
- Part 4: *Rules and Guidelines for the Formulation of Data Definitions*
- Part 5: *Naming and Identification Principles for Data Elements*
- Part 6: *Registration of Data Elements*
- DTR 15452 *Specification of Data Value Domains*

SDC-0002-021-JE-2030

Part 4: Data Definitions

- **Verification Method Code:**
The code that represents the process used to verify the latitude and longitude coordinates.
- **Mailing Address Post Office Box:**
The number of the post office box where mail for the addressee is delivered.
- **International Postal Code:**
The code that represents a postal zone specific to a country.

SDC-0002-021-JE-2030

Data Element Concept

Name: Country Identifiers
Context:
Definition:
Unique ID: 5769
Conceptual Domain: →
Maintenance Org.:
Steward:
Classification:
Registration Authority:
Others

Afghanistan
 Belgium
 China
 Denmark
 Egypt
 France
 Germany

Data Elements

Afghanistan
Belgium
China
Denmark
Egypt
France
Germany

ISO 3166
English Name

AFG
BEL
CHN
DNK
EGY
FRA
DEU

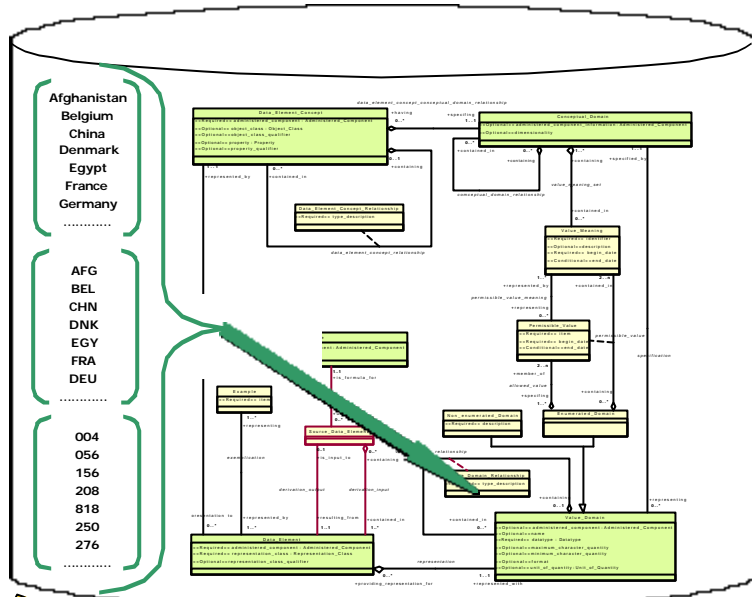
ISO 3166
3-Alpha Code

004
056
156
208
818
250
276

ISO 3166
3-Numeric Code

SDC-0002-021-JE-2030

- In order to reduce costs associated with managing metadata, we want to enable interchange of metadata including terminology between metadata registries.
- Organizations that are responsible for particular terminology and data elements can propagate these changes to other 11179 metadata registries.



11179 Metadata Registry

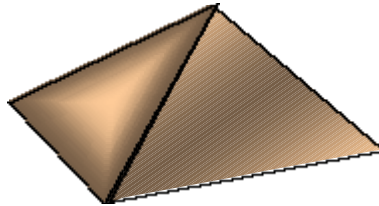
SDC-0002-021-JE-2030

Elements of Terminology

trout
Salmo trutta
brown trout
truite

Term

Concept



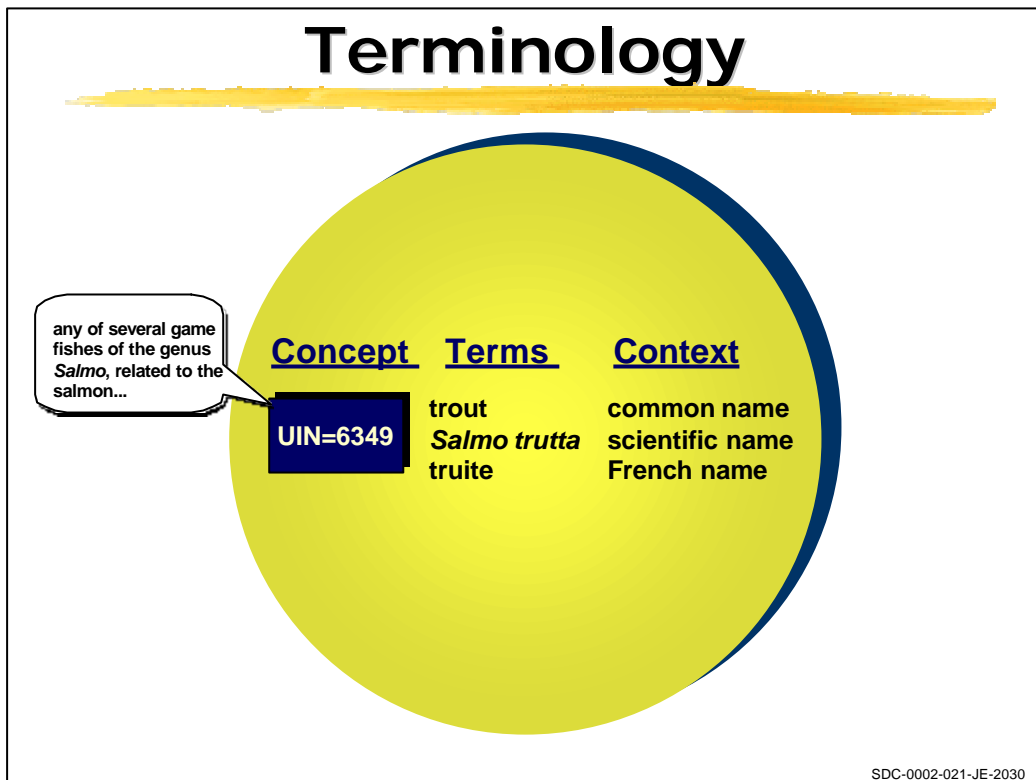
Object

Definition

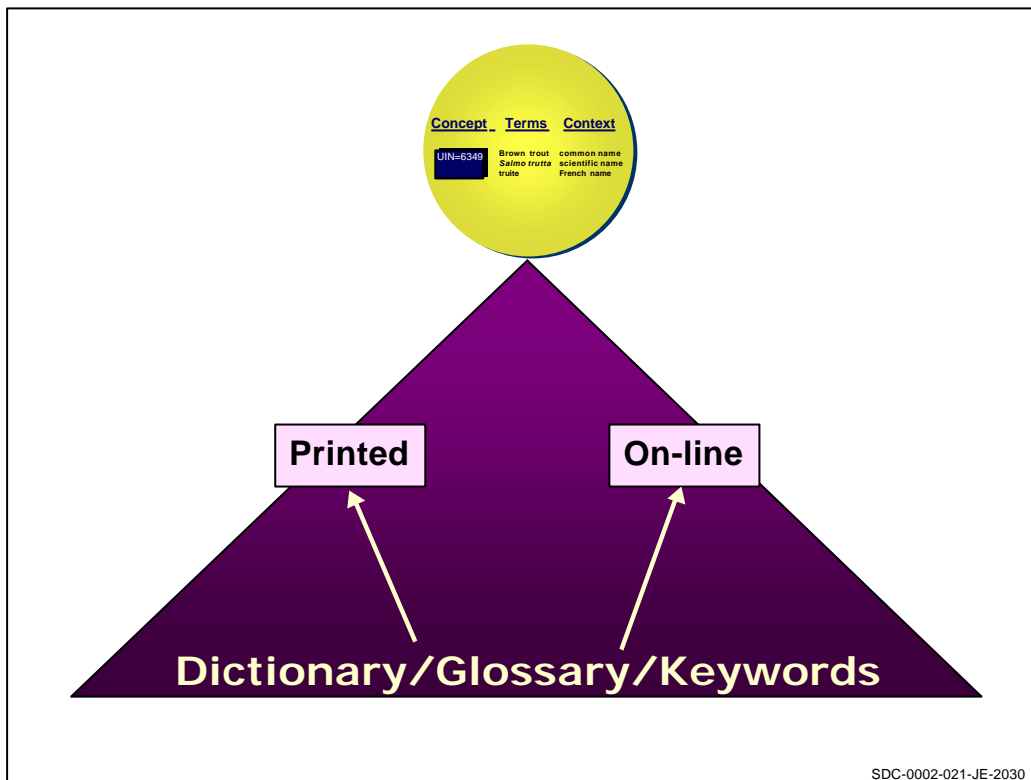
any of several game fishes of the genus *Salmo*, related to the salmon...

SDC-0002-021-JE-2030

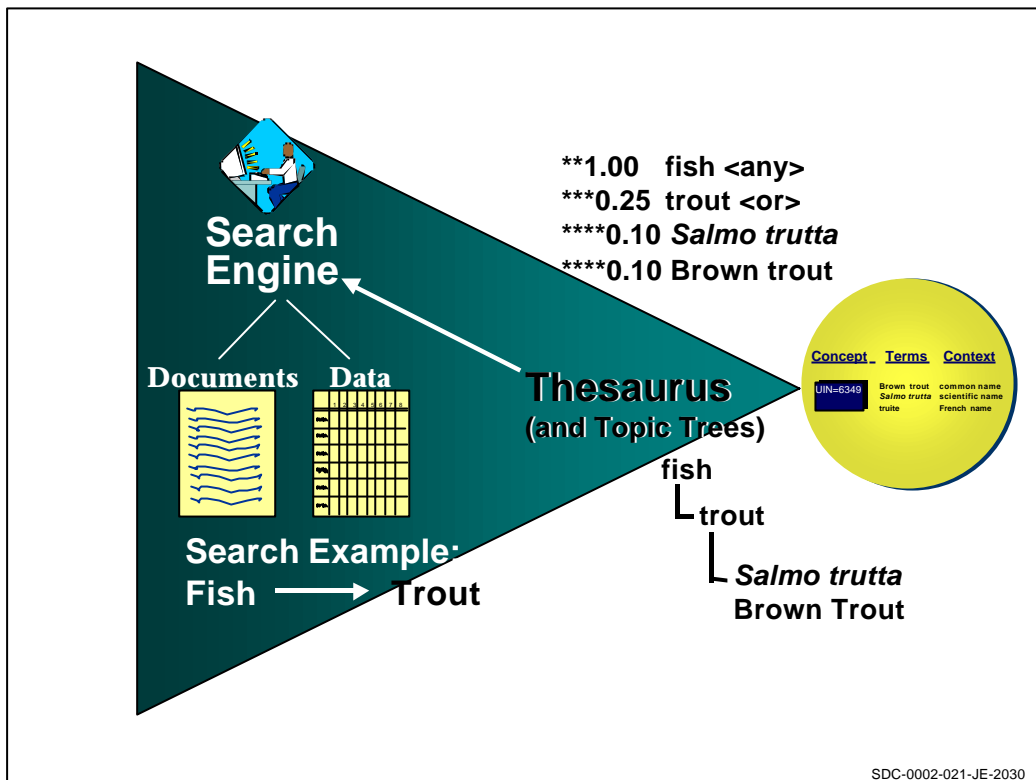
Terminology



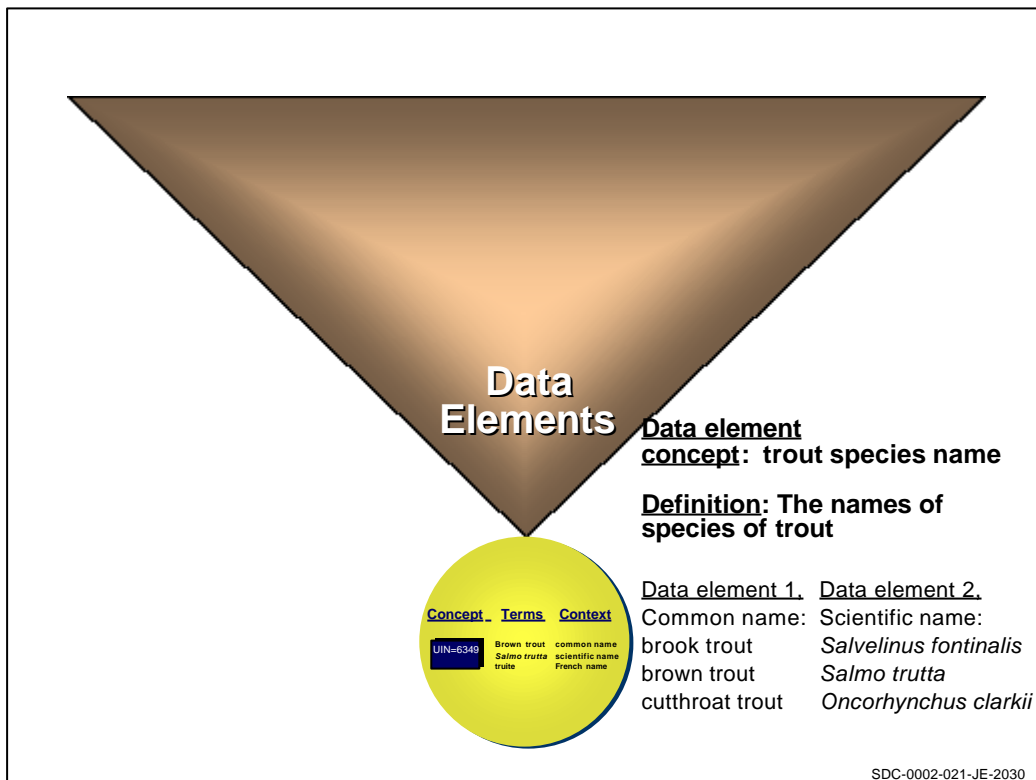
- We capture a concept with a definition.
- We assign unique identifiers to uniquely identify concepts.
- There are potentially many terms associated with any concept, and many concepts associated with any one term.
- Each term has a context from which it is extracted.



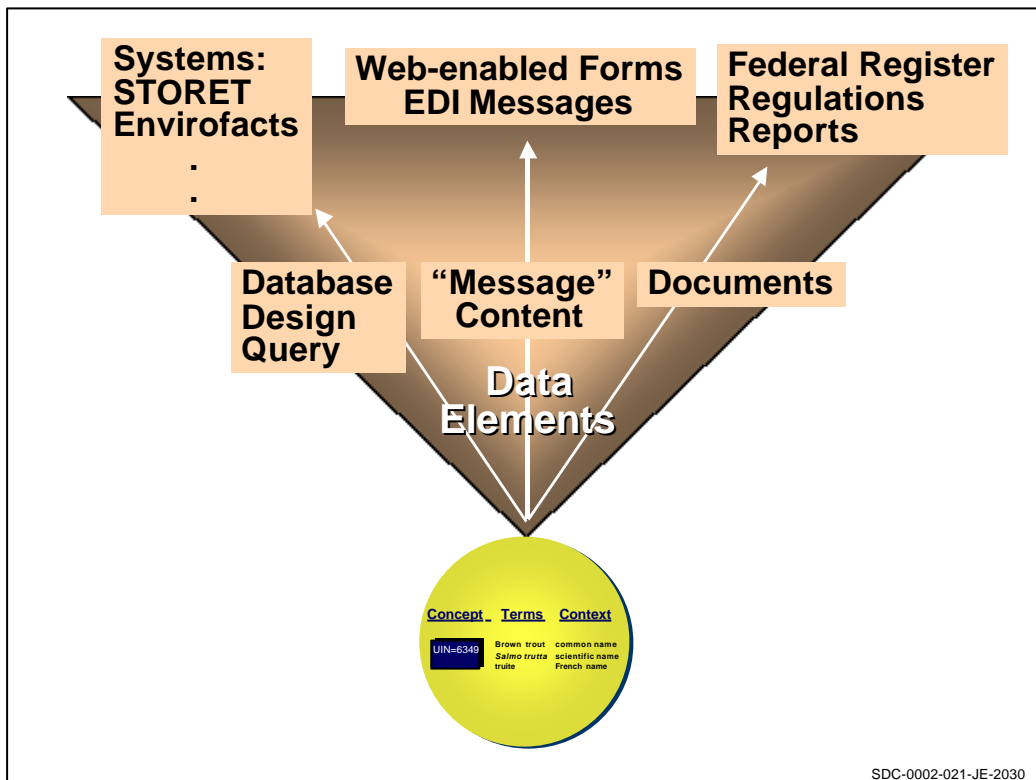
- It is common to alphabetize terms and demonstrate many senses of a term (e.g., dictionary, glossary, key word list).
- Terms within a specific context may also be stored for a particular subject area.
- We can alphabetize terms so people can look up and see all the concepts associated with a particular term.
- Specialized information, or subject matter specialties gather and store terms specific to a particular context (e.g., legal dictionary, medical glossaries, etc.).
- Deployment technologies: printed, on-line



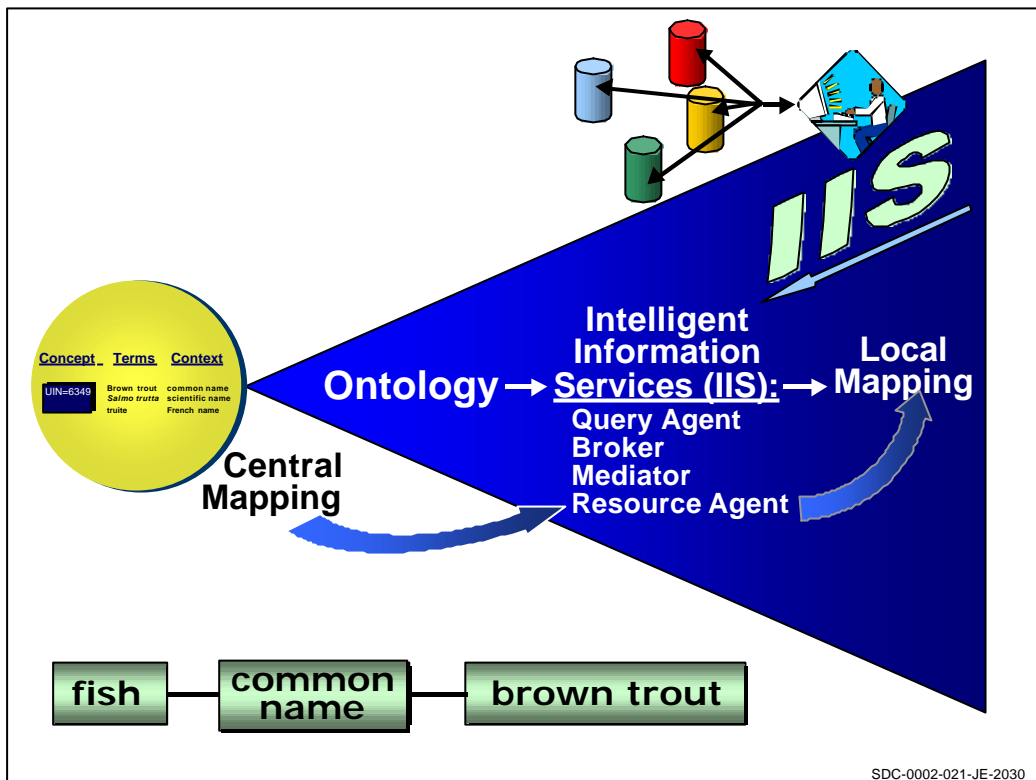
- Another way to structure terminology is into a thesaurus.
- A thesaurus has a hierarchical relationship (NT, BT, UF). We add value by building a structure which is more informative than a dictionary.
- Thesauri can be deployed in the form of published books, software, or used to feed search engine applications.



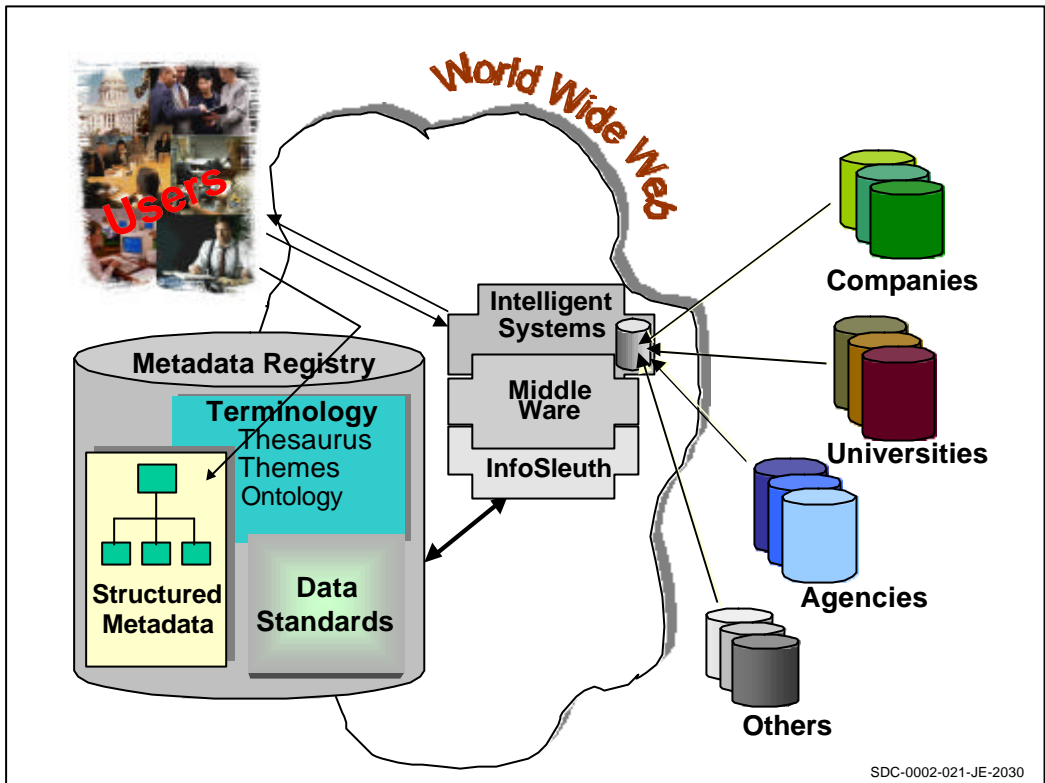
- We organize our terms into data element.
- Terms can be used in names, definitions, and values.
- The way to access all this data is by term.
- In our metadata registry we want to be able to view this information by the term name.



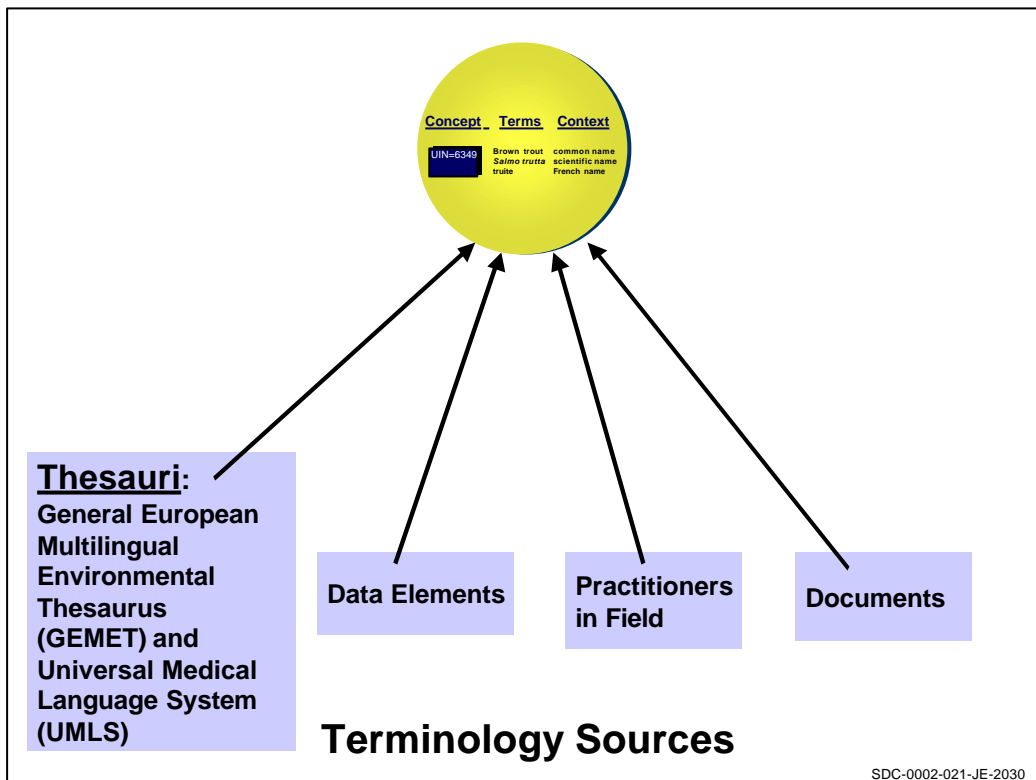
- Data elements are organized into the design of data bases, and can be used to perform queries.
- Data elements can be assembled into message content (I.e., message sets, transaction sets, information bundles, flat files, etc.) and can be deployed using EDI messages and other forms of communication.
- Terms and data elements can be organized into documents and published as reports, regulations, included in legislation, etc.



- A fourth way to structure terminology is into ontologies. Ontologies have rich structures of relationships (generalization, specialization, inheritance).
- Terms can be used to drive IIS's that utilize query agents, wrapping agents, mediators, and similar technology to access data on geographically dispersed, disparate sources. Example: an ontology might be structured to include table, data elements, and data element values, so that queries could be constructed using these ontological terms and resources. Queries are mapped to this ontology, thus enabling data to be integrated without requiring all resources to use the exact same terms in each system.

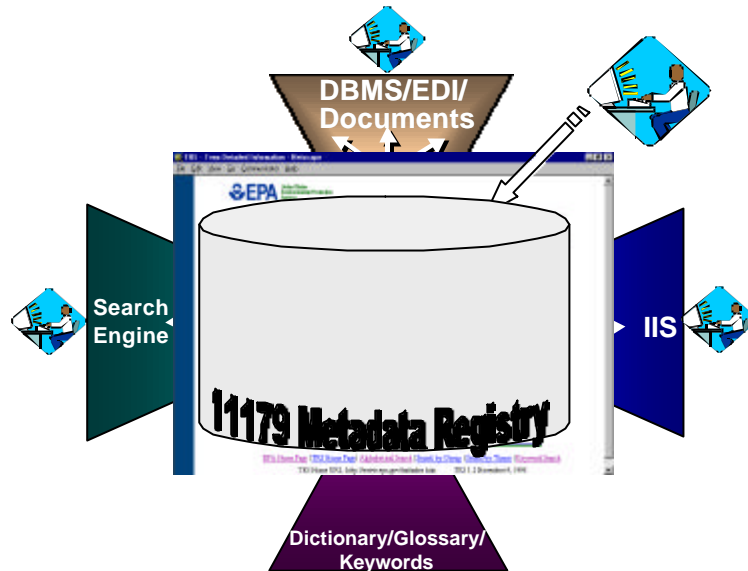


SDC-0002-021-JE-2030



There are many sources to obtain semantic content for our terminology (thesauri like GEMET and UMLS, or other sources like data elements, documents, and practioners of the world).

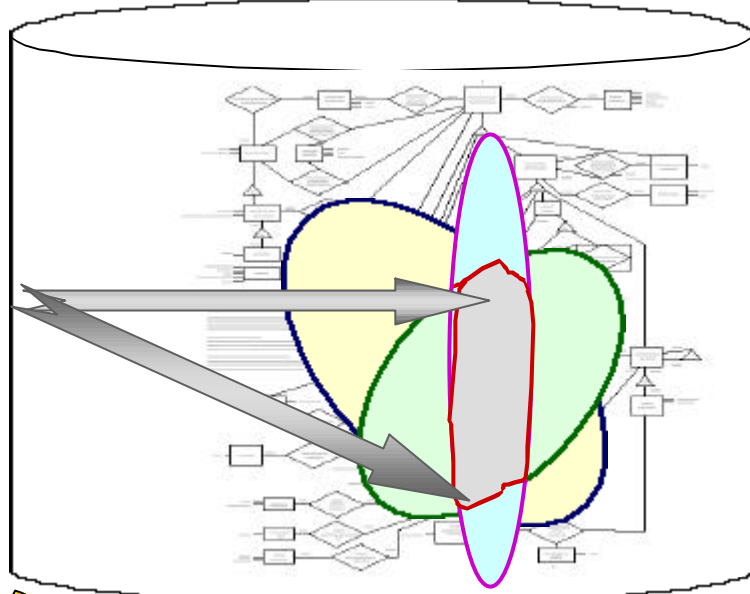
Semantics Management



SDC-0002-021-JE-2030

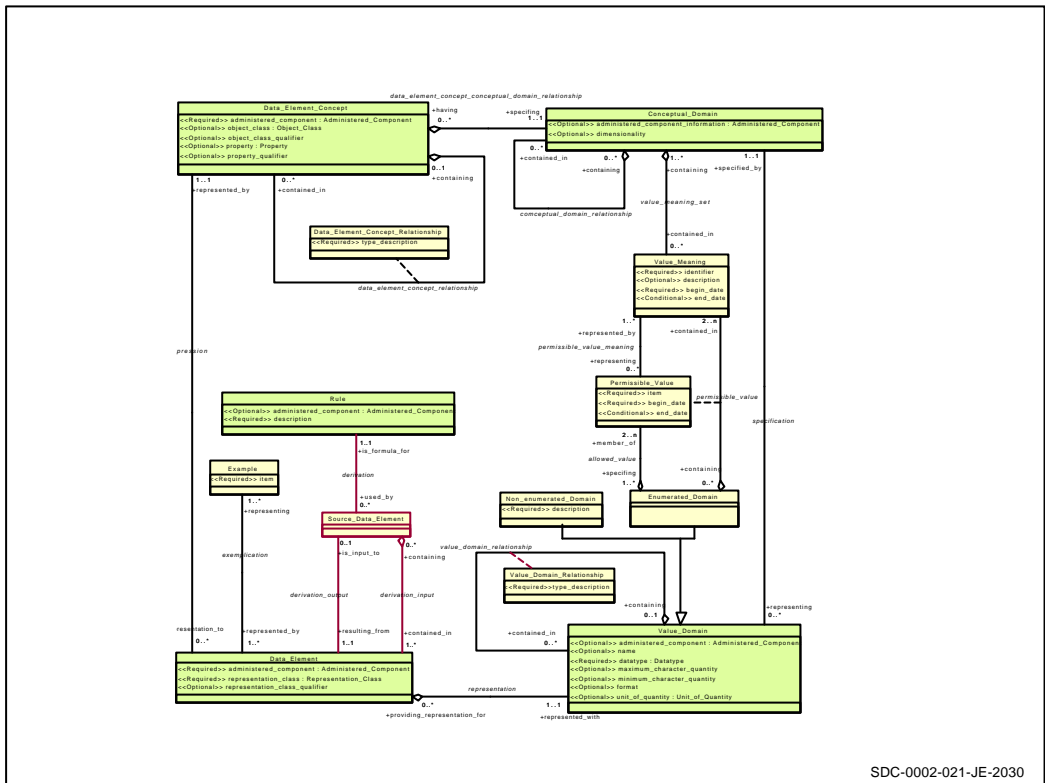
- 11179 New Work Items define overlap of 11179 and relevant standards for managing, organizing, and structuring terminology.
- We want to include modifications needed in 11179, or any additions to it.
- Users may want to access terminology directly in a metadata registry.

Profiles of ISO/IEC 11179

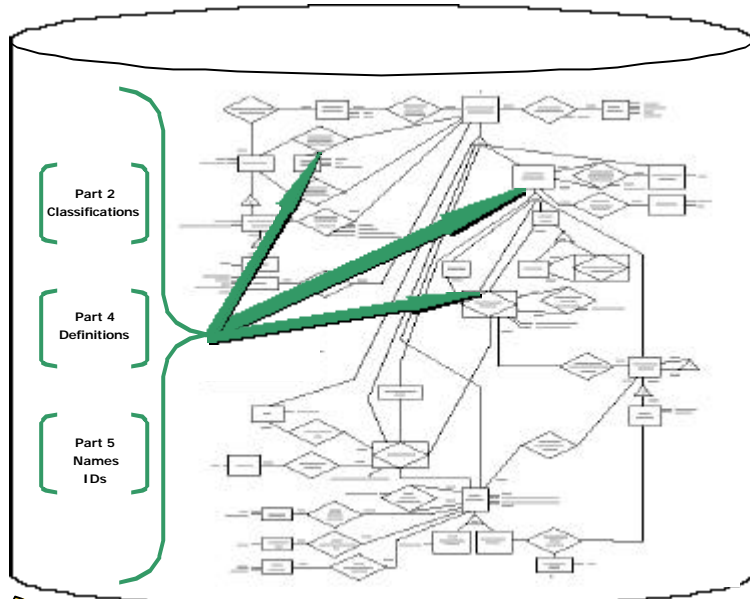


11179 Metadata Registry

SDC-0002-021-JE-2030

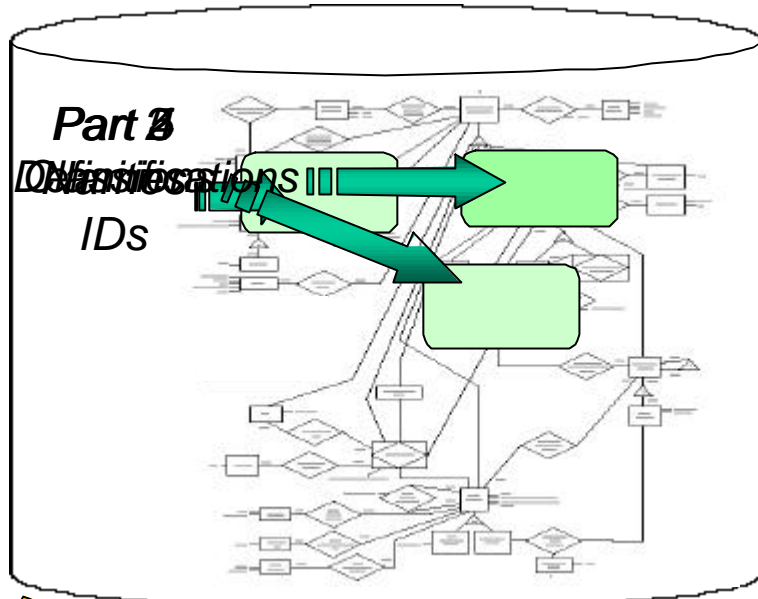


SDC-0002-021-JE-2030



11179 Metadata Registry

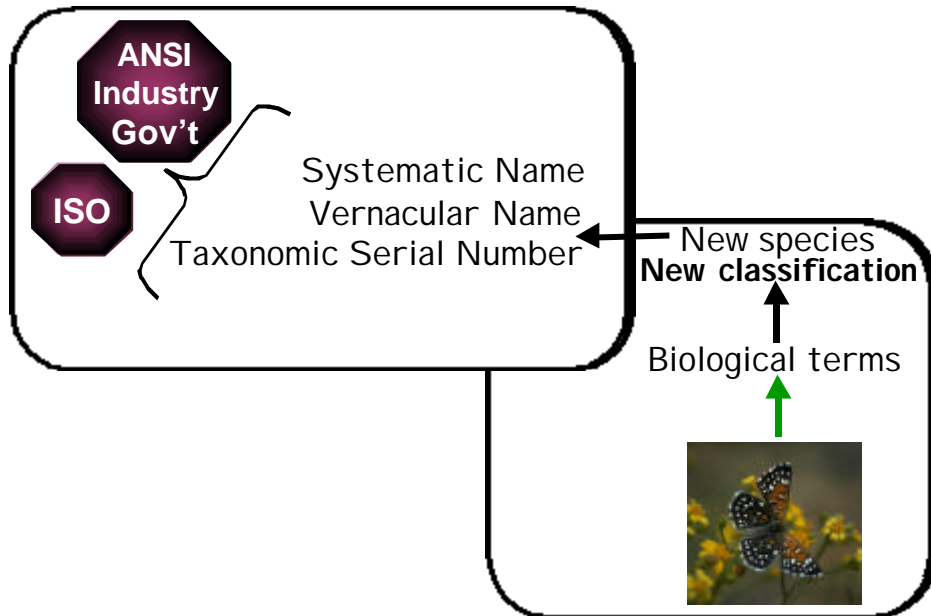
SDC-0002-021-JE-2030



11179 Metadata Registry

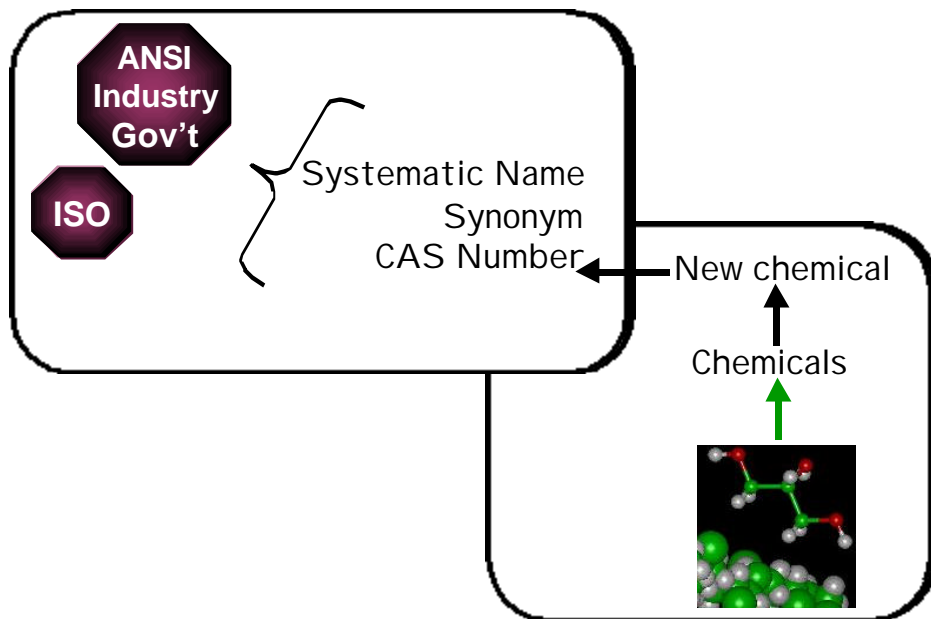
SDC-0002-021-JE-2030

Impacts of Data Change



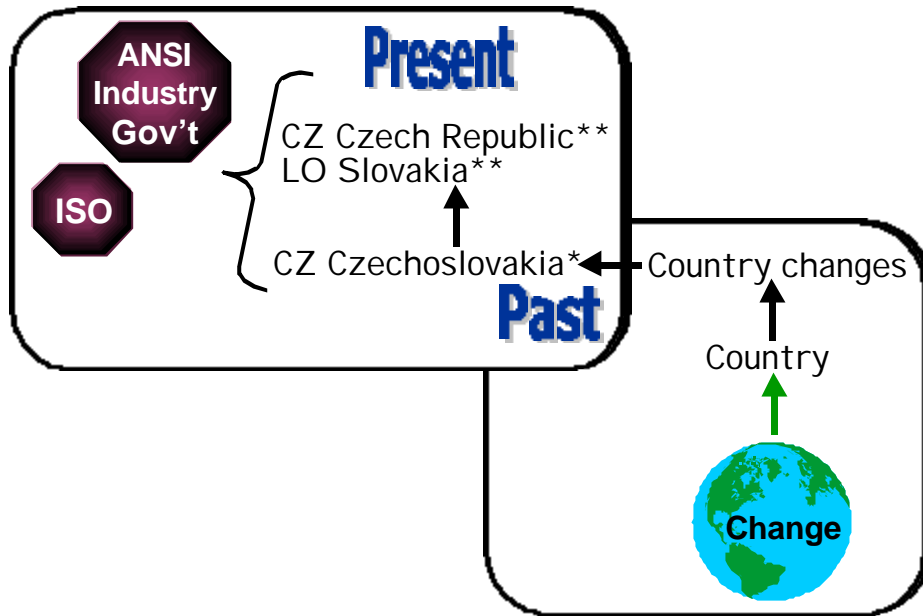
SDC-0002-021-JE-2030

Impacts of Data Change



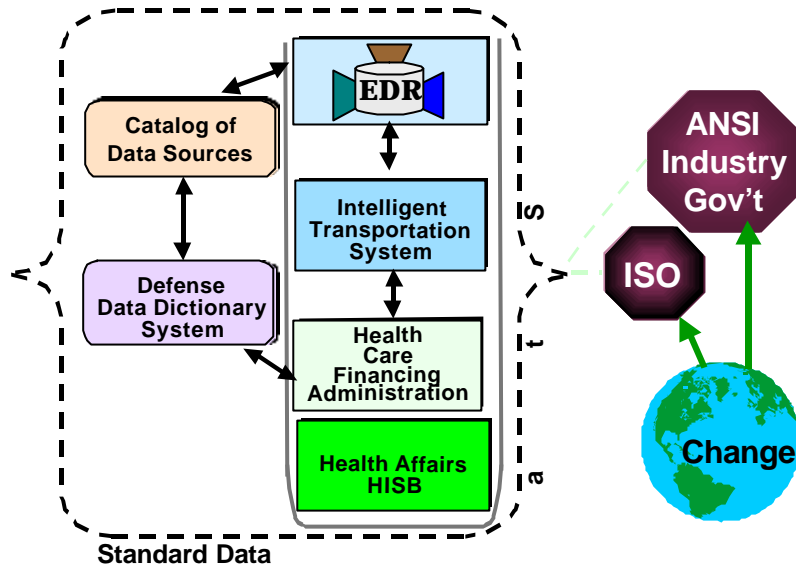
SDC-0002-021-JE-2030

Impacts of Data Change



SDC-0002-021-JE-2030

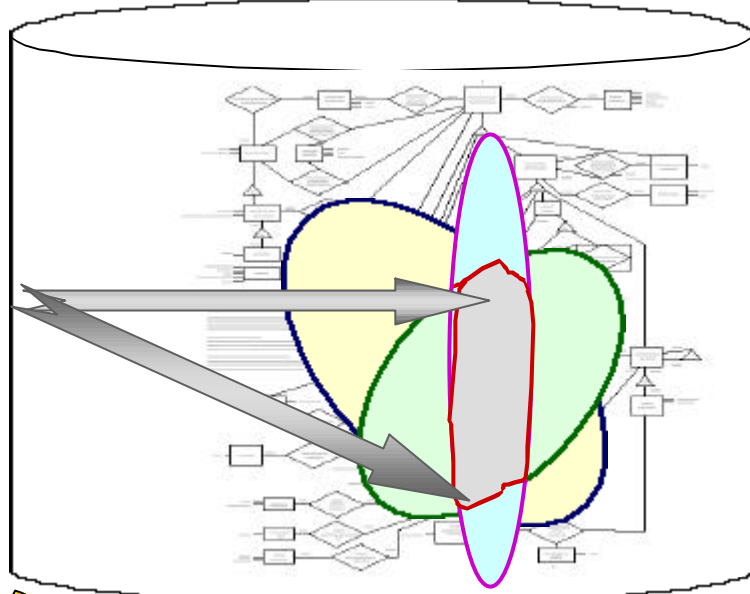
As changes in the world occur, metadata changes will be recorded and effectively communicated between registries.



SDC-0002-021-JE-2030

u

Profiles of ISO/IEC 11179



11179 Metadata Registry

SDC-0002-021-JE-2030

Implementations

- Individual Parts of 11179
- Profiles
- Extensions

SDC-0002-021-JE-2030