

Data Modeling and Data Categories for Terminology Management

Sue Ellen Wright

Kent State University

International Terminology Summer School 2009



Terminology Management

There is unfortunately no cure for terminology; you can only hope to manage it.

↪ (Kelly Washbourne)



Overview

- # Data modelling principles
 - 1) Concept orientation
 - 2) Data categories (datcat)**
 - 3) Term autonomy & datcat repeatability**
 - 4) Elementarity**
 - 5) Dependencies (combinability)
 - 6) Data modelling variance
 - 7) Granularity**
 - 8) Shared resources**

2) Data Categories

Prior to computerization, data categories were not discussed in detail in terminology theory.

Initial approaches to data categories involved describing “fields” in paper forms for recording terminological data offline.

Fachgebiet(e)			Sprache
Teilbestand	Projektcode	Notation	Klassifikations- schlüssel
Benennung			Quelle
(Kurzformen, Abkürzungen, orthographische Varianten)			Grammatische Angaben
Definition(en)			Quelle(n)
Kontext(e)			Quelle(n)
Bemerkungen			
Synonyme (falls nicht als separater Eintrag, dann mit Angabe der Quelle)			Quelle(n)
Erfasser - Datum	Bearbeiter - Datum		Eintragsklasse

TEAM-Datenbank / TermPC

- 00 Eintragsadresse
- 01 frei
- 02 frei
- 03 Qualitätskennung
- 04 Zeitpunkt der Eingabe
- 05 Bestandskennung
- 06 Fachgebietenkennung
- 07 frei
- 08 frei
- 09 frei

- 10 Deutscher Terminus
- 11 Wortart
- 12 Quellenangaben
- 13 frei
- 14 Definition
- 15 Kontextbeispiele
- 16 Synonyme
- 17 Quasisynonyme, Schlagwörter
- 18 Sachgebietenangabe
- 19 frei

- 20 - 29 Sprachblock Englisch
- 30 - 39 Sprachblock Französisch
- 40 - 49 Sprachblock Spanisch
- 50 - 59 Sprachblock Katalanisch / Russisch
- 60 - 69 Sprachblock Italienisch
- 70 - 79 Sprachblock Portugiesisch
- 80 - 89 Sprachblock Niederländisch

Paper Fiche

2) Data Categories

Fachgebiet(e) Subject field		Sprache Language	
Teilbestand Subset	Projektcode Project code	Notation Notation	Klassifikations- schlüssel Classification code
Benennung Term		Quelle Source	
(Kurzformen, Abkürzungen, orthographische Varianten) Short forms, abbreviations, orthographic variants		Grammatische Angaben Grammar	
Definition(en) Definition		Quelle(n) Source	
Kontext(e) Contexts		Quelle(n) Source	
Bemerkungen Comments (note)			
Synonyme (falls nicht als separater Eintrag, dann mit Angabe der Quelle) Synonyms		Quelle(n) Source	
Erfasser – Datum Creator/Date	Bearbeiter - Datum	Editor/Date	Eintragsklasse Entry Class

Packaging Data

Model for visualizing information

- ↪ An amorphous flow of undifferentiated “stuff”
- ↪ An aggregate of individual elements (little packages, components) that can be identified, delimited, organized (modeled), stored, retrieved, manipulated, and reused
- ↪ Stuff people can figure out if they think about it
- ↪ Stuff a computer program can automatically recognize and process

The Shopping Basket Model

- # A shopping basket is a container to put stuff in.
- # The stuff in the basket *can be* tossed in loose without packaging.
- # Stuff gets lost.
- # Stuff gets mixed up.

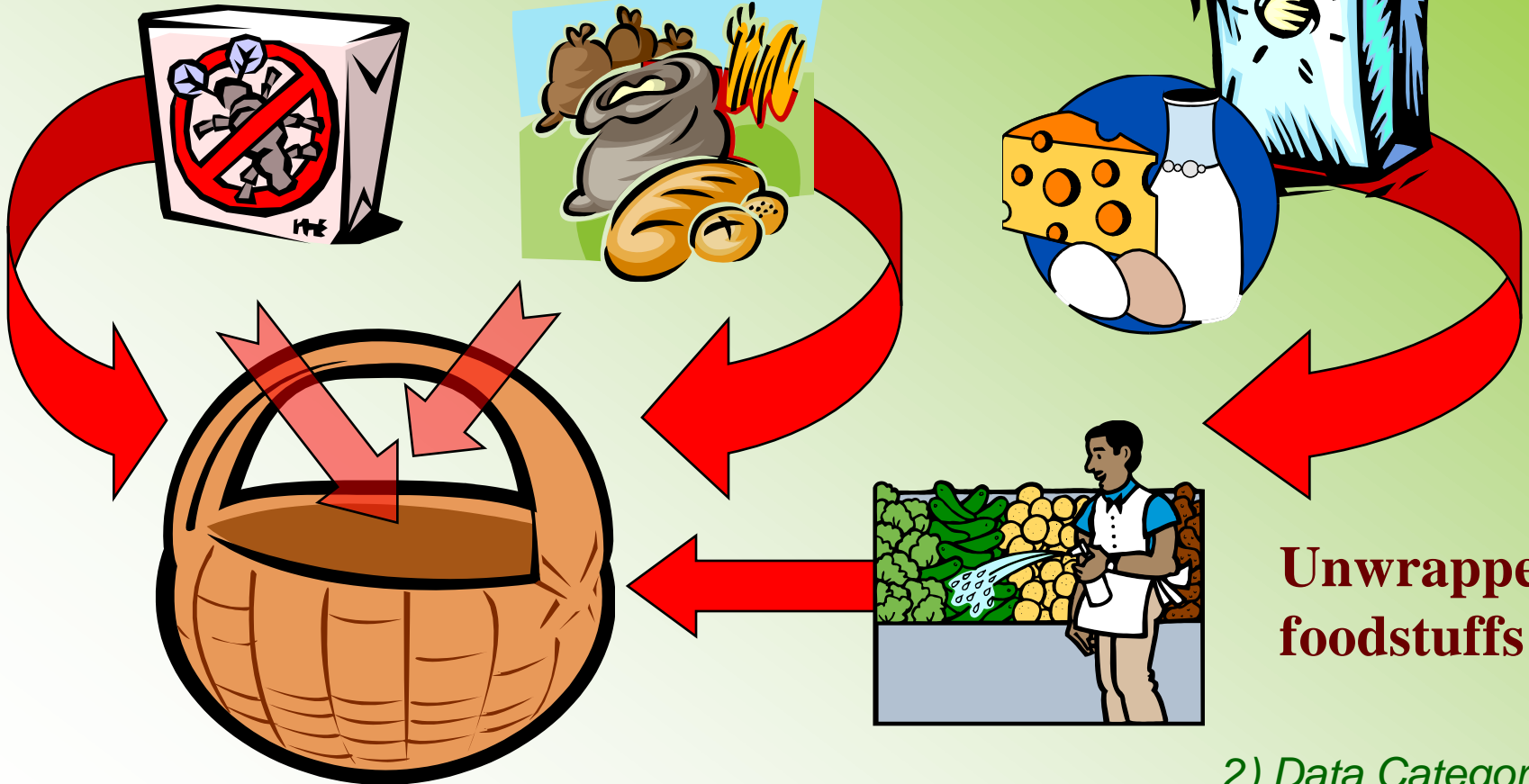
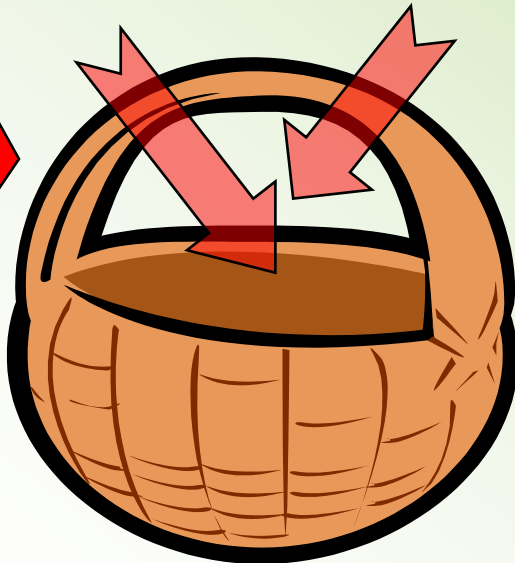


Bulk Product Loading

**Unpackaged
rat poison**

Loose flour

**Unwrapped
foodstuffs**



Packaging Stuff

- # Keeps products clean and uncontaminated
- # Makes them easy to identify
- # Makes them easy to store
- # Makes them easy to reuse



03.02.1997 - 18:16:50 Walther 25.01.2000 - 18:56:54
Wright 34 Network management alias 97/02/10 -
10:29:28 Walther 97/02/10 - 10:29:28 Walther noun
Main Entry Term shorter form of a long name, such as
an email address, a directory, or a command Fahey:7
Actually, all text-type addresses, long or short, are
aliases of the IP address which is numerical only.
Fahey:7 Aliases in real-time chat are usually
referred to as nicknames and handles nickname
noun Synonym When you start an IRC session, you
specify a nickname, up to nine characters, which you
can also change at any time. Osborne94: 413 handle
noun The nickname you assign yourself when conversing
in discussion groups. Falcón: 94 Use only in
colloquial situations. alias noun m un nombre corto o
apodo que se utiliza para enviar mensajes a aquellas
direcciones de uso más frecuente. ~~Carballar94: 112~~
alias representa a un grupo de pe
que se envía un correo a un alias
todo el grupo de personas que lo forman. Carballar94:
112

Unpackaged Data

- air filter
 - bolt
 - brake caliper
 - brake line
 - cam chain
 - camshaft
 - carburetor
 - chain guard
 - chain tensioner
 - clutch
 - connecting rod
 - crankcase
 - crankshaft
 - cylinder
 - cylinder head
 - cylinder head cover
 - exhaust manifold
 - exhaust valve
 - fuel line
 - fuel tank
 - gas tank
 - gaskett
 - handlebar
 - headlight
- TermDel12_Fedorko

Entry number: 26
 Subject: Front Wheel and Brake
 Graphic:



G-Source: <http://www.streetfightersinc.com/SpieglerBrakeLines.jpg>

EN English

Definition: small-diameter rigid steel tubing used to channel hydraulic pressure from the Master cylinder to the wheel cylinder or Caliper in a brake system
 D-Source: http://www.motorera.com/dictionary/BR.htm#Brake_Lines

brake line

Part of Speech: noun
 Grammatical Number: singular
 Context: With the new brake line installed, you'll need to add brake fluid to the system and bleed the brakes.
 C-Source: http://autorepair.about.com/od/fixityourself/ss/brakeline_repl_4.htm

Spanish

Definition: conjunto de cables de acero trenzados utilizado para controlar los frenos
 D-Source: http://www.infovisual.info/05/033_es.html

cable del freno

Part of Speech: noun
 Grammatical Gender: masculine
 Grammatical Number: singular
 Context: El brazo posterior del freno está situado apenas debajo del montaje de choque posterior y tiene un solo perno de 13m m con un agujero a través del centro que sostiene el cable del freno en lugar por la compresión.
 C-Source: <http://www.vespania.net/manuales/QUITAR%20EL%20MOTOR%20DE%20UNA%20VESPA.pdf>

Organized Data

- Project: Untitled.xdp
 - TermDel12_Fedorko
- Hit List Favourites Project

TC 37 Definitions

- # **Data elements and categories are strictly defined.**
- # **data element**
unit of data that in a certain context is considered indivisible
one unit of information in a data entry
- # **data category / data element type**
result of the specification of a given data field

Specifying the Data Category

- # */term/* is a *data category*
- # */xml:lang/* (xml language representation) is a *data category*
- # We specify what */term/* is with a definition: *verbal designation of a general concept in a specific subject field*
- # We can also *constrain* the */term/* category: terms are reported as *plain text* and not as graphics or numbers (although a number can be used as a term).
- # We use the word *constrain* to mean that we limit what the something can do.

Kinds of Data Categories

- # **complex data category (MultiTerm: index, text, picklist)**
 - ↳ **data category** that has content of some sort
- # **open data category (MultiTerm: index, text)**
 - ↳ **complex data category** whose **conceptual domain** is not restricted to a set of values (non-enumerated CD)
- # **closed data category (MultiTerm: picklist)**
 - ↳ **complex data category** whose **conceptual domain** is restricted to a set of identified **simple data categories** making up its **value domain** (enumerated conceptual domain)
- # **simple data category (MultiTerm: picklist value)**
 - ↳ **data category** with no **conceptual domain**; a member of a value domain

Data Categories

open data category

↳ **complex data category** whose **content (conceptual domain)** is not restricted to a set of values (non-enumerated conceptual domain)

↳ Example: */term/*, */definition/*, */note/*

↳ Open data categories are still constrained by their specifications and definitions.

↳ Only a term should go in a term field.

↳ Only a definition should go in a definition field

↳ But no one can say precisely what terms or definitions there are

Sample Open Data Category

Data categories can be combined:

↳ ***en:term*** = a term in English

↳ An instance of a data category is a data element.

en:term=external collateral ligament is an individual data element.

en:term=fibular collateral ligament is another individual data element.

Picklists: Closed Data Categories

- # */grammatical gender/* is a data category
- # */grammatical gender/* has defined values:
 - ↳ masculine
 - ↳ feminine
 - ↳ neuter
 - ↳ other
- # m/f/n/o are also called *picklist values*, *permissible values* or *permissible instances*.

Permissible Values

- # Together, permissible values comprise a *value domain*.
- # No other values are allowed unless a change is made in the system.
- # Terminologists sometimes call these values ***simple data categories***. They are values, but they do not themselves have dependent values.
- # Each has its own *value meaning*.

Data-Element Related Features

Data Element Autonomy (repeatability)

- ↳ Elementarity of data elements
- ↳ Dependency (combinability)

Data Modeling Variance

- ↳ Elementarity
- ↳ Granularity

Shared Resources

Metamodel issues

- ↳ Standards



Constrained Data Categories

Some data categories can hold only certain kinds of data.

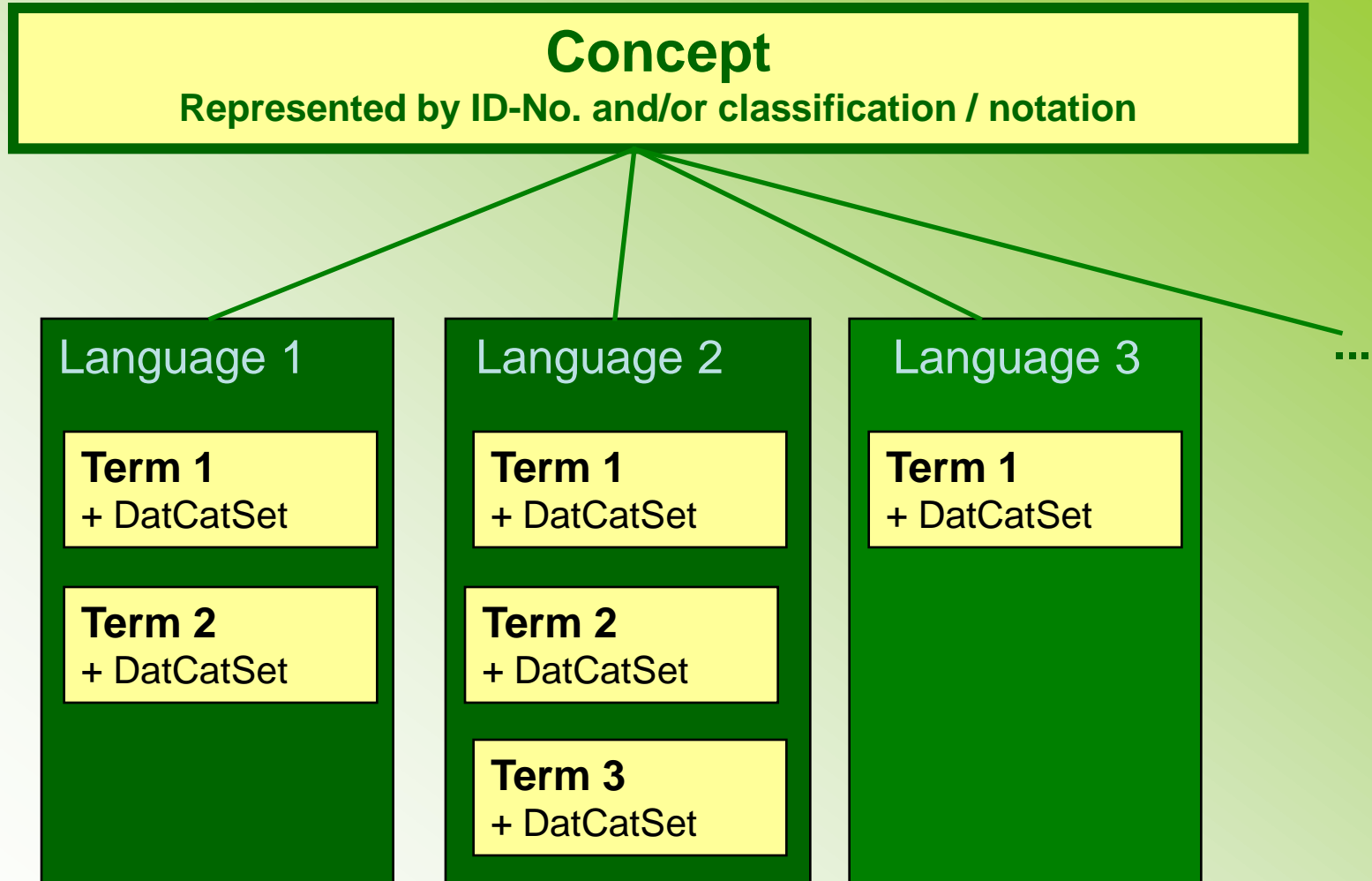
↳ Example: We can create a data category called *Graphic* and set the data type of that datcat to *Media*

↳ Example: We can create a data category called *Date* and set the data type of the datcat to *Date*

Term Autonomy & Repeatability

- # All terms should be managed as autonomous (repeatable) blocks of data categories without any preference for a specific term
 - ↳ Each term documented with the relevant term-related data categories
 - ↳ Term autonomy applicable for preferred term, synonyms, variants, & short forms
 - ↳ Term autonomy not explicitly discussed in other theoretical literature

Term Autonomy



- ductal carcinoma in situ
- edema
- fluoroscopy
- ganglion
- HER-2/neu
- hormone receptor
- IDC
- ILC
- infiltrating cancer
- infiltrating ductal carcinoma
- intraductal carcinoma
- invasive cancer
- invasive ductal carcinoma
- invasive lobular carcinoma
- LCIS
- lesion
- lobular carcinoma in situ
- magnetic resonance imaging
- malignant
- mammography
- margin
- mastectomy
- metastasis
- MRI

Project: Untitled.xdp

TermDel13Carlson

Hit List Favourites Project

Here each term has its own DatCatSet.

Entry number: 49
Subject: Pathology

EN English

Definition: early breast cancer (or precancer) that consists of abnormal cells growing in the lining of a breast duct that have not spread beyond the duct to invade the surrounding breast tissue
D-Source: [Turkington and Kraq 2005](#)
Note: Women who have DCIS are at an increased risk of invasive breast cancer. If the cancer is detected at this stage, there is no risk of its spreading into the blood or lymph system.
N-Source: [Turkington and Kraq 2005](#)

ductal carcinoma in situ

Part of Speech: noun
Grammatical Number: singular
Status: admitted
Type: phrase
Usage Register: technical
Context: Treatment of ductal carcinoma in situ may include the following: breast conserving surgery with or without radiation therapy or hormone therapy; total mastectomy with or without hormone therapy; or clinical trials testing breast-conserving surgery and hormone therapy with or without radiation.
C-Source: [Engel 1996](#)

DCIS

Part of Speech: noun
Grammatical Number: singular
Status: admitted
Type: acronym
Usage Register: technical
Context: Some DCIS cells can change genetically and become true cancers, and women should not be lulled into thinking that a DCIS diagnosis can be ignored or dismissed. We still do not know for sure which DCIS cells will change and become invasive and which will remain DCIS.
C-Source: <http://www.dcis.info/dcis.html>

intraductal carcinoma

Part of Speech: noun
Grammatical Number: singular
Status: admitted
Type: phrase
Usage Register: technical
Context: After a time, the excess cells may begin to look peculiar or atypical; later they may come to resemble cancer, and this is when they are termed intraductal carcinoma in situ, or "cancer confined to the duct."
C-Source: [Engel 1996](#)

French

Definition: prolifération de cellules épithéliales malignes de type canalaire, confinées à l'intérieur des structures canales et pouvant s'étendre aux lobules et à l'épiderme mamelonnaire sans signe d'invasion du tissu mammaire adjacent
D-Source: <http://www.cancer-sein.net/front/index.php?IdSSRubrique=27>
Note: Le carcinome canalaire in situ ou non invasif (DCIS ductal carcinoma in situ) est une forme pré-invasive de cancer du sein. Il s'agit d'un cancer qui est confiné aux cellules de l'épithélium de la lumière des canaux. Il a un excellent pronostic s'il est isolé et découvert fortuitement à la mammographie sous forme de fines microcalcifications.
N-Source: <http://pages.infinit.net/trh/general.html>

carcinome canalaire in situ

Part of Speech: noun
Grammatical Gender: masculine

3) Autonomy, KDS

Here terms are hidden in /variant/ and /synonym/ fields.

Eintragsnummer 486
Fachgebiet Technik
Lenoch-Code #TR:
Autor Sonia Banik

air bag
Gebrauch *norm*
Quelle ISO/DIS 12097-1
Definition A flexible material, forming an enclosed volume that receives the gas from the inflator and restrains the occupant.
Variante airbag
Synonym air cushion, air bag cushion
Kurzform bag
D-Quelle ISO/DIS 12097-1
Kontext The actual air bag in an air bag restraint system; typically made of a polyamide fabric, often coated with neoprene on the inside; newer air bags have no coating to facilitate recycling.
K-Quelle Schmitt.1992

Luftsack
Genus *m.*
Quelle Schmitt.1992
Definition Eigentlicher Luftsack in einem Airbag-Rückhaltesystem: meist ein Gewebe aus Polyamid (Nylon), das entweder innen beschichtet ist (Neopren) oder zwecks besserer Recyclingfähigkeit (Sortenreinheit) unbeschichtet bleibt.
Kurzform Sack
Synonym Luftkissen, Prallkissen, Prallsack, Airbag
D-Quelle Schmitt.1992
Kontext Airbagsysteme sind seit Beginn der 80er Jahre zunächst in Fahrzeugen der Oberklasse eingebaut worden. Auf der mechanischen Seite bestehen sie aus einem porösen oder mit Auslaßöffnungen versehenen Luftsack mit einem Volumen zwischen 35 (Eurobag) und ca. 140 l (US-Beifahrer-Airbag) der von einem Gasgenerator in kürzester Zeit (25-30 ms) aufgeblasen wird.
K-Quelle Ostertag.1996, S. 67

Notation 1-1-2-2-2-2-3-1

3) *Autonomy, KDS*

Embedded or Hidden Terms

- # **en:term=external collateral ligament**
- # **definition=ligament that bla bla bla ...**
- # **note=bla,bla, bla, also sometimes called the “fibular collateral ligament”.**
- # **Problem:**
 - ↳ You can't look up the buried term.
 - ↳ It doesn't become part of any concept system.
 - ↳ It ends up in a */note/* element instead of a */term/* element.

Index Englisch

air bag, airbag, b.. air bag, airbag, b..

Here the synonyms and variants are strung together in one /term/ field & we can't document info about individual terms..

Eintragsnummer 485
 Fachgebiet Technik
 Lench-Code #TR:
 Autor Sonia Banik

air bag, airbag, bag, air cushion, air bag cushion

Quelle ISO/DIS 12097-1

Definition A flexible material, forming an enclosed volume that receives the gas from the inflator and restrains the occupant.

D-Quelle ISO/DIS 12097-1

Kontext The actual air bag in an air bag restraint system; typically made of a polyamide fabric, often coated with neoprene on the inside; newer air bags have no coating to facilitate recycling.

K-Quelle Schmitt.1992

Luftsack, Luftkissen, Prallkissen, Prallsack, Sack, Airbag Genus m.

Quelle Schmitt.1992

Definition Eigentlicher Luftsack in einem Airbag-Rückhaltesystem; meist ein Gewebe aus Polyamid (Nylon), das entweder innen beschichtet ist (Neopren) oder zwecks besserer Recyclingfähigkeit (Sortenreinheit) unbeschichtet bleibt.

D-Quelle Schmitt.1992

Kontext Airbagsysteme sind seit Beginn der 80er Jahre zunächst in Fahrzeugen der Oberklasse eingebaut worden. Auf der mechanischen Seite bestehen sie aus einem porösen oder mit Auslaßöffnungen versehenen Luftsack mit einem Volumen zwischen 35 (Eurobag) und ca. 140 l (US-Beifahrer-Airbag) der von einem Gasgenerator in kürzester Zeit (25-30 ms) aufgeblasen wird.

K-Quelle Ostertag.1996, S. 67

Notation 1-1-2-2-2-3-1

Klaus-Dirk Schmitz

Data Element Elementarity

- # Only one of a ***thing*** can occupy a data element
 - ↳ e.g., only one term in a term field
- # Only one ***kind of thing*** can occupy a data element
 - ↳ e.g., no terms or synonyms listed as such in definition fields



Common Elementarity Errors (1)

Error:

↪ en:term = United Nations (UN)

Correct:

↪ en:term = United Nations

term type = full form

↪ en:term = UN

term type = acronym

Combinability enables us to identify individual units of content: */term/* combines here with */term type/*.

Common Elementarity Errors (2)

Error:

↪ definition = international organization that ... (Merriam Webster, 10th. Edition 2004, p. 256)

Correct:

↪ definition = international organization that ...
source = Webster2005, p. 256

↪ *Webster2005* points to a shared resource
(bibliographical entry)

Combinability: Source can be used with a term or any text or graphics field.

Modelling Dependencies (Combinability)

- # Dependencies between data categories

- # ISO 12620:1999 provides a “simple hierarchy” of data categories

 - ↪ *grammar* = term-related:
grammar is dependent upon *term*

- # Other dependencies

 - ↪ *source* is dependent upon *definition*

 - ↪ for additional definitions, additional sources are needed

 - ↪ Individual source indicated for different items (term, definition, context, etc.)

Granularity

- # The degree of detail that can be achieved by using the available data fields (data categories) to document terminological information
- # Low granularity:
 - ↪ Ex: */grammar/* m,n,s (masculine noun singular)
- # High granularity
 - ↪ */part of speech/* = noun
 - ↪ */grammatical gender/* = masculine
 - ↪ */number/* = singular



Granularity

- # Advantage of granularity: retrievability
- # Disadvantage: more work
- # Best practices:
 - ↳ General agreements on what constitutes an acceptable level of effort in order to achieve a desired level of granularity makes it easier to share data.
 - ↳ How can you automatically convert low granularity to high granularity?

Minimum Granularity

Data Element	Content
Name	John Doe
Address	267 Prospect Street Kent, Ohio 44240

Increased Granularity

Data Element	Content
Last name	Doe
First name	John
Street address	345 North Prospect St.
City	Kent
State	Ohio
Zip Code	44240

Power of Granularity

- # */zip code/* (postal code) can be searched and manipulated
- # Extra effort in entering a separate field yields search power later, and in this case functionality.
- # Sorting by postal code makes mail delivery more efficient & saves money.
- # Know ahead of time what you want to do with your data when deciding on granularity.

Elemental Nature of Data Elements (Violation)

Data Element	Content
Last name	Doe
First name	John and Jane
Street address	345 North Prospect St.
City	Kent
State	Ohio
Zip Code	44240

Shared Resources

Index	Oxford1990
Editor1	Fowler, H.W.
Editor2	Fowler, F.G.
Editor3	Allen, R.E.
Location	Oxford
Press	Clarendon
Year	1990

Entry 1
Source ID: Oxford1990

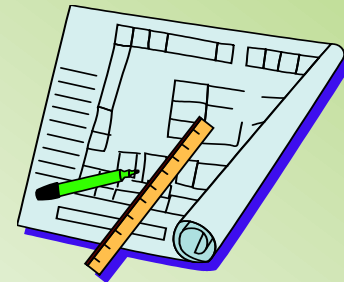
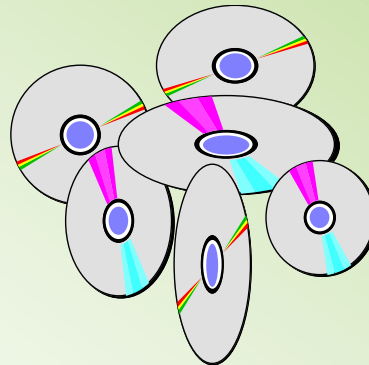
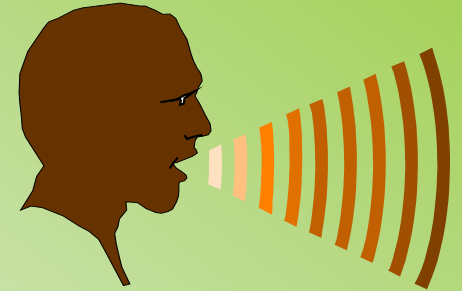
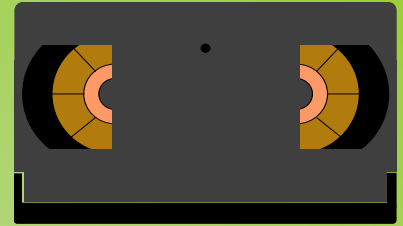
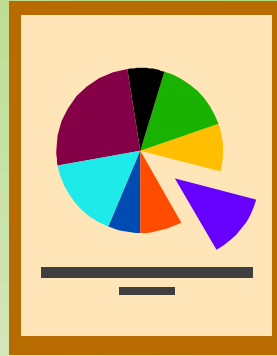


Entry 2
Source ID: Oxford1990



Shared Resources

- # Graphics
- # Charts
- # Audio
- # Video
- # Drawings
- # Disk Archives
- # Responsibility Records



Advantages of Shared Resources

- # Data category configured as a link or with a link
- # Shared items stored only once
- # Multiple references linked to same resource
- # Change or update only one file or entry
- # File formats (images, audio, etc.) preserved externally and objects linked
- # Terminology files of manageable in size

- банкомат
- бухгалтер
- бухгалтерия
- вклад
- выписка с банковского счета
- денежный брокер
- долговая расписка
- дорожный чек
- заем
- залоговое обеспечение займа
- инвестиции
- ипотечный кредит
- кассир
- квитанция
- клиринг
- коммерческий банк
- кредит
- кредитная карта
- кредитный союз
- ликвидность
- наличные средства
- необеспеченный заем
- объединенный счет
- овердрафт
- плательщик

Project: Untitled.xdp

TermDel13_Surkova

Hit List Favourites Project

Entry number: 24
 Subject: Goods and Services

Russian
 Definition: программно-технический комплекс, предназначенный для автоматизированных выдачи и приёма наличных денежных средств как с использованием платёжных карт, а также выполнения других операций, в том числе оплаты товаров и услуг и составления документов, подтверждающих соответствующие операции
 D-Source: <http://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%BD%D0%BA%D0%BE%D0%BC%D0%B0%D1%82>

банкомат
 Part of Speech: noun
 Grammatical Gender: masculine
 Grammatical Number: singular
 Context: Порядок применения банкоматов в осуществлении расчётов по банковским картам в России определяется Положением Банка России Об эмиссии банковских карт и об операциях, совершаемых с использованием платёжных карт № 266-П от 24 декабря 2004 г.
 C-Source: <http://ru.wikipedia.org/wiki/%D0%91%D0%B0%D0%BD%D0%BA%D0%BE%D0%BC%D0%B0%D1%82>

EN English
 Graphic:



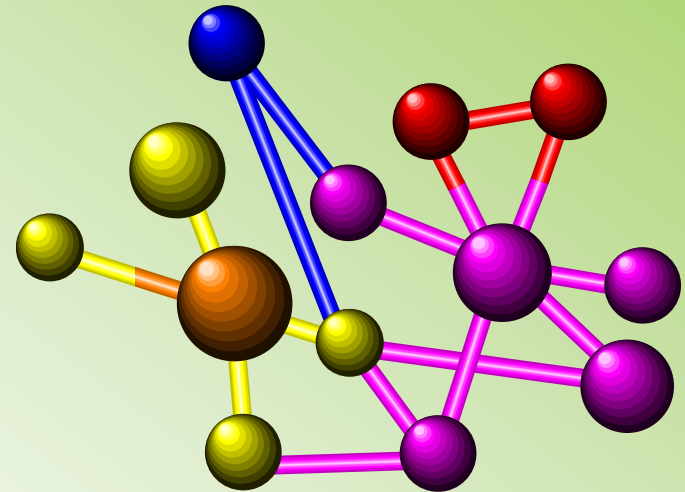
G-Source: www.jcheights.com/news.shtml
 Definition: a computerised self service machine allowing the holder of an ATM card and authorised personal identification number (PIN) to obtain cash and statements etc.
 D-Source: [http://www.fairinvestment.co.uk/automated-teller-machine-\(ATM\)-Banking-glossary-94.html](http://www.fairinvestment.co.uk/automated-teller-machine-(ATM)-Banking-glossary-94.html)

automated teller machine
 Part of Speech: noun
 банкомат

Graphics used as resources for whole entry, for a single language, or even to illustrate a single term.
Wish list: Image maps where the user can jump to various terms by clicking on parts of the image.
Note source of the graphic.

Term Type

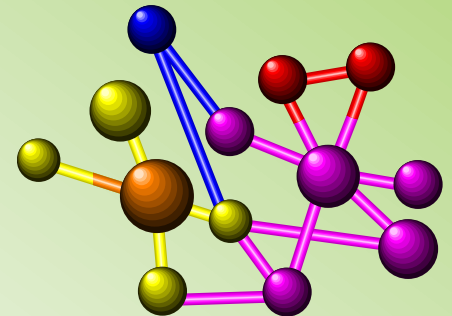
- ~~# Main entry term~~
- ~~# Synonym~~
- # Abbreviation
- # Full form
- # Variant
- # Phrase
- # Collocation



Term-Related DatCats

Term

- ↪ Part of speech
- ↪ Grammatical gender
- ↪ Grammatical number (use when necessary)
 - ☞ **Plural form**
- ↪ Term type (Type) (see next slide)
- ↪ Status
- ↪ Regional label
- ↪ Pronunciation
- ↪ Register (usage register)



9) Metamodel + “Vocabulary”

- actuate
- Belleville spring
- clutch
- clutch assembly
- clutch cover
- clutch master cylinder
- clutch slave cylinder
- conical disk spring
- cover
- cover
- diaphragm spring
- flywheel
- hydraulic actuated cl
- hydraulic clutch

Register	
Context	
C_Source	
Note	
N_Source	
Transfer	
Comment	
Related	
Term	

German

Responsibility	
Graphic	
G_Source	
Definition	
D_Source	
Note	
N_Source	

Picklist Items

Part of	
Speech	
Status	
Grammatical	
Gender	
Number	
Type	
Usage	
Register	
Context	
C_Source	

Permissible values, Term Type

- abbreviation
- acronym
- antonym
- common name
- equation
- formula
- full form
- international scientific term
- internationalism

- internationalism
- orthographical variant
- part number
- phraseologism
- short form
- stock keeping unit
- symbol
- transcribed form
- transliterated form

Project: Untitled.xdp

DriveTrain

- actuate
- Belleville spring
- clutch
- clutch assembly
- clutch cover
- clutch master cylinder
- clutch slave cylinder
- conical disk spring
- cover
- cover
- diaphragm spring
- flywheel
- hydraulic actuated cl
- hydraulic clutch

Project: Untitled.xdp

DriveTrain

- Register
- Context
- C_Source
- Note
- N_Source
- Transfer
- Comment
- Related
- Term

German

- Responsibility
- Graphic
- G_Source
- Definition
- D_Source
- Note
- N_Source

Picklist Items

- Part of
- Speech
- Status: **admitted**
- Gramm
- Gender
- Gramm
- Number
- Type
- Usage
- Register
- Context
- C_Sour
- Note
- N_Sour
- Transfe
- Comme
- Related
- Term

- admitted
- deprecated
- legal
- new
- nonstandardized
- preferred
- proposed
- recommended
- regulated
- standardized
- superseded

Permissible values, Term Status

French

- Responsibility

- actuate
- Belleville spring
- clutch
- clutch assembly
- clutch cover
- clutch master cylinder
- clutch slave cylinder
- conical disk spring
- cover
- cover
- diaphragm spring
- flywheel
- hydraulic actuated cl
- hydraulic clutch

Project: Untitled.xdp

DriveTrain

- Register
- Context
- C_Source
- Note
- N_Source
- Transfer
- Comment
- Related
- Term

German

- Responsibility
- Graphic
- G_Source
- Definition
- D_Source
- Note
- N_Source

- Part of
- Speech
- Status **admitted**
- Grammatical Gender
- Grammatical Number
- Type

- Usage Register:
- bench-level
 - in-house
 - neutral
 - slang
 - technical
 - vulgar

- Context
- C_Source
- Note

Picklist Items

Permissible values, Usage Register

Definition-Related Info

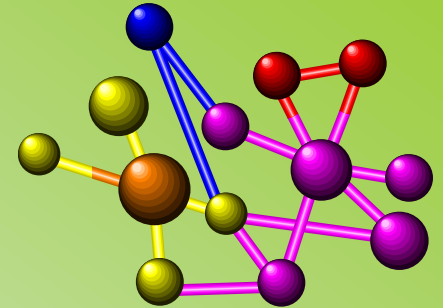
Definition

↪ Source ID

☞ (Points to shared resources)

↪ Definition type

☞ Translation?



Choice of levels

↪ Concept (one def for all languages)

↪ Language (one def for each language)

↪ Term (one def for each term in descriptive work)

Context-Related Info

Context

↪ Source ID

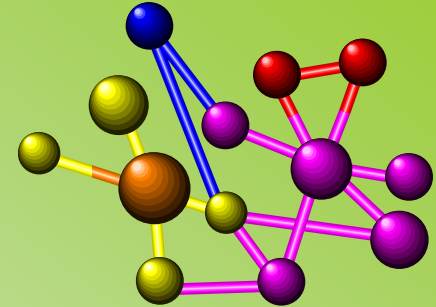
☞ (Points to shared resources)

↪ Context type

☞ Translation?

↪ Both term and concept-related

↪ Always associated at the term-level, never at lang or concept level



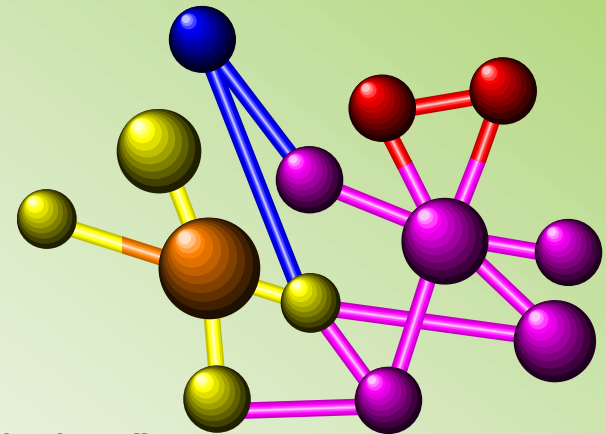
Administrative Information

Administrative info

- ☞ Responsibility
- ☞ Date
- ☞ Authorization
- ☞ Role
- ☞ Various sorting subsets
 - Business unit
 - Customer
 - Computing environment
 - Product class, etc.
- ☞ Applicable at multiple levels

Other Data Categories

- # Concept relations
- # Notes
- # Other administrative information
- # Bibliographical information
- # Special categories, e.g.:
 - ↪ For standardization
 - ↪ For inventory control



9) *Metamodel + “Vocabulary”*

MultiTerm Levels & Types

Index

↳ Terms

↳ Term-like elements

↳ (collocations, boilerplate)

Text (definitions, contexts, notes, etc.)

↳ Free-form data entry

Picklist

Multimedia file

Number

Date



10) MultiTerm Meta-Categories

Language Identifiers

- # xml:lang options (ISO 639 + IETF BPC 47)
- # Simplest level: language code from 639-1, 2, or 3
 - ↳ fr, de, es, en, ru, etc.
 - ↳ fra, deu, esp, eng, rus, etc.
- # Level 2: lang + country code
 - ↳ fr-FR, fr-CA; eng-ENG, eng-USA
- # Level 3: lang + script + country code
 - ↳ zh-Hans-ZH; zh-Hant-TW
 - ↳ Other options: see IETF 4646, IETF 4646bis

ISOcat.org

- # TC 37 Data Category Registry (DCR)
- # Create data category specifications (DC)
- # Select Data Category Selections (DCS)
- # Basis for TMS design
- # Share DCs and DCSs with colleagues
- # Standardize DCs & DCSs
- # <http://www.isocat.org>



Data Category Registry

Defining widely accepted linguistic concepts

ISO 12620 provides a framework for defining data categories compliant with the ISO/IEC 11179 family of standards. According to this model, each data category is assigned a unique administrative identifier, together with information on the status or decision-making process associated with the data category. In addition, data category specifications in the DCR contain linguistic descriptions, such as data category definitions, statements of associated value domains, and examples. Data category specifications can be associated with a variety of data element names and with language-specific versions of definitions, names, value domains and other attributes.

The Registration Authority of the TC 37 DCR is the [Max Planck Institute for Psycholinguistic](#), Nijmegen, The Netherlands.

This new DCR implementation is currently *under development* by the MPI in collaboration with TC 37. However, the core services should now be stable enough for day to day DCR usage. We welcome *any feedback* (bug reports, feature requests, etc.). You can use the following methods to contact us:

- file a [bug report or feature request](#)
- send an [email](#)

Home

Web interface

Web services

News

Publications

Thematic Domain Groups

ISO 12620

Known issues

Planning

Feedback



enter keywords here Search

- IAL Dump
- reliabilityCode
- Wiktionary
- NewL10nDatcats
- language
- Shared
- Public
- Thematic Views
 - Metadata
 - Morphosyntax
 - Semantic Content Representation
 - Syntax
 - Language Resource Ontology
 - Lexicography
 - Language Codes
 - Terminology
 - Terminology
 - Multilingual Information Management
 - Lexical Resources
 - Lexical Semantics

Edit Export Terminology Tree Reload Detach

#	Name	Version	Administration st	Registration statu	Check	Type	Owned by	Scope
338	partNumber	1.0	accepted	standard		simple	Terminology	public
396	partOfSpeech	0.0.0	private	candidate		closed	Sue Ellen Wright	public
397	part-whole relation	1.0	accepted	standard		open	Terminology	public
113	patentDocument	0.0.0	private	candidate		simple	Sue Ellen Wright	public
114	periodical	0.0.0	private	candidate		simple	Sue Ellen Wright	public

Add Explore Edit Scope Check partOfSpeech - 0.0.0 Detach

Change Description 12620-2:2003; 12620-3:2003

2. description

profile Terminology

2.1 Language Section

language en

2.1.1 Definition Section

definition A category assigned to a word based on its grammatical and semantic properties.

Source ISO12620

2.1.2 Example Section

Example ...

Source Mitre; TEI(green text); 1951

2.2 Language Section

language de

2.2.1 Definition Section

definition Eigenschaft eines Wortes ...