

La terminología en las ontologías y otros recursos lexicográficos

Dra. Guadalupe Aguado de Cea

lupe@fi.upm.es

<http://www.oeg-upm.net>

Ontological Engineering Group

Facultad de Informática

Universidad Politécnica de Madrid

Campus de Montegancedo sn,

28660 Boadilla del Monte, Madrid, Spain

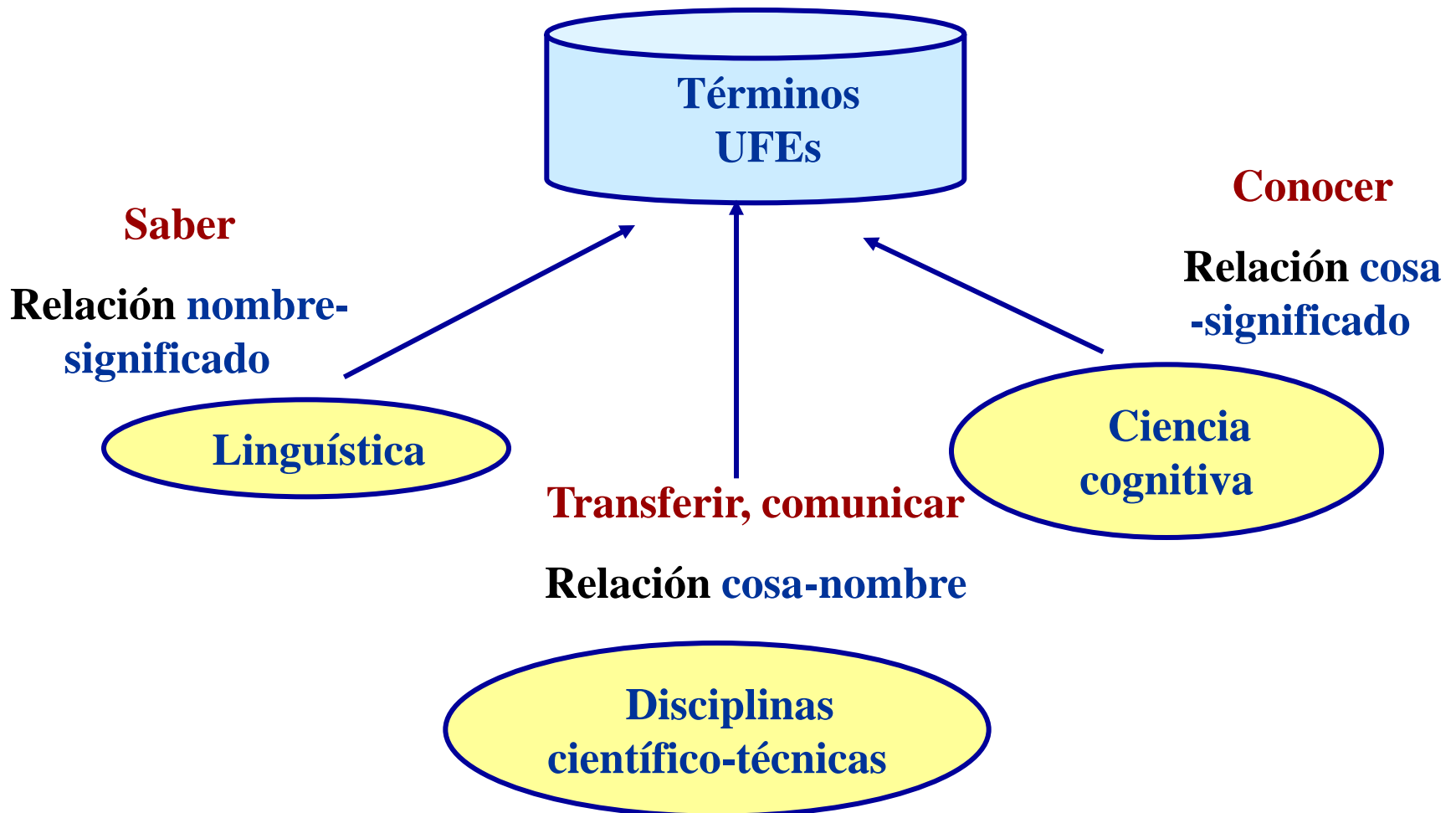
Indice

- Qué es la terminología
- Objeto de estudio y método de trabajo
- Fundamentos de la terminología
- Los conceptos y sus relaciones
- Los términos y su formación
- Los recursos lingüísticos en relación con las ontologías
- Cómo conjugar la terminología con las ontologías
- Enfoques que combinan terminología y ontologías
- Tabla comparativa de recursos

¿Qué es la terminología?

- Como **producto**: conjunto de términos de una materia especializada.
- Como **disciplina**: conjunto de principios y bases conceptuales que rigen el estudio de los términos especializados, su descripción, análisis y relaciones.
- Como **práctica** conjunto de directrices que se utilizan en el trabajo metodológico y que rigen la recopilación de los términos y las relaciones entre estos y los conceptos.

Objeto de estudio de la terminología



Modelización de los sistemas de conceptos: metodología en terminología

1. Identificar el **contexto** o el dominio de conocimiento
2. Identificar los **conceptos** u objetos del dominio
3. **Determinar** las propiedades atribuidas a cada concepto
4. Determinar las características (propiedades) **esenciales** que conforman el concepto
5. Establecer relaciones conceptuales: **mapas de conceptos**
6. **Atribuir** una designación
 1. creación neológica: *folksonomies, servlet, applet, formatear*
 2. análisis contextual y variación en las UT: *ordenador, computador, ordenata, máquina, PC,*
7. **Definir** cada uno de los conceptos
8. **Finalidad:**
 1. Compilar un vocabulario de un campo de especialidad
 2. Normalización (estandarización)
 3. Otros fines

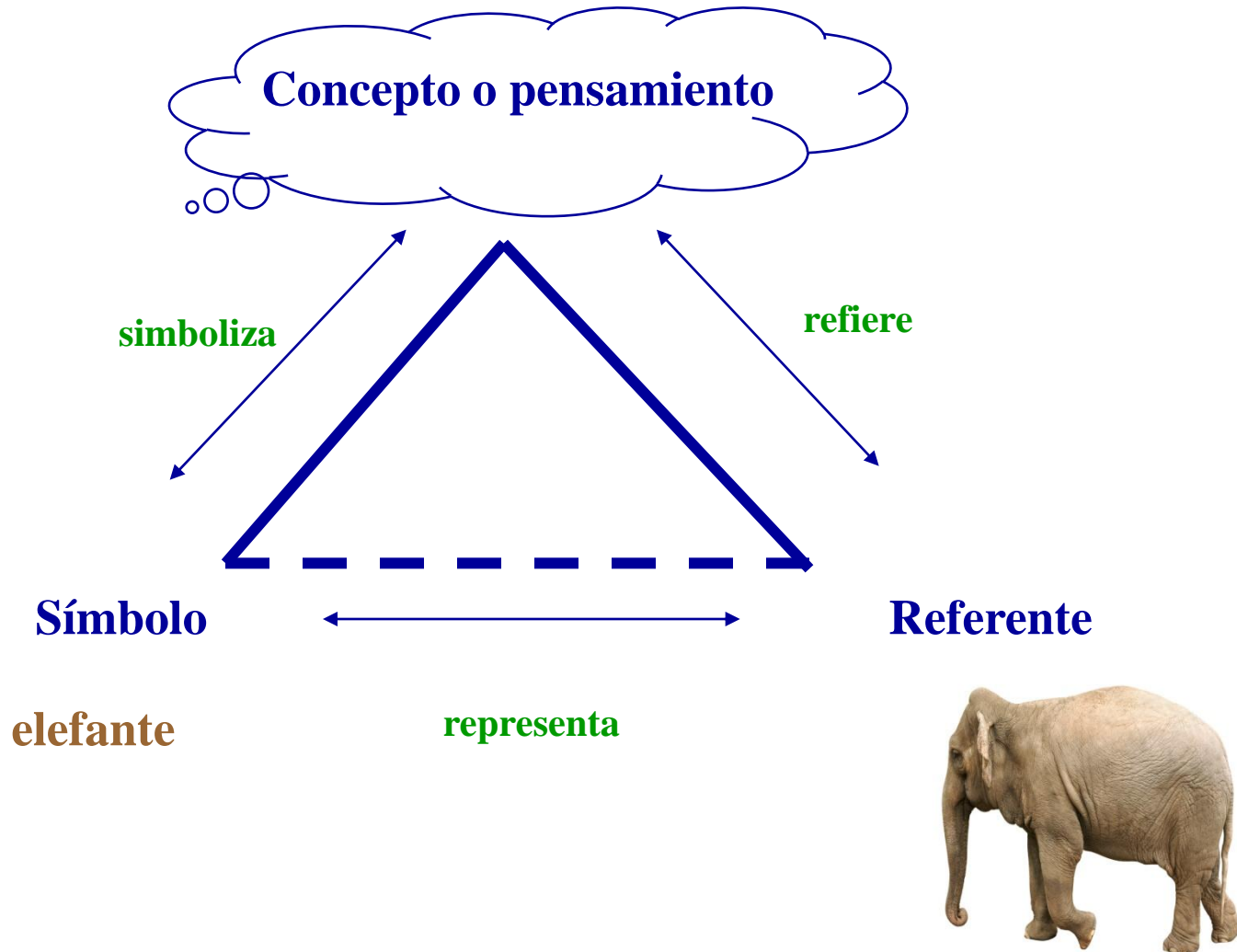
Fundamentos de la terminología

- El concepto
- La denominación: el término
- Relación entre denominación y concepto
- La definición

Definición de concepto

- *In general: Unit of knowledge created by a unique combination of characteristics. ISO 1087- 1 (2000)*
- *In terminology: Concepts shall be considered mental representation of objects within a specialized context or field*
ISO/DIS 704 (2006)
- **PERO** reciben la influencia del trasfondo social y /o cultural de cada momento, lo que puede conducir a clasificaciones distintas.
- Los **conceptos** tienen un papel poliédrico:
 - Unidades de **pensamiento**: sirven para representar y reconocer mentalmente el objeto
 - Unidades de **conocimiento**: representan el conocimiento de cada campo
 - Unidades de **comunicación**: comunican el conocimiento por medio de símbolos lingüísticos

Triángulo semántico de Ogden y Richards



Descripción de un concepto

- Se hace atendiendo a los rasgos comunes, propiedades o características

Por **intensión**

- Conjunto de propiedades esenciales que caracterizan a un concepto.
 - lobo: *mamífero* → *carnívoro* → *fisípedo* → *cánido*
- A mayor número de características comunes la intención es más restringida. En el caso extremo de intención restringida se identifica con un objeto único e individual.

Por **extensión**

- Totalidad de objetos a la que se aplica correctamente un término
- Un concepto genérico tiene una extensión amplia porque abarca muchos tipos de objetos
 - *Halógenos: flúor, cloro, bromo, yodo, astato*
 - *Poliedros regulares: tetraedro, hexaedro, octaedro, dodecaedro e icosaedro*

Características de un concepto

- Las características son en sí mismas conceptos.
- Según **la importancia** en la conformación de un concepto
 - **esenciales**: indispensables para entender y diferenciar un concepto
 - *El respaldo diferencia un taburete de una silla*
 - **complementarias**: *color, material,*
- Según **la relación** con el objeto representado
 - **intrínsecas**, que son propiedades observables:
 - Forma: ovalado, redondo, estrecho, ancho
 - Material: de madera, de piedra, metálico
 - Color: rojo, azul, verde, granate...
 - Posición: suspendido, colgante, vertical, inclinado
 - **extrínsecas**, relación del objeto con otros
 - de uso: aplicación, funcionamiento: *analógico, digital, híbrido.*
 - de origen: productor, inventor, proveedor, país, época



Subject field Computer hardware	Object (visual representation)	Concept Abstraction based on the set of all mechanical mice	Term: mechanical mouse	
Category	Property	Characteristic	Degree of essentiality	
Level of abstraction	concreteness	1. concreteness	essential	
Type of object	Input device	2. input device	essential	
colour	Ivory-coloured	3. Any colour	Non-essential	ISO /DIS 704, 2006
usage	Hand-manoeuvred along a hard, flat-surface	4. Has a ball on its underside on which the device rolls	essential	
composition	Has a metal ball on its underside on which the device rolls	5. Has a ball on its underside on which the device rolls	essential	
composition	Has three buttons	6. Has at least one button	essential	
connection	Connected to a computer by a connecting wire	7. May be connected to a computer by a wire	Non-essential	
function	Rollers detect the movement of the ball	8. Rollers (mechanical sensors) detect the movement of the ball	essential	
function	The ball controls the movement of a cursor on a computer display screen	9. Ball movement controls the movement of a pointer on a computer display screen	essential	

Relaciones entre conceptos: relaciones jerárquicas

Estrecha relación entre concepto y característica (predicado, rasgo, faceta)

A.RELACIONES GENÉRICAS (género-especie) IS_A

- Se basan en la semejanza de los conceptos (uno de los conceptos incluye al otro)
 - **vertical:** hiperonimia- hiponimia; superordinado –subordinado
 - **horizontal:** dos nociones específicas del mismo genérico, con alguna característica que los diferencia
- En los tesauros se limitan a:
 - *Broader than* (BT)
 - *Narrower than* (NT)
 - *Associated to* (AT)

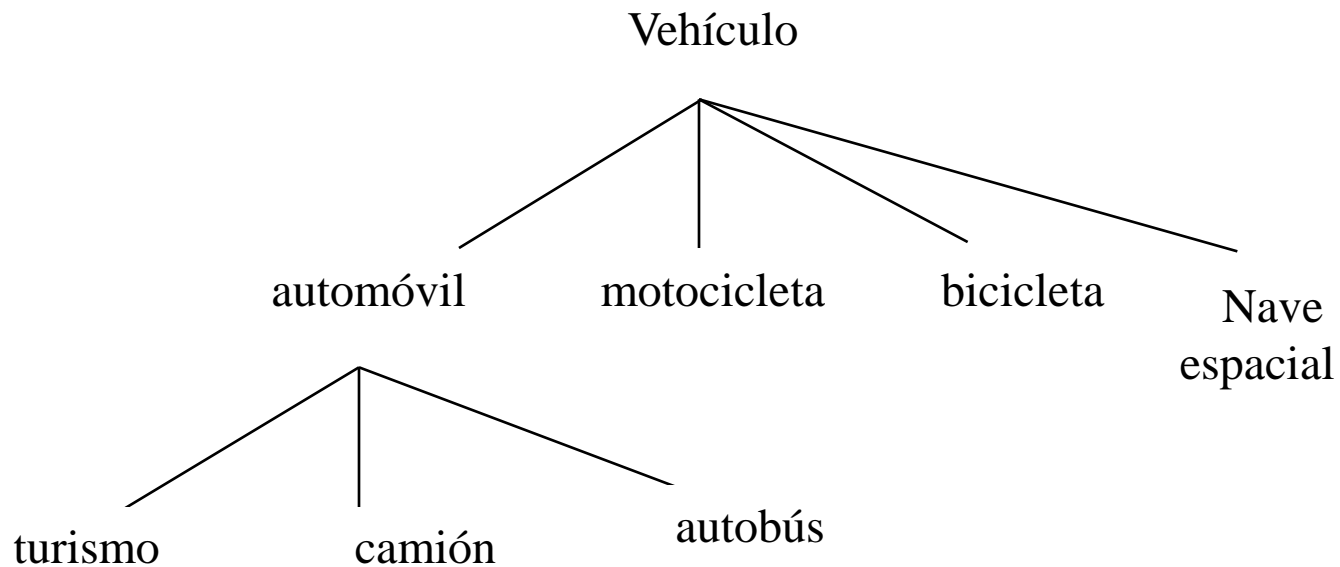


B. RELACIONES PARTITIVAS

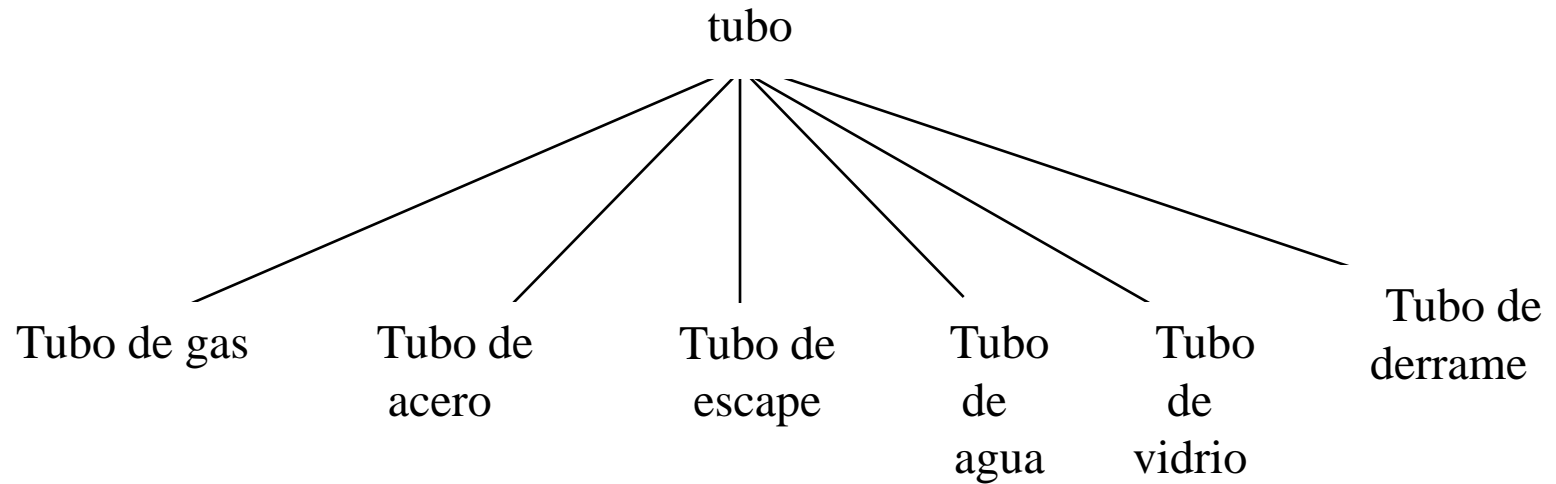
- Se llaman también relaciones **meronímicas** o parte-todo. PART_OF
 - *Coche: ruedas, asientos, volante, puertas, caja de cambios...*
- Distintos tipos de relaciones meronímicas



Relación vertical



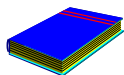
Relación horizontal: cohipónimos



Relaciones meronímicas

Relación	Ejemplo
componente - objeto	<i>pedal - bicicleta</i>
miembro - colección	<i>barco - flota</i>
porción - masa	<i>rebanada - pan</i>
material - objeto	<i>acero - coche</i>
fase - actividad	<i>pagar - comprar</i>
lugar - área	<i>oasis - desierto</i>

Tabla II.2: Modelo de Winston *et al.* (1987)



Climent, S. 1999 *Individuación e información parte-todo. Representación para el procesamiento computacional del lenguaje*

Relaciones no jerárquicas o complejas (relaciones *ad-hoc*)

- *Causado por*: (lluvia ácida- explosión nuclear,)
- *Producto de*: (papel- pasta de madera)
- *Propiedad de* (compresibilidad –gas)
- *Medida cuantitativa* (temperatura-calor)
- *Instrumento para* (computador- procesamiento de datos)
- *Contra-agente para* (insecticida-insectos)
- *Contenedor de* (caja de herramientas, herramientas)
- *Método de* (perforación al diamante- perforación)
- *Material para* (hierro-construcción de puentes)
- *Lugar para* (mina de carbón- explotación hollera)
- *Asociado con* (producción-consumo)

¿Cómo expresamos los conceptos?

- En **lenguaje natural**:
 - Términos (simples o multipalabra): Designación o denominación de una o más palabras que representa un concepto de un dominio de conocimiento
 - Definiciones
 - Glosas, etc.
- En **lenguaje artificial**
 - Códigos
 - Fórmulas
- En un **recurso multimedia**:
 - Iconos
 - Fotos
 - Diagramas
 - Gráficos
 - Video-clips
 - Audio-clips
 - Otras representaciones multimedia

Tipos de términos

- Según el **origen**:

- **PRÉSTAMOS** de otra lengua: *hardware, software, football, cookies, folksonomies*
- **ADAPTACIONES**: *formatear, inicializar, fútbol, etc*
- **CALCOS de estructura**: *inteligencia artificial, lógica difusa, programación orientada a objetos, anotación social*
- **CALCO semántico**: *aplicación, utilidades, editar, icono, ratón, menú,*

- Según la **formación**:

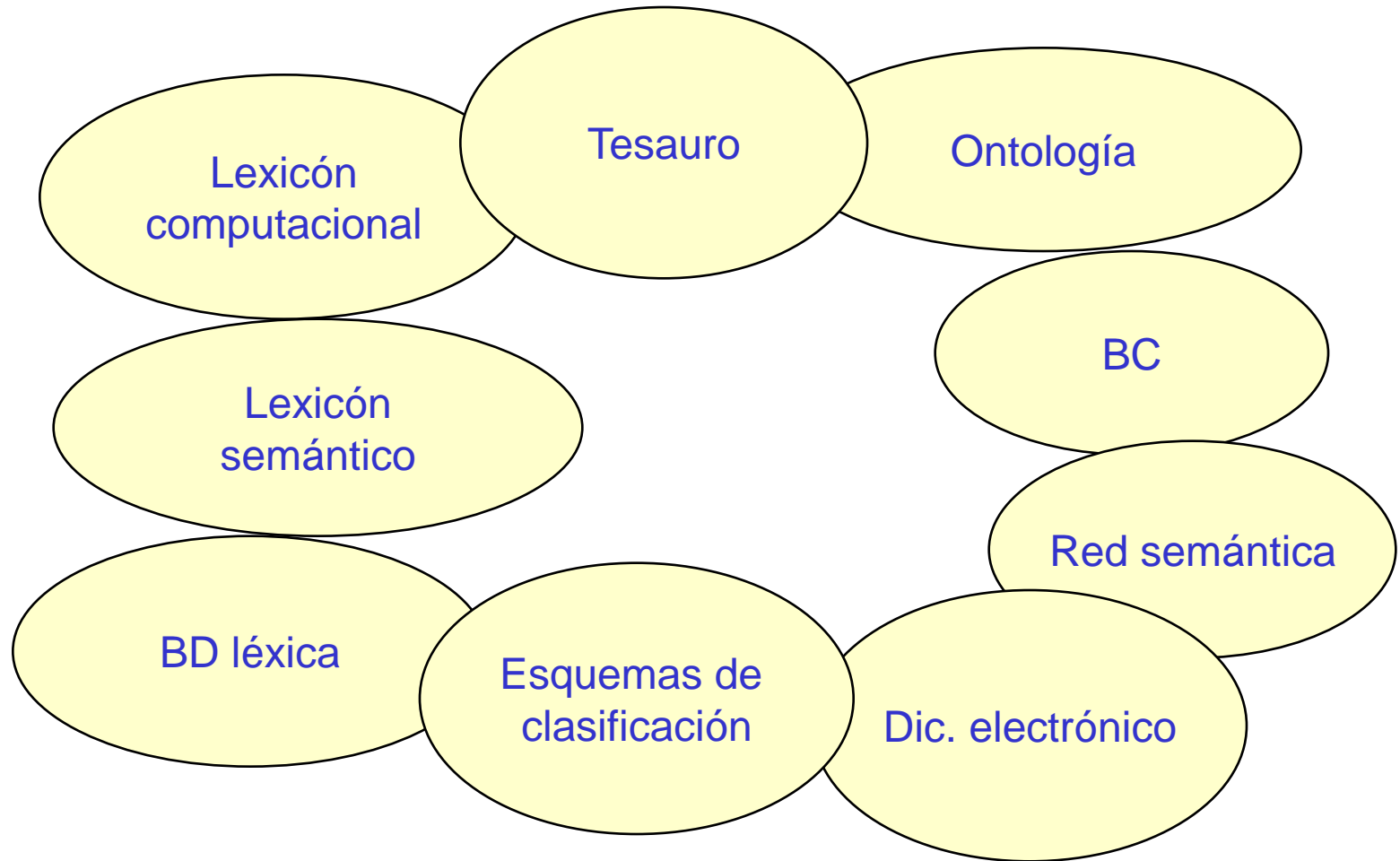
- Unidades léxicas **simples** o unidades terminológicas: *Programa, aplicación, icono, menú, ratón*
- Unidades polilexemáticas, o **unidades multipalabra**: *lenguaje de programación, lenguaje de alto nivel, programación orientada a objetos*

- Según los **componentes**:

- **SUFIJACIÓN**:
 - *teca/tica*: *animática, burórica, indumática, ofimática, robótica, telemática, turismática*
 - *- itis*: a) inflamación: *bronquitis, faringitis, amigdalitis, otitis*
 - b) obsesión: *madriditis, mamitis, futbolitis*
- **PREFIJACIÓN**: *ciber: cibercafé, cibernauta, cibermedicina, ciberdelito, etc*
- **COMPOSICIÓN**: *salvapantallas, reposapiés, sujetamanos, radiotelevisión,*
- **ABREVIACIÓN**: *PC, PDF, TCP/IP, MS-DOS, MP3, wysiwyg,*
- **CONVERSIÓN**: *: download, input, output,*
- **NEOLOGIZACIÓN**: *autoedición, “gustomizar”,*
- **METAFORIZACIÓN**: *paquete de mejoras salariales, navegar, autopistas de la información, papelera, escritorio, bajar de la red, machacar un fichero, caerse el sistema, etc*

Recursos lingüísticos y ontológicos

¿caos terminológico?



Algunas definiciones

- **Thesaurus:** Vocabulario controlado y organizado de términos con sus relaciones que se aplican a un dominio del conocimiento.
 - **Lexicon:** colección de palabras y su información léxica elaborado para diversas aplicaciones de procesamiento de lenguaje humano.
 - **Base de datos:** conjunto de datos que posibilita su almacenamiento y recuperación.
 - **Ontología:** representación formal de la semántica de las unidades gramaticales y de la información que se maneja.
- word-sense-entry** →
[- ORTHOGRAPHY : string
- WORD-MEANING : word-meaning-id+
SYNONYMS : word-meaning-id*
NEAR-SYNONYMS : word-meaning-id*
HYPONYMS : hyponym*
HYPERONYMS : hyperonym*
ANTONYMS : antonym*
MERONYMS : meronym*
HOLONYMS : holonym*
QUANTIFICATION : quantification*
COLLOCATIONS : collocation*
SEMANTIC-FRAME : sem-frame
ACTIONALITY : actionality
ENTRY-CREATOR: (HUMAN | MACHINE)
IS_VALIDATED: Boolean]

(fonológica, morfológica, sintáctica, semántica y pragmática)

Diccionario terminológico enriquecido con información ontológica

- Dancette, J. & C. Réthoré 2000
- *Dictionnaire Analytique de la Distribution. Analytical Dictionary of Retailing.* (Inglés-francés) Les presses de l'université de Montréal
- Diccionario sobre « Venta al por menor »
- Objetivos:
 - Estimular la creatividad del traductor con **información enriquecida ontológicamente** sobre el tema, en francés, lengua meta
 - Optimizar la comprensión, **estimulando la red semántica** en el cerebro del traductor

Entrada para “Label”, etiqueta

definición

relaciones inter-conceptuales

información lingüística

referencias cruzadas

precisiones semánticas

información complementaria

contextos y ejemplos

LABEL tag étiquette

DÉFINITION

Document d'identification du produit qui lui est apposé ou y est attaché et qui en décrit les caractéristiques (nature, prix, provenance, marque, etc.).

PRÉCISIONS SÉMANTIQUES

Les informations contenues sur l'étiquette sont destinées au caissier (prix, codification) et au consommateur (ingrédients, mode d'emploi, etc.). L'étiquette peut présenter un graphisme plus ou moins élaboré.

Depuis les années 1970, l'étiquette comprend généralement un code-barres (⇒ BAR CODE). Le code-barres contient des informations telles que la description et le prix du produit, qui seront lues à l'aide d'un lecteur optique (⇒ OPTICAL READER).

On distingue l'étiquette adhésive (*self-sticking label*), ou étiquette autocollante, et l'étiquette mobile (*tag*) attachée au produit.

Sur l'étiquette de prix (*price tag*) ne figure que le montant à payer.

Aujourd'hui, certaines grandes surfaces n'indiquent plus le prix sur le produit mais plutôt sur la gondole (⇒ GONDOLA), avec des étiquettes de gondole (*shelf labels*, *shelf edge labels*) ou des étiquettes électroniques de gondole (⇒ ELECTRONIC SHELF LABEL).

RELATIONS INTERNATIONNELLES

Le terme anglais TAG désigne une étiquette que l'on peut facilement enlever, ce qui n'est pas le cas de LABEL.

Ne pas confondre l'anglais LABEL avec son homonyme label, qui a le sens de marque (⇒ BRAND), comme dans le terme PRIVATE LABEL (marque de distributeur).

Le terme anglais *quality label* (label de qualité), ou *seal of quality*, ne désigne pas l'étiquette

mais la conformité du produit à des caractéristiques ou à des qualités définies.

COMPLÉMENTS D'INFORMATION

L'idée de remplacer le marquage (*marking*) des prix par des codes-barres sur les produits vendus en magasin a suscité beaucoup de craintes chez les consommateurs. Ils y voyaient un préjudice à l'exercice de leur droit à l'information. Pour répondre à cette inquiétude certains distributeurs mettent des lecteurs de codes-barres (⇒ BAR-CODE READER) à la disposition de leurs clients, dans les divers rayons de leurs magasins. Par contre, il est admis que le code-barres présente des avantages par rapport à l'étiquette de prix, dont l'exactitude et la rapidité des enregistrements à la caisse.

Les producteurs ont l'obligation, en vertu de la Loi sur la protection du consommateur (*Consumer Protection Act*), de répertorier sur l'étiquette tous les ingrédients contenus dans le produit alimentaire.

INFORMATIONS LINGUISTIQUES

Marquage a été normalisé par l'OLF (1983) et l'ISO (1987).

étiquette, tag, label
étiqueteuse, labeteur, label machine

CONTEXTES

But it wasn't until 1900 that [he] put the first Polar label on a bottle of cool, naturally purified water taken directly from one of these springs on his property.

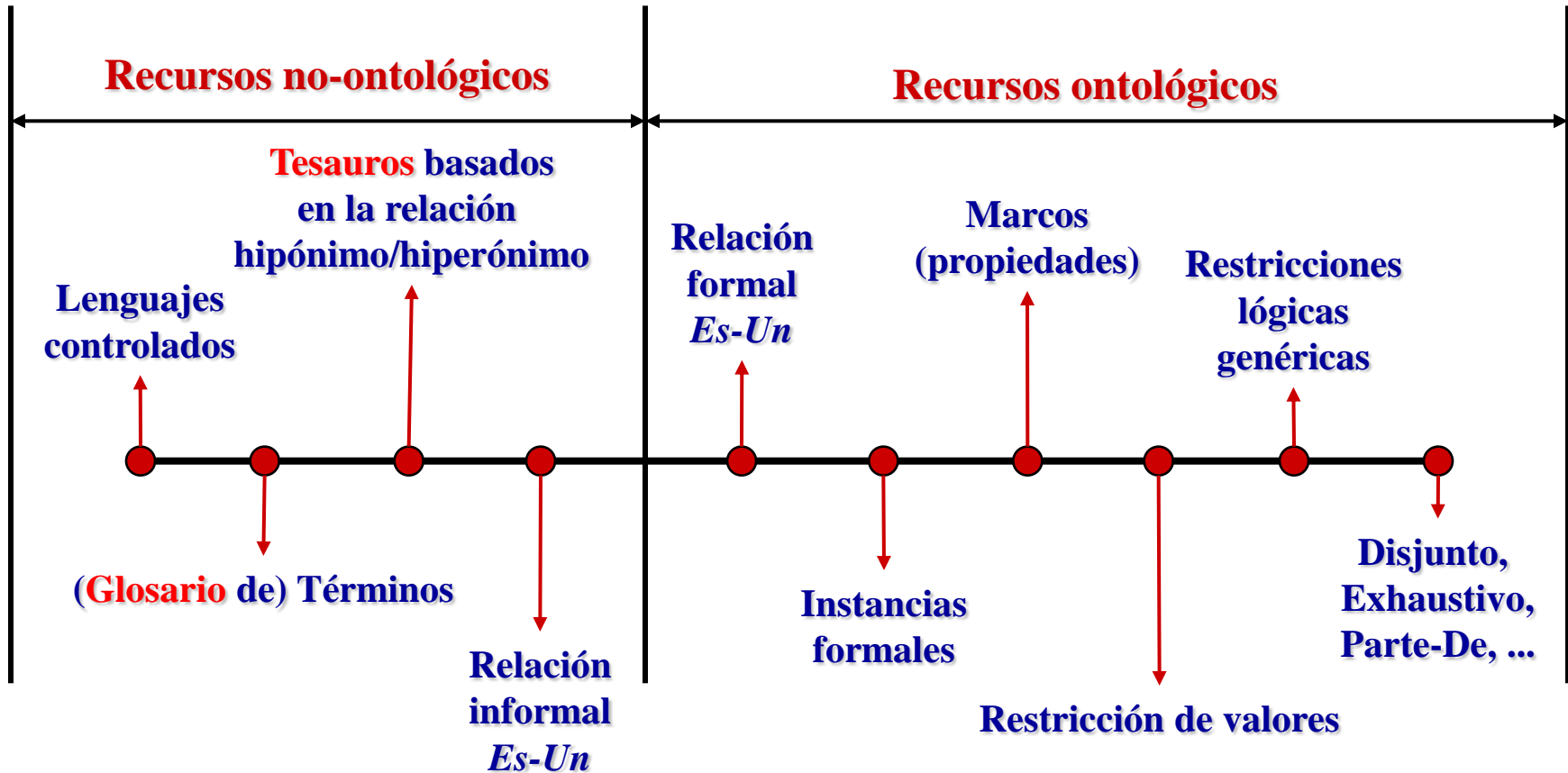
<<http://www.water.com/polar/index.html>> (30-3-99)

Dans ce but, la réglementation mise au point par les organismes de la CEE et par l'administration française prévoit sur chaque étiquette la présence d'un certain nombre de mentions obligatoires, en fonction de la catégorie du vin.

<<http://www.vin-champagne.com/etiq.htm>> (30-3-99)

LADS Voir LIMITED ASSORTMENT DISCOUNT STORE

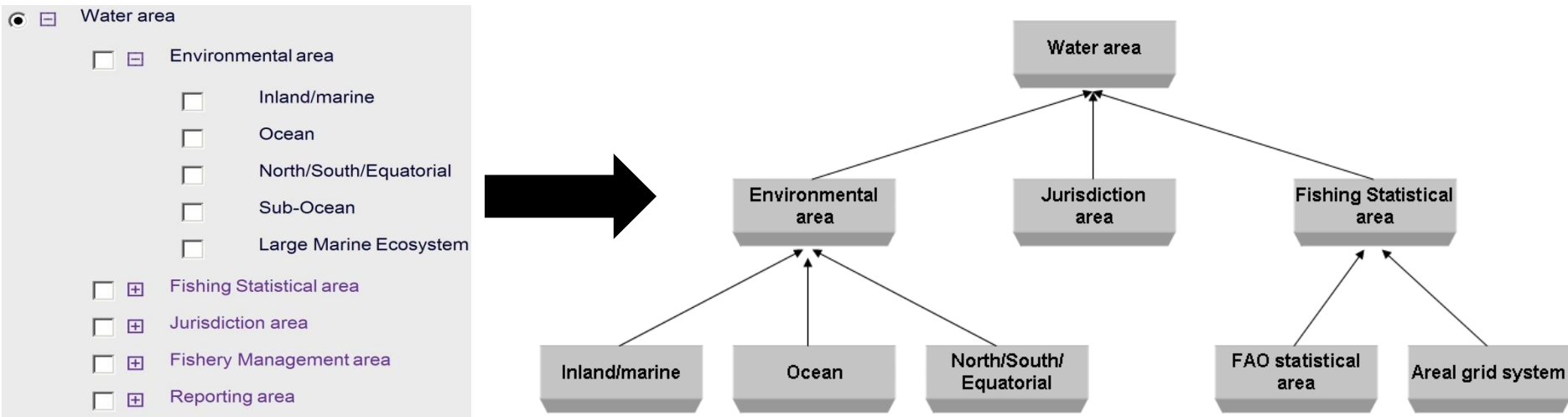
Clasificación de ontologías y otros recursos no ontológicos (Lassila & McGuinness)



- Gómez-Pérez, A., Fernandez-Lopez, M., Corcho, O. (2003) *Ontological engineering: with examples from the areas of knowledge management, e-commerce and the Semantic Web*. Londres:Springer Verlag London Ltd.

Esquemas de clasificación

- **Esquema de clasificación**¹: Información descriptiva de una organización o división de los objetos en grupos atendiendo a las características comunes de los objetos/conceptos”: *water area classification scheme*².
- Se conocen como **Sistemas de la Organización del Conocimiento, SKO** (Systems of Knowledge Organization)
- Promueven la gerencia del conocimiento.



1. International Standard Organization (ISO). Information technology – Metadata registries – Part 1: Framework Report ISO/IEC FDIS 11179-1.
2. <http://www.fao.org/figis/servlet/RefServlet>

Glosarios

- Lista de términos, generalmente con definiciones.
- Los términos pueden ser de un campo específico
- Se definen dentro de un dominio específico
- Se trata de evitar la polisemia.

U.S. ENVIRONMENTAL PROTECTION AGENCY

List of all Topics

 [Bookmark](#)

[Recent Additions](#) | [Contact Us](#) Search: [Advanced search](#)

You are here: [EPA Home](#) » [Browse EPA Topics](#) » List of all Topics

List of all Topics

Alphabetical List of All Topics

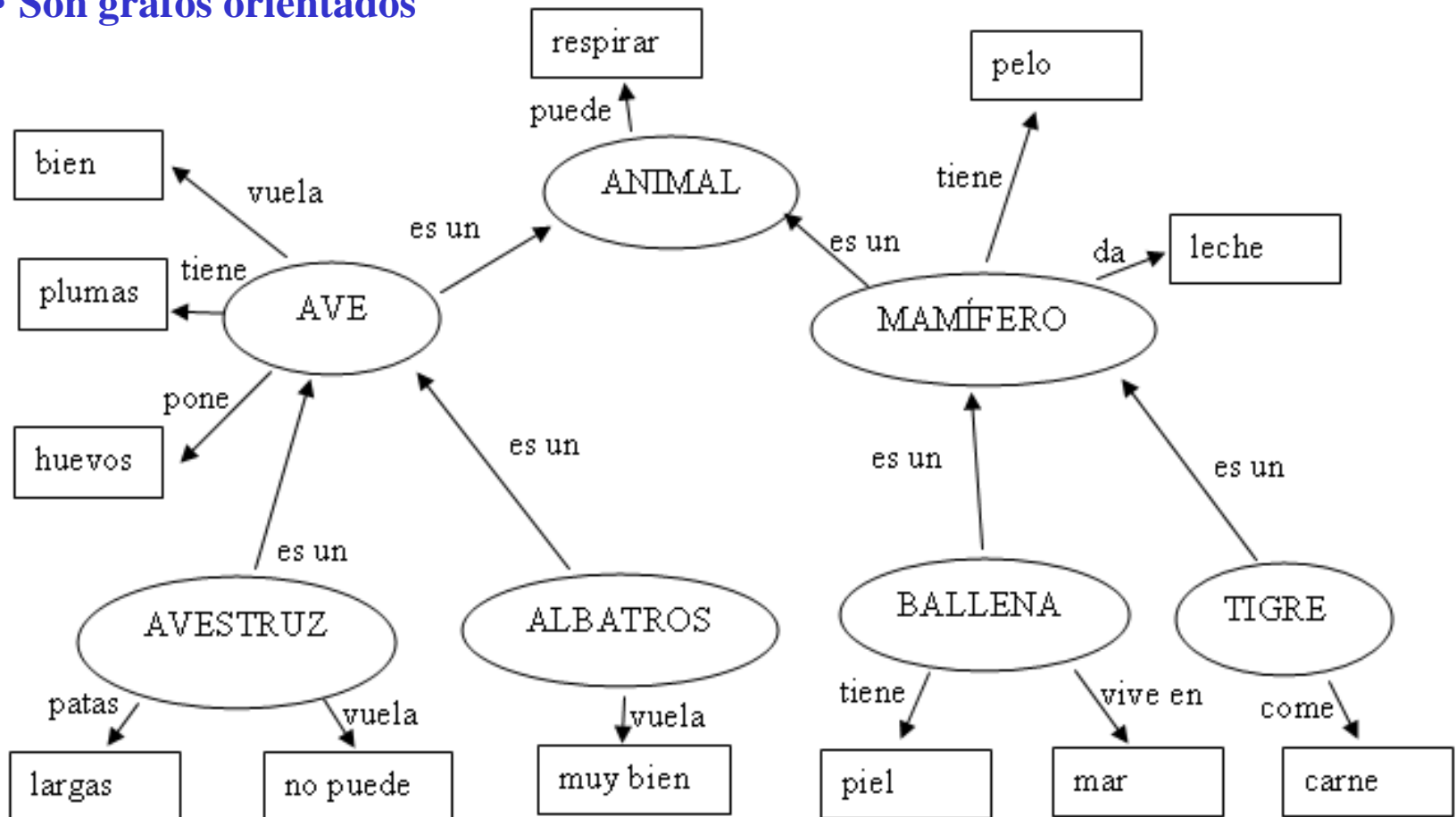
This page organizes topics into alphabetical order, like a book's index. You can also browse topics organized into [broad categories, like a book's table of contents](#).

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

A	Back to Top
Abandoned Mine	Ecosystems > Mines > Abandoned Mine
Abandoned Properties	Cleanup > Storage Tanks > Abandoned Properties
Abatement	Air > Air Pollution Control > Abatement
Abatement	Treatment/Control > Air Pollution Control > Abatement
Above Ground Tanks	Industry > Storage Tanks > Above Ground Tanks
Accident Preparedness	Emergencies > Accidents > Accident Preparedness
Accident Prevention	Emergencies > Accidents > Accident Prevention
Accidents	Emergencies > Accidents
Accomplishments	Cleanup > Accomplishments
Acetone	Pollutants/Toxics > Soil Contaminants > Acetone
Acid Rain	Air > Air Pollution Effects > Acid Rain
Acute Exposure	Human Health > Exposure > Acute Exposure
Administrative Civil Enforcement	Compliance And Enforcement > Civil Enforcement > Administrative Civil Enforcement
Administrator	Environmental Protection Agency > Administrator
Advisories	Human Health > Advisories
Advisory Committees	Environmental Protection Agency > Science Advisory Board (SAB) > Advisory Committees

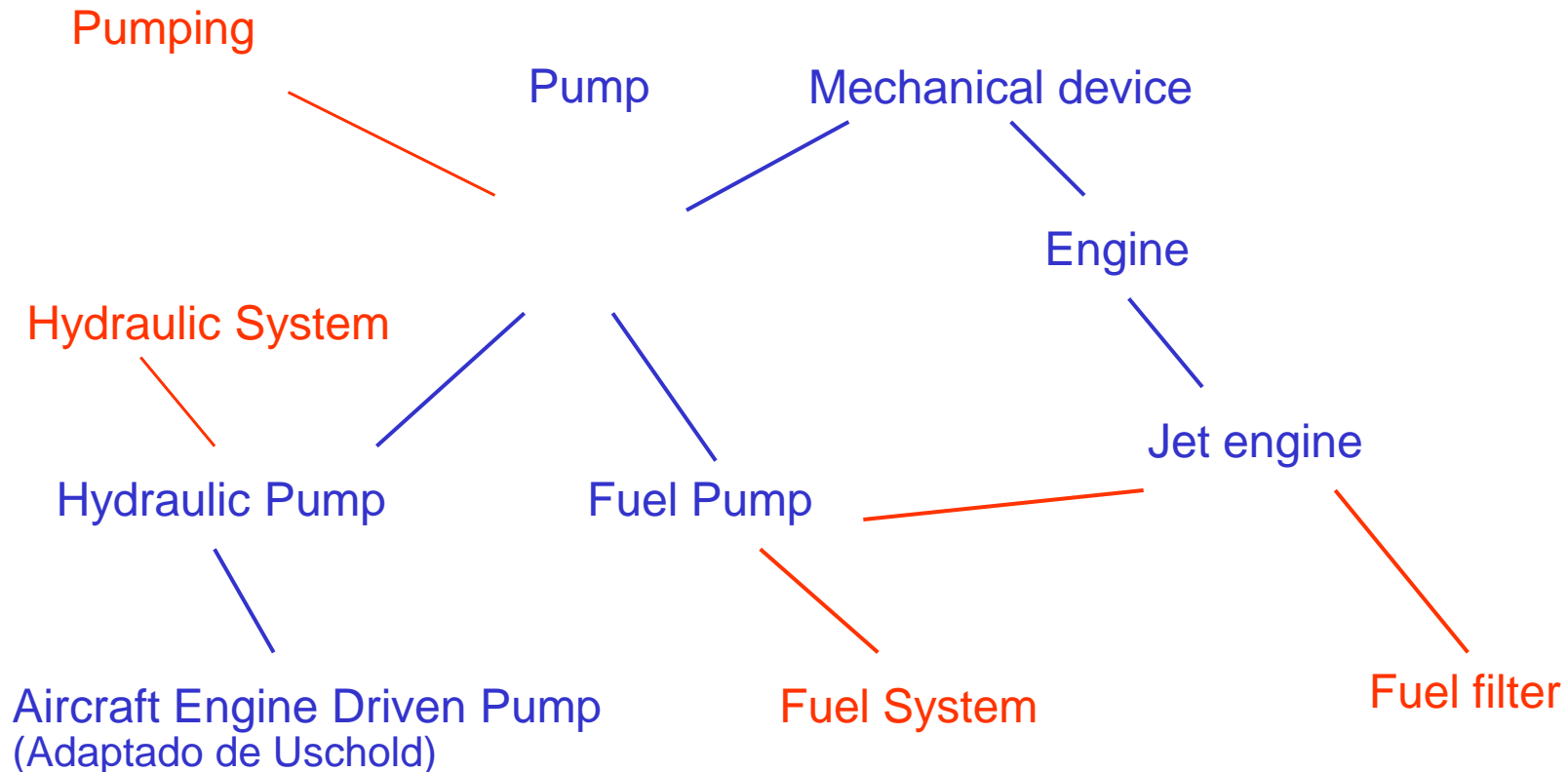
Red semántica

- Estructura de conceptos en una red, con nodos y relaciones, no jerarquía
- Pueden incluir relaciones BT, NT, RT u otras asociativas o jerárquicas
- Son grafos orientados

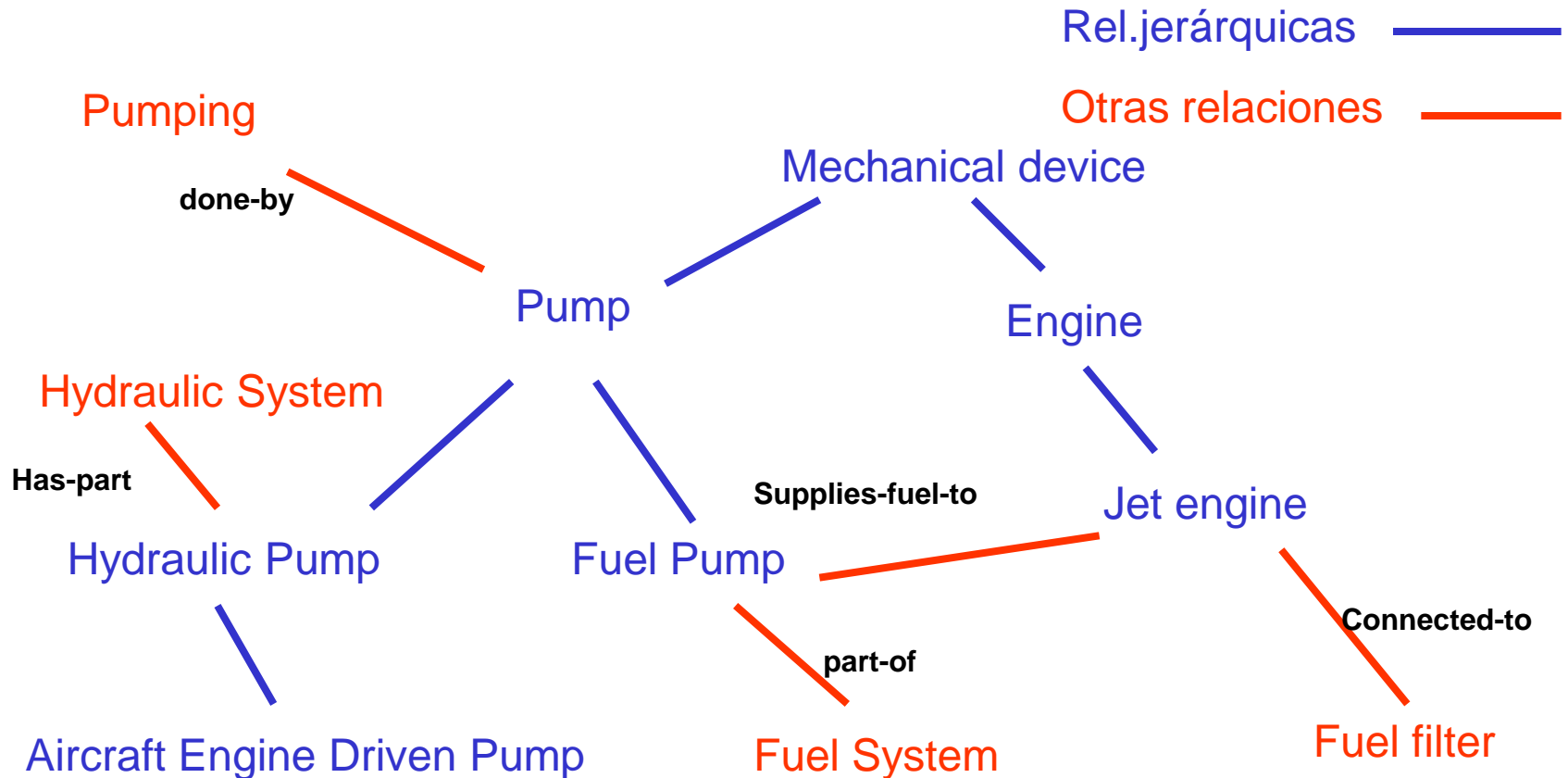


Ejemplo de tesauro

- Estructura conceptual
 - Organización semántica mediante
 - Explicitación de relaciones conceptuales
 - Restricción del significado de los términos que los representan
 - Se utilizan para indexar y recuperar información en grandes BD: AGROVOC
- Término relacionado: —



Ontología “ligera”: Taxonomía estricta + relación formal



Adaptado de Uschold

nouns

round object that is hit or thrown or kicked in games; "the ball travelled 90 mph on his serve"; "the mayor threw out the first ball"; "the ball rolled into the corner pocket"

- ☐ Hypernyms (... is kind of)
- ☐ Hyponyms (kinds of ...)
- ☐ Antonyms (opposites of ...)
- ☐ Meronyms (parts of ...)
- ☐ Holonyms (... is part of)
- ☐ Related Verbs
- ☐ Related Adjectives

a solid ball shot by a musket; "they had to carry a ramrod as well as powder and ball"

an object with a spherical shape; "a ball of fire"

verbs

adjectives

SMART THESAURUS MUSIC is organized with respect to the elements of the lexicon, and is supported by the lexical reference system: nouns, verbs, and adjectives.

Thus, three different tabs are presented to you. A simple click opens a certain tab, and, offers its content: a list of meanings, each representing a certain **synset** of the search term. In order to find out which element of the web refers to which meaning or synset, please click on it. Two things happen:

The meaning gets marked (with red color) and so do the corresponding elements of the web. A certain circle or sphere, representing a specific synset, becomes marked red, and also all of the edges that point to the set of synonyms (representing the synset). In addition, the 'meaning' opens its content and presents a list of lexical pointers associated with the selected part of speech. A click on one of these pointers, e.g. hypernym, lets you explore the broader terms associated with the selected synset.

SMART THESAURUS MUSIC supports the following lexical relationships:

Noun

- [1] Hypernym or broader term (...is a kind of)
- [2] Hyponym or narrower term (kinds of ...)
- [3] Antonym (opposites of ...)
- [4] Meronym (parts of ...)
- [5] Holonym (... is a part of)
- [6] Related verbs
- [7] Related Adjectives

Verb

- [1] Hypernym or broader term (...is a kind of)
- [2] Hyponym or narrower term (kinds of ...)
- [3] Related verbs
- [4] Related nouns

Tipos de relaciones

Cómo representar las relaciones de meronimia

iglesia

<Entrada> (2)
<Synset> : parroquia-#1
<Atributos>
 <Categoría> : Sustantivo
 <Tipo> : Clase
 <Léxico> : General
 <Glosa> : Iglesia que tiene a su cargo administrar los sacramentos a los fieles que habitan en determinado territorio o distrito.
<Relaciones>
 <Relaciones de Clase>
 <Clave semántica> : iglesia, eclesiástico
 <Hiperónimo> : iglesia-#1
 <Relaciones de Meronimia>
 <Meronimia Tipo> : mero-1
 <Merónimo> : abadía-#1
 <Meronimia Tipo> : mero-1
 <Merónimo> : anejo-#3
 <Meronimia Tipo> : mero-1
 <Merónimo> : anteiglesia-#1
 <Meronimia Tipo> : mero-1
 <Merónimo> : feligresía-#1
 <Meronimia Tipo> : mero-1
 <Merónimo> : iglesia-#1
 <Meronimia Tipo> : mero-1
 <Merónimo> : pila-#4
 <Meronimia Tipo> : mero-1
 <Merónimo> : rectoral-#1
 <Meronimia Tipo> : mero-1
 <Merónimo> : sagrario-#1

<Entrada> (3)
<Synset> : parroquia-#3, feligresía
<Atributos>
 <Categoría> : Sustantivo
 <Tipo> : Clase
 <Léxico> : General
 <Glosa> : El territorio, los fieles o ambas cosas juntas, de cada parroquia.
<Relaciones>
 <Relaciones de Clase>
 <Clave semántica> : territorio, fieles
 <Hiperónimo> : fiel-#1
 <Relaciones de Holonimia>
 <Holonimia Tipo> : holo-0
 <Holónimo> : parroquia-#1
 <Relaciones de Meronimia>
 <Meronimia Tipo> : mero-6
 <Merónimo> : parroquia-#1

Territorio,
los fieles o
ambos

Parroquia y sus
relaciones

Díez Orzas(1999)
*LA RELACIÓN DE MERONIMIA
EN LOS SUSTANTIVOS DEL LÉXICO
ESPAÑOL:
CONTRIBUCIÓN A LA
SEMÁNTICA COMPUTACIONAL*

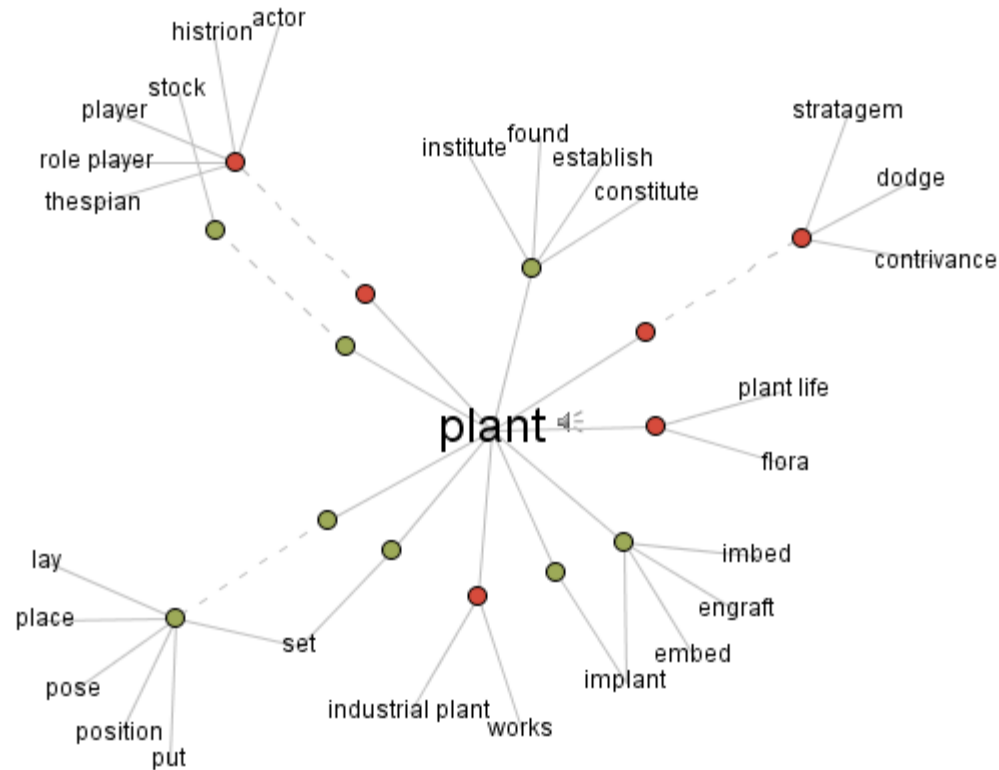
Entrada> (4)
<Synset> : parroquia-#2, feligresía
<Atributos>
 <Categoría> : Sustantivo
 <Tipo> : Clase
 <Léxico> : General
 <Glosa> : El territorio, los fieles o ambas cosas juntas, de cada parroquia.
<Relaciones>
 <Relaciones de Clase>
 <Clave semántica> : territorio, fieles
 <Hiperónimo> : territorio-#1
 <Relaciones de Meronimia>
 <Meronimia Tipo> : mero-6
 <Merónimo> : parroquia-#1

<Entrada> (5)
<Synset> : parroquia-#4
<Atributos>
 <Categoría> : Sustantivo
 <Tipo> : Clase
 <Léxico> : General
 <Glosa> : División administrativa, dentro del municipio.
<Relaciones>
 <Relaciones de Clase>
 <Clave semántica> : municipio, Galicia
 <Hiperónimo> : división administrativa-#1
 <Relaciones de Holonimia>
 <Holonimia Tipo> : holo-1
 <Holónimo> : municipio-#1

División
administrativa

Ejemplo de Tesauro visual

NOMBRES



VERBOS

<http://www.visualthesaurus.com/store/>

Lexicones

Dos tipos generalmente

- *De carácter general*
 - Cubren la lengua utilizada en todos los contextos
- *De carácter específico*
 - Cubre un dominio con información más precisa en la entradas principales

Consecuencias de ambos tipos de lexicones

- Cantidad y calidad de la información (granularidad)
- Complejidad en el diseño
- Complejidad en el proceso de creación

¿Cómo se utilizan los lexicones en PLN?

- Contienen toda la información lingüística necesaria para construir representaciones de significado

Lexico

n Account *n.* Domain [**financial**]

Account *v.* ...

Bank_1 *n* domain [**financial**]

Bank_2 *n* domain [geography]

Money *n.* domain [**financial**]

bank.....
.....account
.....
money.....



Went *v*past GO

Go *v.* (NP_SUNJ ((role AGENT) (sem + animate))

(VP ((verb GO)

(PP ((prep TO)

(NP ((role TARGET) (sem +loc))))

John *n.* sem: human

Store *n.* sem: loc

Topic=
financial

John went to the store

GO
AGENT John TARGET store

Adaptado de Nancy Ide

Información en lexicones

- Diversos tipos:

- Información **morfosintáctica**:



- <http://www.mat.upm.es/~aries/description.html>

- <http://www.ims.uni-stuttgart.de/projekte/CorpusWorkbench/CQP-HTMLDemo/PennTreebankTS.html>

- Información **semántica**

- Rasgos semánticos: ser vivo, animado, humano, objeto físico, objeto mental, .
 - Entradas diferentes para sentidos distintos
 - Relaciones semánticas: sinónimos, cuasi-sinónimos, antónimos, etc
 - Relaciones jerárquicas: part_of, kind_of, etc...

- Información **sintáctico-semántica**: colocaciones



- Información sobre un **dominio**

- **Definiciones**



Lexicón sobre “Existencia” (Faber y Mairal, 1999)

- 1.1 General: To exist / to continue to exist.
- 1.2 To begin to exist [*be, live*]
 - 1.2.1 To cause something to exist [*create, make*]
- 1.3 To exist in the perception of others [*appear*]
 - 1.3.1 To cause something to exist in the perception of others [*show*]
- 1.4 To exist in time (becoming real) [*happen*]
 - 1.4.1 To cause something to exist in time [*induce, provoke*]
 - 1.4.1.1. To cause something to exist in time in a particular way [*precipitate, hasten*]
 - 1.4.1.2 . To cause something to happen, making it possible [*allow, permit*]
 - 1.4.1.3. To cause something not to happen [*prevent, avoid, stifle, smother*]
- 1.5 To exist as something
 - 1.5.1 To exist as the representation of something else [*represent, express*]
 - 1.5.1.1 To cause something to exist as a representation of something [*copy, reproduce*]
 - 1.5.2 To exist as a part of something [*comprise, constitute*]
- 1.6 To begin to exist [*start, commence, be born*]
 - 1.6.1 To cause to begin to exist [*start, commence*]
 - 1.6.1.1 To cause to be born [*abort*]
 - 1.6.2 To begin to exist in the perception of others [*arise, form*]
 - 1.6.3 To begin to exist in time (becoming real) [*start, originate*]
 - 1.6.3.1 To cause something to begin to exist in time [*start, initiate*]
- 1.7 To continue to exist [*last, endure*]
 - 1.7.1 To stop something from continuing [*interrupt*]
- 1.8 To stop existing [*die*]
 - 1.8.1 To cause somebody/something to stop existing [*kill, murder*]
 - 1.8.2 To stop existing in the perception of others [*disappear, vanish*]
 - 1.8.2.1 To cause something to stop existing in the perception of others [*erase, delete*]
 - 1.8.2.2 To stop existing in time [*end, finish, cease*]
 - 1.8.2.3 To cause something to stop existing in time [*end, finish, cease*]





ARIES Natural Language Tools

The *ARIES Natural Language Tools* make up a lexical platform for the Spanish language. These tools can be integrated into NLP applications. They include: a large Spanish lexicon, lexical maintenance and access tools and morphological analyser/generator.

Non-exclusive, non-transferable licenses are available for the following components:

[The Prolog GRAMPAL analyser/generator](#)

A public domain demonstration system written in Prolog of our morphological treatment and lexicon. It includes a small demo lexicon, a DCG grammar for word formation and some predicates to test both analysis and generation. It runs under Sicstus Prolog 2.1.9.

The Prolog GRAMPAL dictionary

A collection of Prolog predicates suitable for use with the public domain GRAMPAL DCG grammar. It is capable of generating/recognizing well formed inflected forms for verbs, nouns and adjectives. It has no adverbs, determiners, conjunction, prepositions, etc. It does not treat clitic pronoun attachment nor derivatives.

The expanded ARIES dictionary

A collection of expanded entries (allomorphs) with morphological information. It contains a full set of morphemes dealing with clitic pronoun attachment (but without verb marking for correct attachments). It includes information about some derivative morphological processes (inflected adjectives from past participles and adverbs ended in "-mente" from adjectives).

The source ARIES lexical base

A collection of inflectional models, rules for off-line computing of allomorphs, unexpanded lemma entries, lexicalized irregular words. It is the most complete source of information we have available and the most useful for dictionary maintenance. A tool for expanding the source dictionary to the expanded dictionary is also provided. The current size of this lexicon is 38,500 lemma entries (21,000 nouns, 10,000 adjectives, 7,500 verbs and 500 auxiliary words) plus more than 600 inflectional morphemes.

Access tools

The C/C++ programming interface for lexical access to the ARIES dictionary: It is a set of tools and libraries to build trie indexes to the allomorph dictionary and to retrieve them by an application.

Morphological analyser

The C/C++ morphological analyzer that makes use of the lexical interface mentioned above. This permits to improve efficiency by integrating word segmentation with lexical access also. By now, it is a (pseudo)-unification chart based parser for context-free morphological grammars.

Glosario del CINDOC

[Tesauros](#)[Inicio](#)[Alfabético](#)[Búsquedas](#)

Glosario de Máquinas Herramienta

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Listado alfabético de terminos [#1] *(no-descriptores en cursiva)*

[1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [Siguiete] [Fin]

a prueba de empleo incorrecto
abrazadera para tubos
accionamiento de la mesa
accionamiento del avance
accionamiento eléctrico
accionamiento forzado
accionamiento hidráulico
accionamiento individual
accionamiento neumático
accionamiento por cuerda
accionamiento por fricción
accionamiento por grupos
accionamiento por poleas escalonadas
accionamiento por trinquete
aceleración
acoplador roscado para tuberías
acoplamiento de desembrague
acoplamiento de ejes

¿Qué es una ontología?

- “An ontology is similar to a dictionary or glossary, but with greater detail and structure that enables computers to process its content. (IEEE Standard Upper Ontology Working Group)
- “An ontology consists of a set of concepts, axioms, and relationships that describe a domain of interest.” SUMO ontology
<http://ontology.teknowledge.com/>

Las ontologías en los lexicones

- Representar el significado de una palabra requiere, al menos:
 - Diferenciar entre los diferentes sentidos:
 - *cabo de una vela, cabo geográfico, cabo del ejército, etc.*
 - Captar las inferencias semánticas posibles:
 - *piedra **implica** inanimado, ser humano **implica** animado*
 - Captar las implicaciones posibles
 - *Caminar **supone** movimiento, roncar **supone** dormir*
 - Representar la semejanza de significado con otras palabras:
 - Glosas, conjuntos semánticos,

Enfoques que incorporan terminología en las ontologías

□ Tres tipos de enfoques:

□ Enfoque 1:



□ Enfoque 2:

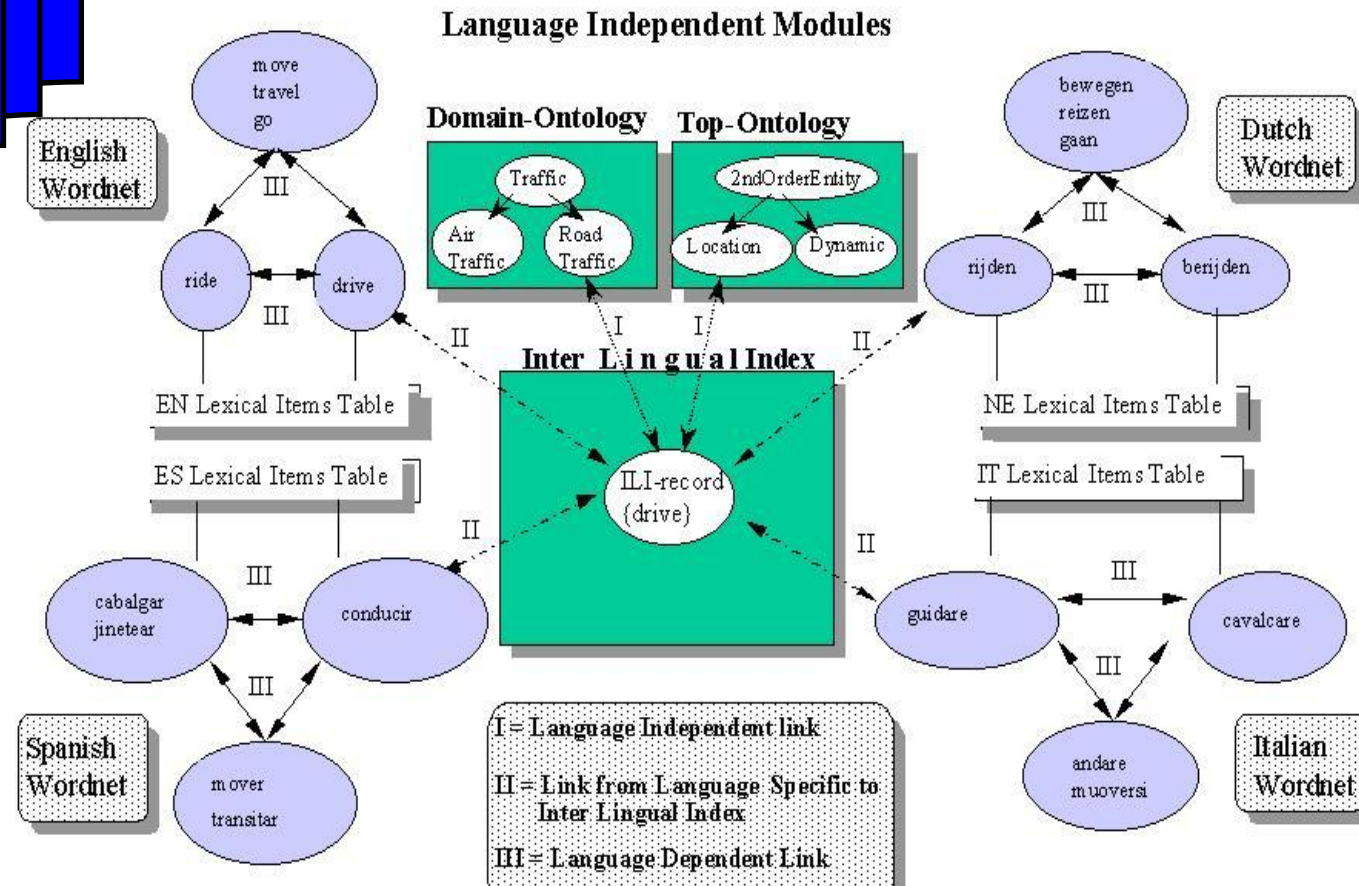


□ Enfoque 3:



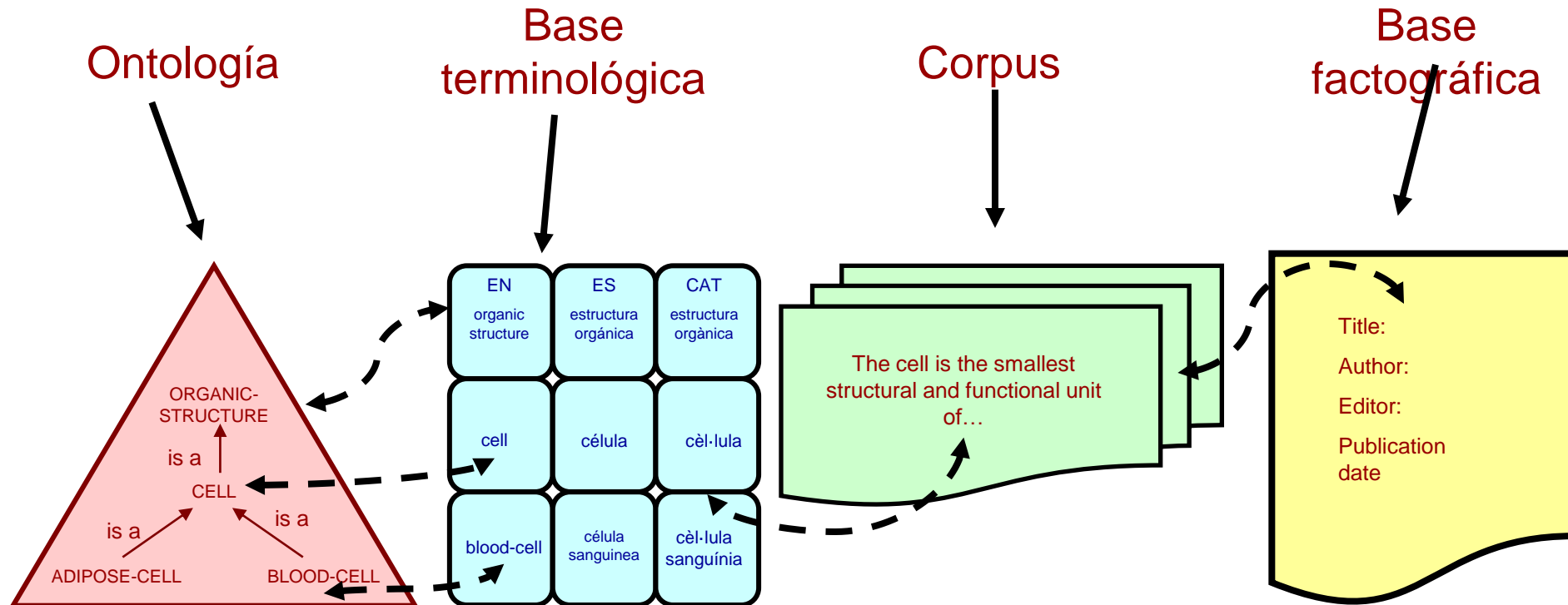
Enfoque 1: EuroWordNet

Lexicón
multilingüe en 8
lenguas



Enfoque 2: GENOMA-KB

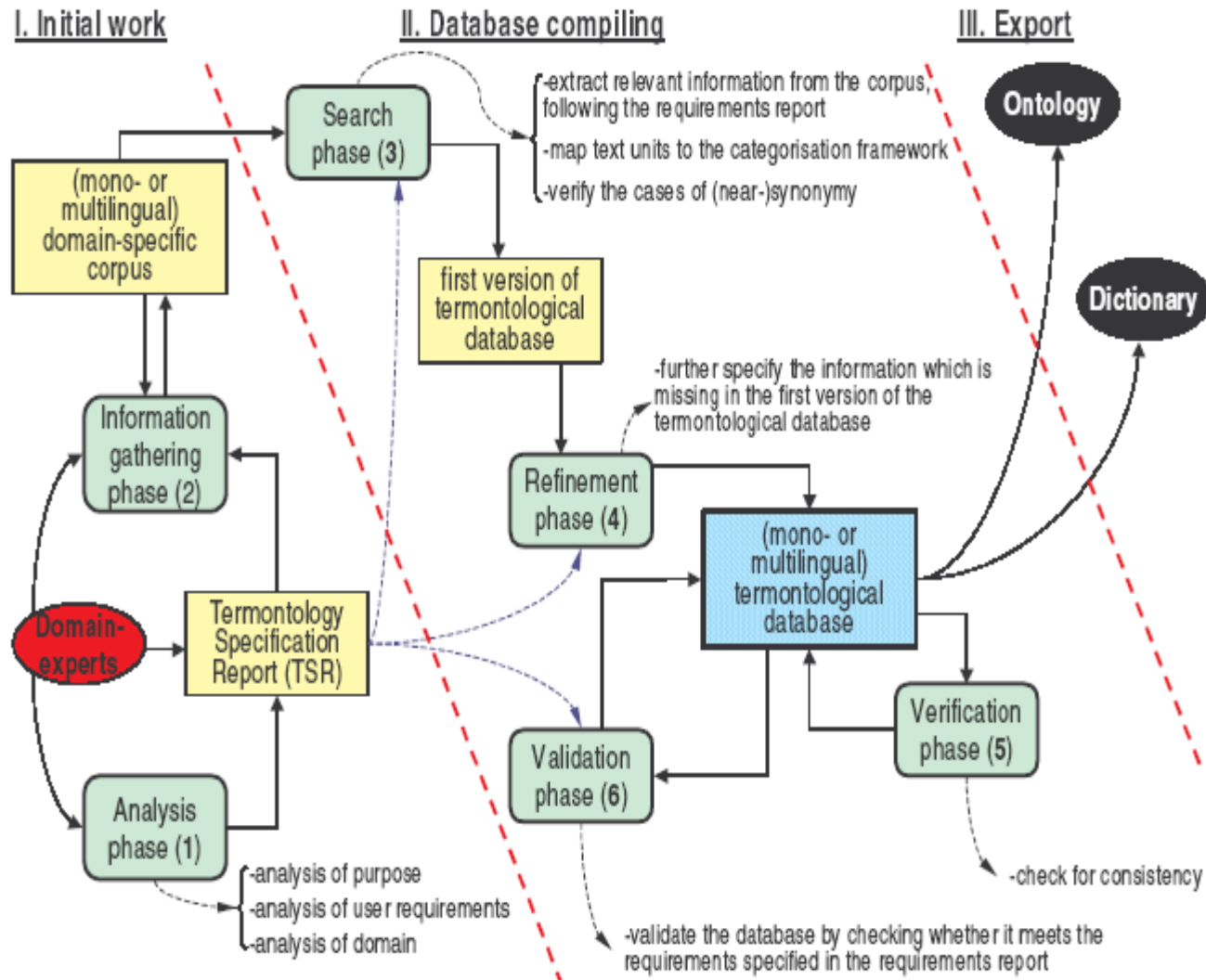
Base de conocimientos
multilingüe sobre el genoma
humano en EN, ES, and
CAT



Enfoque 2: Termontography

- Método interdisciplinar que combina la ingeniería ontológica y la multilingualidad. POIROT (IST 2001-38248).
- Resultado de la colaboración entre terminólogos y ontólogos.
- Desarrollado para la gestión de conocimiento mediante la representación de un dominio con datos en lenguaje natural
- Combina inglés, francés, holandés e italiano
- Los pasos que sigue son:
 - Análisis (1)
 - Recogida de información (2)
 - Búsqueda (3)
 - Refinamiento (4)
 - Verificación (5)
 - Validación (6)

Termonography workflow



Herramientas utilizadas en Termontography

- **Web crawler:** recupera online documentos de un dominio,
- **Keyword extractor:** informa al usuario del contenido del documento
- **Text converter:** almacena texto en cualquier formato
- **Automatic aligner:** alinea textos paralelos y sólo se procesa una versión en la fase de búsqueda
- **Similarity measuring tool:** elimina versiones redundantes
- **Automatic term identifier:** reconoce las unidades terminológicas extraídas en otros textos
- **Smart concordancer:** resalta la importancia del contexto
- **Term extractor:** propone candidatos a término, basándose en “mappings”
- **Translation extractor:** encuentra la traducción equivalente en el corpus bilingüe

Enfoque 3: LabelTranslator

EuroWordNet
Wikipedia
Babelfish

Herramienta para
una traducción
semi-automática
de las etiquetas
de una ontología

The screenshot displays the LabelTranslator web interface. On the left, a tree view shows a hierarchy of terms under 'Person', including 'Academic Staff', 'Company Staff', 'Administrative Staff', 'Manager', 'Technical Staff', 'Project Officer', 'Student', 'Master Student', 'PhD Student', and 'Undergraduate'. The 'Academic Staff' term is selected. Below the tree, there are language selection options: 'Preferred language: Spanish', 'Second language: German', and 'Fallback language: English'. A 'Show debugging information' link is at the bottom left. The main panel on the right is titled 'Translation: Academic Staff'. It shows the 'Word-by-word translation by EuroWordNet:' section with two rows: 'Academic' (translated to 'abstracto' and 'académico') and 'Staff' (translated to 'without practical purpose or intention: abstract reasoning ; an academic question'). Below this, the 'Other choices:' section shows 'Translation from Babelfish:' (Personal Académico) and 'Direct input:' (empty). The 'Original terms:' section shows 'Academic Staff'. A 'Submit' button is at the bottom right.

Terminología y ontologías

- La Terminología ayuda en la **organización del conocimiento** al establecer relaciones entre términos y conceptos
- Una ontología es una conceptualización o representación de una parte de un dominio, consensuada y formalizable computacionalmente.
- Enfoque orientado hacia la **comunicación** entre usuarios de una organización.
- **Cambio terminológico**: de BD terminológicas a BDTC terminológicas y de conocimiento / ontologías: términos en contexto y relaciones entre ellos.
- Los **términos extraídos** de los textos pueden convertirse en el **punto de partida** en la creación de ontologías

Las ontologías en la terminología

- Tecnologías que permiten:
 - Adquirir y modelar el conocimiento
 - Navegar y recuperar información
 - Representar y visualizar la información
 - Interpretar y compartir conocimiento
- Sistemas que incorporan semántica y contexto mediante ontologías y corpus textuales (Termonotography)
- Las herramientas actuales facilitan la interoperabilidad y las aplicaciones a la web semántica

La terminología en otras aplicaciones informáticas

- sistemas de corrección ortográfica
- sistemas de indización de textos
- sistemas de extracción de términos
- sistemas de recuperación de información
- sistemas de ayuda a la traducción
- sistemas expertos de toma de decisiones: medicina, bolsa
- sistemas de generación de textos especializados: resúmenes, etc.
- sistemas basados en el conocimiento: ontologías, etc.

Tabla comparativa de recursos léxicos

CLASSIFICATION CRITERIA	GLOSSARY	DATABASE	DICTIONARY	THESAURUS	LEXICON	ONTOLOGY
Organization	alphabetical order	alphabetical order	alphabetical order	semantically + generically related lexical entries	semantically related lexical entries	semantically related lexical entries
Semantic information	definition in NL	definition + other kinds of info. in NL	definition + pos + etymologies + derivation + usage examples in NL	hierarchical, associative, equivalent relationships	explicit hierarchy (synonymy, antonymy, meronymy...) + grammatical + contextual information	explicitly defined hierarchy relationships around a unique concept
Physical format	paper + electronic format	electronic format	paper + electronic format	paper + electronic format	electronic format	electronic format (readable also by machines)
Domain of knowledge	general + specific	general + specific	general + specific	specific	general + specific	general + specific (agreed by domain experts)

Factores de comparación

	Terminología	Ontologías
Grado de formalidad de la definición	Texto en LN	Lenguaje formal, sin ambigüedades
Soporte informático	Bancos terminológicos, con pocas relaciones entre conceptos	Lenguajes sólidos de representación de conocimiento, con algunas relaciones entre conceptos
Usuarios	Traductores Expertos del dominio Mediadores lingüísticos Redactores técnicos	Intercambio de información entre personas y máquinas Proyección para la web semántica
Lenguaje	LN para expresar el conocimiento con precisión	Menor importancia de las etiquetas para denominar conceptos

Colaboración entre terminólogos y ontólogos

Constructores de ontologías	Terminógrafos
Especificar requisitos	Especificar requisitos
Extraer conocimiento	Extraer conocimiento
Conceptualizar conocimiento	Conceptualizar conocimiento
Formalización con SRC basados en marcos o en logica descriptiva	Descripción en LN controlado, para satisfacer los requisitos del ontólogo
¿multilingüidad?	Análisis multilingüe

La terminología en las ontologías y otros recursos lexicográficos

Dra. Guadalupe Aguado de Cea

lupe@fi.upm.es

<http://www.oeg-upm.net>

Ontological Engineering Group

Facultad de Informática

Universidad Politécnica de Madrid

Campus de Montegancedo sn,

28660 Boadilla del Monte, Madrid, Spain