

Curso sobre lenguajes de indización

Construcción de tesauros

Programa General de Información y UNISIST

**Organización de las Naciones Unidas para
la Educación, la Ciencias y la Cultura**

**París,
noviembre de 1981**

CURSO SOBRE
LENGUAJES DE INDIZACION
CONSTRUCCION DE TESAuros

Realizado por el
Grupo de Trabajo sobre Lenguajes de Indización
con sede en el Centro Argentino de Información
Científica y Tecnológica del Consejo Nacional
de Investigaciones Científicas y Técnicas.
Buenos Aires — Argentina

Basado en el
Seminario Regional sobre Lenguajes de Indización
UNESCO — CAICYT
Buenos Aires, 14 al 26 de Agosto de 1978

INDICE DE CONTENIDO

1	INTRODUCCION	5
1.1.	Presentación	
1.2.	Antecedentes	
2	CURSO	6
2.1.	Objetivos	
2.2.	Organización	
2.3.	Metodología	
2.3.1.	Organización de contenidos	
2.3.2.	Recursos didácticos	
2.3.3.	Evaluación	
	a) Evaluación del alumno	
	b) Evaluación del curso	
2.4.	Bibliografía analítica	
2.5.	Clases desarrolladas:	
	UNIDAD I: TEORIA DE LA RECUPERACION DE LA INFORMACION	13
	Clase 1: Sistema de almacenamiento y recuperación de información.	
	Clase 2: Organización del archivo. Pre y post-coordinación.	
	Clase 3: Factores que afectan la eficacia de un sistema de almacenamiento y recuperación de información.	
	Clase 4: Vocabulario controlado.	
	UNIDAD II: ESTRUCTURA DE LOS TESAUROS	33
	Clase 1: Introducción. Partes constitutivas de los tesauros. Los descriptores.	
	Clase 2: Presentación de los tesauros.	
	UNIDAD III: CONSTRUCCION DE TESAUROS	45
	Clase 1: Consideraciones preliminares a la construcción de tesauros.	
	Clase 2: Recolección de términos.	
	Clase 3: Organización de los términos.	
	Clase 4: Proceso de revisión de los términos.	
	Clase 5: Introducción a la construcción práctica del tesoro.	
	Clase 6: Prueba, desarrollo y actualización del tesoro.	
	UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION	63
	Clase 1: Los vocabularios de indización en computadora.	
	Clase 2: Lenguaje natural.	
	Clase 3: Vocabularios híbridos. Compatibilidad de tesauros. Tesauros monolingües y multilingües.	
	Clase 4: Evaluación de tesauros.	
	Clase 5: Prospectiva de los lenguajes de indización.	
3	GLOSARIO	83
4	ANEXOS	101

Nota: para la confección de las transparencias (Unidad I: N° 1 a 59; Unidad II: N° 1 a 35; Unidad III: N° 1 a 24; Unidad IV: N° 1 a 60) se suministran por separado las láminas correspondientes. Recorte cada original y céntrelo en la hoja sensible, a los efectos de su reproducción.

1 INTRODUCCION

1.1. PRESENTACION

En los países iberoamericanos las listas de encabezamientos de materia y los sistemas de clasificación tradicionales son conocidos y aplicados por muchas instituciones. Sin embargo, no ocurre lo mismo con otras herramientas más modernas, los tesauros, cuya adopción, adaptación o elaboración es aconsejable cuando se desea utilizar sistemas de registro multidimensionales, semi-mecánicos o automáticos, para el tratamiento de información científica y técnica, en constante evolución.

En atención a la necesidad de adoptar nuevos métodos en el campo de la información, y facilitar la formación de recursos humanos capaces de asumir esa tarea, se presenta este Curso sobre lenguajes de Indización, preparado con la intención de satisfacer esa demanda.

1.2. ANTECEDENTES

El proceso previo a la preparación del presente trabajo se inició en 1976 con una sugerencia del Centro Argentino de Información Científica y Tecnológica (CAICYT) a la Secretaría del Programa General de Información de la UNESCO, de incluir en el programa ordinario un curso regional sobre lenguajes de indización, que se realizaría en la Argentina.

En 1977, el CAICYT procuró reunir a los especialistas de información interesados en el problema que nos ocupa, mediante un cursillo de introducción a la indización coordinada, del que surgió la disposición de algunos participantes para formar grupos de trabajo por áreas comunes de interés.

A raíz de una consulta del Centro Latinoamericano de Documentación Económica y Social de la CEPAL (CLADES), oficina que tiene a su cargo la actualización en español del Macrotesauro sobre desarrollo económico y social de la OCDE, en marzo de 1978 se invitó a un conjunto de especialistas. Los mismos constituyeron un Grupo de Trabajo que viene reuniéndose periódicamente desde entonces en el CAICYT para desarrollar una serie de actividades vinculadas con los lenguajes de indización (Anexo 4).

Algunos integrantes de este Grupo participaron en el Seminario Regional sobre Lenguajes de Indización, realizado en Buenos Aires del 14 al 26 de agosto de 1978, con el auspicio de la UNESCO y dictado por el profesor Lancaster de la Universidad de Illinois, Urbana, Estados Unidos.

Los participantes del Seminario, integrantes del Grupo de Trabajo, se habían propuesto tomar la mayor cantidad de elementos de este curso a fin de elaborar otro semejante en español. Desde marzo de 1979 el Grupo encaró la traducción, análisis y discusión de los temas y materiales didácticos proporcionados por el Profesor Lancaster y la preparación del curso respetando la metodología y el programa iniciales.

El "Curso sobre Lenguajes de Indización" se realizó en Buenos Aires entre los días 17 de Setiembre y 5 de Octubre de 1979 para 28 participantes. Se efectuaron evaluaciones a distintos niveles teórico, práctico, de participación, etc.

Esta experiencia permitió probar el Curso de referencia y efectuar las mejoras necesarias para la elaboración de un instrumento didáctico destinado a posibilitar un efecto multiplicador, que facilitará la capacitación de recursos humanos en este campo.

2 CURSO

2.1. OBJETIVO

Proveer información teórico-práctica sobre la construcción de tesauros a bibliotecarios especializados y otros especialistas de la información, que tuvieran la responsabilidad de clasificar o indizar documentos, en países de habla hispana.

2.2. ORGANIZACION

2.2.1. Duración

Se considera que se debe desarrollar en 42 horas, distribuidas de la siguiente manera:

- clases teóricas (46 o/o)
- trabajos prácticos (46 o/o)
- evaluaciones y mesa redonda (8 o/o)

Según la experiencia realizada, puede indicarse que la modalidad residencial posibilita un mejor aprovechamiento y dedicación: un curso de dos semanas, de lunes a viernes, con 5 horas diarias.

2.2.2. Docentes

Profesional especialista en información científica y tecnológica con dominio de las técnicas de clasificación e indización y preferentemente con experiencia docente.

Observación: si no existieran estas condiciones, podría hacerse cargo del curso un profesional con experiencia en construcción o uso de tesauros. Se recomienda un mínimo de dos profesores.

2.2.3. Participantes

Se aconseja un máximo de 25. Deberían reunir los siguientes requisitos:

- 1) Tener estudios completos a nivel superior o universitario.
- 2) Tener a su cargo tareas de clasificación o indización de documentos en un servicio de biblioteca y/o información, o ejercer la docencia en el área referida.
- 3) Idioma inglés y francés (lectura y traducción).

2.2.4. Comité de Selección de los participantes

Se recomienda su constitución para examinar y evaluar a los candidatos en los aspectos mencionados.

2.3. METODOLOGIA

2.3.1. Organización de contenidos

Según la proporción establecida en el punto 2.2., en las horas requeridas, el curso tiene una parte teórica y una aplicación práctica consistente en la simulación de la construcción de un tesoro.

Parte teórica:

Los aspectos teóricos están agrupados en cuatro unidades que, con continuidad entre sí, completan individualmente un tema. Esta organización tiene por objeto permitir la gradualidad de los contenidos, y un mejor control de su desarrollo.

Unidad I: Teoría de la recuperación de la información

- Clase 1:** Sistema de almacenamiento y recuperación de información
- Clase 2:** Organización del archivo. Pre y post-coordinación
- Clase 3:** Factores que afectan la eficacia de un sistema de almacenamiento y recuperación de información
- Clase 4:** Vocabulario controlado

Unidad II: Estructura de los tesauros

- Clase 1:** Introducción. Partes constitutivas de los tesauros. Los descriptores.
- Clase 2:** Presentación de los tesauros

Unidad III: Construcción de tesauros

- Clase 1:** Consideraciones preliminares a la construcción de tesauros
- Clase 2:** Recolección de términos
- Clase 3:** Organización de los términos
- Clase 4:** Proceso de revisión de los términos
- Clase 5:** Introducción a la construcción práctica del tesoro
- Clase 6:** Prueba, desarrollo y actualización del tesoro

Unidad IV: Aplicación de la automatización a los lenguajes de indización

- Clase 1:** Los vocabularios de indización en computadora
- Clase 2:** Lenguaje natural
- Clase 3:** Vocabularios híbridos. Compatibilidad de tesauros. Tesauros monolingües y multilingües
- Clase 4:** Evaluación de tesauros
- Clase 5:** Prospectiva de los lenguajes de indización

Parte práctica: se simulará individualmente la construcción de un tesoro, cumpliendo con todas las etapas indicadas en las clases teóricas. Cada alumno podrá escoger el tema, dentro del área de la bibliotecología o las ciencias de la información.

Cada participante seleccionará términos significativos o palabras claves de los resúmenes sobre ciencias de la información "Library and Information Science Abstracts", editado en Londres por The Library Association, y los "Information Science Abstracts", publicado en Washington por el American Documentation Institute - American Chemical Society, y también el "Bulletin Signaletique" Sección 101. Cada término será escrito en una ficha estándar.

La selección de términos podría hacerse también de artículos de revistas sobre el tema, publicadas en español, como por ejemplo "Revista de la UNESCO en ciencia de la información, bibliotecología y archivología". Por este método la búsqueda se hace mucho más lenta debido a que los artículos no traen resúmenes que, en las publicaciones antes mencionadas, incluyen terminología normalizada.

Se recomienda la reunión de un mínimo de 300 descriptores por participante, y un máximo de 500.

Luego se procederá a separar y agrupar los descriptores para determinar los campos o clases principales y formar las facetas. Se establecerán las relaciones de equivalencia, asociativas y jerárquicas.

Por último se ordenará el trabajo en forma alfabética.

Se recomienda que la selección de términos no exceda del 30 o/o del tiempo total del trabajo práctico.

2.3.2. Recursos didácticos

a) Recursos didácticos proporcionados con el paquete

– **Transparencias:** 174 láminas que sintetizan el desarrollo del curso y que sirven como elemento motivador y de fijación de contenidos. Además de su proyección en una pantalla*, cada alumno debe poseer una copia del juego completo, lo que permitirá seguir las explicaciones y tomar las notas que considere necesarias. Esto agiliza el ritmo de las clases y permite un mejor aprovechamiento de las mismas. Tanto las transparencias para uso del docente como las copias para cada participante, pueden reproducirse a partir de las láminas adjuntas al paquete.

– **Planes de clase:** los contenidos a impartir se encuentran distribuidos en 17 clases. Para cada una de ellas están determinados los objetivos, las ideas principales y las ideas accesorias a desarrollar, las transparencias correlativas y los textos complementarios para cada una de ellas.

– **Textos complementarios:** textos breves que desarrollan el contenido de cada transparencia. Las palabras del texto señaladas mediante un asterisco están definidas en el Glosario. Cada término está marcado sólo una vez en cada clase.

– **Bibliografía analítica:** referencias bibliográficas de la bibliografía básica correspondiente a cada unidad.

– **Glosario:** se definen 112 términos significativos del contenido del curso.

b) Recursos didácticos que deben proveerse a los participantes

– **Bibliografía completa del curso**

– **Revistas de resúmenes para los trabajos prácticos**

– **Tesauros**

– **Glosarios, diccionarios, etc.**

– **Juego completo de las láminas.**

– **Aproximadamente 500 fichas tamaño estándar (12,5 x 7,5 cm) para el trabajo práctico.**

2.3.3. Evaluación

La evaluación se realiza en varios niveles:

a) **Trabajos prácticos:** durante las prácticas y una vez finalizada la construcción individual de un micro-tesauro en el tema establecido, se revisará cada trabajo. Esta revisión permitirá evaluar aparte de los aspectos específicos del trabajo práctico, la adquisición de los conocimientos impartidos.

b) **Evaluación del curso:** puede realizarse por medio de los siguientes elementos:

– **Cuestionario previo:** con la finalidad de evaluar el interés de participación de los alumnos, y compararlo con un cuestionario final. De esta manera se comprobará si el grado de satisfacción expresado al fin del curso tiene relación con las necesidades expresadas inicialmente (Anexo 3).

* **Nota:** Utilizar un retroproyector de transparencias.

- Evaluación y control diarios (didácticos): es importante que este control se realice para cada clase, ya que permite efectuar los ajustes necesarios por parte del docente.
 - Evaluación del trabajo práctico: se realizará teniendo en cuenta los siguientes criterios:
 - calidad de los términos seleccionados
 - construcción de facetas
 - relaciones correctas entre los términos
 - Evaluación final: cuestionario que recaba las opiniones de los alumnos sobre aspectos generales del curso y el grado de satisfacción obtenidos en el mismo (Anexo 3).
 - Mesa Redonda: se realiza con la idea de finalizar el curso con un diálogo de conjunto en el cual cada participante pueda expresar sus opiniones en relación con la aplicación de lo aprendido en sus tareas habituales.
- c) Requisitos de aprobación del curso: los requisitos mínimos de aprobación del curso son:
- 80 o/o de asistencia a las clases teóricas y 80 o/o de asistencia a las clases prácticas
 - Trabajo práctico aprobado.

2.4. BIBLIOGRAFIA ANALITICA

1. RECUPERACION DE LA INFORMACION

Lancaster, F.W. Vocabulary control for information retrieval. Washington, D.C., Information Resources Press, 1972. Cap. 1: p. 1-4; cap. 2: p.5-7; cap. 13: p.107-114.

Lancaster, F.W. Vocabulary control in information retrieval systems. (En: ADVANCES IN LIBRARIANSHIP, 1977, v.7: p. 1-40).

Aitchison, Jean - Gilchrist, Alan. Thesaurus construction; a practical manual. Londres, ASLIB, 1972. B-B3.1: p. 2-6.

2. ESTRUCTURA DE TESAUROS

Lancaster, F.W. Vocabulary control for information retrieval. Washington, D.C., Information Resources Press, 1972. Cap. 5: p. 22-26; cap. 9: p. 70-76.

Aitchison, Jean - Gilchrist, Alan. Thesaurus construction; a practical manual. Londres, ASLIB, 1972. B B3.1: p. 2-6; C: p. 14-19.

3. CONSTRUCCION DE TESAUROS

Lancaster, F.W. Vocabulary control for information retrieval. Washington, D.C., Information Resources Press, 1972. Cap. 6: p. 27-37.

Aitchison, Jean - Gilchrist, Alan. Thesaurus construction; a practical manual. London, ASLIB, 1972. 95 p.

Dym, Eleanor D. A new approach to the development of a technical thesaurus. (En: Proceedings of the American Documentation Institute, 4, 1967. p. 126-131.

Pickford, A.C.A. An objective method for the generation of an information retrieval language. (En: The information scientist, March 1968, 2, p. 17-37.

Vickery, B.C. Faceted classification. Londres, ASLIB, 1968.

Lancaster, F.W. Vocabulary control for information retrieval. Washington, D.C., Information Resources Press, 1972. Cap. 7: p. 38-65; cap. 10: p. 77-89; cap. 11: p. 90-97; cap. 12: p. 98-106.

Willems, Margaret. An investigation of the nature of the relation between terms in thesauri. (En: Journal of Documentation, September 1975, v. 31, N° 3, p. 158-184).

4. AUTOMATIZACION

Lancaster, F.W. Vocabulary control for information retrieval. Washington, D.C., Information Resources Press, 1972. Cap. 8: p. 66-69; cap. 11: p. 90-97; cap. 15: p. 121-124; cap. 16: p. 135-152; cap. 17: p. 153-160; cap. 18: p. 161-176.

Aitchison, Jean. The thesaurifacet; a multipurpose retrieval language tool. (En: Journal of Documentation, September 1970. V. 26, N° 3, p. 187-203)

Vickery, B.C. Faceted classification. Londres, ASLIB, 1968.

Smith, Linda C. Systematic searching of abstracts and indexes in interdisciplinary areas. (En: Journal of the American Society for information science, November-December 1974, p. 343-353).

UNESCO. Guidelines for the establishment and development of multilingual thesauri. Paris, UNISIST, 1976. 49 p. (UNISIST Program SC.76/WS/102)

UNESCO. Guidelines for the establishment and development of monolingual thesauri. Paris, UNISIST, 1973. 37 p. (UNISIST Program SC/WS/555)

Bureau Marcel Van Dijk. Definition of thesauri characteristics. Bruselas, 1976. 2 v.

2.5. CLASES DESARROLLADAS

2.5.1. UNIDAD I

UNIDAD I: TEORIA DE LA RECUPERACION DE LA INFORMACION

CLASE 1: SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION

Objetivo:

Presentar el proceso de un Sistema de almacenamiento y recuperación de información y ubicar dentro de ese proceso al vocabulario, mostrando su importancia.

Idea principal:

El objetivo fundamental de un Sistema de almacenamiento y recuperación de información es el de recuperar información que satisfaga las necesidades de los usuarios.

Ideas accesorias:

Para cumplir con ese objetivo fundamental el Sistema debe realizar las etapas de entrada y salida en forma eficiente y coherente.

Una buena recuperación de información depende de la indización y de la estrategia de búsqueda, y en consecuencia del vocabulario.

TRANSPARENCIA I-1:

- A) *Entrada al Sistema**: Proceso de entrada de la información, que corresponde a la etapa del almacenamiento de la información*.

Existe un *universo de documentos* publicados (patentes, informes, libros, artículos, informes, etc). El Sistema de Almacenamiento y Recuperación de Información* *seleccionará* de ese universo aquellos documentos* que sean de más interés y valor para los usuarios del Sistema. Ese conjunto de documentos constituye la entrada al Sistema. Una vez seleccionados, los documentos son procesados.

En un Sistema de Almacenamiento y Recuperación de Información especializada se *indizan* los documentos en dos etapas:

- a) se realiza el *análisis conceptual** del documento por el cual se determina el tema que trata.
- b) luego se *traduce* ese contenido conceptual en un *vocabulario** elegido por el Sistema. Este vocabulario puede ser un lenguaje de indización*, tal como el *tesauro**.

Luego de indizar los documentos tendremos entonces: por una parte, los documentos que irán a sumarse a la *colección* del Sistema y, por otra parte, el resultado de la indización* de los documentos que integrará la *base de datos* (índice de acceso a los documentos) para permitir la posterior recuperación de los mismos.

* Las palabras señaladas con asterisco están definidas en el glosario.

- B) *Salida del Sistema:* Proceso de salida de la información, que corresponde a la etapa de recuperación de la información.

Los usuarios del Sistema se dirigen al mismo solicitando información. Para satisfacerlos es necesario hacer compatible su interés con las características de los documentos. Para ello es conveniente utilizar el mismo vocabulario empleado en la indización. La *preparación de las estrategias de búsqueda** consta de dos partes: el *análisis conceptual de la pregunta* y la *traducción** de ésta a los términos del vocabulario del Sistema.

TRANSPARENCIA 1-2:

Cuando se indiza, se representan los conceptos asignando etiquetas* a los documentos, que se extraen de un lenguaje de indización. Un lenguaje de indización está formado por etiquetas con sus definiciones y sus relaciones.

Una etiqueta puede variar su forma (nombre), pero debe precisarse por medio de una definición su contenido conceptual.

Las relaciones facilitan la ubicación de las etiquetas precisas.

Ante una pregunta se prepara una estrategia de búsqueda mediante etiquetas que puedan coincidir con las asignadas previamente en la entrada a los documentos que quieran recuperarse.

TRANSPARENCIA 1-3:

La recuperación de la información depende de la indización y de la estrategia de búsqueda. En el ciclo de almacenamiento y recuperación de información no interesa tanto la modalidad del Sistema elegido (automático, semiautomático o manual), como los factores intelectuales que intervienen en la indización y en la búsqueda de información.

UNIDAD I: TEORIA DE LA RECUPERACION DE LA INFORMACION

CLASE 2: ORGANIZACION DEL ARCHIVO. PRE Y POSTCOORDINACION

Objetivo:

Hacer conocer los sistemas postcoordinados de almacenamiento y recuperación de información y compararlos con los sistemas tradicionales precoordinados.

Idea principal:

Un sistema de recuperación mediante la postcoordinación permite una accesibilidad múltiple a los conceptos contenidos en los documentos.

Ideas accesorias:

En los sistemas precoordinados se combinan los términos en el momento de la indización.

En los sistemas postcoordinados se combinan los términos en el momento de la búsqueda.

Los tesauros están concebidos para los sistemas postcoordinados.

La evolución acelerada de los campos científicos y tecnológicos, el surgimiento de nuevas áreas, el desarrollo de las interdisciplinas y la aparición de los sistemas electrónicos para el procesamiento de datos, forzaron el cambio de los sistemas tradicionales de tratamiento de la información. Estos sistemas demostraron ser rígidos e incapaces de adecuarse al ritmo de desarrollo de las distintas áreas científicas y tecnológicas.

A continuación se presentan algunos conceptos sobre los sistemas de clasificación* de documentos que se aplican en nuestras bibliotecas.

TRANSPARENCIA I-4

En la *Clasificación Decimal Universal* las nociones específicas derivan de las nociones más generales que las engloban. Se tienen inclusiones sucesivas en clases más o menos rígidas.

Cuando se hace la presentación de esquemas de clasificación*, se desarrollan en forma horizontal, figurando a la izquierda el tronco y a la derecha las sucesivas divisiones. Los esquemas son eminentemente jerárquicos.

TRANSPARENCIA I-5

Otro sistema que han usado las bibliotecas son las listas de encabezamientos de materia*.

En esta transparencia se ven dos ejemplos de asientos de encabezamientos de materia* correspondientes a la *Lista de Encabezamientos de Materia para Bibliotecas*, publicada por la Unión Panamericana.

En los asientos existe una clasificación implícita*. También está indicada la posibilidad de combinar los encabezamientos en el momento de la indización (precoordinación)* por medio de subdivisiones geográficas.

NOTA: Aunque en la Lista de Encabezamientos de la Unión Panamericana éstos aparecen en letra minúscula, en las fichas catalográficas se deben transcribir con letra mayúscula.

TRANSPARENCIA I- 6

Si las bibliotecas usaran para la indización descriptores* de un tesoro*, tendrían asientos semejantes a los de las listas de encabezamientos de materia.

En esta transparencia aparecen los mismos conceptos de la transparencia anterior, tal como se presentan en el *Macrothesaurus* (*Macrotesauro**) de la OCDE (uno de los tesauros más usados en la actualidad).

En el momento de indizar deben elegirse los descriptores que expresen los conceptos* contenidos en el documento* sin hacer ninguna combinación entre ellos (Precoordinación). En el momento de recuperar la información, por ejemplo DERECHO DE LA PROPIEDAD EN LA ARGENTINA, se combinan los tres conceptos (postcoordinación*) para encontrar los documentos correspondientes a dicho tema específico.

TRANSPARENCIA I- 7

En esta transparencia se muestra la representación del tema DERECHO DE LA PROPIEDAD EN LA ARGENTINA en CDU, encabezamientos de materia y descriptores, como una síntesis de lo explicado en las tres transparencias anteriores.

TRANSPARENCIA I - 8

La organización de los archivos dependerá de la forma pre o postcoordinada del lenguaje que se emplee. Los archivos* correspondientes a sistemas precoordinados se organizan en forma *lineal, unidimensional*. A, luego B, luego C, luego D, luego E, etc.

Por ejemplo, el concepto: Fatiga de metales (A, B) estará entrado sólo por Fatiga y no por Metales. Fatiga, de paneles de aluminio, en avioneta, a velocidad supersónica: Concepto complejo expresado con muchos términos, entra sólo por el término FATIGA y no se puede recuperar por los otros términos.

TRANSPARENCIA I- 9

En un archivo correspondiente a un sistema precoordinado* se tendrá un orden establecido, alfabético o numérico. Detrás de cada ficha que se usa como separador, se acumulan todas las fichas que representan los documentos.

TRANSPARENCIA I- 10

Hay que considerar que los conceptos de los documentos no son lineales sino *multidimensionales*. Hay múltiples relaciones implícitas entre los conceptos de un documento que no pueden ser contempladas en su totalidad en un sistema lineal. Este carácter multidimensional de los conceptos es reducido a un carácter lineal en los archivos precoordinados. Lo multidimensional en secuencia lineal no es posible.

TRANSPARENCIA I- 11

La expresión de este carácter multidimensional es permitida en los sistemas postcoordinados*. Si se representa la colección de documentos con una matriz y se hace corresponder a cada número de documento una serie de características* o descriptores*, se ve con claridad que se incrementan las posibilidades de combinaciones.

TRANSPARENCIA I- 12

En los sistemas *postcoordinados* se hace la combinación de los términos en el momento de la búsqueda de información. Al *indizar*, se asignan descriptores a cada documento, identificado mediante un número, que es secuencial según el ingreso. Por ejemplo, el primer documento que ingrese al sistema llevará el número 1, el siguiente el número 2, etc. Los descriptores o características están representados por fichas que se ordenan alfabéticamente. En cada una de las fichas se anotan los números de documentos que poseen esa característica.

En el momento de realizar la *búsqueda*, se retiran las fichas cuyas características interesan, al compararlas aparecerá la coincidencia de un número o varios, correspondientes a un documento o a varios.

TRANSPARENCIA I- 13

Las fichas, tipo *unitermino**, tienen columnas numeradas del 0 al 9. En cada columna se anotan los números de documentos cuyo último dígito coincida.

TRANSPARENCIA I- 14

El mismo criterio se sigue en los sistemas de selección óptica* y en los sistemas automatizados.

Los sistemas *postcoordinados* se inician a fines de la década del 40. Se trata de encontrar una forma de búsqueda que permita establecer relaciones y combinaciones para cada uno de los temas.

El tipo de vocabulario* que se utilice estará en relación con la *precoordinación* o la *postcoordinación* que permita el sistema.

Los tesauros están previstos para sistemas *postcoordinados*.

UNIDAD I: TEORIA DE LA RECUPERACION DE LA INFORMACION

CLASE 3: FACTORES QUE AFECTAN EL RENDIMIENTO DE UN SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION.

Objetivo:

Presentar el ciclo de recuperación de información a partir de la demanda de un usuario y terminando con la respuesta a esa demanda.

Describir cada etapa del ciclo y explicar los factores que inciden en el rendimiento de un sistema de almacenamiento y recuperación de información, con especial referencia a los problemas del lenguaje de indización.

Ideas principales:

El éxito final del ciclo de recuperación de información depende de la eficacia con que se cumplen cada una de sus etapas.

La elección del lenguaje de indización es un factor esencial para el rendimiento del sistema.

Ideas accesorias:

La mayor o menor eficacia con que se cumplen cada una de las etapas se va acumulando para el resultado final.

La especificidad y exhaustividad en la indización están directamente relacionados con las tasas de acierto y precisión en la recuperación de la información.

TRANSPARENCIA I- 15

El ciclo se inicia con el *usuario* que tienen una necesidad de información pero no siempre sabe expresarla. Requiere, por lo tanto, la ayuda del especialista en información para formular la *solicitud*, que es la expresión de esa necesidad. El especialista en información selecciona la base de datos, que puede ser automatizada (computadora), impresa (bibliografía) o un archivo de acceso manual (fichero).

Con la solicitud del usuario y de acuerdo con la base de datos seleccionada, se prepara la *estrategia de búsqueda** utilizando el vocabulario* elegido por el Sistema*.

Se realiza una *búsqueda en la base de datos* mediante la comparación de los *descriptores** determinados en la *estrategia de búsqueda** con los de la base de datos.

Finalmente, se filtra la respuesta, es decir, se seleccionan aquellos documentos que pueden ser pertinentes* al usuario.

En este ciclo puede haber silencio, falta de información, o ruido*, es decir, información que no esté de acuerdo con las necesidades particulares de un determinado usuario. En cada paso de este ciclo se pueden cometer errores, que por un fenómeno acumulativo inciden en el rendimiento* de un sistema de almacenamiento y recuperación de información.

TRANSPARENCIA I- 16

El usuario comienza un proceso desde el momento que reconoce que tiene una *necesidad* de información, se siente *motivado* para satisfacerla, y la *expresa*.

Al conocer la *existencia* de un Centro de Información, también se conocen sus posibilidades de *acceso* (físico, intelectual, psicológico), el *costo* en dinero y en esfuerzo, el *tiempo* que demanda la búsqueda, si fue buena alguna experiencia anterior, si los servicios son *rápidos*, si la solución que ofrece el Centro es válida para el problema del usuario, si el Centro tiene la posibilidad de resolver el problema o si el usuario tiene que dirigirse a otro centro.

TRANSPARENCIA I- 17

Una vez que el usuario sabe con claridad cuál es su necesidad de información y adonde debe dirigirse para satisfacerla, debe redactar la *demanda*. El servicio debe saber *interpretar las capacidades y limitaciones del usuario* para expresar la demanda. Por su parte, el usuario debe *indicar en forma clara* su necesidad para que ésta sea interpretada correctamente por el especialista que lo atiende. Es muy importante que la interacción usuario/especialista sea fluida, y si fuera posible, sin interrupciones. Ayudar al usuario con formularios y entrevistas, es un factor que contribuye a mejorar esa interacción.

TRANSPARENCIA I- 18

Una vez realizada la demanda del usuario, se selecciona la base de datos según: 1) posibilidades de acceso, 2) conocimiento sobre su alcance (campos, tipos de documentos y período que cubre), 3) calidad de la base, por ej. en cuanto a la selección de documentos y en cuanto a la indización* de los mismos.

TRANSPARENCIA I- 19

Para preparar la *estrategia de búsqueda* el especialista debe entender claramente lo que el usuario necesita. Tratará de cubrir todas las *vías de acceso* a la información, necesitará de un *vocabulario** adecuado en cuanto a su contenido y estructura, deberá tener *habilidad* para construir las estrategias de búsqueda de modo lógico. En esta etapa, influyen además las limitaciones que puedan tener los programas* de búsqueda en sistemas automatizados.

TRANSPARENCIA I- 20

Si la demanda de un usuario es la producción de maíz en Egipto y el vocabulario se adecúa a la búsqueda de ese tema, la estrategia consistirá en combinar MAIZ+ EGIPTO. Pero si el vocabulario es muy general y sólo ofrece CEREAL y MEDIO ORIENTE, se tendrá que buscar por estos términos*, que incluyen los conceptos* específicos buscados y muchos otros más.

TRANSPARENCIA I- 21

Se supone que la pregunta del usuario se refiere al método de enseñanza del concepto matemático en la educación primaria.

Si el vocabulario permite coordinación* entre conceptos y varios niveles de especificidad* se podrán hacer búsquedas precisas.

La transparencia muestra diferentes grados de especificidad (nivel de profundidad en el tema) en el plano vertical y diferentes grados de exhaustividad* (nivel de coordinación) en el plano horizontal.

TRANSPARENCIA I- 22

Al preparar la estrategia de búsqueda se debe saber el nivel de especificidad que ofrece el vocabulario del sistema.

TRANSPARENCIA I- 23

La estrategia de búsqueda debe tener en cuenta los niveles de especificidad y de exhaustividad a los que se quiere llegar.

Si se tienen las facetas A, B y C relacionadas se hará una búsqueda exhaustiva. Si esta búsqueda exhaustiva se realiza a un nivel más específico (A_1 , B_1 , C_1) se tiene una búsqueda específica.

TRANSPARENCIA I- 24

Buscar es como pescar; se puede usar una red pequeña o una grande. Si se busca algo específico se puede usar una red chica con la cual se pescan los tres pececitos marcados (temas específicos), pero regularmente se pierden otros peces más grandes que quizás puedan ser útiles (temas generales que incluyen esos temas específicos).

Si se usa una red grande para esa misma búsqueda específica se pescan peces chicos y grandes. De esta manera se recuperaría todo lo que se necesita pero también mucho de lo que no se necesita.

TRANSPARENCIA I- 25

En la búsqueda deben cumplirse correctamente los pasos siguientes: el análisis conceptual* de la demanda y la traducción*. Ambos están dirigidos a satisfacer las necesidades del usuario.

TRANSPARENCIA I- 26

Los factores que influyen en el momento en que se realiza la búsqueda en una base de datos, son los ya comentados: el vocabulario utilizado (que deberá ser el mismo en la indización y en la búsqueda), la exhaustividad con que se realiza la indización, la consistencia* de la indización a fin de eliminar errores y la posibilidad de evitar relaciones ambiguas o falsas.

TRANSPARENCIA I- 27

Las citas bibliográficas resultantes, deben ser filtradas* antes de la entrega de los documentos correspondientes. El éxito de esta etapa depende de cómo se fueron desarrollando las anteriores etapas del ciclo de búsqueda.

La eficacia de un sistema de recuperación de información, está determinada por el lenguaje de indización* utilizado, el proceso de indización, el proceso de búsqueda y la interacción usuario-sistema.

TRANSPARENCIA I- 28

Criterios para evaluar una búsqueda bibliográfica

Para poder evaluar una búsqueda bibliográfica debemos considerar el grado de satisfacción o insatisfacción del usuario. Los documentos pueden ser positivos o relevantes* (aquellos que contribuyen a la satisfacción de una necesidad de información del usuario) o negativos o irrelevantes.

El *acierto** corresponde a la cantidad de documentos relevantes recuperados en relación con la totalidad de documentos relevantes del sistema. La *precisión** se refiere a los documentos relevantes recuperados en relación con todos los documentos recuperados.

En el ejemplo un lector concurre a un servicio de información en busca de seis documentos, que forman parte de la colección de ese centro, y que él sabe satisfacerían su necesidad de información.

En el primer caso se siente muy defraudado porque no se ha recuperado ningún documento relevante: *acierto* y *precisión* = 0.

En el segundo caso no está muy contento pues sólo se recuperan dos documentos relevantes. El *acierto* = 33 (relación 2 en 6) y la *precisión* = 100 pues los documentos recuperados son relevantes.

En el tercer caso está un poco más satisfecho pues de los seis documentos recuperados la mitad son relevantes. Finalmente su satisfacción es completa pues recuperó los seis documentos relevantes.

El rendimiento del sistema de recuperación se evalúa mediante las tasas de *acierto** y de *precisión*, que se calculan con fórmulas.

Hay una diferencia entre los conceptos de relevancia y pertinencia, que está en estrecha relación con el tipo de persona que decide si los documentos son positivos o negativos (+ ó -). Cuando esta determinación la realiza un especialista de información en el tema, se considera que los documentos son *relevantes*. Cuando es el usuario el que juzga el resultado de la búsqueda, se considera a los documentos como *pertinentes**. Los documentos pueden ser relevantes según un pedido, o pertinentes para una necesidad de información. La *relevancia** de un documento puede ser determinada por cualquier persona que conozca el tema. La *pertinencia** la determina el usuario y corresponde a una necesidad informativa suya. La *relevancia* está conformada por las respuestas adecuadas a una consulta. La *pertinencia* es la relación entre documentos recuperados y la conformidad de los usuarios correspondientes.

TRANSPARENCIA I- 29

En esta tabla pueden verse las distintas cantidades correspondientes a la búsqueda de documentos en una colección. Con estas cifras se puede hallar, por medio de fórmulas, la tasa de *precisión** y la tasa de *acierto* de una búsqueda bibliográfica.

TRANSPARENCIA I- 30

Esta transparencia es la presentación gráfica de los datos correspondientes a la transparencia I-29.

TRANSPARENCIA I- 31

En esta transparencia se dan las fórmulas que permiten obtener las tasas de *acierto* y de *precisión* de una búsqueda bibliográfica.

TRANSPARENCIA I- 32

Las fallas de acierto y precisión son factores de fracaso de los sistemas de información y pueden originarse en: 1) el lenguaje de indización, 2) en la indización, 3) la búsqueda, o 4) la interacción usuario-sistema. Se ejemplifica con el estudio que realizó Lancaster en el Sistema Medlars, sobre la base de 300 búsquedas, en las que detectó estas fallas.

TRANSPARENCIA I- 33

Los factores que afectan la eficacia de un Sistema de almacenamiento y recuperación de información son la política de indización y la precisión de la indización.

La política de indización está definida por la institución y fuera del control del indizador. La precisión de la indización está bajo el control del indizador.

TRANSPARENCIA I- 34

Ante la indización de un documento, hay que decidir si se indiza en forma selectiva o exhaustiva y determinar el tiempo que se dedicará a la indización del documento. Cuando se reduce el tiempo se afecta la exhaustividad.

En el ejemplo, si se eligen hasta tres términos se tendrá una indización selectiva, que originará una alta precisión y un bajo acierto. Si se eligen hasta diez términos se obtendrá una indización exhaustiva que producirá un alto acierto y una baja precisión.

TRANSPARENCIA I- 35

Cuanto más exhaustiva sea la indización mayor será el porcentaje de necesidades satisfechas pero aumentará el costo de la indización.

TRANSPARENCIA I- 36

El artículo de la transparencia fue indizado con varios términos que, en forma general, expresan su tema. Sin embargo, se omitió el concepto específico concerniente.

Si por este concepto específico, un usuario solicita información al sistema, no se recupera el artículo y por lo tanto la necesidad de información queda insatisfecha.

TRANSPARENCIA I- 37

Los problemas que se presentan al traducir a términos del lenguaje de indización son los siguientes:

1. **Término incorrecto:** por desconocimiento del tema el indizador no selecciona el término correcto. El borohidruro de aluminio es un combustible líquido.
2. **Falta de especificidad en el vocabulario:** este factor se relaciona con los niveles del vocabulario y con los niveles de indización determinados por una política.

TRANSPARENCIA I-38

3. Falta de especificidad en la indización: puede deberse al mal uso del vocabulario por desconocimiento por parte del indizador o porque el vocabulario no posea las referencias suficientes.

TRANSPARENCIA I-39

El rendimiento de un Sistema de Almacenamiento y Recuperación de Información está también afectado por otros factores, tales como el productor que controla la base de datos y el lenguaje; el distribuidor comercial que controla el software* y el hardware*; el Centro de información que controla la interacción con el usuario y las estrategias de búsqueda. Todos estos factores están interrelacionados.

TRANSPARENCIA I-40

El Centro de información tiene una o más terminales* y utiliza el servicio en línea*. El servicio en línea recibe las bases de datos de los productores de bases de datos. El productor de bases de datos las produce. El servicio en línea hace accesibles las bases de datos. El centro de información ofrece el servicio a los usuarios. El vocabulario de indización está controlado por el productor de las bases de datos. El Centro de información tiene control sobre parte del ciclo en cuanto a la selección de las bases de datos que se consultarán y a la preparación de las estrategias de búsqueda.

En este punto, se puede hablar de los servicios comerciales en línea de los Estados Unidos, tales como Bibliographic Retrieval System, el SDC, el Lockheed y de la Red Europea Diane.

UNIDAD I: TEORIA DE LA RECUPERACION DE LA INFORMACION

CLASE 4: VOCABULARIO CONTROLADO

Objetivo:

Explicar qué es un vocabulario controlado y analizar algunos de los más usados.

Idea principal:

El control de los lenguajes de indización contribuye a dar mayor eficacia a un Sistema de Almacenamiento y Recuperación de Información.

Ideas accesorias:

Cuando no se utilizan vocabularios controlados se pueden originar diferentes problemas, tales como: inconsistencia en el tratamiento de la información, fallas en la búsqueda, aumento de tiempo y costo en el Sistema.

Para usar eficientemente vocabularios controlados hay que prever las fallas que pueden presentar y conocer los métodos para solucionarlas.

La calidad de un sistema de recuperación de información* depende en gran medida del *vocabulario** utilizado para indizar los documentos.

En bibliotecología el uso de vocabularios controlados* ha sido una herramienta fundamental desde hace ya un siglo. Un esquema de clasificación* es una forma de vocabulario controlado, así como también lo es una lista de encabezamientos de materia*.

En la década del 50 Mortimer Taube crea el sistema Unitérmino* por el cual se indiza por medio de palabras únicas tomadas del texto analizado, sin controlar el vocabulario.

Los inconvenientes producidos por falsas coordinaciones entre términos al recuperar la información llevó a sustituir las palabras únicas por conceptos formados por una o varias palabras según las necesidades, es decir dio lugar a un cierto grado de precoordinación*. De "unitérminos" se pasó a "uni-conceptos" que podían estar formados por uno o varios términos. Como resultado de esta evolución surgieron los tesauros*, que son vocabularios totalmente controlados.

Actualmente, en aquellos países altamente industrializados, donde se usa habitualmente computadora, se vuelve al uso del lenguaje natural* pues es posible, con métodos automatizados, indizar documentos mediante la extracción de términos* significativos de los títulos, resúmenes o textos completos de dichos documentos.

TRANSPARENCIA I- 41

Un tesoro, una lista de encabezamientos de materia y un esquema de clasificación son diferentes tipos de vocabularios controlados.

En los tres casos se trata de un *conjunto de signos (términos o números)* tomados de un conjunto mayor de signos del lenguaje natural, que se utiliza para la indización* y la recuperación de información.

TRANSPARENCIA I- 42

Si no se utiliza un vocabulario controlado habrá:

- *inconsistencia*. El uso de un vocabulario controlado permite el tratamiento consistente y uniforme (coherente) de la información en varios niveles:
 - en la indización: *consistencia en la representación de un tema de un documento.*
coherencia entre los distintos indizadores al tratar un mismo tema.
 - en la búsqueda: *consistencia en la representación de una demanda de un usuario.*
coherencia entre los distintos indizadores al elaborar la estrategia de búsqueda* para un mismo tema.
 - coherencia entre la indización y la búsqueda de un tema.
- *separación*. Si hay inconsistencia habrá separación del material relacionado con un tema bajo distintos términos. El vocabulario controlado indica la relación de los conceptos* cuando existe dicha relación. En cambio, si se emplea lenguaje natural, el establecimiento de esas relaciones queda librado al conocimiento del tema por parte del indizador.
- *fallas en la búsqueda* producidas precisamente por existir los problemas de inconsistencia y separación.

Si no se tiene un vocabulario controlado aumentará la tarea de quien busca información pues las fallas le demandarán más tiempo de búsqueda, lo que implica también mayores costos.

Se puede decir que, a mayores costos de entrada se tendrán menores costos de salida. Esto significa que al trabajar con un vocabulario controlado aumentarán los costos en la entrada pero los costos de salida serán menores pues el vocabulario controlado facilita la búsqueda.

En cambio, si se usa un vocabulario libre bajarán los costos en la entrada, pero aumentarán en la búsqueda justamente por las fallas que hemos visto aparecer al trabajar sin vocabulario controlado. Según la política de indización aplicada a un servicio de información, variará la decisión respecto a los mayores costos en la entrada al sistema o en la salida, o buscar un equilibrio entre ambos. Si se tienen, por ejemplo, pocas consultas, no preocuparán mayores costos en la búsqueda, pero si el número de consultas es muy importante es más conveniente hacer el gasto necesario para indizar con un buen vocabulario controlado (aunque demande mayores gastos) que nos permita indizar bien y facilitar la tarea de búsqueda reduciendo los costos en la salida.

TRANSPARENCIA I- 43

Las funciones del vocabulario controlado en el sistema de recuperación de información se consideran tanto en la entrada al indizar como en la salida al efectuar la búsqueda.

1. Permite la representación consistente del tema al indizar y al buscar información pues el vocabulario controlado evita la dispersión de un tema bajo distintos encabezamientos* o descriptores*. Esto lo logra mediante el control de los sinónimos*, los casi sinónimos*, los antónimos* y los homónimos*.
2. Facilita búsquedas amplias, generales y una términos relacionados semánticamente. El vocabulario controlado permite las búsquedas generales pues las relaciones paradigmáticas* (de estructura) se especifican usando la referencia TG y TE en los tesauros y en las listas de encabezamientos de materia por medio de una o dos x. Los términos relacionados sintagmáticamente quedan indicados en el vocabulario por medio de la referencia véase además en las listas de encabezamientos de materia y por TR en los tesauros.

Un vocabulario bien estructurado debe mostrar todas estas relaciones verticales (paradigmáticas) y horizontales (sintagmáticas)*.

TRANSPARENCIA I- 44

En esta transparencia se muestran ejemplos de cómo un vocabulario controlado trata los sinónimos. Si se busca Ala Triangular se indicará que debe usarse Alas Delta. En el vocabulario dirá USE o véase. Significa que el encabezamiento o descriptor Alas Delta es la forma preferida y ha sido elegida para expresar el concepto de Ala Triangular.

La forma elegida está en relación con el tipo de usuario. Si se trata de un servicio de información para médicos muy probablemente bajo Pie de Atleta se vea la referencia USE o véase Tinea Pedis, que es la forma conocida científicamente.

En el caso de las siglas, se prefiere su forma abreviada cuando: 1) ésta es universalmente reconocida; 2) si el significado no depende de un contexto determinado; 3) cuando no se produzcan dificultades en la recuperación automática por diferencias de puntuación o uso de mayúsculas-minúsculas; 4) su forma desarrollada es casi desconocida (Ej. LASER (1), RADAR (2)). En todos los casos se harán las referencias a los términos elegidos.

TRANSPARENCIA I- 45

En un vocabulario controlado los casi sinónimos se consideran como los sinónimos. Los antónimos también pueden considerarse de la misma manera. Los antónimos son conceptos opuestos de una misma cualidad. La aspereza es la ausencia de lisura y la lisura es la ausencia de aspereza, existe una continuidad entre un término y otro. Son extremos de una línea continua en la cual a medida que nos acercamos de un extremo al otro se va perdiendo una cualidad y se va incorporando la otra.

TRANSPARENCIA I- 46

Problemas de los lenguajes de indización

Los problemas más comunes en los lenguajes de indización o, si se quisiera decirlo de otra forma, las fallas más frecuentes en los vocabularios controlados que afectan el rendimiento* del sistema de recuperación de la información son: en primer lugar, la falta de especificidad* en el vocabulario (Transparencias nros. 46/48), o sea el uso predominante de términos genéricos en el vocabulario; en segundo, las relaciones ambiguas (Transparencias nros. 49/52) que pueden darse por coordinaciones falsas y relaciones incorrectas y en tercer lugar, falta de una estructura consistente (Transparencias nros. 53/56) en el lenguaje por carencia de vinculaciones jerárquicas y de términos relacionados.

TRANSPARENCIA I- 47

Este es un ejemplo de un vocabulario con falta de especificidad, o sea un vocabulario muy general. Para indizar un documento sobre ratones ciegos se encontrará lo siguiente:

	RATONES	use	ROEDORES
y	CIEGOS	use	INCAPACITADOS

o sea que por falta de especificidad del vocabulario, los conceptos ratones y ciegos pierden su individualidad y pasan a clases más genéricas: ROEDORES o INCAPACITADOS en las cuales están contenidas.

(1) Light amplification by the stimulated emission of radiation (Amplificación de la luz por emisión estimulada de radiación).

(2) Radio detection and ranging (Descubridor e indicador radioeléctrico).

Al querer recuperar documentos sobre ratones se recuperarán documentos sobre todos los roedores, en consecuencia se pierde precisión* en la búsqueda. Si se tuviera un vocabulario más específico se podría recuperar con precisión documentos sobre ratones ciegos y si se deseara una búsqueda más general se recorrería la jerarquía hasta cubrir todos los documentos que hablan de roedores incapacitados. Por esta razón es importante que el vocabulario guíe hacia términos relacionados jerárquicamente para lograr una búsqueda general o específica según las necesidades.

TRANSPARENCIA I- 48

Por lo que se acaba de decir, se deduce que de la especificidad del vocabulario controlado depende en forma directa la precisión de la recuperación. Cuanto más específico es el vocabulario que se utiliza, mayor precisión se obtendrá en la indización* y en la búsqueda.

Si el vocabulario es muy general (en este ejemplo sería un vocabulario con el descriptor BIBLIOTECAS solamente) se perderá precisión al indizar y al recuperar. En cambio, si se tiene vocabulario más específico, BIBLIOTECAS ACADEMICAS, BIBLIOTECAS ESTUDIANTILES, BIBLIOTECAS ESCOLARES, etc. no se perderá precisión ni en la indización ni en la búsqueda pues se puede indizar y recuperar el concepto preciso del documento original, y se puede hacer al mismo tiempo, si se desea, búsquedas generales en la estructura jerárquica.

La especificidad tiene su precio porque:

- hace más dificultosa la recuperación exhaustiva, o sea el acierto* alto.
- hace la indización consistente más dificultosa pues la persona que realiza el trabajo tiene que elegir en el vocabulario el encabezamiento o descriptor más adecuado para este concepto. Si se tuviera un vocabulario con poco grado de especificidad la elección sería más simple ya que se optaría por un término general. Cuantos más términos tiene un vocabulario, más difícil se hace su uso en forma consistente.
- será más costoso el sistema pues el vocabulario tendrá muchos más términos.

A pesar de estos inconvenientes es importante la precisión en el vocabulario controlado en áreas especializadas del conocimiento.

TRANSPARENCIA I- 49

Si en estas coordenadas se grafica el acierto y la precisión se observa que a mayor acierto, haciendo una indización exhaustiva, se obtiene menor precisión. O sea cuanto más exhaustiva es la indización menor es la especificidad, lo que repercute en una menor precisión.

En última instancia es una decisión relacionada con la política de indización del servicio que indicará cuan exhaustiva o específica debe ser la indización. El ideal (100 o/o exhaustividad* y 100 o/o especificidad, punto derecho arriba en el cuadro) no existe.

TRANSPARENCIA I- 50

El problema de la *falsa coordinación* se da cuando en la búsqueda se coordinan términos que no están relacionados en el documento original. Cuanto mayor es el número de descriptores usados para indizar, mayores son las posibilidades de falsa coordinación.

En el ejemplo se muestran varios descriptores utilizados para indizar un documento pero se desconoce cuáles están relacionados entre sí en dicho documento. Esto puede dar lugar a falsas coordinaciones en el momento de la búsqueda. Las relaciones existentes en el documento son: ratones ciegos, ratas sordas y conejos mudos. Si se recupera por conejos ciegos o ratones sordos se haría una coordinación* falsa.

TRANSPARENCIA I- 51

Otro problema es el de las *relaciones incorrectas*. En este caso los términos asignados están relacionados en el documento pero cuando se realiza la búsqueda se pueden establecer relaciones incorrectas.

COMPUTADORA, DISEÑO y AVION son conceptos relacionados entre sí en un documento, pero no es lo mismo el *diseño de computadoras para manejar aviones* que el *diseño de aviones por medio de computadoras*.

LECTURA y EPILEPSIA son conceptos relacionados entre sí en un documento, pero no es lo mismo la *influencia de la lectura en casos de epilepsia* que la *lectura de documentos sobre epilepsia*.

TRANSPARENCIA I- 52

Para solucionar estos problemas se adoptan diferentes técnicas, entre las cuales las más comunes son las *uniones**, la *precoordinación** y los *indicadores de rol**. Las *uniones* de los términos se efectúa en el momento de indizar. A cada unión se le asigna una letra. Por ej.: ratones ciegos es la unión A en el documento n° 3002. Al recuperar documentos sobre ratones mudos obtendré el documento n° 3002 pues en su indización figuran estos dos términos, pero inmediatamente observaré que dice 3002/A para ratones y 3002/C para mudos, con lo cual se deduce que a pesar de estar en el mismo documento, estos dos términos no están relacionados. Una cierta *precoordinación* de los términos también ayuda a solucionar el problema de las relaciones ambiguas. Por ej. elegimos SOLDADURA ULTRASONICA en lugar de SOLDADURA y ACUSTICA ULTRASONICA.

TRANSPARENCIA I- 53

Otro recurso utilizado para solucionar problemas de relaciones ambiguas por coordinación falsa o relaciones incorrectas, es el de asignar indicadores de rol a términos, como por ejemplo: objeto de una acción, herramienta de esa acción, modo de realización. Al indizar un documento sobre diseño de aviones mediante computadoras asigno el rol de objeto de acción al avión y el de herramienta a modo de realización a computadora. De esta manera no existirá duda si se trata de un documento sobre diseño de computadoras para el manejo de aviones o de un documento sobre computadoras para el diseño de aviones.

TRANSPARENCIA I- 54

Debe asegurarse que el vocabulario tenga una *estructura*, o sea que los términos con relaciones semánticas* estén relacionados en el vocabulario para permitir búsquedas amplias o específicas según las necesidades del caso.

En este ejemplo particular, un buen tesoro* relacionará FALIAS (mecánicas) con estabilidad y con corrosión, fatiga y tensión, además de tener como términos específicos (TE) FRACTURA y RUPTURA.

TRANSPARENCIA I- 55

El *vocabulario de entrada** está formado por los términos que permiten acceder al vocabulario controlado o sea al descriptor elegido para expresar un concepto. El vocabulario de entrada debe estar formado por mayor cantidad de términos que el vocabulario de descriptores ya que, contiene sinónimos, casi sinónimos, antónimos, etc. Todos los términos del vocabulario (preferidos y no preferidos) pueden considerarse como términos de entrada.

TRANSPARENCIA I- 56

En general se considera el vocabulario de entrada a aquel que remite de una palabra no usada como descriptor a una usada.

El vocabulario de entrada, con todas sus referencias:

- evita inconsistencia en la indización y en la búsqueda.
- indica a quien busca dónde buscar.
- asegura a quien busca que ese tema está cubierto pues si no existiera algún documento on ese tema no figuraría el término en el vocabulario (cuando éste ha sido construido a partir de los documentos existentes)
- reduce la tarea del indizador, pues cuanto más elaborado y más vocabulario de entrada exista menos decisiones deberá tomar el indizador. Por cualquier término que entre al vocabulario, el mismo vocabulario lo guiará hasta el descriptor o encabezamiento elegido, donde le mostrará todas las relaciones.

TRANSPARENCIA I- 57

En este caso el vocabulario de entrada nos remite mediante la palabra USE a una entrada específica. Si no se tuviera vocabulario de entrada, el indizador podría elegir cualquiera de las tres posibilidades pues SOLDADURA DE ARCO ARGON no figura en el vocabulario y tiene que decidir con qué descriptor indizarlo.

TRANSPARENCIA I- 58 y I-59

Parámetros del sistema*

Antes de elegir un vocabulario controlado, se deben analizar los parámetros del sistema que afectan las decisiones respecto a los lenguajes de indización.

1. El *ámbito o alcance del sistema** indica cuáles son los temas centrales (núcleo) y cuáles los periféricos en el sistema. El sistema deberá establecer las áreas temáticas en las cuales el vocabulario deberá permanecer a nivel general (periferia) y en cuáles deberá ser más específico (núcleo). El alcance está en relación con el ámbito geográfico del servicio, acceso a otras bases de datos, políticas nacionales, problemas regionales, etc.
2. El *tipo de documentos** que ingrese al sistema de información* determinará las políticas para su tratamiento. Si la mayoría son libros, que pueden indizarse en forma más amplia ya que frecuentemente tratan de un solo tema, se puede elegir un vocabulario más general. En cambio, si el material consiste principalmente en informes o artículos de revistas que necesitan una indización más especializada, se elegirá un vocabulario más específico. El nivel de indización de informes internos y externos depende de la política de cada institución.
3. De acuerdo con el volumen actual de la literatura y la tasa de su crecimiento se decidirá qué tipo de indización se hará (específica o general). Además debe tenerse en cuenta el personal disponible, presupuesto, costos, etc.
4. El mayor o menor control del vocabulario estará determinado por el tipo de usuario que utiliza directamente el sistema. Si se trata de especialistas en campos temáticos, se preferirá el lenguaje natural, pero se optará por lenguaje controlado* en el caso que sean especialistas en información.
5. El *tipo de consultas*: Si estas consultas son muy generales no se justifica un vocabulario muy específico.
6. Los responsables de la política de indización del sistema deben decidir cómo equilibrar los costos de entrada y de salida. Cuanto más específica es la indización más altos son los costos de entrada pero más bajos los de salida.
7. La implementación física del sistema, sea manual o por computadora, dependerá de variables tales como los costos, los usuarios del sistema, el tipo de servicio, etc. El vocabulario que se elegirá dependerá de dicha implementación física.

2.5.2.

UNIDAD II

UNIDAD II: ESTRUCTURA DE LOS TESAUROS

CLASE 1: INTRODUCCION. PARTES CONSTITUTIVAS DE LOS TESAUROS. LOS DESCRIPTORES.

Objetivo:

Analizar la estructura de los tesauros, para llegar al concepto de lo que es un tesoro.

Idea principal:

El tesoro es un lenguaje controlado y dinámico de términos relacionados semántica y jerárquicamente que cubre un campo específico del conocimiento.

Ideas accesorias:

Un tesoro debe tener por lo menos dos partes: una alfabética y otra sistemática. El descriptor es la unidad significativa de un tesoro.

TRANSPARENCIA II- 1

Moore y Taube fueron dos precursores de los tesauros*.

Moore proponía el uso de vocabularios muy reducidos formados por términos muy amplios, para los que creó la denominación descriptor*. También a él se debe el concepto de recuperación de la información.

Mortimer Taube fue el creador en 1951 del Sistema Unitérmino*. En la Unidad I clase 4 se presentan las características del sistema y su evolución que dio lugar al nacimiento de los tesauros.

En el año 1959 la Compañía Du Pont construyó el primer tesoro, aunque no lo publicó. Era un tesoro para indexar documentos sobre química e ingeniería química.

En 1960 el Armed Services Technical Information Agency (ASTIA), ahora llamado Defense Documentation Center (DDC), produjo el *Thesaurus of ASTIA descriptors*.

A partir de estas experiencias iniciales se desarrolló el tesoro como una de las herramientas más eficientes de control terminológico en los sistemas modernos de almacenamiento y recuperación de información. En 1961 el American Institute of Chemical Engineers, editó el *Chemical Engineering Thesaurus*. En 1964 el Engineers Joint Council, editó el *Thesaurus of Engineering Terms*, con la asistencia de la National Science Foundation.

El interés y el esfuerzo común del sector industrial y técnico, representado por las sociedades de ingenieros, y del sector de defensa, representado por el Departamento de Defensa, dieron como resultado el monumental Proyecto LEX auspiciado por la Office of Naval Research. Del trabajo conjunto surgió el *Thesaurus of Engineering and Scientific Terms (TEST)*, 1967. Tenía más de 23.000 términos de acceso, se tardó un año en construirlo y participaron 300 especialistas organizados en equipos.

La realización de esta obra permitió acumular una gran experiencia en la construcción de tesauros. Con el fin de fijar dicha experiencia se establecieron normas y convenciones para la construcción del *Thesaurus of Engineering and Scientific Terms*. Se llamaron *LEX Conventions* para la construcción de tesauros y fueron adoptadas por el Committee on Scientific and Technical Information (COSATI) que recomendó su uso a todos los organismos oficiales. Estas convenciones constituyeron el material básico sobre el cual trabajó el USA Standards Institute (USASI) para la creación de las normas nacionales de los Estados Unidos.

La UNESCO también elaboró directrices para la construcción de tesauros monolingües* y multilingües*. Son completamente compatibles con las de USASI.

TRANSPARENCIA II- 2

En un tesoro, los términos están agrupados, en general, en dos formas que son complementarias:

- a) una parte alfabética
- b) una parte sistemática o clasificada

En la parte alfabética, la sistematización de los conceptos se da en forma encubierta. Esta sistematización se presenta en forma abierta en la parte clasificada.

TRANSPARENCIA II- 3

En esta transparencia se muestra un ejemplo de clasificación explícita*. Existe una relación jerárquica* entre el término BIBLIOTECAS y BIBLIOTECAS PUBLICAS, y a su vez entre BIBLIOTECAS PUBLICAS y BIBLIOTECAS MUNICIPALES Y POPULARES. Este mismo criterio se evidencia en las demás relaciones.

Esta clasificación explícita es la equivalente a la presentación arbórea de las clasificaciones, es decir que se podría representar de la siguiente manera:



TRANSPARENCIA II - 4

La transparencia 4 muestra un ejemplo de clasificación implícita *. Las relaciones entre los términos de esta transparencia son las mismas que las relaciones de la transparencia número 3, pero no se manifiestan gráficamente. En lugar de existir sangría para indicar las diferentes jerarquías, los términos van precedidos por abreviaturas que indican cuál es el grado de relación con el descriptor que encabeza el asiento. Así, BIBLIOTECAS es el término genérico que abarca BIBLIOTECAS ACADEMICAS y BIBLIOTECAS ESTUDIANTILES y BIBLIOTECAS UNIVERSITARIAS son los términos específicos o sea "clases*" de BIBLIOTECAS ACADEMICAS.

EDUCACION SUPERIOR es un término relacionado (vinculación asociativa) de BIBLIOTECAS ACADEMICAS.

La clasificación implícita se encuentra generalmente en la parte alfabética de cualquier tesauro donde figuran todos los descriptores con sus relaciones.

En resumen, un tesauro se compone básicamente de una parte alfabética, donde la sistematización de los conceptos está implícita, y una parte sistemática, donde el esquema conceptual se presenta en forma explícita. Esta parte sistemática, aunque muy importante, no siempre se incluye.

TRANSPARENCIA II- 5

El asiento de un tesauro puede tener hasta cinco elementos después del descriptor. En esta transparencia se pone énfasis en los términos de entrada (referencias UP) que remiten al descriptor y en los términos relacionados (TR).

TRANSPARENCIA II- 6

Esta transparencia presenta el asiento completo de un descriptor. El ejemplo fue tomado del *Macrotesauro de Desarrollo Económico y Social* de la OCDE.

El asiento está encabezado por el descriptor SERVICIOS SOCIALES. Primero aparece la Nota de Alcance*, NA (en inglés *Scope Note*, SN); a continuación figuran dos términos que, en este tesauro, pueden considerarse sinónimos* del descriptor que encabeza el asiento, precedidos por "Usado por", UP (en inglés *Used for*, UF); luego aparecen el o los términos genéricos, TG (en inglés *Broader Term*, BT), los términos específicos, TE (en inglés *Narrower term*, NT) y los términos relacionados TR (en inglés *Related terms*, RT).

Cada descriptor tiene tres tipos de relaciones: relaciones de equivalencia*, relaciones jerárquicas o paradigmáticas* y relaciones asociativas o sintagmáticas*.

TRANSPARENCIA II- 7

Al Crear Moers el término descriptor lo concibió integrado por dos elementos:

- a) la *etiqueta** que representa el concepto (o sea su forma o continente).
- b) la determinación del *alcance* del concepto.

TRANSPARENCIA II- 8

Con respecto a la forma de presentación del término o términos que constituyen un descriptor, la Unesco en sus Directrices ha contemplado el problema y ha elaborado una serie de normas.

TRANSPARENCIA II- 9

Las directrices indican que se prefiere: a) la forma directa sobre la indirecta; b) la forma sustantiva sobre la forma adjetiva, aunque se aceptan excepciones; c) la forma plural sobre la singular, siempre que el plural y el singular no tengan significados diferentes, por ejemplo, no es lo mismo ESTADISTICA como disciplina que ESTADISTICAS como datos resultantes de una investigación estadística.

El singular se usa en general para mencionar procedimientos y propiedades.

Una ayuda para saber si un nombre va en singular o plural es preguntar cuántos o cuánto. Los nombres que responden a la pregunta cuántos irán en plural ya que estarán relacionados con cantidad. Los nombres que responden a la pregunta cuánto irán en singular ya que están relacionados con volumen.

TRANSPARENCIA II- 10

Las abreviaturas deben evitarse excepto cuando su significado está reconocido por el grupo de usuarios al cual va dirigido el tesaurus.

La puntuación debe evitarse aunque muchas veces es necesario incluir el guión y el paréntesis.

TRANSPARENCIA II- 11

Los diferentes significados de los homógrafos* se aclaran mediante calificadores que se escriben entre paréntesis y forman parte del descriptor.

TRANSPARENCIA II- 12

En esta transparencia se muestra una síntesis de los criterios que deben seguirse para lograr la normalización de los descriptores en un tesaurus.

Con respecto a la cantidad de términos que deben constituir un descriptor, las normas de la Unesco recomiendan el menor número posible.

Presentación de los tesauros

Es común que un tesaurus presente los descriptores ordenados alfabéticamente con notas de alcance y todas sus relaciones con el fin de permitir el acceso en la forma más rápida y sencilla posible.

En general, la parte alfabética constituye el cuerpo de la mayoría de los tesauros existentes pero debe incluirse además la parte sistemática.

A continuación se muestran transparencias con ejemplos de distintas formas de presentación de los tesauros (parte alfabética y parte sistemática).

En primer lugar se da un ejemplo completo del *Macrothesaurus*, edición de 1979, presentando sus 3 partes principales.

Este Macrotesaurus tiene ediciones en tres idiomas: español, inglés y francés. En cada edición los descriptores aparecen en un idioma con sus equivalentes en los otros.

TRANSPARENCIA II- 13

Parte alfabética del *Macrothesaurus* de la OCDE en español.

Aparecen los asientos completos de los descriptores; incluye códigos que remiten a la parte sistemática.

UNIDAD II: ESTRUCTURA DE LOS TESAUROS

CLASE 2: PRESENTACION DE LOS TESAUROS

Objetivo:

Analizar ejemplos de diferentes presentaciones de tesauros.

Idea principal:

Los tesauros pueden tener distintas formas de presentación.

Ideas accesorias:

La forma más generalizada de presentación es la alfabética.

Parte sistemática del Macrotesauros de la OCDE en español.

Esta parte tiene dos presentaciones: por temas y por ordenamiento jerárquico.

TRANSPARENCIA II- 14

Tesoro* por temas: Grupos de descriptores* correspondientes al tema Medio Ambiente, Recursos Naturales.

TRANSPARENCIA II- 15

Ordenamiento jerárquico

Los términos que encabezan las jerarquías, en el ejemplo IRRADIACION, ISOTOPOS, JUBILACION, etc., están ordenados alfabéticamente. Debajo de cada uno de ellos aparecen los términos específicos correspondientes, en diferentes niveles de jerarquía, indicados por sangrías.

Para hacer este ordenamiento se toman los descriptores que no tienen términos más amplios y se van buscando los descriptores específicos en diferentes grados de profundidad.

TRANSPARENCIA II- 16

Presentación alfabética del Tesoro de la NASA (1967).

Debajo de cada descriptor aparecen los códigos correspondientes a la parte sistemática.

Debe señalarse que bajo el descriptor ANATOMIA aparecen todos los términos específicos correspondientes a diferentes niveles, en un solo orden alfabético, sin ninguna indicación de esas diferencias.

TRANSPARENCIA II- 17

Presentación alfabética del Water Resources Thesaurus (1971).

Como este tesoro no incluye la parte sistemática, en la presentación alfabética muestra -por medio de sangrías- los diferentes niveles jerárquicos de los términos específicos.

TRANSPARENCIA II- 18

La transparencia muestra un mismo asiento de tesoro con dos presentaciones diferentes. En la primera, los términos específicos aparecen en un solo orden alfabético sin tener en cuenta los niveles jerárquicos. En la segunda, se diferencian los niveles jerárquicos por medio de sangrías.

Son muchas las posibles *presentaciones sistemáticas de los tesauros*, por lo tanto se hace una selección de las más interesantes.

En líneas generales estas formas de presentación pueden ser:

- a) por grupos temáticos.
- b) en gráficos: circulares
 flechados
 poligonales
 ruteros
- c) por facetas*

Los tesauros presentados gráficamente tienen gran aceptación y uso en Europa mientras que en los Estados Unidos se tiende a editar solo la parte alfabética. Una experiencia sumamente interesante es la que se está haciendo con modelos físicos (tridimensionales) que muestran el sistema de relaciones dentro de los tesauros.

TRANSPARENCIA II- 19

Este es un ejemplo de una de las formas más usadas de presentación de categorías* temáticas. Debajo de la categoría aparecen todos los descriptores en orden alfabético.

TRANSPARENCIA II- 20

Este es el mismo grupo de descriptores del ejemplo anterior con la indicación de diferentes niveles por medio de sangrías. Se tiene así una presentación temática jerárquica.

TRANSPARENCIA II- 21

Presentación en gráficos

Presentación circular: Este es un ejemplo de una página del tesoro TDCK (*Technisch Documentatie Centrum voor der Krijgsmacht, Netherland Armed Services Technical Documentation and Information Center*), Agencia de Información Técnica de las Fuerzas Armadas de los Países Bajos, 1963. Los términos están ubicados en círculos concéntricos. Cada círculo representa un nivel de jerarquía. En el centro se encuentra el término que encabeza la jerarquía.

Cada gráfico del tesoro representa una clase. Se complementa con un índice alfabético donde cada término va acompañado del número correspondiente a la clase. Una de las ventajas de este tipo de presentación es que ofrece una visión panorámica de cada clase*. Fuera del círculo aparecen los términos relacionados que pertenecen a otras clases con los números correspondientes.

TRANSPARENCIA II- 22

El tesoro de EURATOM (*European Atomic Energy Community*) es otra forma de presentación gráfica. El

ejemplo de esta transparencia corresponde a la primera edición.

Es un tesoro muy amplio, cada gráfico flechado del tesoro representa un grupo de descriptores encabezados por una palabra-clave* (clase). Las flechas indican las relaciones entre los términos*.

El ejemplo muestra el grupo 85: PARTICULAS ELEMENTALES. En la página del gráfico se incluyen los descriptores del grupo en orden alfabético. La dirección de las flechas va del descriptor genérico al o a los descriptores específicos. Los términos de igual nivel jerárquico están vinculados con doble flecha.

TRANSPARENCIA II- 23

Este ejemplo corresponde a la segunda edición (1966-67) del Tesoro de EURATOM. Se ha cambiado la presentación ya que se encuadran los sub-grupos de términos relacionados semánticamente en racimos alrededor de las palabras clave que aparecen en mayúsculas. A esta presentación se la denomina poligonal. La clase PARTICULAS (en general) número 47 es más amplia que la anterior. Ha desaparecido el listado alfabético y aparecen líneas de diferente grosor que señalan la fuerza de la relación semántica* entre los grupos de conceptos*.

Los números que figuran en cada racimo indican cuántas veces ha sido utilizada la palabra clave en la indización* del documento*. Los términos relacionados pertenecientes a otras clases se ubican en el recuadro.

TRANSPARENCIA II- 24

El ejemplo de esta transparencia muestra una experiencia de presentación gráfica concebida por Loren Doyle como un atlas o mapa rutero semántico. En el mapa el tamaño de la letra indica la importancia relativa de cada término dentro de la clase. Los números indican la frecuencia de concurrencia de los términos señalados por las flechas, por lo tanto indican la vinculación entre esos términos. Doyle concibió este tipo de tesoro para un sistema de recuperación en línea* en el cual se puede tocar el término en la pantalla mediante un dispositivo especial y aparecen sus relaciones. Esto permitía acceder en forma inmediata a cualquier término del vocabulario.

TRANSPARENCIA II- 25

El Tesoro SPINES fue construido entre 1972 y 1976 por miembros de la Unesco, con la asistencia de expertos. Se creó como herramienta del sistema SPINES (Sistema Internacional para Intercambio de Información en Ciencia y Tecnología en los aspectos de Política, Administración y Desarrollo) de acuerdo con las normas que la Unesco determinó para el UNISIST. El asiento del descriptor tiene los elementos comunes. Está encabezado por el descriptor y un código numérico (B).

Incluye las referencias comunes a todos los tesauros (de un término no usado al descriptor elegido). En el ej.:

- A. R & D
C. uf research and development
I. research and development
USE R&D

Pero además, presenta otras referencias, tales como:

- 1) de un término a más de uno, en el ej.:
A. R & D
D. uf+ rd&e
I. rd&e
USE ENGINEERING
AND R&D

- 2) De un término a más de una opción, en el ej.:
- A. R & D
- E. sf (tomado por) innovation processes (referencia con opción)
- H. innovation processes
 see (véase) R & D
 or (o) TECHNOLOGY TRANSFER

Esto significa que el concepto contenido en el término *innovation processes*, no elegido como descriptor, puede ser expresado por dos descriptores: R & D o TECHNOLOGY TRANSFER. El tesoro ofrece la posibilidad de opción.

En el asiento del descriptor aparecen todos los niveles de términos específicos, con la indicación del grado de nivel, En el ej.:

- F. NT1 EXPERIMENTAL DEVELOPMENT
 NT1 RESEARCH
 NT2 APPLIED RESEARCH
 NT3 PROCESS RESEARCH
 NT3 PRODUCT RESEARCH
 NT2 FUNDAMENTAL RESEARCH

Obsérvese que a APPLIED RESEARCH, ubicado en el nivel 2, le siguen sus dos específicos NT3. A continuación aparece el otro descriptor de nivel 2: FUNDAMENTAL RESEARCH, que no tiene específicos.

En el asiento del descriptor intermedio G. APPLIED RESEARCH aparecen sus dos términos genéricos, bt1 y bt2 (nivel 1 y nivel 2).

Finalmente, en el asiento del descriptor aparecen los términos relacionados (RT).

TRANSPARENCIA II- 26

Parte sistemática del SPINES: Presentación de gráficos terminológicos.

El campo temático total del SPINES se dividió en 34 subcampos. En un primer gráfico terminológico los 34 subcampos y sus relaciones se visualizan en un gráfico global. A continuación se presentan los 34 gráficos que representan cada subcampo.

TRANSPARENCIA II- 27

En el anexo, aparece un número de polígonos conectados por líneas. Cada polígono representa una clase de conceptos que se individualiza por un número (de dos dígitos) (véase C) y un descriptor principal impreso con letras negritas y mayúsculas). Los descriptores de menores niveles se agrupan en polígonos más pequeños (véase A). Todas estas relaciones son jerárquicas*. Los no descriptores (sinónimos o cuasi-sinónimos*) están precedidos por el signo - (véase E). Un asterisco indica que ese término tiene una nota de alcance* en la lista estructurada alfabética (véase D). Las líneas conectan polígonos mostrando relaciones de asociación (véase B).

El gráfico tiene un título (véase C) y un código (véase F).

En el reverso se encuentra:

- 1) una nota técnica sobre el uso de los gráficos. Esta se repite, por razones didácticas, exactamente igual en los 34 gráficos.
- 2) una lista de relaciones asociativas (TR) entre los descriptores que encabezan los polígonos de ese gráfico con los descriptores que encabezan los polígonos de otros gráficos.

- 3) una reducción del gráfico global donde aparece sombreado el número y nombre del gráfico en cuestión, destacando su posición en el conjunto.

TRANSPARENCIA II- 28

El Faceted Subject Classification for Engineering es un ejemplo de presentación facetada.

La clase o categoría switches (interruptores) se ha dividido en dos facetas de acuerdo con dos principios diferentes: por método de conexión y por mecanismos de operación. Dentro de cada faceta los descriptores aparecen en orden jerárquico indicado por sangrías. Cada descriptor tiene una notación. Los términos relacionados están indicados a continuación de cada descriptor entre paréntesis por medio de los códigos correspondientes. Por ej.: Air Break (OB, OBX).

Aitchinson desarrolló tesauros facetados* en el campo de la ingeniería. El Thesaurofacet es la cuarta edición de la Faceted Classification for Engineering.

En cada una de las partes del tesoro facetado, alfabética y sistemática, se da información complementaria sobre cada descriptor y por este motivo deben consultarse ambas.

TRANSPARENCIA II- 29

Este es un ejemplo de un tesoro que presenta los asientos completos (relaciones TG, TE, USE, UF, TR y Notas de Alcance) en la parte jerárquica. Cada faceta tiene un código numérico (Ej.: 31.01 WAGE THEORY, Teoría del salario). Como es un tesoro generado por computadora incluye el número de cada asiento dentro del archivo.

La parte alfabética de este tesoro remite por el código numérico a la parte jerárquica. Incluye los términos de entrada que remiten a los descriptores. Se presenta como un índice KWOC.

TRANSPARENCIA II- 30

Este es un ejemplo del Tesoro de la formación profesional de CINTERFOR. Fue editado en 1979. La parte sistemática está estructurada por grupos temáticos y por sub-grupos. Dentro de los sub-grupos se encuentran los descriptores con su asiento completo. Se siguen los criterios del tesoro de la OIT de la cual CINTERFOR es un organismo especializado. La parte alfabética se presenta como un índice KWOC.

TRANSPARENCIA II- 31

Los tesauros, en general, llevan índices que complementan la parte principal. En el ejemplo se presenta un índice jerárquico que complementa la parte alfabética. Los niveles jerárquicos están representados por sangrías, en este caso marcadas por puntos.

TRANSPARENCIA II- 32

Este es un índice presentado en forma permutada (Key Word out of Context, KWOC).

TRANSPARENCIA II- 33

Este ejemplo corresponde a un tesoro generado por computadora mediante el programa* de expansión AVOCON, en tres etapas:

- 1) Entrada inicial de los términos (Fig. 2.2A). Los términos se ingresan con sus correspondientes Términos Genéricos (BT).
- 2) Asientos del tesoro generados por el AVOCON (Fig. 2.2B). Completa cada asiento con los términos específicos correspondientes.
- 3) Índice jerárquico de un asiento del tesoro (Fig. 2.2C). Los términos específicos de los asientos anteriores están ordenados por nivel de jerarquías. Las sangrías se indican con puntos.

TRANSPARENCIA II- 34

Presentación del asiento de un tesoro en un sistema en línea.

En la pantalla aparece el asiento desarrollado correspondiente al término principal o descriptor que fuera solicitado. En el ej.: SOLAR CONCENTRATOR (Concentrador solar). La diferencia con la presentación estándar de tesauros se da al agregar una numeración corrida a los descriptores genéricos, a los específicos y a los relacionados.

Esta numeración permite solicitar al sistema el asiento completo de otros descriptores, mediante una orden codificada que incluye ese código. (ej.: NT5).

TRANSPARENCIA II- 35

Una vez construido un tesoro se procede a su edición con el fin de difundir la obra terminada. Veamos entonces cuáles serían las partes constitutivas de la obra:

- 1) la portada: donde se disponen los datos que identifican el trabajo (autor, título, edición, lugar, editor, fecha, serie).
- 2) El sumario o contenido: donde se consignan las distintas partes de la obra con su paginación correspondiente.
- 3) La introducción: donde se hace una breve historia del tesoro, los motivos de su creación, la metodología seguida, los criterios adoptados, los aportes recibidos, los especialistas e instituciones que colaboraron, etc. Es muy importante señalar las innovaciones introducidas, los aportes de interés para los especialistas en información.
- 4) El cuerpo de la obra: parte alfabética y parte sistemática.
- 5) Los anexos. Por lo general son índices* de distinto tipo: de siglas, de abreviaturas, de fórmulas, de símbolos, de nombres geográficos, de equivalencias a otros idiomas, permutados*, etc.).

2.5.3. UNIDAD III

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 1: CONSIDERACIONES PRELIMINARES A LA CONSTRUCCION DE TESAuros.

Objetivo:

Hacer comprender que antes de construir un tesoro se deben considerar varios problemas relacionados con el sistema de información al que se aplicará y con la existencia de otros tesauros similares.

Ideas principales:

Antes de construir un tesoro se debe analizar el sistema de información al que servirá y considerar las decisiones alternativas que se presentan.

Ideas accesorias:

Es necesario averiguar si existen tesauros en el campo temático de interés.

Si existiera un tesoro aceptable como marco general del que se quiere construir, se lo puede adaptar ampliando o restringiendo algunas clases.

Se presentan las transparencias I-53 y I-54. Parámetros del Sistema. Se sugiere la ejemplificación con el caso de una unidad de información que requiera un tesoro para procesar su fondo documental. En relación con esta situación se analizarán los parámetros del sistema.

TRANSPARENCIA III- 1

Adoptar significa elegir el tesoro* que más conviene entre los existentes. Para ello será necesario realizar una búsqueda bibliográfica para saber qué tesauros cubren el campo temático de interés (Anexo 1). Existen bibliografías sobre tesauros. (Anexo 2).

Hay un inventario sobre Lenguajes Documentarios no Tradicionales en Español y/o Portugués, elaborado por el Grupo de Trabajo sobre Lenguajes Documentarios, de Reuniber (Conferencia Iberoamericana sobre Información y Documentación Científica y Tecnológica). Este grupo tiene su sede en el Centro Argentino de Información Científica y Tecnológica (CAICYT).

Adaptar. Se puede adaptar ampliando o restringiendo algunas clases*. Si existiere algún tesoro aceptable como marco general se lo puede adaptar mediante el desarrollo en detalle de ciertas áreas.

Traducir. Actualmente es fácil encontrar en otros idiomas, especialmente en inglés, tesauros que respondan a la necesidad de una unidad de información.

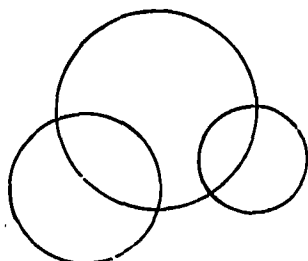
Es aconsejable aprovechar los ya existentes, pues traducir un tesoro resulta un esfuerzo menor que construir uno nuevo aunque en cierto sentido represente una tarea de adaptación.

Hacer un microtesauro compatible.* Cuando se necesita un tesoro sobre un tema específico incluido en un tesoro sobre un campo más amplio, conviene profundizar ese tema en el marco del tesoro mayor.

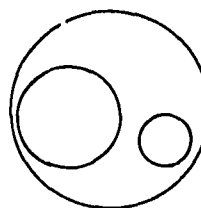
Microtesauro. Un tesoro puede ser usado para desarrollar otro más especializado o más técnico, actuando como fuente para la elección de términos generales y como guía para las relaciones.

Un tesoro puede formar también la base de una red de información* que opera con un conjunto de términos técnicos y con numerosos subconjuntos que, a su vez, podrían considerarse tesauros.

Tanto los más especializados como los subconjuntos, pueden ser microtesauros, pero son construidos en forma diferente según los propósitos que se quieran alcanzar. Las dos formas se pueden ilustrar con los siguientes diagramas:



Los términos del microtesauro sólo están incluidos parcialmente en el tesoro general.



Los términos del microtesauro son subconjuntos del tesoro general.

Dos ejemplos ilustran esta interdependencia entre tesauros.

Thesaurus of Engineering and Scientific Terms

Usado como fuente de términos para tesauros muy técnicos que superaban en cierta forma el vocabulario de ingeniería o caían enteramente dentro del alcance del tesoro padre.

Thesaurus for ERIC Descriptors

Usado fundamentalmente como vocabulario base para la indización* y recuperación de la literatura informativa en el campo de la educación. Es también un punto de coordinación para una serie de vocabularios más técnicos o tesauros desarrollados por los centros de referencia que forman la red del ERIC.

Adoptar partes (periferia). Es posible adoptar partes de otros tesauros para la periferia del que se quiere construir. Algunas secciones pueden adoptarse fácilmente, por ejemplo: propiedades físicas, regiones geográficas, instituciones, etc.

Construir un nuevo tesoro desde el comienzo

TRANSPARENCIA III- 2

Existen dos formas básicas para construir un tesoro: *aposteriori de la indización* (método empírico o de abajo hacia arriba) y *apriori de la misma* (de arriba hacia abajo).

En la primera, el vocabulario surge de la libre indización de un grupo de documentos. Los términos candidatos resultantes de este procedimiento son revisados, agrupados y estructurados en una organización útil.

Fue el método usado en la construcción del tesoro Dupont. Wooster se refiere a él como el "método de las estalagmitas". Se comienza desde abajo y se va construyendo la estructura.

En la segunda, los términos se extraen de glosarios y otras fuentes aunque surgen, fundamentalmente, de las discusiones del comité de especialistas en la materia, quienes consideran las clases principales del campo temático y las subdividen.

Recién cuando el vocabulario llega a su forma final o semifinal se usa para indizar literatura. Este es el método empleado para la construcción del Thesaurus of Engineering Terms del Engineers Joint Council (EJC). Wooster lo llama "método de las estalactitas".

Aunque Lancaster reconoce que los especialistas pueden jugar roles importantes como asesores, particularmente para la estructura del vocabulario, en la construcción de tesauros no es conveniente confiar sólo en los comités. Es preciso, además, contar con la rica verificación empírica de la indización de documentos representativos.

El método empírico es especialmente apto para las áreas muy especializadas del conocimiento. El segundo método puede aplicarse a los campos temáticos amplios que cubren varias disciplinas. Es deseable una combinación de ambos métodos aunque Lancaster recomienda empezar de abajo hacia arriba.

Cabe mencionar que para lograr un tesoro con buena consistencia se deben tener en cuenta ciertas normas o directrices como las indicadas en la Unidad II.

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 2: RECOLECCION DE TERMINOS

Objetivo:

Enseñar a sistematizar la recolección y registro de términos.

Idea principal:

Un tesauo debe ser capaz de describir conceptos que se encuentran en la literatura sobre ese campo y también ser capaz de interrogar sobre dichos conceptos al sistema.

Ideas accesorias:

Es fundamental extraer términos de la literatura primaria*.

La inclusión de un término* se justifica si es utilizada por los usuarios.

El registro de los términos que se van recolectando debe abarcar toda la información a incluir en el tesauo.

Una vez limitado el campo temático que se va a cubrir comienza la recolección de términos. Un tesauo es esencialmente una herramienta práctica. El primer requerimiento que debe cumplir es el de ser capaz de describir conceptos* que se encuentran en la literatura sobre ese campo y también debe ser capaz de interrogar sobre dichos conceptos al sistema. Por lo tanto, un tesauo necesita tener como base las características* de los documentos y de las preguntas y ceñirse continuamente a ellas.

TRANSPARENCIA III- 3

Justificación literaria*. El clasificador británico Wyndham Hulme creó en 1911 el término "literary warrant" (justificación literaria), también conocida como "bibliographic warrant" (justificación bibliográfica). De acuerdo con Hulme las clases* en las que se organizan los documentos no deberían fundamentarse en una "clasificación teórica del conocimiento" sino en grupos que los documentos parecen formar por sí mismos; es decir, las clases deben constituirse en función de la literatura existente. En otras palabras, las características de la literatura en sí misma determinarán las clases definidas en el sistema. El vocabulario del sistema* debería, entonces, apoyarse en un estudio real de la literatura existente en los varios campos temáticos cubiertos por el sistema. Esto en los términos de Hulme es la "trama de las áreas preexistentes en la literatura".

Obviamente, un vocabulario derivado empíricamente de la indización de los documentos tiene considerable justificación literaria, mientras que un vocabulario construido por el método de los comités puede tener poca.

TRANSPARENCIA III- 4

Para conocer los términos que existen en la bibliografía se puede recurrir a varias fuentes:

- a) diccionarios, enciclopedias y glosarios.
- b) índices de manuales, textos y publicaciones periódicas.
- c) revistas de resúmenes (los resúmenes analíticos presentan los términos más importantes de un trabajo).
- d) listados de computadora que registran la frecuencia de aparición de los términos o la frecuencia de la concurrencia en los textos, dado que se supone que los términos que aparecen siempre juntos o cercanos están relacionados semánticamente.

La ventaja de estas fuentes de referencia reside en que los términos están controlados. Sin embargo, no se debe excluir la extracción directa de términos de la literatura primaria: títulos, resúmenes hechos por el mismo autor, conclusiones de los trabajos. También es muy efectivo el marcado de palabras significativas registrando su frecuencia.

TRANSPARENCIA III- 5

La inclusión de un término se justifica si es utilizado por los usuarios. Las fuentes para seleccionar los términos justificados por el usuario son las siguientes:

1. pedidos de búsquedas (historia de los pedidos de una unidad de información).
2. cuestionarios que se envían a los usuarios para que los completen con los siguientes datos:
 - a) perfil de interés (sobre la base de los proyectos de investigación, o los temas que interesan particularmente a los usuarios).
 - b) descripción de tareas.
 - c) títulos de publicaciones consultadas.
 - d) resúmenes de publicaciones consultadas.
3. Los métodos de Dym y de Pickford combinan las fuentes literarias con las del usuario.

Se solicita a los usuarios que señalen los términos más importantes en la literatura correspondiente a sus campos temáticos, que ordenen los términos según su importancia y que los definan.

TRANSPARENCIA III- 6

Registro de términos. Los términos se van registrando en una ficha o en un formulario. Cada registro debe incluir la información que figura en la transparencia.

Es importante indicar la fuente de donde se tomó el término: diccionario, otro tesoro, etc.

La ficha puede hacerse por duplicado para permitir el ordenamiento sistemático de un juego. La información acerca del término y de sus relaciones se agregará gradualmente durante el proceso de compilación. La información inicial será muy completa en el caso de términos extraídos de fuentes de referencia en las que las relaciones ya están registradas.

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 3: ORGANIZACION DE LOS TERMINOS

Objetivo:

Enseñar a sistematizar los conceptos, por facetas y por jerarquías.

Idea principal:

El análisis por facetas, aunque se hace en forma empírica, debe tener como marco de referencia ciertas categorías fundamentales que se determinan previamente de acuerdo con el conocimiento del campo temático que se está cubriendo.

Las subfacetas se organizan sobre la base de un sólo principio de división.

Ideas accesorias:

La organización de los términos registrados empieza por su agrupamiento en pilas de acuerdo con el reconocimiento de analogías y diferencias.

TRANSPARENCIA III- 7

Organización de los términos.

Una vez que se registran los términos* en fichas, éstas se agrupan de acuerdo con las relaciones que se observan entre los términos. Se forman así pilas que en forma empírica representarán diferentes facetas*.

El ejemplo de esta transparencia corresponde a facetas del campo de la bibliotecología.

Aunque Lancaster recomienda hacer esta tarea en forma empírica, se debe tener como marco de referencia ciertas categorías* fundamentales que se determinan previamente de acuerdo con el conocimiento del campo temático que se está cubriendo.

Todo aquel que hace una clasificación* enfrenta el problema de determinar cuáles son las categorías importantes o facetas de un campo temático particular.

TRANSPARENCIA III- 8

Durante años se han hecho intentos para establecer categorías fundamentales aplicables a todos los campos: Kaiser en 1911, Ranganathan en 1933, Vickery en 1958 y Campbell en 1963.

Las categorías fundamentales no deben forzar el campo al que se aplican. Son un marco de referencia, pero cada campo presenta un conjunto especial de problemas al que esas categorías deben adaptarse. En última instancia, cada campo debe dividirse en la forma más útil para los usuarios y para el sistema de información.

TRANSPARENCIA III- 9

Una vez que los términos están agrupados en facetas se empieza a reconocer las subfacetas* y a establecer relaciones jerárquicas* entre dichos términos.

Cada subfaceta se determina sobre la base de un solo principio de división.

En la faceta "materiales bibliotecarios" se reconocen dos subfacetas: por forma de publicación y por contenido.

Dentro de cada subfaceta los términos se organizan jerárquicamente en una presentación estricta de género-especie (en el ejemplo por medio de sangrías).

TRANSPARENCIA III- 10

Las diferentes facetas pueden tener una longitud variable. No es necesario que todas las facetas tengan la misma cantidad de términos. Depende de las necesidades de los usuarios y de las características de la colección. Una faceta puede consistir en un solo término.

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 4: PROCESO DE REVISION DE LOS TERMINOS

Objetivo:

Enseñar a decidir cuál es el término más apropiado para cada concepto y a revisar exhaustivamente los términos recopilados antes de tomar una decisión.

Idea principal:

Para lograr la consistencia del tesoro se deben revisar cuidadosamente los términos registrados para verificar si se ha aplicado siempre el mismo criterio.

Ideas accesorias:

La especificidad de un tesoro depende de la habilidad de sus términos para expresar la materia con la profundidad y con el detalle que requieran los usuarios a los cuales va dirigido.

El nivel de precoordinación de un tesoro depende de que los conceptos se expresen por una sola palabra o por varias.

Para lograr la consistencia de un tesoro se deben revisar cuidadosamente los términos registrados a fin de verificar si se han aplicado siempre los mismos criterios.

Para la forma de entrada de los términos y forma de la palabra puede presentarse nuevamente la transparencia II- 8.

Para la elección de singular o plural se recomienda presentar la transparencia II- 9., para abreviaturas y puntuación la II-10., y para homógrafos, la II-11. Además, en todos los casos, es conveniente ejemplificar en el pizarrón.

TRANSPARENCIA III- 11

Especificidad* relacionada con las necesidades de los usuarios.

A un médico le puede interesar el término genérico perros: en cambio al veterinario le interesa la especie y la familia.

A los médicos les interesa la especificidad de la arteria carótida y la generalidad del análisis matemático. A los matemáticos les interesa la especificidad de las Series de Fourier y la generalidad de su aplicación en medicina.

La especificidad tiene la ventaja de influir en el rendimiento del lenguaje ya que obliga a definir con exactitud los conceptos y, consecuentemente, excluye el ruido*.

La desventaja de un vocabulario* altamente específico consiste en que el número de términos de indización* aumenta y por lo tanto, la compilación, el mantenimiento y la operación son más costosos.

El nivel de precoordinación* de un lenguaje* de recuperación depende de que los conceptos* se expresen por una sola palabra o por varias.

Un sistema en el que predominan los unitérminos* tiene un nivel bajo; uno que usa conceptos expresados por dos o tres palabras tiene un nivel medio, y uno que tiene una gran proporción de conceptos expresados por dos, tres o más palabras tiene un alto nivel de precoordinación.

El problema de los términos compuestos está relacionado estrechamente con la especificidad del vocabulario. Cuanto más complejos son los términos de indización, más específico y más preciso es el vocabulario y más extenso el número total de términos. Pero, desmembrando los términos precoordinados en términos simples, el tamaño del vocabulario se reduce y el acierto* potencial se incrementa a través de referencias del vocabulario de entrada*, desde los términos compuestos a la combinación apropiada de términos simples.

Se usan descriptores* precoordinados en los siguientes casos:

1. Cuando los descriptores simples tienen un significado distinto según se utilicen solos o se incluyan en el descriptor precoordinado.

Ej.: ACIDO MINERAL.

2. Cuando los descriptores simples se utilizan en relaciones jerárquicas distintas a las de los descriptores precoordinados.

Ej.: ARCOIRIS ARCO IRIS

3. Cuando el descriptor precoordinado es un nombre propio o bien si es dudoso que la asociación de descriptores simples refleje exacta y exclusivamente el contenido conceptual.

Ej.: PSICOLOGIA DE LOS ANIMALES = ANIMALES + PSICOLOGIA

pero ARQUITECTO INDUSTRIAL \neq ARQUITECTO + INDUSTRIA

ni PIE DERECHO \neq PIE + DERECHO

Los términos pueden precoordinarse antes de ser incorporados al sistema y deben recuperarse de manera idéntica, o coordinarse durante la búsqueda (post-coordinación)* para representar el concepto buscado.

La post-coordinación debe hacerse en los casos en que las divisiones morfológicas y semánticas sean idénticas.

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 5: INTRODUCCION A LA CONSTRUCCION PRACTICA DEL TESAuro. NOTAS DE ALCANCE.

Objetivo:

Enseñar los pasos del método empírico y establecer las relaciones entre los términos con sus respectivas referencias. Explicar el sentido y la importancia de las notas de alcance.

Idea principal:

Para construir un tesauro se recomienda seguir el método empírico.

La nota de alcance debe precisar el significado de un término dentro del sistema al que el tesauro servirá.

TRANSPARENCIA III- 13

El método empírico parte de la recolección de términos* justificados por la literatura y por el usuario; en una segunda etapa, agrupa dichos términos en clases* de acuerdo con sus afinidades semánticas; luego, los organiza en jerarquías y reconoce relaciones entre esas jerarquías, y, finalmente, vuelca todo el esquema armado en una organización alfabética.

TRANSPARENCIA III- 14

Para establecer las relaciones jerárquicas* debe trabajarse con las jerarquías inmediatamente superiores e inferiores (un paso hacia arriba, un paso hacia abajo).

En el ejemplo, BIBLIOTECAS es la jerarquía inmediatamente superior de BIBLIOTECAS ACADEMICAS (un paso hacia arriba) y BIBLIOTECAS ESTUDIANTILES es la jerarquía inmediatamente inferior (un paso hacia abajo).

TRANSPARENCIA III- 15

Este es un ejemplo de lo que se debe evitar, pues no se ha seguido el principio de un paso hacia arriba y un paso hacia abajo.

TRANSPARENCIA III- 16

La relación jerárquica entre los términos, es decir TG—TE, es una relación *género-especie* o *cosa-clases de*.

La relación *todo-parte* no es relación *género-especie*; aunque las directrices de la UNESCO la permiten, no es recomendable excepto para relaciones geográficas y partes del cuerpo.

TRANSPARENCIA III- 17

La mayor dificultad de las relaciones *TG—TE* reside en que muchos términos pueden pertenecer legítimamente a varias jerarquías y, por lo tanto, hay que elegir una de ellas.

Algunos tesauros incluyen un mismo término en más de una clase o jerarquía.

En un tesoro* muy general es conveniente colocar el término en la clase que le corresponde según su característica o propiedad inherente.

En el ejemplo, ORO corresponde a la clase METALES PRECIOSOS perteneciente a la clase más general METALES aunque en un tesoro de odontología podría estar dentro de la clase PROTESIS DENTAL.

TRANSPARENCIA III- 18

Los *TR* están vinculados por una relación semántica* que no es jerárquica ni de sinonimia.

TRANSPARENCIA III- 19

La relación *TR* es más ambigua que la *TG—TE* porque, en algún sentido, todos los términos de un tesoro están relacionados y el problema reside en decidir qué términos están lo suficientemente conectados para justificar las relaciones *TR*.

No hay normas para establecer las relaciones *TR*, sólo se dan sugerencias orientadoras.

TRANSPARENCIA III- 20

Las referencias *USE* dirigen de un término que no puede ser usado en la indización* y búsqueda a un término que puede ser usado.

La referencia *USE* puede remitir desde sinónimos*, abreviaturas, casi-sinónimos y otras palabras alternativas; en esos casos los dos términos (referido y referido a) son esencialmente equivalentes.

En el ejemplo, AGUA SALADA y AGUA SALINA son sinónimos. Se eligió como descriptor AGUA SALINA: por lo tanto hay una referencia que remite de AGUA SALADA a AGUA SALINA.

HERENCIA y GENETICA son casi-sinónimos*.

RUGOSIDAD y LISURA son antónimos*.

La referencia *USE* es empleada también para remitir desde términos específicos no usados en la indización o búsqueda a los términos genéricos más apropiados.

En el ejemplo se remite de DESTRUCTORES (Término específico) a BUQUES DE GUERRA (Término genérico), que incluye ese específico.

La razón de esta opción es que los parámetros del sistema* no exigen una gran especificidad en ciertos sectores del campo temático.

Estos términos son tratados en el sistema como equivalentes aunque no lo sean en la realidad.

TRANSPARENCIA III- 21

En algunos tesauros la referencia *USE* se emplea también para remitir de un descriptor a varios (el término de entrada no elegido como descriptor CATALOGACION COOPERATIVA remite a CATALOGACION y a COOPERACION). Esto puede indicarse de tres maneras como en el ejemplo.

En la referencia *UP* correspondiente a cada uno de los descriptores (CATALOGACION, COOPERACION) usados por ese término (CATALOGACION COOPERATIVA) se indica por medio de un signo (en el ejemplo el signo+) que no es el único descriptor de dicho término.

TRANSPARENCIA III- 22

Las notas de alcance* pueden ser instrucciones para usar el descriptor en la indización y recuperación o definiciones tipo diccionario (especialmente para términos que no son familiares y cuyo significado puede ser ambiguo).

En la transparencia hay dos notas que indican limitaciones en el alcance del término (una positiva y otra negativa); una definición de un término técnico y una combinación de definición y limitación positiva.

UNIDAD III: CONSTRUCCION DE TESAuros

CLASE 6: PRUEBA, DESARROLLO Y ACTUALIZACION DEL TESAuro.

Objetivo:

Enseñar procedimientos para probar un tesoro antes de su presentación. Fijar criterios para el desarrollo y actualización.

Idea principal:

Antes de presentar el tesoro se recomienda hacer pruebas prácticas mediante la indización de un número suficiente de documentos.

Hay que probar el tesoro con preguntas que puedan señalar defectos y mostrar la necesidad de introducir nuevos descriptores.

Ideas accesorias:

Un vocabulario controlado para la indización y la búsqueda no puede permanecer estático: debe crecer.

La extensión del vocabulario depende no sólo del campo temático sino también de la especificidad de los términos y del tipo de términos usados.

El crecimiento continuado del vocabulario depende del crecimiento del sistema al que sirve.

La actualización del tesoro puede implicar la cancelación de un término.

Las estadísticas del uso de un término en la indización son indicadores muy valiosos en las operaciones de actualización.

La adición de nuevos términos, aunque sea sobre una base provisional, debe ser permanente; la cancelación de términos exige el análisis de estadísticas.

Se recomienda empezar esta clase con una revisión de las posibles presentaciones de los tesauros con ejemplos dados en la Unidad II.

Para considerar terminada la estructura de un tesoro es aconsejable indizar 1.000 documentos por lo menos. A partir de entonces seguirá creciendo mediante la continua indización, pero se mantendrá la misma estructura.

Las pruebas mostrarán si:

- a) la estructura propuesta toma en cuenta las necesidades más recientes.
- b) los descriptores elegidos son útiles para la indización y la búsqueda.
- c) las relaciones son correctas y suficientes.
- d) es capaz de responder a las preguntas que se formulen.

TRANSPARENCIA III- 23

Por supuesto, un vocabulario* desarrollado a través de las operaciones de indización* crecerá muy rápidamente al principio, pero alcanzará una cierta estabilidad a partir de la que aumentará muy poco.

El gráfico ilustra una típica curva de crecimiento en un vocabulario de campo temático muy limitado (anestesiología).

La curva crece rápidamente durante los primeros 200 ó 300 documentos.

En la práctica, los campos más especializados llegarán a tener hasta 8.000 términos* y el total de la ciencia y la tecnología requerirá alrededor de 20.000. El vocabulario del MEDLARS, que cubre la totalidad de la biomedicina, tiene alrededor de 8.000 descriptores, mientras que el Thesaurus of Engineering and Scientific Terms contiene 23.000 términos. La extensión del vocabulario dependerá no solamente del campo temático sino también de la especificidad* y tipo de términos usados. Un vocabulario precoordinado será más extenso que un vocabulario de unitérminos.

El otro gráfico presenta la curva de crecimiento de un tesoro* sobre ciencias de la administración. Este vocabulario ha llegado a un nivel de alrededor de 800 términos después de la indización de 2.000 documentos.

En el desarrollo de un vocabulario es importante tener alguna forma objetiva de evaluar el grado de estructuración antes de ponerlo en uso y ofrecerlo a la revisión de otros individuos u organizaciones.

TRANSPARENCIA III- 24

Snow (Oficina de Patentes del Reino Unido) describió un procedimiento de evaluación conocido como la Prueba de Snow, cuyos tres aspectos fundamentales (objetivos que hay que alcanzar) son:

- 1) Examen riguroso de la naturaleza y extensión de los aspectos pertinentes del campo temático;
- 2) Relación entre la profundidad de la indización prevista y el estado al que ha llegado la especialidad;
- 3) Razonable equilibrio de la lista.

Actualización. Una vez que el vocabulario llega a un cierto nivel puede producirse un crecimiento gradual para aumentar la especificidad y adecuarse a los nuevos tópicos, pero no debe esperarse ningún crecimiento brusco (a menos que el Centro se proponga abarcar áreas completamente nuevas). El crecimiento continuado del vocabulario provendrá del mismo sistema. Se fundamentará en la justificación literaria*; los términos nuevos surgirán de las operaciones de indización y búsqueda. Serán términos necesarios para expresar aspectos del campo temático en cuestión, originados en los documentos* y en las respuestas y que no han podido ser expresados por los términos ya establecidos del vocabulario. Este será el principal criterio para la adición de nuevos términos.

Normalmente, en un centro de información o en una red de centros de información, se crea un grupo de control que revisa coherentemente el vocabulario y, en particular, considera las recomendaciones de nuevos términos formulados por los indizadores y por quienes realizan las búsquedas.

En la red del ERIC cada participante tiene la responsabilidad de controlar y actualizar el vocabulario. Si un indizador encuentra un tópico que no ha podido ser representado adecuadamente por el vocabulario establecido, registra el nuevo término que podría ser un descriptor. Debe estar justificado como los descriptores originales: diccionarios, manuales, directrices y convenciones. El nuevo término se registra en un formulario en el que se establecen las relaciones con los descriptores del tesoro y si fuera necesario se agrega una nota de alcance. El grupo de control revisa los nuevos términos. Si éstos son aceptados pasan al "input" en máquina.

En algunos sistemas es costumbre introducir un nuevo descriptor por un período de ensayo, observar su uso y retenerlo o descartarlo después de un período de 12 meses, por ejemplo (Experiencia en el MEDLARS).

Además de la inclusión de nuevos términos, la actualización del tesoro puede implicar también la cancelación de un término. Un descriptor* puede ser reemplazado por un sinónimo*, casi-sinónimo* o expresión sinónima. Esto reflejará un cambio en la terminología aceptada dentro del campo. Un término puede ser cancelado, y referido, por una referencia *USE*, a otro término más genérico. Esta decisión puede surgir si un término específico no ha tenido mucho uso en la indización.

Las estadísticas del uso de un término en la indización son indicadores muy valiosos en las operaciones de actualización. Estas estadísticas se mantienen en la mayor parte de los sistemas. Estadísticas sobre el uso de los términos en las búsquedas no se registran casi nunca, aunque, en cierto sentido, son más importantes que las de indización. Si un término empleado en la indización no es usado nunca en las búsquedas por un período de dos años, su permanencia no se justifica, ya que esto podría indicar una especificidad* innecesaria. Los indizadores lo han usado porque existe un documento sobre ese tópico específico. Sin embargo, las preguntas nunca han llegado a ese nivel de especificidad que, por lo tanto, no se justifica.

Es importante que los usuarios pregunten en lenguaje natural para que esas preguntas se registren y analicen en el "input" de los procedimientos de actualización de un tesoro. Todas las preguntas formuladas al sistema* deben registrarse inicialmente con las propias palabras de quien busca y la traducción a la terminología del sistema debe ser posterior.

La adición de nuevos términos, aunque sea sobre una base provisional, debe ser continua. La cancelación de términos exige el análisis de estadísticas sobre el uso de los términos, que sólo puede ser efectuado mediante procedimientos de revisión completa del vocabulario. Algunas veces esta revisión se hace anualmente.

El uso de la computadora para mantener la historia de los términos y sustituirlos automáticamente cuando es necesario, minimiza los problemas de las búsquedas.

2.5.4. UNIDAD IV

UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION

CLASE 1: LOS VOCABULARIOS DE INDIZACION EN COMPUTADORA

Objetivo:

Explicar las posibilidades de la computadora en el uso y control de los lenguajes de indización.

Ideas principales:

Dado que la computadora permite realizar las tareas documentarias con un alto grado de rapidez, confiabilidad y gran caudal de información, con su aplicación es posible optimizar el uso de los tesauros e, incluso, generarlos.

Ideas accesorias:

La optimización de tesauros se puede lograr con una estructura facetada, y mediante el uso de recursos de acierto y de precisión. Ambas posibilidades pueden realizarse en forma eficaz con la computadora.

El uso de los sistemas automatizados nos permiten dos posibilidades fundamentales: generar tesauros por computadora y que éstos puedan ser consultados directamente en máquina.

TRANSPARENCIA IV-1

Si el tesoro se encuentra en máquina hay seis buenas razones que demuestran sus ventajas:

- 1) El sistema puede controlar y verificar la consistencia y aceptabilidad de los términos usados por los indizadores y de los empleados en la estrategia de búsqueda. Se trata simplemente de un proceso de comparación que permite rechazar un término resultante de la indización cuando no se encuentra en la lista maestra.
- 2) El sistema puede sustituir automáticamente un término de entrada por un descriptor (por ej. un sinónimo). Además, cuando un término deja de tener vigencia es sustituido por otro más actualizado.
- 3) El sistema puede realizar estadísticas sobre el uso de un término tanto en la indización como en la búsqueda. Puede eliminarlo en caso de poco uso o, de lo contrario, profundizar el concepto mediante términos más específicos. Es muy difícil mantener estas estadísticas en forma manual; en general se realizan con sistemas automatizados.
- 4) El sistema permite mejorar las estrategias de búsqueda y, de esta manera, disminuir los costos.
- 5) El sistema permite generar índices en forma automática con las referencias "véase y véase además".
- 6) El VMC ⁽¹⁾ tiene un registro de la estructura jerárquica completa. Esto facilita la conducción de las búsquedas genéricas. En la búsqueda, solamente los términos genéricos necesitan ser precisados; los términos específicos son aportados automáticamente por el VMC. En el sistema Medlars se hace este tipo de búsqueda mediante la mención del término y entre paréntesis la palabra EXPLOSION. Automáticamente se dan todas las divisiones que abarca esa clase.

(1) VMC: Vocabulario Maestro de Control.

TRANSPARENCIA IV- 2

En esta transparencia se muestra un ejemplo de los cambios de un concepto en un tesoro. Desde 1970 hasta 1975 se había registrado el término PERROS ZORREROS, desde 1976 hasta 1978 PERROS ZORREROS IRLANDESES. El sistema permite sustituir automáticamente una terminología antigua por una terminología más reciente.

TRANSPARENCIA IV- 3

El control del vocabulario en máquina puede realizarse por la frecuencia de uso de un término en dos momentos: la indización y la búsqueda. El sistema puede producir listados con la indicación de cuántas veces fue usado cada término. Se logra, así, que el vocabulario sea representativo de una realidad.

TRANSPARENCIA IV-4

Una estrategia de búsqueda puede ser optimizada automáticamente del siguiente modo:

En la ecuación: $(A \cup B)$ y $(C \cup D)$ y E tenemos que:

a) de $A \cup B$, esquematizados con conjuntos  , resultan 3.029 citas bibliográficas;

b) de $C \cup D$,  , resultan 69 citas bibliográficas y

c) E corresponde a 500.000 citas bibliográficas.

Para hacer la búsqueda seleccionamos la ecuación $C \cup D = 69$, porque corresponde a un número bajo de referencias. Separamos en un archivo intermedio las 69 referencias (luego de buscarlas en un archivo con 2 millones de registros) y las comparamos con $(A \cup B)$ y E, con un resultado de 11 citas bibliográficas que luego se imprimen.

Para reducir la búsqueda se ha elegido el camino más simple: la ecuación cuyo resultado era más pequeño. Con este sistema de optimización se abarata el costo de las búsquedas.

TRANSPARENCIA IV- 5

Es posible generar índices en forma automática a partir de documentos indizados con un tesoro en máquina. En el ejemplo, RUGOSIDAD tiene TE, TR y UP; en el índice impreso aparecerán todas las relaciones con la indicación véase y véase además. En el primer caso hay documentos indizados bajo RUGOSIDAD y RUGOSIDAD SUPERFICIAL. En el segundo, los documentos están indizados bajo RUGOSIDAD y bajo ASPEREZA. En el último caso sólo hay documentos bajo ASPEREZA.

TRANSPARENCIA IV- 6

Esta transparencia muestra el ejemplo de una hoja de entrada de términos a un tesoro en máquina. Cada ítem corresponde a un campo asignado.

Cuando se utilizan vocabularios controlados en el tratamiento de la información con equipos electrónicos se trata de combinar los mejores criterios de los esquemas de clasificación con los mejores criterios de los tesauros. Como consecuencia de esta tendencia aparecen los tesauros facetados, que están contruídos con mayor rigor, son más concisos y consistentes. La idea de los tesauros facetados surgió de Aitchinson, quien debió elaborar un esquema de clasificación sobre ingeniería para la compañía English Electric de Gran Bretaña. Estos tesauros no son considerados de tipo convencional, aunque la mayor parte de sus relaciones son similares a las convencionales.

TRANSPARENCIAS IV-7 y IV-8

En los ejemplos vemos que se utilizan las relaciones USE, UP, TG y TR. Cuando se agrega la nominación (A) a los TG significa que es un término genérico adicional, agregado al TG correspondiente. Este TG (A) se encuentra en general en otra faceta. Por ej. el TG (A) SERVICIOS DE INFORMACION EMPRESARIAL corresponde a la faceta R, mientras que BIBLIOTECAS corresponde a la faceta L. Cada término tiene una notación, en este caso la notación es alfabética.

En el tesauro facetado, la parte alfabética del tesauro reemplaza el índice alfabético de materias que normalmente acompaña el esquema de una clasificación facetada convencional. En la parte alfabética no se eliminan sinónimos. En esta parte no aparecen las relaciones indicadas en las facetas, ya que están aclaradas en la parte facetada.

TRANSPARENCIA IV-9

Para mejorar la aplicación de los lenguajes de indización, se emplean algunos recursos adicionales:

- 1) recursos de acierto, que reducen la cantidad de descriptores elegidos y amplían el tamaño de las clases;
- 2) recursos de precisión, que aumentan la cantidad de descriptores elegidos y reducen el tamaño de las clases.

Los recursos de acierto se usan para el control de sinonimia y de la forma de las palabras y para el agrupamiento jerárquico. También se utilizan para la asociación estadística, pues se considera que las palabras que se presentan juntas con cierta frecuencia, están de algún modo relacionadas.

Se consideran recursos de precisión: la coordinación (pre y post coordinación), la unión (que es un tipo de pre-coordinación), los indicadores de rol o de relaciones sintácticas y la ponderación de la indización.

TRANSPARENCIA IV-10

En esta transparencia se grafican el *acierto* (A, E y G se combinan en una clase mayor para no mantener una distinción) y la *precisión* (M se subdivide en tres términos más específicos M1, M2 y M3, formando tres clases más reducidas).

TRANSPARENCIA IV-11

El *control de los sinónimos* es un ejemplo de recursos de acierto por el que se elige uno de los dos términos que expresan un concepto. En este caso se elige PUBLICACIONES PERIODICAS.

Se puede realizar un *control de la forma de las palabras* del siguiente modo: el término SOLDADURA es elegido porque su raíz representa los aspectos semánticos de la palabra. SOLDADURA es la clase mayor. Se considera que las palabras que derivan de una misma raíz etimológica están relacionadas semánticamente. Esta conversión se puede hacer por medios automatizados.

TRANSPARENCIA IV-12

Agrupamiento jerárquico. Es posible imponer una estructura jerárquica en un vocabulario por medio de: a) una organización clasificatoria formal (clasificación explícita), como en el caso de MICROFORMAS, MICROFICHAS, MICRO-

TARJETAS y MICROFILMES, términos reunidos en un conjunto; o b) una red de referencias cruzadas (clasificación implícita), ejemplo de **MICROFORMAS** con los **TE**, **MICROFICHAS**, **MICROFILMES** y **MICROTARJETAS**. En el caso de una clasificación abierta, es posible, -desde una terminal de computadora- tocar uno de los términos y recuperar el conjunto de todos los términos relacionados.

Asociación estadística. En un sistema automatizado, se puede formar un racimo (del inglés clumb o cluster) de términos vinculados entre sí. Interesa conocer cuántas veces A, Q, V y S tienden a concurrir en un texto. De este modo se amplía el tamaño de las clases. Esto es virtualmente imposible de implementar en un sistema manual.

Entre los recursos de precisión se ejemplifican los siguientes:

TRANSPARENCIA IV- 13

Coordinación. Al realizar la intersección de dos conjuntos (**ULTRASONIDO** y **SOLDADURA**), resulta un sólo término (**SOLDADURA ULTRASONICA**) más específico.

TRANSPARENCIA IV- 14

Uniones. Se tienen artículos que tratan sobre **TUBERCULOSIS**, **DIABETES** y **AMILOIDOSIS**, otros artículos se refieren solamente a **TUBERCULOSIS** y **DIABETES**, y los terceros a **TUBERCULOSIS** y **AMILOIDOSIS**. Para evitar una falsa coordinación se utilizan las uniones (links).

TRANSPARENCIA IV- 15

Indicadores de rol. En esta transparencia se muestra un ejemplo de aplicación de este recurso de precisión. Dado que no es lo mismo "un acceso epiléptico ocurrido durante la lectura" que "las habilidades de lectura de niños epilépticos", cuando se indiza un documento habrá que aclarar si **LECTURA** es (2) causa o (3) actividad, para lograr una buena recuperación en el momento de la búsqueda.

TRANSPARENCIA IV- 16

La construcción de un lenguaje de indización por computadora se realiza sobre la base de la concurrencia de los términos, pues la máquina sólo puede reconocer términos que se presentan juntos en un texto.

Este reconocimiento puede realizarse en dos niveles de asociación: directa o por métodos estadísticos.

La directa es la más simple: interesa saber cuántas veces aparece el término A con los demás términos. Todo el grupo de términos que aparecen junto a A forman una clase.

En el otro ejemplo se deduce que A y E pueden estar asociados entre sí dado que tanto A como E aparecen con Q en los documentos. Por ello se infiere que A y E pueden tener significados similares.

Por medio de la computadora es factible conocer la frecuencia de concurrencia de términos y además registrar la frecuencia más alta. Esto puede ser una ayuda eficaz para la construcción de las clases.

TRANSPARENCIA IV- 17

También es posible formar clases almacenando los términos en forma de racimos (clusters) e indicando la fuerza de conexión entre los mismos por frecuencia de relación.

L no ha podido pertenecer al mismo racimo debido a que no tiene concurrencia con C, Q, R y V, sólo lo tiene con A.

TRANSPARENCIA IV- 18

Indización ponderada. Si se tienen tres documentos sobre gatos interesa saber cuánta información sobre este tema tiene cada uno de los documentos. La parte que abarca la llave es la proporción del tema que contiene cada documento.

En el primer documento se indica un peso 3, es decir, todo el documento trata sobre gatos: es la máxima ponderación.

El segundo documento pesa 2, incluye menos de la mitad del tema a indizar.

El tercer documento pesa 1, sólo un 10 o/o se refiere al tema. Se considera la mínima información aceptable para indizar: es la mínima ponderación.

Cuando se procede a la recuperación, para preguntar si el sistema posee documentos completos que traten el tema, se dará ponderación 3 y el resultado será de pocos documentos.

Pero si se desea la mayor cantidad de documentos sobre el tema, se indicará una ponderación 3 y 2, ignorando el 1, y se obtendrá una recuperación casi completa, pues se habrán desechado los documentos menos significativos.

UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION

CLASE 2: LENGUAJE NATURAL

Objetivo:

Explicar el uso del lenguaje natural para indizar y recuperar información en sistemas automatizados.

Idea principal:

El lenguaje natural puede ser utilizado como lenguaje de indización

Ideas accesorias:

Para optimizar el uso del lenguaje natural se usan recursos, tales como la lógica de Boole, la distancia de las palabras, la frecuencia, los fragmentos de palabras, etc.

*Lenguaje natural**: Su uso en computadora data de los años 50, es decir, es anterior a la creación de los tesauros*. Se inició sobre textos legales, por ej. en el sistema LITE (Legal Information Through Electronics) del Departamento de Defensa de los EE.UU.

Posteriormente, se utilizó en sistemas comerciales tales como el LEXIS. Luego se empleó en sistemas en línea*, y finalmente para servicios de información secundaria como el Chemical Abstracts.

Cuando se usa lenguaje natural en sistemas de almacenamiento y recuperación de información automatizados, la computadora recorre el texto, palabra por palabra de acuerdo con un perfil del usuario, que se encuentra registrado en un índice* que contiene todas las palabras que ese usuario indicó como términos descriptivos de su necesidad de información. Estas palabras son comparadas con las contenidas en los registros de la base de datos, y cuando hay coincidencia significa que se ha recuperado la información contenida en los asientos bibliográficos que responden a la necesidad del usuario. Esto es posible pues las computadoras pueden realizar las búsquedas a velocidades muy altas. Recordemos que estos equipos trabajan a velocidades medidas en nanosegundos.

TRANSPARENCIA IV-19

Algunos de los inconvenientes que derivan del uso de lenguaje natural para los sistemas de almacenamiento y recuperación de información* son los siguientes:

1. Coordinación falsa, en la recuperación debido a la falta de control del vocabulario y la carencia de recursos de acierto* y precisión*.
2. Relación incorrecta entre términos, por el mismo motivo indicado en 1.

Pero por otra parte, el lenguaje natural, presenta ventajas tales como:

3. La indización* es exhaustiva, pues al no existir un control del vocabulario, los documentos* son representados por la mayor cantidad de términos* significativos.
4. La búsqueda es genérica o total, debido a que se ha realizado una indización exhaustiva*.

Dentro de los sistemas que utilizan el lenguaje natural hay una recuperación irrelevante de un 30 a un 40 o/o. Lancaster considera que no es un mal porcentaje, pues es similar al de un sistema que utiliza un vocabulario controlado*. Estos problemas pueden ser solucionados con la preparación de una estrategia de búsqueda* precisa.

TRANSPARENCIA IV- 20

Los elementos que posibilitan el uso eficiente de un lenguaje natural son los siguientes:

a) *Lista de detención** (en inglés Stop List). Es una lista de palabras no significativas (artículos, preposiciones, conjunciones, etc.). Estas palabras no ingresan al archivo invertido*, ni forman parte del perfil de búsqueda.

Cada vez que se realiza la entrada de un registro, se van eliminando por comparación, todas las palabras contenidas en la lista de detención, e ingresando aquellas significativas como términos de indización. De este modo, el texto primario se reduce en un 50 o/o.

b) *Archivo invertido*: Señala la ubicación exacta de las palabras en el archivo. Por ej. EPILEPSIA está en el documento N° 1001, en la línea número 7, en la palabra número 4.

TRANSPARENCIA IV- 21

*Lógica de Boole**. Para solucionar las dificultades que presenta el uso de los lenguajes naturales, se deben realizar estrategias de búsqueda precisas. Para ello se utiliza el álgebra de Boole con los operadores lógicos $\ast=y$; $\ast=\acute{o}$; NO.

Si se desea realizar una búsqueda sobre las pérdidas de las Compañías de aviación en la Argentina, se debe preparar la ecuación de búsqueda como sigue:

Se incluyen todos los términos que puedan corresponder a líneas aéreas, tales como TRANSPORTES AEREOS, LINEAS AEREAS, etc., y todos los términos relacionados con BENEFICIOS, FINANZAS,

En cada conjunto se combinan entre sí los términos con el operador* lógico $\ast=\acute{o}$ y se combinan ambos conjuntos con el operador lógico $\ast=y$.

De este modo se obtendrá una respuestas pertinente a la necesidad de información indicada.

TRANSPARENCIA IV- 22

En lenguaje natural pueden utilizarse los *operadores métricos** o *indicadores de distancia de palabras*, considerando el principio de que, cuanto más cerca esté una palabra de la otra en un texto, hay mayor probabilidad de que estén relacionadas entre sí. Al hacer una búsqueda, por ej. LEGISLACION ESCOLAR, se indica un límite de distancia entre ambas palabras (por ej. hasta 5). Se recuperarán solamente aquellos documentos en los cuales la distancia entre esas dos palabras no sobrepase el límite establecido.

TRANSPARENCIA IV- 23

Este es un ejemplo de salida de computadora que indica la frecuencia de aparición de cada palabra. Es un elemento muy útil para el armado de perfiles y para la búsqueda de información, así como también es una excelente fuente de términos que pueden servir para comenzar a elaborar un vocabulario.

Además, es una ayuda para la búsqueda de información pues sirve para calcular aproximadamente cuántos documentos sobre el tema que interesa, existen en el sistema.

TRANSPARENCIA IV-24

Esta es una de las técnicas más importantes para la recuperación de la información en un sistema con lenguaje natural.

1. El truncado a la derecha es la búsqueda por prefijo o eliminación del sufijo. En esta técnica, se corre el riesgo de obtener términos que no correspondan al concepto requerido. Es más difícil de aplicar en ciencias sociales que en ciencia y técnica.
Es la búsqueda más común. Compensa la carencia de un vocabulario controlado, pues al aplicarlo se encuentran todas las palabras que tienen el mismo origen semántico. Sin embargo, puede recuperar información no pertinente, por ej.: Río Colorado, colorete (un elemento de maquillaje).
2. El truncado a la izquierda es la búsqueda por sufijo, que es muy útil en ciencia. Si buscamos por MICINA, recuperamos automáticamente toda la información registrada en el sistema sobre antibióticos (estreptomina, pantoicina), sin tener necesidad de conocer o de darle al sistema la información completa sobre el tema.
3. El truncado a izquierda y derecha es una combinación de las dos anteriores, llamado truncado en ambas direcciones (Ej.: MAGNET: electromagnetismo, geomagnetismo).
4. En el truncado por prefijo y por sufijo nos interesa el infijo, es decir, la palabra intermedia.
Es muy útil en los campos científicos (triclorocobaltato, triiodocobaltato).

UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION

**CLASE 3: VOCABULARIOS HIBRIDOS.
COMPATIBILIDAD DE TESAuros.
TESAuros MONOLINGUES Y MULTILINGUES.**

Objetivo:

Enseñar cómo se puede obtener la compatibilidad de lenguajes de indización con especial referencia al lenguaje natural y al controlado, a tesauros correspondientes a diferentes sistemas y a tesauros en diferentes idiomas.

Idea principal:

A pesar de que un tesoro está previsto para un sistema determinado de información se puede llegar a una compatibilidad universal superando las barreras de los idiomas, de las diferencias de los sistemas y aún de la diferencia entre el lenguaje natural y el lenguaje controlado.

TRANSPARENCIA IV- 25

*El Vocabulario híbrido** utiliza una combinación entre el vocabulario controlado* y el lenguaje natural*.

Existen muchos sistemas que lo utilizan. Por ejemplo el INSPEC del Institution of Electrical Engineers de Gran Bretaña. Este sistema utiliza descriptores* de un tesoro* y además, palabras clave * tomadas del lenguaje natural de los documentos*.

De este modo, cuando se realiza una búsqueda de información en una base de datos formada con estas características, se puede preparar una estrategia de búsqueda* combinando descriptores del tesoro, y otros términos* del lenguaje natural.

La búsqueda se inicia por los descriptores y luego se hace un barrido (lectura secuencial) de los títulos y los resúmenes, detectando aquellas palabras conceptualmente significativas del documento.

Ejemplo: si buscamos heridas en la cabeza de los jugadores de fútbol, y el tesoro posee solamente descriptores tales como HERIDAS EN LA CABEZA y HERIDAS POR EL DEPORTE, lograremos un grado mayor de especificidad*, si el término FUTBOL se encuentra en un título, o en un resumen.

TRANSPARENCIA IV- 26

En esta transparencia se ejemplifica cómo se aplican los vocabularios híbridos.

Si se diseña el sistema con un vocabulario controlado formado por pocos términos y se combina con una cantidad mayor de términos correspondientes al lenguaje natural, se logra que los términos controlados den contexto a las palabras clave no controladas, que posibilitan una mayor especificidad.

Ejemplo: Se tienen los términos generales TRABAJO y GUERRA correspondientes al vocabulario controlado y se les agrega la palabra clave HUELGA, tomada de los títulos de los documentos.

En el ejemplo, el vocabulario controlado de este sistema híbrido, tiene un reducido número de códigos temáticos (amplios), aproximadamente 250. El sistema utiliza también códigos geográficos (en el ejemplo DJO, correspondiente a Jordania) y las palabras clave tomadas del lenguaje natural aumentan la especificidad.

TRANSPARENCIA IV- 27

Los vocabularios son herramientas que se utilizan para favorecer la cooperación y el intercambio entre sistemas. De este modo, toda la indización que se realiza en el centro A, puede ser usada por el centro B, D y C y viceversa.

Con los medios modernos de comunicación: telex, satélites, microondas, etc., las posibilidades son infinitas. Si estos centros estuvieran conectados entre sí mediante una red de comunicación en línea y se utilizaran vocabularios diferentes en cada uno de ellos, existen técnicas que permiten que estos sistemas diferentes sean compatibles.

TRANSPARENCIA IV- 28

Los microtesauros permiten obtener la compatibilidad* de sistemas.

Los microtesauros se usan en campos altamente especializados. Por ejemplo, dentro de un sistema de información en agricultura, el cual utilizará un macrotesauro*, se puede desarrollar un microtesauro sobre café para ser utilizado en un centro especializado en ese tema.

El macrotesauro sobre agricultura dará el marco genérico y el microtesauro sobre café será muy específico.

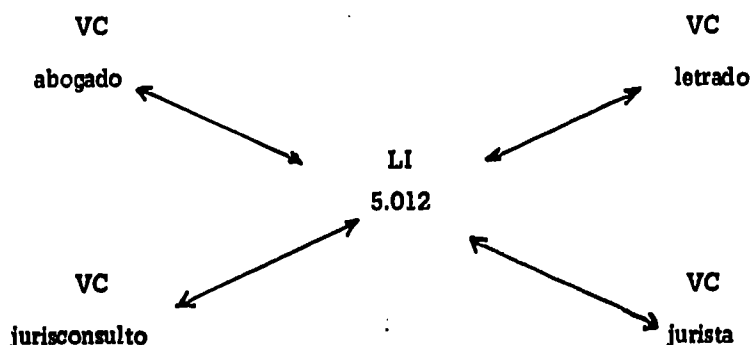
Si el sistema utiliza medios automatizados*, los términos específicos del microtesauro pueden ser relacionados automáticamente con el término genérico correspondiente al macrotesauro.

TRANSPARENCIA IV- 29

Se puede utilizar también un *léxico intermedio*.*

Por ejemplo si se tienen cuatro centros que utilizan cada uno de ellos un vocabulario controlado diferente, se comparan los vocabularios de a dos, buscando las equivalencias. Se confecciona una lista con las equivalencias y se les asigna un código a cada término. Estos códigos forman el léxico intermedio. Luego se realiza la misma tarea con un tercer vocabulario y así sucesivamente.

Ejemplificando la transparencia:



Vemos que la conversión automática es simple.

Cuanto más centros colaboran en un sistema, éste resulta más económico. Con sistemas en línea*, la conversión automática de vocabularios es factible y permite a cada servicio que participa de un sistema realizar búsquedas de información utilizando su propio vocabulario.

TRANSPARENCIA IV-30

El léxico intermedio puede conmutar términos de una misma lengua o términos de diferentes lenguas.

Por medio de sistemas automatizados se pueden realizar algunas equivalencias para la construcción de un léxico intermedio, pues la computadora puede reconocer:

1. la equivalencia directa, o sea términos idénticos.
2. la variación en la escritura.
3. la forma de las palabras.
4. las inversiones
5. las referencias cruzadas.
6. la jerarquía.

Ejemplo: En un tesoro existe el término NIÑO y como término específico, PARVULO. En otro tesoro aparece solamente el término NIÑO. Al comparar, la computadora puede realizar la conversión automática, reduciendo el esfuerzo intelectual al mínimo y estableciendo la similitud entre NIÑO y PARVULO.

En base a un estudio comentado por H.H. Neville en Feasibility study of a scheme for reconciling thesauri covering a common subject. Journal of Documentation. 26 (1970): 313-336, se proyectarán una serie de transparencias con indicación del tipo de concordancia de vocabularios y el modo de asignar códigos para formar un léxico intermedio.

TRANSPARENCIA IV-31 Tabla 37

Existe una correspondencia exacta del término AIRFIELDS (campos aéreos) del tesoro fuente y los tesoros B y C. Por ello, si se asigna a cada término el código 0101, se realizará en forma automática la concordancia entre los tres vocabularios.

TRANSPARENCIA IV-32 Tabla 38

En el caso de sinónimos*, UNDERGROUND STRUCTURES (estructuras subterráneas), BURIED STRUCTURES (estructuras enterradas) en el tesoro fuente y SUBSURFACE STRUCTURES (estructuras subsuperficiales) en el Tesoro B, llevan el código 0201, en los términos elegidos como descriptores. Por medio de este código se realiza la concordancia de los sinónimos.

TRANSPARENCIA IV-33 Tabla 39

Cuando en el Tesoro B no existe un término, en el ejemplo, SNOWDRIFTS (ventiscas), se agrega como descriptor con el código correspondiente para posibilitar la concordancia.

TRANSPARENCIA IV-34 **Tabla 40**

Cuando el tesoro C no tiene el término específico SNOWDRIFTS (0301), se puede agregar haciendo referencia al término más amplio SNOW (0300).

TRANSPARENCIA IV-35 **Tabla 41**

Cuando el tesoro fuente contiene términos en sus formas sustantivadas y adjetivadas y el Tesoro B no las tiene, se incluyen ambas formas en el Tesoro B con la indicación de los códigos que posibilitan la concordancia.

TRANSPARENCIA IV-36 **Tabla 42**

En esta transparencia se muestra el caso de un término que no existe en el Tesoro D, pero se lo incluye como término no usado con el código correspondiente del Tesoro fuente.

TRANSPARENCIA IV-37 **Tabla 43**

Cuando en el tesoro fuente se tienen descriptores simples, tales como CONTRACTION (contracción) y EXPANSION (expansión) y en el tesoro B se combinan los dos antónimos EXPANSION/CONTRACTION, se utilizan códigos diferentes para cada uno de los términos y se realizan las referencias necesarias a la forma elegida en cada tesoro.

TRANSPARENCIA IV-38 **Tabla 44**

Cuando el tesoro fuente tiene términos precoordinados FROST PENETRATION (penetración de la helada) y el tesoro B utiliza palabras únicas, con la indicación de que deben post coordinarse (signo +), (FROST + PENETRATION) (helada + penetración), se indica el código en el término de entrada frost penetration.

TRANSPARENCIA IV-39

Es el mismo caso que la transparencia IV-38, con la diferencia que las palabras únicas del Tesoro B son diferentes aunque conceptualmente signifiquen lo mismo.

STIFFNESS METHODS (métodos de rigidez) en el tesoro fuente es sinónimo de STRUCTURAL ANALYSIS + DISPLACEMENT (análisis estructural + desplazamiento).

TRANSPARENCIA IV-40

En este caso el mismo término TERMOMETER (termómetro), puede encontrarse desmembrado en el Tesoro B, TEMPERATURE + MEASUREMENT INSTRUMENT (temperatura + instrumento de medición).

Se asigna un código (0503) al término de entrada termómetro del Tesoro B y el mismo código al descriptor termómetro del tesoro fuente.

TRANSPARENCIA IV- 41

Este ejemplo presenta la concordancia de vocabularios mediante modificadores con equivalencias indirectas.

En el tesoro fuente HIGH (alto o altura) es un descriptor que puede combinarse con otros, tales como TEMPERATURE, BUILDINGS (edificios), STRENGTH (fuerza), para expresar otros conceptos. En el Tesoro B esos conceptos están representados por descriptores más específicos, tales como HIGH TEMPERATURE (alta temperatura), MULTI STOREY BUILDINGS (edificios de muchos pisos) y HIGH STRENGTH (alta tensión).

En ambos tesauros se asignan los códigos que establecen la concordancia.

TRANSPARENCIA IV- 42

Para producir tesauros multilingües se utilizan en forma más estricta los métodos de concordancia ya descritos.

Existen varias posibilidades:

- 1.- Recolectar los términos en diferentes lenguas y luego determinar las equivalencias.
- 2.- Traducir directamente los tesauros existentes. En ese caso puede ocurrir que no sea posible traducir algunos términos en forma equivalente y por lo tanto habría que cambiar las relaciones entre los términos.
- 3.- Fusionar dos o más tesauros aplicando el principio de concordancia antes descrito y obtener un tesoro multilingüe.

TRANSPARENCIA IV- 43

Cuando hay una equivalencia exacta tal como física-physics-physique y physik, se trata como un caso de sinonimia simple.

Cuando no hay equivalencia exacta se complica la concordancia ya que hay que evitar la utilización de conceptos incorrectos, pues no es lo mismo programa en español, software en inglés y programme en francés.

TRANSPARENCIA IV- 44

El ejemplo presenta un tesoro en inglés con la traducción al francés en cada nivel.

TRANSPARENCIA IV- 45

El ejemplo corresponde al mismo tesoro invirtiendo los idiomas.

TRANSPARENCIA IV- 46

Este ejemplo corresponde a la parte sistemática de un tesoro con los asientos completos de los descriptores en tres idiomas.

UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION

CLASE 4: EVALUACION DE TESAUROS

Objetivo:

Mostrar una metodología para evaluar tesauros.

Ideas principales:

Para evaluar un tesauro hay que tener en cuenta las siguientes características: consistencia, estructura, estética, enriquecimiento, flexibilidad y nivel de precoordinación y especificidad de los descriptores.

Ideas accesorias:

Algunos aspectos de un tesauro son observables en el mismo tesauro y otros hay que comprobarlos en su aplicación.

TRANSPARENCIA IV- 47

No es posible evaluar un tesauro* si no se utiliza. De cualquier modo hay algunas características que pueden ser consideradas a priori.

Es posible observar si el tesauro posee todas sus partes: la introducción y las presentaciones alternativas (sistemática, alfabética, gráfica, etc.)

Para poder comprobar si el tesauro es válido desde el punto de vista terminológico es necesario probarlo indizando documentos* y realizando búsquedas.

Al revisar los términos* y sus características nos interesa saber: a) cuántos términos son precoordinados, este aspecto es observable; b) nivel de especificidad* de los términos; se necesita comprobarlo por lo menos por muestreo; c) relación entre los términos de entrada y los descriptores, si es proporcionada y correcta; d) la presentación correcta de los términos genéricos y específicos y e) la cantidad y calidad de las notas de alcance* y de las definiciones.

TRANSPARENCIA IV- 48

Para la evaluación de la *consistencia** de los tesauros hay que considerar los siguientes aspectos:

- la reciprocidad de las relaciones de equivalencia* jerárquicas* y asociativas* (UP-USE, TG-TE, TR).
- la forma de las palabras aplicando la norma referida a la entrada invertida.
- el uso del *singular* y del *plural* de acuerdo con las normas.

En cuanto a la *estructura* interesa conocer en qué grado están vinculados los términos.

Con relación a la *estética* se debe contemplar si tiene una presentación clara y una tipografía adecuada.

TRANSPARENCIA IV-49

La evaluación del grado de *conexión general* de los descriptores de un tesoro, surge de la relación entre la cantidad de descriptores conectados con otros términos y el total de descriptores. La fórmula para el cálculo es la siguiente:

$$\frac{(b - a)}{a} \quad \text{donde } a = \text{cantidad de descriptores sin conexiones, y } b = \text{total de descriptores.}$$

El valor aconsejado es aproximadamente = 1. En un estudio realizado se encontró que de 20 tesauros, 11 tenían el valor de conexión aconsejado.

TRANSPARENCIA IV-50

Con relación a las *definiciones* de los términos se aplica la misma fórmula, donde a = descriptores que puedan ser ambiguos. El resultado debe ser mayor, o igual que 1. En el estudio se comprobó que de 20 tesauros, 16 reunían esa cualidad.

TRANSPARENCIA IV-51

En este ejemplo figuran los resultados del estudio en relación con la tasa de equivalencia*, la tasa de reciprocidad* y la tasa de enriquecimiento*.

TRANSPARENCIA IV-52

La *flexibilidad* es la relación entre las palabras componentes de descriptores que, a su vez, pueden ser descriptores y las que no pueden ser descriptores.

Es conveniente que de esta relación resulte un valor de 0,6 o más.

En el estudio, de los 20 tesauros examinados, 9 excedían esa proporción.

Si tenemos por ejemplo:

- Ruido de máquina jet,
- Ruido,
- Máquinas jet,
- Paneles vibrantes,
- Alas delta,

Hay en este grupo algunas palabras que son descriptores por sí mismos (jets, ruido, paneles, alas) y otras son componentes de descriptores (delta, vibrantes, máquinas).

Es decir que, de estas 7 palabras, 4 aparecen como descriptores por sí solas. La tasa de flexibilidad es de 4.

Para el *tamaño de los grupos de términos* (categorías* o clases* amplias) la norma ISO recomienda grupos de 30 a 40 términos. Lancaster opina que no hay fundamentos para elegir estas cifras y no otras.

TRANSPARENCIA IV-53

El *nivel de precoordinación** se calcula en base a la cantidad de palabras utilizadas por descriptor.

Las tasas deseadas serían de 1.5 a 2 para los tesauros en inglés y en francés. Para el alemán se requiere una tasa menor, de 1.1 a 1.2, debido a que el idioma alemán está naturalmente precoordinado. En el estudio resultaron tasas que indican un correcto nivel de precoordinación.

UNIDAD IV: APLICACION DE LA AUTOMATIZACION A LOS LENGUAJES DE INDIZACION

CLASE 5: PROSPECTIVA DE LOS LENGUAJES DE INDIZACION

Objetivo:

Presentar un panorama de la tendencia actual en el uso de los lenguajes para el almacenamiento y recuperación de Información.

Idea principal:

Mientras que los países tecnológicamente más desarrollados muestran una fuerte tendencia hacia el uso del lenguaje natural en los sistemas de almacenamiento y recuperación de la información, en los de menor desarrollo se advierte que el tesoro constituye la herramienta de uso más frecuente.

TRANSPARENCIA IV-54

Históricamente ha sido demostrado que, como lo puntualiza Lancaster, se tiende hacia el uso del lenguaje natural* para el almacenamiento y recuperación de información. En un comienzo se utilizaron vocabularios controlados*. En los años 50 se adoptó el uso de los unitérminos*, que son palabras únicas del lenguaje natural. A partir de los años 60 se vuelve a utilizar lenguajes controlados con la aparición de los tesauros* y con el mejoramiento del hardware*. Esto posibilita una mayor rapidez en la recuperación y un menor costo en el procesamiento. En la actualidad, se tiende nuevamente al empleo de lenguajes naturales.

Encontramos que:

1. existen bases de datos secundarias en lenguaje natural, en las cuales incluyen los resúmenes de los documentos con el lenguaje natural de los mismos.
2. existen bases de datos primarias en lenguaje natural. Tal es el caso de las bases de datos de legislaciones.
3. son cada vez más factibles debido a: a) el avance de la electrónica, y consecuentemente de la capacidad de las computadoras para almacenar y procesar información. b) a la comprensión intelectual más clara de la terminología correspondiente a los distintos campos del conocimiento.
4. el científico prefiere utilizar los sistemas de información en línea*, desde su lugar de trabajo, con una comunicación directa sin intervención del especialista en información

TRANSPARENCIA IV-55

El especialista en información prefiere utilizar el lenguaje controlado*

TRANSPARENCIA IV- 56

Se pueden dar cuatro posibilidades:

1. Control en la indización* y en la búsqueda.
2. Control sólo en la indización. Se consideran sistemas híbridos.
3. No hay ningún tipo de control.
4. El control se realiza en la búsqueda.

TRANSPARENCIA IV- 57

En sistemas híbridos se utilizan ciertos artificios para recuperar la información. Tal es el caso de las tablas de sinónimos* y los grupos conceptuales que controlan en parte el lenguaje natural.

TRANSPARENCIA IV- 58

Otro ejemplo de tabla de sinónimos y de un grupo conceptual.

TRANSPARENCIA IV - 59

Otro ejemplo de tabla de sinónimos correspondiente a LINEAS AEREAS.

TRANSPARENCIA IV- 60

El asterisco es un artificio que se utiliza en la búsqueda para el truncado de términos.

Conclusión:

A pesar de esta tendencia que se presenta en los países con tecnologías más avanzadas, en países en desarrollo es conveniente el empleo de lenguajes controlados ya que éstos se adaptan tanto al uso de sistemas intermedios (manuales y semiautomáticos) como automatizados. Esto no impide ni dificulta una futura conversión al lenguaje natural en un sistema automatizado.

3. GLOSARIO

GLOSARIO

ACCESO

En el mundo de la información existen por un lado, *documentos* y por otro *citas o asientos* que representan los *documentos* en catálogos, bibliografías o archivo de computadora y que constituyen el medio para acceder a ellos.

Acceso es la facilidad, mediante técnicas especiales, con que se ayuda al usuario a encontrar una cita o asiento en un catálogo, bibliografía o archivo de computadora, o un documento en las estanterías de una biblioteca.

ACCESO DIRECTO O INMEDIATO

Es el modo de acceso a un archivo en discos magnéticos, donde la información requerida se ubica directamente mediante un algoritmo de búsqueda, sin un barrido previo de los demás registros de la memoria.

ACCESO EN LINEA véase SISTEMA EN LINEA

ACCESO MULTIPLE

Cuando a una misma información se puede llegar a través de varios puntos de entrada.

ACCESO SECUENCIAL

Es el tipo de acceso que se obtiene al recorrer un archivo en el cual debe accederse consecutivamente a todos los registros anteriores al que se busca.

ACIERTO

Es la cantidad de *documentos relevantes recuperados*. Es la capacidad del sistema para *recuperar documentos relevantes*.

ALCANCE DEL SISTEMA

Es el área temática que abarca el sistema, determinando los temas núcleo y los temas periféricos.

ALGEBRA DE BOOLE

Disciplina basada en el análisis y síntesis de circuitos. Emplea la notación algebraica para expresar relaciones lógicas en la misma forma que se utiliza el álgebra convencional para expresar relaciones matemáticas.

De la expresión convencional del álgebra, sustituye las variables algebraicas por sentencias (frases) y los símbolos de operaciones matemáticas (+ ; - ; x ; etc.) por símbolos de operaciones lógicas ("y"; "o").

Estas operaciones son base de:

- 1) La lógica simbólica.
- 2) Ciertas estrategias de búsqueda en sistemas de recuperación de información.
- 3) La lógica interna de dispositivos internos de un computador digital.

ALGORITMO

Es un procedimiento matemático que indica los pasos a seguir para resolver un tipo de problema. Se deriva de un matemático árabe: Al-Karismi.

ALMACENAMIENTO

(de información). Es el proceso acumulativo de ubicación y retención de datos (en ficheros o computadoras) o documentos (en bibliotecas) para su posterior utilización.

AMBITO DEL SISTEMA véase **ALCANCE DEL SISTEMA**

ANALISIS POR FACETAS

Es la separación de las distintas facetas de un concepto, clase, categoría o campo temático, por medio de características.

ANALISIS CONCEPTUAL

Es la determinación de los distintos conceptos o ideas existentes en documentos al indizar, y en la pregunta al preparar la estrategia de búsqueda.

ANTONIMOS

Palabras que expresan ideas opuestas o contrarias.

ARCHIVO INVERTIDO

Es aquél que se ordena por alguno de los elementos o características contenidas en los registros (p. ej. descriptor, autor, etc.) y remite a los registros que los incluyen.

ARCHIVO LINEAL

Es aquél en el cual cada documento está representado por un registro o varios. El conjunto de registros se ordena secuencialmente.

AUTOMATIZACION

Control de procesos mediante equipos automáticos.

CALCULO DE BOOLE

véase

ALGEBRA DE BOOLE**CARACTERISTICA**

Es la propiedad o atributo que distingue conceptos, clases, categorías o facetas, y que permite separarlos, agruparlos, o colocarlos en sucesión.

CASI SINONIMOS

Términos que tienen diferente forma, pero un significado similar, dentro de un campo temático específico.

CATEGORIA

Es una *clase de alta generalidad*, que agrupa conceptos más específicos, o una *faceta general* que se aplica a todo un campo temático.

CLASE

Es un *conjunto de conceptos* con una o varias *características comunes*. En una clasificación jerárquica, es cualquier *subdivisión* que comprende más de un concepto.

CLASIFICACION

- 1) Proceso de reconocimiento de clases en el contenido temático de un documento, de acuerdo con un esquema o tabla de clasificación.
- 2) Es un vocabulario controlado precoordinado, cuyas clases forman un *esquema* ordenado jerárquicamente con una *notación* que traduce ese orden y con un *índice alfabético* que remite a la notación.

CLASIFICACION EXPLICITA

Es aquella en que la clasificación se evidencia en su presentación jerárquica.

CLASIFICACION IMPLICITA

Es aquella presente en un ordenamiento alfabético de términos, incluida en las relaciones jerárquicas.

CLASIFICACION FACETADA

Es la **CLASIFICACION 2)** que aplica con rigor el análisis facetar en su estructuración.

CLASIFICACION JERARQUICA

véase

CLASIFICACION 2)

COMPATIBILIDAD

Es la cualidad que tienen dos o más lenguajes documentarios de campos idénticos o próximos, de poder usarse en forma complementaria, o poder llegar a fusionarse.

CONCEPTO

Es una idea, una imagen mental de una cosa (algo, cualquier cosa). En un sistema de almacenamiento y recuperación de información, el concepto se representa mediante una etiqueta.

CONSISTENCIA DEL SISTEMA DE RECUPERACION DE INFORMACION

Es la coherencia, es decir, que en la *indización o entrada* un mismo tema, sea almacenado por el mismo o distintos indizadores, siempre con el mismo término; y que en la *búsqueda o salida* los indizadores o los usuarios realicen la estrategia de búsqueda con los mismos términos asignados en la indización.

COORDINACION

véase además

SISTEMA PRECOORDINADO – SISTEMA POSTCOORDINADO

Es la combinación de conceptos. Se puede realizar en la *indización o entrada*, o en la *búsqueda o salida*. En el 1^{er} caso es un sistema precoordinado; en el 2^o caso es postcoordinado, también llamado coordinado.

CONVERSION

Es la fusión de varios tesauros compatibles.

DESCRIPTOR

Es el término o etiqueta de un concepto, clase, etc. en un tesoro. Lo ideal es que esté definido (de-finido), es decir, delimitado exactamente con una nota de alcance. Esta es una de las diferencias fundamentales con las listas de encabezamiento de materia.

DOCUMENTO

Tiene tres significados:

- 1) *En general:* es todo soporte que tiene información registrada. Por ej. libro, revista, artículo, informe, documento

(en especial), disco, película, cinta magnética, etc. Los distintos sistemas de información tratan de recuperar estos documentos o parte de ellos.

- 2) *En especial*: es un tipo de publicación, en general no impresa tipográficamente y publicada por organismos internacionales, que tiene como característica su identificación con una sigla y un número propio de cada organismo.
- 3) *Como sinónimo de publicación oficial - gubernamental*: se emplea sobre todo en E.E.U.U. En ese país, en la Biblioteca del Congreso, el Departamento de Publicaciones Oficiales se denomina Departamento de Documentos, y la editorial de Publicaciones Oficiales del país se denomina Superintendencia de Documentos.

DOCUMENTO PERTINENTE véase **PERTINENCIA**

DOCUMENTO RELEVANTE véase **RELEVANCIA**

EFFECTIVIDAD DEL SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION véase **RENDIMIENTO DEL SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION**

ENCABEZAMIENTO DE MATERIA

Es el término o etiqueta de un concepto, tema o materia en una lista de encabezamientos de materia.

EN LINEA véase **SISTEMA EN LINEA**

ENTRADA DIRECTA

(De un término) cuando el encabezamiento de materia o descriptor, formado por más de una palabra, se presenta en el orden del lenguaje común. Es la forma preferida para los tesauros o listas de descriptores. Ej.: **INTENSIDAD DE LLUVIA**.

ENTRADA INVERTIDA

(De un término) cuando en el encabezamiento de materia o descriptor formado por más de una palabra se invierte el orden del lenguaje común. Ej.: **LLUVIA, INTENSIDAD DE**.

ESQUEMA DE CLASIFICACION

Representación de un sistema de clasificación

ESQUEMA JERARQUICO véase **ESQUEMA DE CLASIFICACION**

ESPECIFICIDAD

Es la capacidad de un sistema de información para representar los diferentes conceptos de un documento con un grado adecuado de profundidad.

ESTRATEGIA DE BUSQUEDA

Es el análisis y formulación del método óptimo mediante el cual una demanda particular de información puede ser contestada por el sistema de recuperación de información.

ETIQUETA

Es la *representación* de conceptos por medio de *signos*. Cuando estos signos son *números o letras*, la etiqueta recibe el nombre de *notación*, y cuando son *palabras* recibe el nombre de *término, encabezamiento de materia, palabra clave o descriptor*.

EXHAUSTIVIDAD

Es la capacidad de un Sistema de Información para representar las diferentes facetas de un tema.

FACETA

Es un aspecto o punto de vista desde el cual puede ser tratado un concepto, clase, categoría o campo temático. Por ej. una *obra literaria* tiene 2 aspectos o facetas: la lengua en la que está escrita y la forma literaria a la que pertenece; el *oro* puede ser tratado desde el punto de vista minero, o de su producción, o de sus propiedades metalúrgicas, o como un objeto de orfebrería.

FILTRADO

Revisión y selección de los documentos recuperados mediante una estrategia de búsqueda a fin de entregar al usuario los más relevantes.

HARDWARE

Todo el equipamiento físico (dispositivos mecánicos, magnéticos, eléctricos y electrónicos) que constituye una computadora. También se lo llama *Ferretería*.

HOMOGRAFOS véase **HOMONIMOS**

HOMONIMOS

Palabras que siendo iguales por su forma, tienen distinta significación.

INDICE

Lista ordenada de datos contenidos en un documento o texto -que puede ser un fichero o salida de computadora, o lista impresa o mecanografiada- destinada a localizar esos datos en el documento.

INDICE PERMUTADO

Es una lista de términos en que cada una de las palabras significativas de cada grupo (temático o de título), va desplazándose al primer lugar como *término de acceso*, ubicándose en su nuevo orden alfabético. Se generan por computadora.

INDIZACION

- a) Es la asignación de etiquetas (signos o términos) a los documentos, de acuerdo con su tema, concepto, clase, etc.
- b) Es la asignación de un término de entrada para localizar un documento por autor o por título (indización de autor o de título) o por materia (indización temática). Este concepto es utilizado por LISA (Library and Information Science Abstracts) en su afán de unificar la terminología bibliotecológica, documentaria y de la ciencia de la información.
- c) Es el proceso de esa asignación solamente en la entrada del Sistema, reservándose el concepto de *indización de búsqueda* para la salida del Sistema.
- d) Es el proceso de esa asignación en la entrada y la salida del Sistema; por lo tanto, en este caso será sinónimo de *recuperación de la información* (también en su sentido más general, porque ésta en su sentido más estricto, sería sinónimo de *búsqueda de la información*).

INDIZACION AUTOMATICA

Es la indización que realiza la computadora, es decir, es el procesamiento automático de la indización.

INDIZACION COORDINADA véase SISTEMA POSTCOORDINADO

INDIZACION PONDERADA

Cuando a los términos se les agrega un artificio cuantificador que los distingue como más o menos importantes, con más o menos peso en el texto del documento. Se la denomina también ponderación o pesos.

INDIZACION POSTCOORDINADA véase SISTEMA POSTCOORDINADO

INDIZACION PRECOORDINADA véase SISTEMA PRECOORDINADO

INFORMATICA

Generalmente se lo limita al procesamiento automático de la información. A veces se lo equipara a Ciencia de la información.

JUSTIFICACION LITERARIA

La incorporación de una etiqueta (descriptor, encabezamiento de materia, código sistemático, etc.) a un vocabulario de indización, debe estar justificada por la existencia de bibliografía en el tema.

LENGUAJE

Es un sistema de signos para comunicación que utiliza generalmente un vocabulario y normas (gramática).

LENGUAJE CONTROLADO

Es un conjunto de términos normalizados, seleccionado de un conjunto mayor de términos del lenguaje natural; y que se utiliza para la indización y la recuperación de información. También se lo acostumbra llamar vocabulario controlado.

LENGUAJE DE BUSQUEDA

Lenguaje de indización utilizado en la búsqueda de información.

LENGUAJE DE INDIZACION

Lista de etiquetas (signos y términos) que forman el vocabulario de indización y el que con reglas que forman su sintaxis, se utiliza para la *indización de la información*. Este concepto puede usarse indistintamente como lenguaje de entrada solamente, o como lenguaje de entrada y salida (véase INDIZACION c) y d). Algunos autores denominan al lenguaje de salida, lenguaje de búsqueda.

LENGUAJE DOCUMENTARIO

véase

LENGUAJE DE INDIZACION

LENGUAJE INTERMEDIO O CONMUTADOR

Lista de etiquetas con sus distintos equivalentes en los distintos lenguajes controlados, con el objeto de traducir las etiquetas de un lenguaje a otro. Todos los lenguajes tienen de común el análisis conceptual y difieren en las etiquetas. Mediante el lenguaje intermedio se intenta utilizar el análisis conceptual de un lenguaje aplicándolo a los demás.

LENGUAJE NATURAL

Es el lenguaje utilizado por los autores de los documentos. Puede usarse en la indización o en la búsqueda de información, o en ambas.

LISTA DE DETENCION

En indexación automática, con lenguaje natural, es una lista de palabras excluidas como encabezamiento. Son palabras no significativas (temáticamente) como artículos, preposiciones, comparaciones, etc., que no deben utilizarse en la búsqueda.

LISTA DE ENCABEZAMIENTOS DE MATERIA

Es un vocabulario controlado para sistemas precoordinados de recuperación de información, presentado en forma alfabética, mostrando una clasificación implícita por medio de relaciones jerárquicas. El término utilizado como unidad se denomina encabezamiento de materia.

LITERATURA PRIMARIA

Documento original que no es resultado de un proceso de documentación.

LITERATURA SECUNDARIA

Documento que es resultado de un proceso de documentación, y por lo tanto, da información de documentos primarios, ya sea indicando su existencia (bibliografía), su localización (catálogos), su resumen (abstracts), o reelaborando su contenido en obras de referencia (enciclopedias, diccionarios, revisiones, avances, etc.).

LOGICA DE BOOLE

véase

ALGEBRA DE BOOLE**MACROTESAURO**

Es un tesoro que incluye solamente términos o descriptores básicos comunes a varios campos temáticos, los cuales pueden tener a su vez un vocabulario más especializado denominado tesoro o microtesoro.

MICROTESAURO

Es un tesoro que incluye algunos descriptores básicos comunes de un macrotesoro y los propios de su campo o especialidad.

NANOSEGUNDO

Unidad de tiempo equivalente a 10^{-9} segundos.

NOTA DE ALCANCE

Es la aclaración del significado exacto de la "etiqueta" en un lenguaje de indización, que explica no sólo *qué quiere decir*, sino también *qué no quiere decir*.

ON-LINE véase SISTEMA EN LINEA

OPERADOR

- a) Técnico que maneja una computadora.
- b) En matemática y lógica, signo que indica operación (matemática o lógica) que debe realizarse con los operandos. Por ej., el signo+ en matemática y "o" en álgebra de Boole.
- c) En indización, signo que representa la relación entre dos términos.

OPERACION LOGICA

Operación no aritmética, tal como: seleccionar, clasificar, comparar, etc. Es aquella en que el resultado de asignar uno de los valores (1 ó 0; "sí" o "no") a cada una de una serie de variables, representa en sí uno de los valores.

OPERADOR METRICO

Es un signo indicador de la distancia entre dos términos.

PALABRA CLAVE

Palabra significativa seleccionada de un documento para ser usada como término de recuperación.

PARAMETROS DEL SISTEMA

Aspectos constantes que determinan las características del sistema de almacenamiento y recuperación de información.

PERFORMANCE DEL SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION véase RENDIMIENTO DEL SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION

PERTINENCIA

Es la conformidad *dada por el usuario*, de que el documento recuperado es adecuado a su necesidad de información.

PRECISION

Es la capacidad del sistema de no recuperar documentos no relevantes.

PRE-COORDINACION véase SISTEMA PRECOORDINADO

PROGRAMA

Conjunto de instrucciones escritas en un lenguaje de programación, necesario para la ejecución automática de un proceso.

RED DE INFORMACION

Sistema de información compuesto por más de 2 bibliotecas o centros de información, relacionados por operaciones continuas -generalmente procesos y servicios comunes- con el fin de mejorar éstos y reducir costos.

RELACIONES DE AFINIDAD véase RELACIONES DE ASOCIACION

RELACIONES DE ASOCIACION

Vinculación sintáctica entre conceptos, dentro de un mismo campo temático específico. En un tesoro, se identifican con la sigla TR (término relacionado).

RELACIONES DE EQUIVALENCIA

Relaciones horizontales entre conceptos que, dentro de un mismo campo temático, pueden ser tratados como sinónimos. En un tesoro se indican con USE y UP (use por).

RELACIONES JERARQUICAS

Relación semántica entre dos conceptos donde uno se encuentra subordinado al otro. En un tesoro se indican con las siglas TG (término genérico) y TE (término específico).

RELACIONES PARADIGMATICAS

Relaciones entre conceptos, manifestados en la estructura jerárquica de un lenguaje documentario y que existen independientemente del proceso de indización. Se las llama también relaciones analíticas o verticales.

RELACIONES SEMANTICAS

Relaciones entre conceptos según su significado.

RELACIONES SINTAGMATICAS

Relaciones entre conceptos, establecidas durante la indización y no contenidas en la estructura jerárquica del lenguaje documentario. Se las denomina también relaciones sintéticas u horizontales.

RELEVANCIA

Es la conformidad, dada por el especialista en el tema o el especialista en información en ese campo específico, de que el documento es adecuado a la necesidad de información del usuario.

RENDIMIENTO DEL SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION

Es la calidad del funcionamiento del Sistema de Almacenamiento y Recuperación de Información.

ROL

Indicador que se agrega a un término capaz de asumir diferentes funciones, a fin de identificar esa función. También se lo denomina indicador de función.

RUIDO

- 1) *En comunicación:* cualquier disturbio en un mensaje, después de la emisión y antes de la recepción.
- 2) *En documentación:* es la cantidad de documentos no pertinentes o no relevantes recuperados.

SERVICIO EN LINEA véase **SISTEMA EN LINEA**

SINONIMOS

Vocablos y expresiones que tienen diferentes formas y una misma significación.

SISTEMA véase además **SISTEMA DE INFORMACION DOCUMENTARIA**

Conjunto de elementos interrelacionados con el objeto de producir efectos determinados, o de cumplir una o varias funciones.

SISTEMA DE ALMACENAMIENTO Y RECUPERACION DE INFORMACION véase **SISTEMA DE INFORMACION DOCUMENTARIA**

SISTEMA DE CLASIFICACION véase **CLASIFICACION 2)**

SISTEMA DE COINCIDENCIA OPTICA

Sistema semi-mecánico de almacenamiento y recuperación de información que utiliza fichas de entrada de términos en las que los Nos. de los documentos están registrados como perforaciones en posiciones establecidas mediante un sistema de coordenadas, y en el cual la recuperación de la información se obtiene mediante la exposición de las fichas frente a una fuente de luz.

SISTEMA DE INFORMACION DOCUMENTARIA

Es un sistema que tiene por fin almacenar, procesar, recuperar y difundir información contenida en documentos.

SISTEMA DE SELECCION OPTICA

véase

SISTEMA DE COINCIDENCIA OPTICA

SISTEMA EN LINEA

Sistema en el cual los datos de entrada son llevados al computador directamente desde terminales y/o, los datos procesados son transmitidos directamente adonde son utilizados: se omiten las etapas intermedias tales como perforación de datos en tarjetas, cintas de papel.

SISTEMA PRE-COORDINADO

Es aquél en que la combinación de conceptos se realiza durante la indización o entrada de información al sistema.

SISTEMA POSTCOORDINADO

Es aquél en que la combinación o coordinación de conceptos se realiza durante la búsqueda de la información. Se lo denomina también sistema coordinado.

SISTEMA UNITERMINO

Indización por palabras únicas tomadas del texto analizado sin controlar vocabulario. Originalmente empleado por Mortimer Taube para designar las palabras únicas usadas como entrada en un sistema postcoordinado de indización. Actualmente se emplea para indicar un índice coordinado de fichas unitérmino.

SOFTWARE

véase además

PROGRAMA

Conjunto de programas y sistemas operativos que forman el material intelectual necesario para el uso y funcionamiento de una computadora.

SUBFACETA

Es la separación de conceptos, etc. dentro de una faceta, por medio de una sola característica.

TASA DE ACIERTO

Es la relación entre los *documentos relevantes recuperados* y la totalidad de *documentos relevantes* del sistema.

TASA DE CRECIMIENTO

Es la relación entre el número de *descriptores* de un *tesauro* en un momento dado y su variación en un período de tiempo determinado.

TASA DE ENRIQUECIMIENTO

De un *tesauro*: es la relación entre el número de *relaciones jerárquicas o asociativas* y el número de *descriptores*.

TASA DE EQUIVALENCIA

De un *tesauro*: es la relación entre los *descriptores* y los *no descriptores* (sinónimos).

TASA DE PRECISION

Es la relación entre los *documentos relevantes recuperados* y la totalidad de los *documentos recuperados*.

TASA DE RECIPROCIDAD

En un *tesauro*: es el grado en que todos los *términos* y sus *relaciones* sean *recíprocos*.

TERMINAL

Es una unidad de entrada o salida por medio de la cual el usuario se comunica con el o los computadores asociados a la red.

TERMINO

Palabra o símbolo que denota un concepto.

TESAURO

Según la UNESCO un *tesauro* puede definirse de acuerdo a su función o de acuerdo a su estructura.

Desde el punto de vista de su función, un *tesauro* es una herramienta de control terminológico utilizada para traducir a un lenguaje más restringido (lenguaje documentario, lenguaje de información) el lenguaje natural usado en los documentos y utilizados por los indizadores y usuarios.

Desde el punto de vista de su estructura, el *tesauro* es un vocabulario controlado y dinámico de términos que guardan entre sí relaciones semánticas y genéricas, que se aplica a un campo específico del conocimiento.

TESAURO FACETADO

(Thesaurus Facet). Es un sistema de lenguaje controlado, elaborado por primera vez en Inglaterra por la Compañía Inglesa de Electricidad en 1970, combinando una clasificación facetada (sistemática) con un tesoro (alfabético) que reemplaza al índice alfabético de la clasificación. La CDU reformada planifica para la década del 80, en forma similar, un índice tesaúrico de la CDU.

TESAURO MONOLINGUE

Es el que contiene términos o descriptores en una lengua.

TESAURO MULTILINGUE

Es el que contiene términos o descriptores con sus equivalentes en más de una lengua.

TIEMPO COMPARTIDO véase además SISTEMA EN LINEA

Sistema automatizado que permite a múltiples usuarios operar o usar un sistema en una forma simultánea o aparentemente simultánea.

TIEMPO REAL

Forma de procesamiento de la información, especialmente explotado desde terminales, donde el tiempo de respuesta del sistema es considerado inmediato, y está disponible para influir en la toma de decisiones en una operación continua.

TRADUCCION

Es la conversión de un lenguaje natural a otro natural, o de un lenguaje natural a un lenguaje controlado.

UNION

Símbolo usado en la indización para indicar la vinculación entre dos descriptores que se hallan relacionados en el texto de un documento. También se lo denomina indicador de enlace.

UNITERMINO véase SISTEMA UNITERMINO

VOCABULARIO

Conjunto de términos o signos utilizados en un lenguaje.

VOCABULARIO CONTROLADO

véase

LENGUAJE CONTROLADO

VOCABULARIO DE BUSQUEDA

véase

LENGUAJE DE BUSQUEDA

VOCABULARIO DE ENTRADA

Conjunto de términos que dan entrada al vocabulario controlado, formado por los descriptores elegidos para expresar los conceptos, más los sinónimos, antónimos, etc. que son los términos no usados que remiten a los descriptores.

VOCABULARIO HIBRIDO

Es aquel vocabulario que utiliza una combinación del vocabulario controlado y del lenguaje natural.

VOCABULARIO DE INDIZACION

véase

LENGUAJE DE INDIZACION

VOCABULARIO INTERMEDIO O CONMUTADOR

véase

LENGUAJE INTERMEDIO O CONMUTADOR

4. ANEXOS

ANEXO 1

INSTITUCIONES QUE INFORMAN SOBRE LENGUAJES DE INDIZACION EXISTENTES

- **ASLIB**
3 Belgrave Square
London SW 1X 8PL
Gran Bretaña

- **Centralny Instytut Informacji Naukowej Technicznej i Ekonomicznej (CIINTE)**
Clearinghouse al.
Niepodlegosci 188
Warsaw
Polonia

- **Centro Argentino de Información Científica y Tecnológica (CAICYT)**
Moreno 431/33
1091 Buenos Aires
Argentina
(especialmente para lenguas española y portuguesa)

- **Centro de Documentação e Informação**
Ministerio da Educação e Investigação Científica
Rua Yan 47
Lisboa 3
Portugal

- **Committee FID/CR "Classification Research"**
President
Ingetraut Dahlberg
Woostr 36a.
D-6000 Frankfurt -M50
República Federal de Alemania

- **Centro de Documentação e Informação (CDI)**
Junta de Investigações Científicas do Ultramar
Rua Jau 47
Lisboa 3
Portugal

- **Bibliographic Systems Center**
Case Western Reserve University
School of Library Science
Cleveland, Ohio 44106
U.S.A.

Las anteriores son algunas de las instituciones que en el mundo proveen información sobre tesauros y lenguajes de indicación en general y sobre temas específicos.

ANEXO 2

BIBLIOGRAFIAS DE TESAUROS

AA, H.J. VAN DER.- Thesaurus bibliography.- Amsterdam: Studiecentrum NOVI, 1977.- 86 p.

BIBLIOGRAPHIC BULLETIN OF THE CLEARINGHOUSE AT IINTE.- Naukowej, Technicznej i Ekonomicznej. Dom Clearingowy.- Warszawa, 1969.

BUREAU NATIONAL D'INFORMATION SCIENTIFIQUE ET TECHNIQUE (FRANCIA).- Annuaire des thesaurus, suivi d'un Index des domaines et de sous-domaines couverts par les thésaurus, et d'une Liste des Organismes Auteurs. 3 éme. éd.- Paris, 1976.

GILBERT, VALERIE.- A list of thesauri and subject headings held in the ASLIB Library.

En: ASLIB proceedings 31(6), 1979.- p. 264-267.

INSTITUTO BRASILEIRO DE INFORMACAO EM CIENCIA E TECNOLOGIA (IBICT).- Bibliografia sobre thesaurus / organizada por Suly Cambrala Alves.- Rio de Janeiro: IBICT, 1978.- 37 p.

INTER-ORGANIZATION BOARD FOR INFORMATION SYSTEM (IOB).- Bibliography of United Nations: Thesauri, Classifications, Nomenclatures.- Geneva: IOB, 1979.- 252 p.

JUNTA DE INVESTIGACOES CIENTIFICAS DO ULTRAMAR. CENTRO DE DOCUMENTACAO E INFORMACAO (PORTUGAL).- Thesauri: relacao bibliográfica, vol. 5, 1971 a 1975.- Lisboa, 1977.- 72 p.

KUTTEN, A.- Thesauri bibliography.- Haifa: Technion Israel Institute of Technology, Elyachan Library, 1975.- 62 p.

LANCASTER, F. WILFRID.- Some Controlled vocabulaires for Study or examination.

En: Vocabulary control for information retrieval.- Washington, D.C.; Information Resources Press, 1972.- p. 227-228.

LAUREILHE, MARIE THERESE.- Bibliographie des thesauri et index par matiere parus depuis 1960.

En: Bulletin des Bibliothèques de France.- Paris, 1969-1978.

MAC CAFFERTY, MAXINE.- Thesauri & Thesauri construction / compiled by Maxine Mac Cafferty.- London: ASLIB, 1977.- 191 p.- (ASLIB Bibliography N° 7).

WALKLEY, JANET; HAY, BARBARA.- An annotated list of thesauri held in the ASLIB Library.

En: ASLIB Proceedings 23(6), 1971.- p. 292-300.-

GRUPO DE TRABAJO SOBRE LENGUAJES DOCUMENTARIOS. REUNIBER '78.- Inventario de lenguajes documentarios no tradicionales en lengua española y/o portuguesa.- Buenos Aires. 95 p. (en publicación)

ANEXO 3

CUESTIONARIO PREVIO

Sobre el tema del curso Ud. posee:

1) Conocimientos teóricos

amplios ☐

básicos ☐

escasos ☐

ninguno ☐

2) Conocimientos prácticos

amplios ☐

básicos ☐

escasos ☐

ninguno ☐

3) Sobre este tema ha realizado Ud. otros cursos?

SI ☐

NO ☐

Cuáles?

4) Qué se propone obtener de este curso?

5) Cómo se enteró del curso?

ANEXO 3

EVALUACION FINAL

1) El Curso ha respondido a sus expectativas (INDICAR CON UN CIRCULO)

5	4	3	2	1
Completamente	Casi completamente	Parcialmente	Poco	Muy poco

2) Qué grado de utilidad tuvo el material presentado para su trabajo o interés profesional?

5	4	3	2	1
Extremadamente útil	Muy útil	Útil	De poca utilidad	De ninguna utilidad

3) La calidad del curso fue:

5	4	3	2	1
Muy bueno	Bueno	Regular	Pobre	Muy pobre

4) Fue suficiente el tiempo dedicado a los siguientes aspectos del curso?

	Tiempo insuficiente	Tiempo suficiente	Demasiado tiempo
Clases teóricas			
Intercambio de ideas			
Trabajos prácticos			

5) El grupo de participantes fue demasiado numeroso para la eficiencia del curso?

☐ SI

☐ NO

6) El nivel en que se trató el tema fue para Ud.:

Demasiado fácil ☐

Correcto ☐

Demasiado simple ☐

7) Indique su grado de satisfacción con cada uno de los siguientes aspectos del Curso

	Bueno	Satisfactorio	Insatisfactorio
Intervalos			
Salas			
Transparencias			
Facilidades para Trabajos prácticos			

8) En su opinión ¿en qué medida la serie de clases teóricas y prácticas han tratado el tema de tesauros?

Estudio completo ☐

Estudio bastante completo ☐

Estudio incompleto ☐

9) Elija la frase que indique con mayor exactitud la medida en que el contenido del curso fue nuevo para usted.

La mayor parte del contenido era nueva para mí ☐

La mitad, aproximadamente, era nueva para mí ☐

Menos de la mitad era nueva para mí ☐

El curso no contenía nada que fuese nuevo para mí ☐

10) ¿Cuáles fueron los mejores aspectos del Curso?

11) ¿Cuáles fueron los peores aspectos del Curso?

12) ¿Tiene Ud. otros comentarios que hacer acerca del Curso?

ANEXO 4

GRUPO DE TRABAJO SOBRE LENGUAJES DE INDIZACION CAICYT - ARGENTINA

Artola, María O.

**Biblioteca
Consejo Federal de Inversiones
San Martín 871 - 2º E
1004 - Buenos Aires
TE: 221-2472 / 2782 / 2142 / 2202
int. 139 - 140**

Domicilio particular:

**Moreno 624 - 6º M
TE: 34-7847**

García Acosta, Araceli

**Biblioteca
Instituto Nacional de Planificación Económica
Ministerio de Economía
Hipólito Irigoyen 250
1310 - Buenos Aires
TE: 33-1121 ; 34-6411 int. 605 - 272**

Hepburn de Santacápita, Cristina

**Centro de Investigación Documentaria
Instituto Nacional de Tecnología Industrial
Gral. Paz y Nazca
1650 - Miguelete
Prov. de Buenos Aires
TE: 755-3660 Correo Central 1000 Bs. As.

Av. L.N. Alem 1067
1001 Capital**

*** Martínez, José María**

**Biblioteca
Instituto para la Integración de América Latina
Cerrito 264 - 1º Piso 119
1010 - Buenos Aires
TE: 35-5513**

- * **Molina, Celia**
- Instituto Nacional de Ciencias y Técnicas Hídricas
(INCYTH)
Centro Informática Hídrica
Empalme ruta 205 Kd, 2,5
C.C. 23 - 1802 Aeropuerto Ezeiza
TE: 295-1167
- Domicilio particular:**
- Gral. Agustín P. Justo 551
1153 Buenos Aires
- Moritz de Tay, Gina**
- Biblioteca
Instituto de Investigaciones Científicas y Técnicas
de las Fuerzas Armadas (CITEFA)
Zufriategui y Varela
1603 - Villa Martelli
TE: 761-0081 / 31 / 0231 int. 199
- Domicilio particular:**
- Larrea 1042 - 10° C
TE: 825.7263
- * **Patalano, Mercedes**
- Confederación Médica de la República Argentina
Hipólito Irigoyen 2038
Buenos Aires
TE: 47-9704 / 3892
- Domicilio particular:**
- Manuel Galvez 159 - 9° 2
1155 - Buenos Aires
TE: 221-4859
- * **Carballo, Graciela:**
- Centro de Investigación Textil
INTI
Av. L.N. Alem 1067
1001 - Capital
- * **Pesce, Laura**
- Biblioteca
Instituto Torcuato Di Tella
11 de Setiembre 2139
1428 - Buenos Aires
TE: 784-8225 / 8264; 781-5013 / 15
- * **Saugy de Babini, Dominique**
- Centro de Información sobre Estudios de Post-
grado y Becas (CIESBE)
Fundación José María Aragón
Perú 263 - 3° piso
1067 - Buenos Aires
TE: 34-8643 / 5178

* Zítara de Ríbezzo, Ethel
Domicilio particular:

San Martín 996 - 14º D
1004 Buenos Aires
TE: 31-5981

COORDINACION Y SECRETARIA DEL GRUPO

* Allmand de Gallo, Mónica

Capacitación y Estudios
Centro Argentino de Información Científica y
Tecnológica (CAICYT)
Moreno 431/33
1091 - Buenos Aires
TE: 30-5749 34-6637

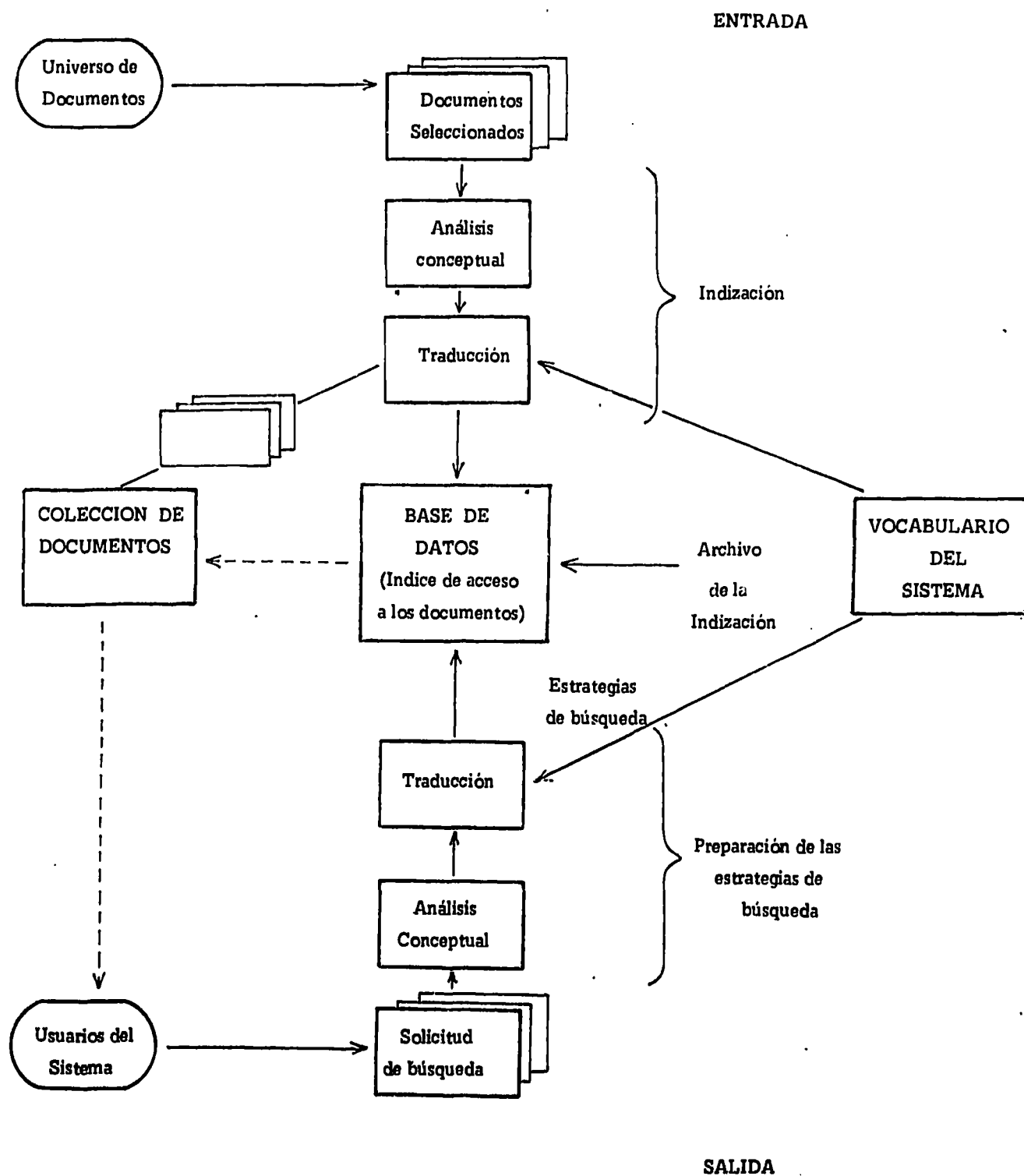
* Sanllorenti, Ana María

Secretaría Técnica
CAICYT

* Los señalados con asterisco, son los miembros actualmente activos del grupo.

TRANSPARENCIAS

TRANSPARENCIA I-1



TRANSPARENCIA I-2

ENTRADA

INDIZACION = ASIGNACION DE ETIQUETAS A LOS DOCUMENTOS

LENGUAJES DE INDIZACION = LISTA DE ETIQUETAS QUE FORMAN EL VOCABULARIO DEL SISTEMA

ETIQUETAS (NOMBRES)

DEFINICIONES

ORGANIZADAS PARA FACILITAR SU USO

SALIDA

ESTRATEGIA DE BUSQUEDA = ASIGNACION A LAS PREGUNTAS DE POSIBLES ETIQUETAS QUE PUDIERON HABERSE DETERMINADO PARA LOS DOCUMENTOS QUE SE QUIEREN RECUPERAR

TRANSPARENCIA I-3

LA RECUPERACION DE LA INFORMACION

DEPENDE DE LA

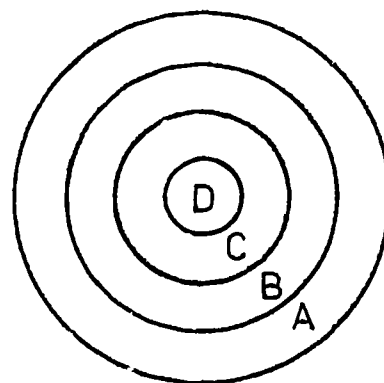
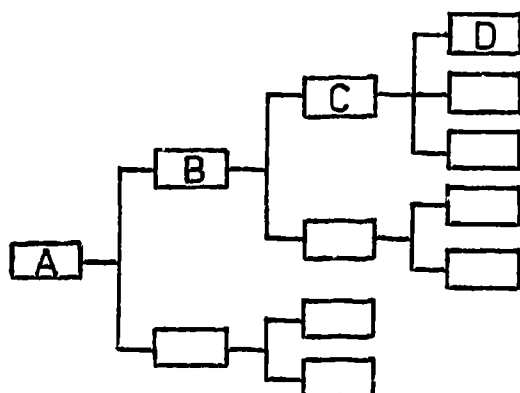
- INDIZACION (ENTRADA)

Y LA

- ESTRATEGIA DE BUSQUEDA (SALIDA)

CLASIFICACION (DECIMAL UNIVERSAL)

A	34	DERECHO
B	347	DERECHO CIVIL
C	347.2	DERECHOS REALES
D	347.23	DERECHO DE LA PROPIEDAD

Estructura jerárquica

ENCABEZAMIENTOS DE MATERIA

Argentina

x República Argentina

Derecho de propiedad (Subdv. geogr.)

va. Dominio eminente

x Propiedad privada

xx Propiedad

TRANSPARENCIA I-6

ARGENTINA

BT AMERICA LATINA
 CONO SUR
 RT ARGENTINO

DERECHO

SN -----

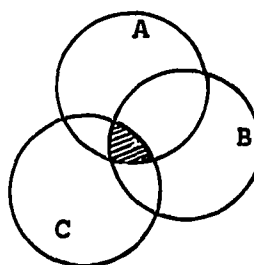
 BT CIENCIAS JURIDICAS
 NT DERECHO CIVIL
 . . .
 . . .
 RT ABOGADOS

PROPIEDAD

NT PROPIEDAD COLECTIVA
 . . .
 . . .

TEMA: Derecho de la propiedad en la Argentina

- A. DERECHO
 B. PROPIEDAD
 C. ARGENTINA



TRANSPARENCIA I- 7

SISTEMAS PRECOORDINADOS Y POSTCOORDINADOS

Ej.: DERECHO DE LA PROPIEDAD EN LA ARGENTINA

CDU	347.23(82)	}	CONCEPTOS PRECOORDINADOS EN EL SISTEMA Y EN EL MOMEN- TO DE LA INDIZACION
ENCABEZAMIENTOS DE MATERIA	DERECHO DE LA PROPIEDAD- ARGENTINA		
DESCRIPTORES	DERECHO PROPIEDAD ARGENTINA	}	CONCEPTOS POSTCOORDINADOS EN EL MOMENTO DE LA BUSQUEDA

TRANSPARENCIA I- 8

ORGANIZACION DEL ARCHIVO

Pre-coordinado

Post-coordinado

Pre-coordinado

UNIDIMENSIONAL - LINEAL

A, B, C, D, E

Ej.: FATIGA (MECANICA)

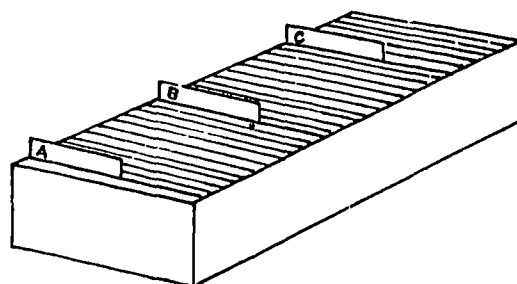
FATIGA DE METALES

FATIGA, DE PANELES DE ALUMINIO, EN AVIONES,
A VELOCIDAD SUPERSONICA

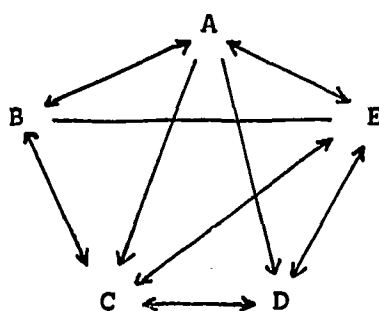
ORDENAMIENTO DEL ARCHIVO CORRESPONDIENTE A UN SISTEMA PRE-COORDINADO

UN ORDEN ESTABLECIDO

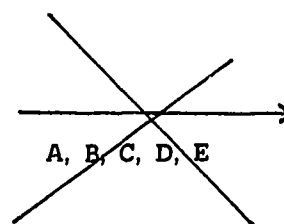
- ALFABETICO
- NUMERICO

Fichero Lineal

MULTIDIMENSIONAL



NO LINEAL



No es posible

TRANSPARENCIA I- 11

COLECCION DE DOCUMENTOS

		Características (Descriptores)									
NUMEROS DE LOS DOCUMENTOS		A	B	C	D	E	F	G	H	I	
	1			X			X				
	2		X		X						
	3		X				X				
	4				X				X		
	5		X				X			X	
	6				X				X		X
	7			X			X				
	8		X		X						

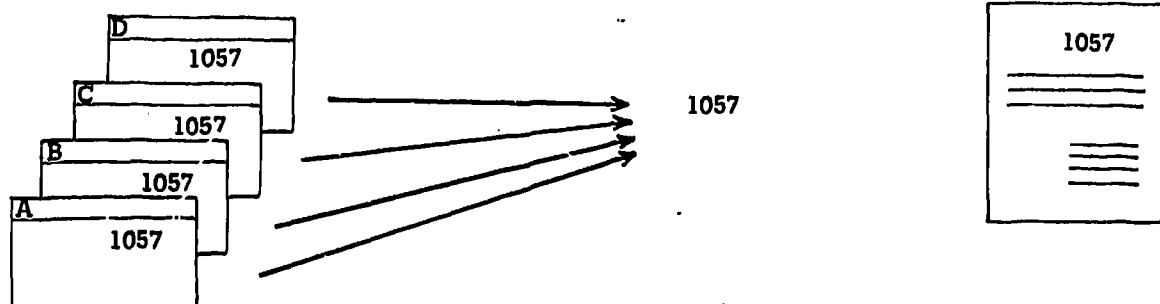
TRANSPARENCIA I- 12

SISTEMA POSTCOORDINADOARCHIVO INVERTIDO (MULTIDIMENSIONAL)

Indización



Búsqueda



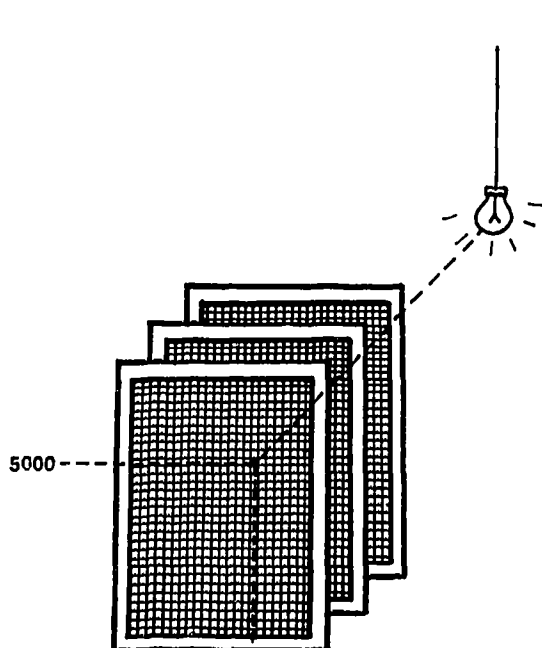
TRANSPARENCIA I- 13

DESCRIPTOR									
0	1	2	3	4	5	6	7	8	9
	21 1071				35 175 1085			238	

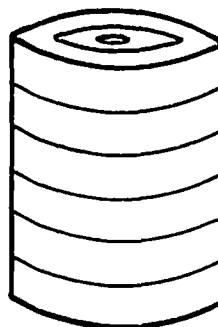
TRANSPARENCIA I- 14

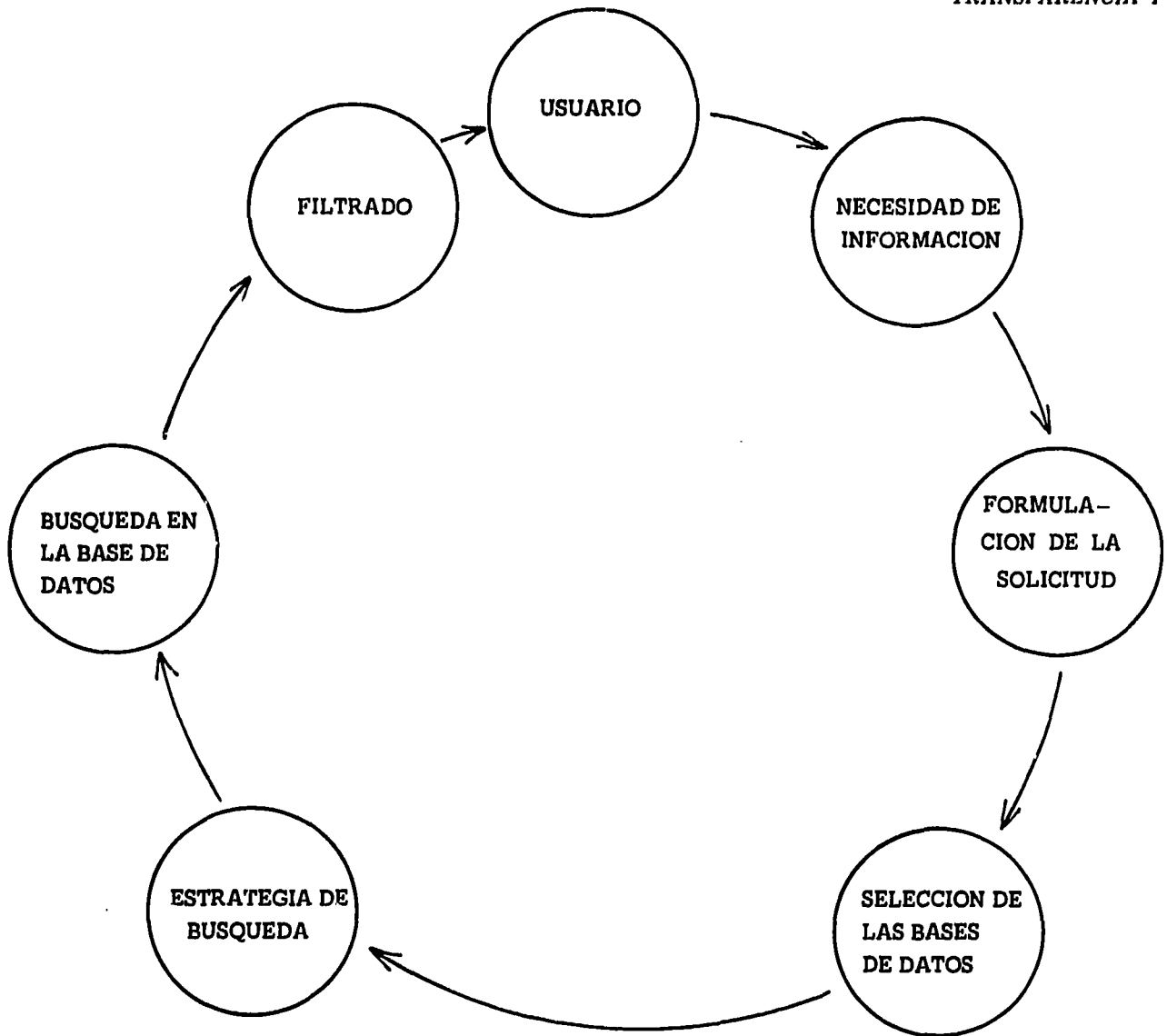
SISTEMA POSTCOORDINADO – ARCHIVO MULTIDIMENSIONAL – INVERTIDO

SELECCION OPTICA

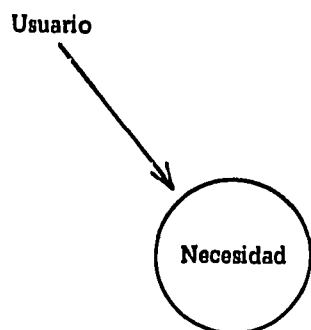
ARCHIVO INVERTIDO
AUTOMATIZADO

Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº	Nº
tér.	Doc.	Doc.	Doc.	Doc.	tér.	Doc.	Doc.	Doc.	Doc.	tér.	Doc.	Doc.

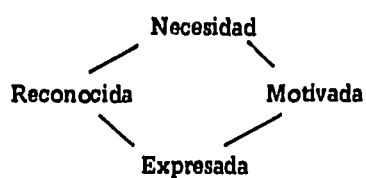




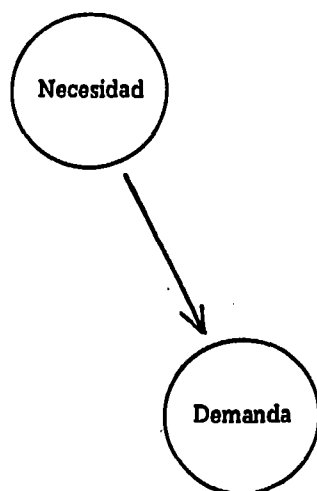
TRANSPARENCIA I- 16



Necesidad → pedido
 Existencia de Centros
 Accesibilidad
 Costo
 Facilidad de uso (p. ej. Tiempo)
 Experiencia previa
 Rapidez del servicio
 Validez de la solución
 Probabilidad de que exista una solución



TRANSPARENCIA I- 17



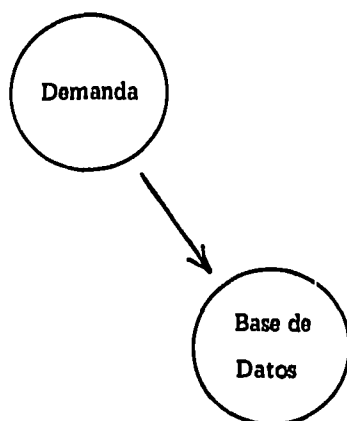
Interpretación de las
 capacidades/limitaciones (del usuario)

 Habilidad para expresar claramente
 las necesidades

 Claridad de percepción de la necesidad

 Modo de interacción

 Provisión de asistencia

TRANSPARENCIA I- 18

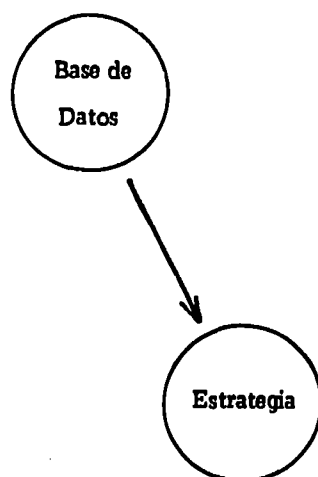
Selección de la Base de Datos

Conocimiento de la Base de Datos y su alcance

Cobertura

Calidad (p. Ej.: Indización)

Accesibilidad

TRANSPARENCIA I- 19

Interpretación de las Necesidades del usuario

**Habilidad para cubrir todas las aproximaciones
razonables (Marco de referencia)**

Proyección del vocabulario y otras ayudas

**Adecuación del vocabulario para describir temas
que se necesitan**

Habilidad para construir estrategias lógicas

Capacidad de los programas de búsqueda.

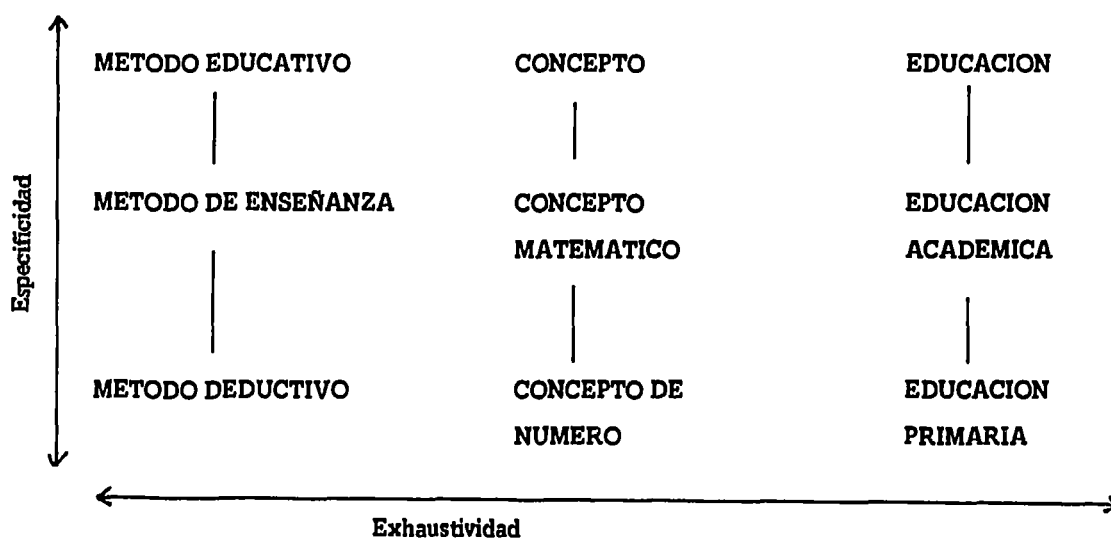
TRANSPARENCIA I- 20

ESPECIFICIDAD

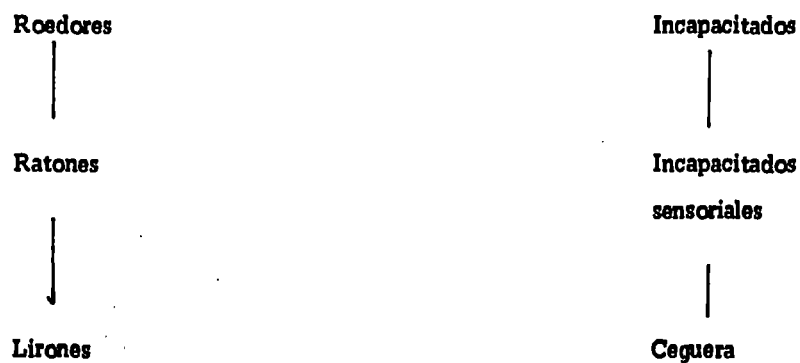
Producción de maíz en Egipto

CEREAL	MEDIO ORIENTE
TRIGO	EGIPTO
ARROZ	JORDANIA
MAIZ	SIRIA
CEBADA	LIBANO

TRANSPARENCIA I- 21



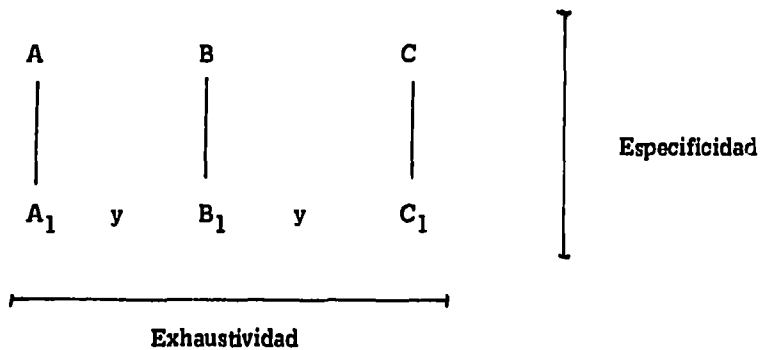
TRANSPARENCIA I- 22

NIVEL DE ESPECIFICIDAD

TRANSPARENCIA I- 23

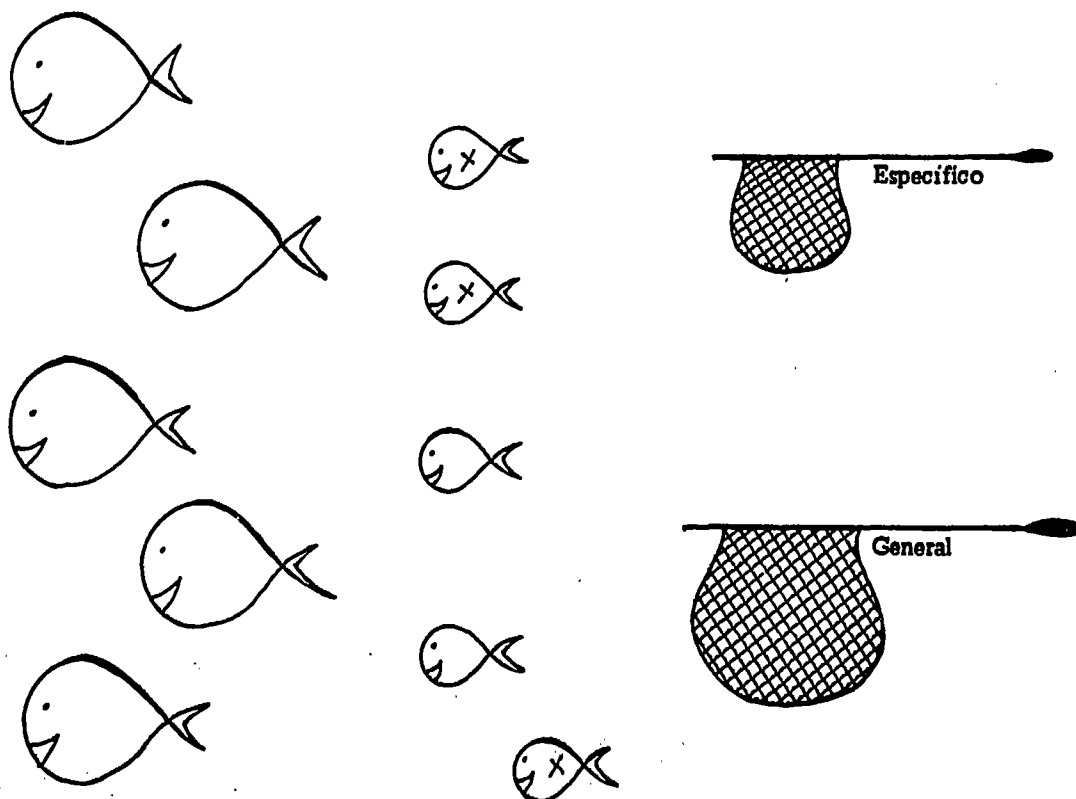
Búsqueda

Estrategias



(Nivel de coordinación)

TRANSPARENCIA I- 24

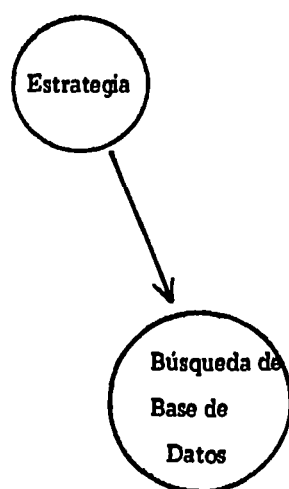


Análisis Conceptual

1. ¿Cuál es el tema?
2. ¿Por qué le interesa a nuestro usuario?

Traducción

¿Qué términos representan mejor el tema seleccionado?



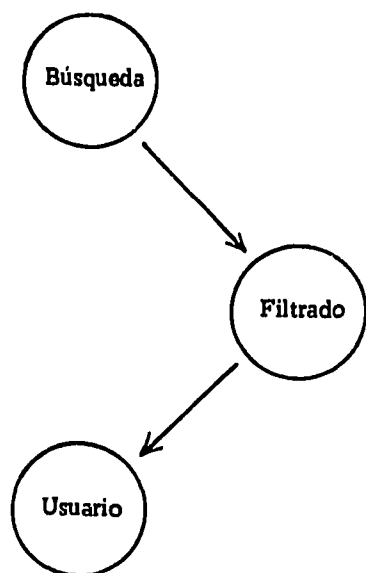
Adecuación del vocabulario

Exhaustividad en la indización

Errores de indización / consistencia

Relaciones ambiguas o falsas

TRANSPARENCIA I- 27



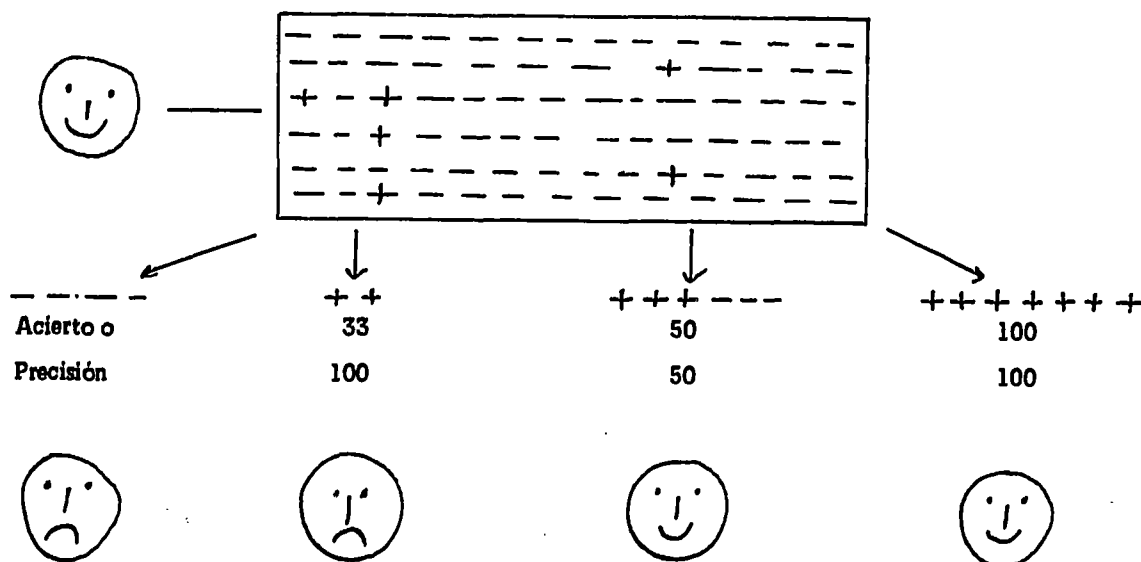
Calidad de la solicitud del usuario

Calidad de reemplazo de términos o traducción

Interpretación de las necesidades del usuario para la búsqueda

Tiempo de filtrado

TRANSPARENCIA I- 28

Búsqueda Bibliográfica

TRANSPARENCIA I- 29

	RELEVANTES	NO RELEVANTES	TOTAL
RECUPERADO	15	10	25
NO RECUPERADO	5	499,970	499,975
TOTAL	20	499,980	500,000

$$\text{ACIERTO: } \frac{15}{20} = 0,75$$

$$\text{PRECISION: } \frac{15}{25} = 0,60$$

TRANSPARENCIA I- 30

RELEVANTES	NO RELEVANTES
20	499,80
5	
15	10

$$\text{TASA DE ACIERTO} = \frac{15}{20} \cdot 100 = 75 \text{ o/o}$$

$$\text{TASA DE PRECISION} = \frac{15}{25} \cdot 100 = 60 \text{ o/o}$$

TRANSPARENCIA I-31

TASA DE ACIERTO =

$$\frac{\text{DOCUMENTOS RELEVANTES RECUPERADOS}}{\text{TODOS LOS DOCUMENTOS RELEVANTES}} \times 100 = \frac{15}{20} \cdot 100 = 75 \text{ o/o}$$

TASA DE PRECISION =

$$\frac{\text{DOCUMENTOS RELEVANTES RECUPERADOS}}{\text{TODOS LOS DOCUMENTOS RECUPERADOS}} \times 100 = \frac{15}{25} \cdot 100 = 60 \text{ o/o}$$

TRANSPARENCIA I-32

	Fallas de Acierto	Fallas de Precisión
Lenguaje de indización	81 (10.2 o/o)	1094 (36 o/o)
Indización	298 (37.4 o/o)	393 (12.9 o/o)
Búsqueda	279 (35.0 o/o)	983 (32.4 o/o)
Interacción	199 (25 o/o)	503 (16.6 o/o)
Otros	11 (1.4 o/o)	78 (2.5 o/o)

Factores de la indización que afectan el rendimiento de un sistema de almacenamiento y recuperación de información

1. Política de indización
2. Precisión de la indización
 - Todos los conceptos importantes
 - Representación correcta

Selectivo

Alta Precisión

Bajo Acierto

1.	_____
2.	_____
3.	_____
4.	_____
5.	_____
6.	_____
7.	_____
8.	_____
9.	_____
10.	_____

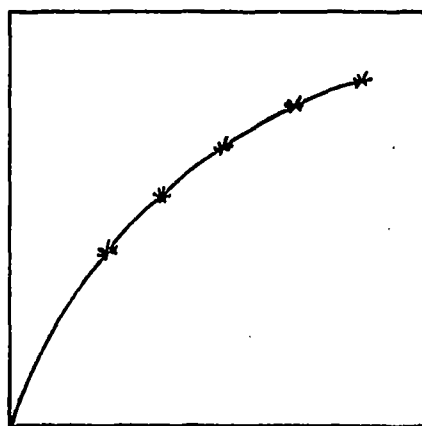
Exhaustivo

Alto acierto

Baja precisión

c/o de
Necesidades
satisfechas

100



0

Número de términos

50

Título del artículo

"Efecto de la pérdida visual
en el crecimiento de la corteza
cerebral en los ratones

Indización

Corteza cerebral

Visión

Ceguera

Ratones

Pérdida sensorial

Solicitud:

Factores que afectan el crecimiento ,
regeneración y degeneración en el
sistema nervioso central

Crecimiento ?Problemas de Traducción

1. Término incorrecto

Propiedades de ignición del Borohidruro de Aluminio

Combustibles
sólidos

Combustibles
líquidos

2. Falta de especificidad en el vocabulario

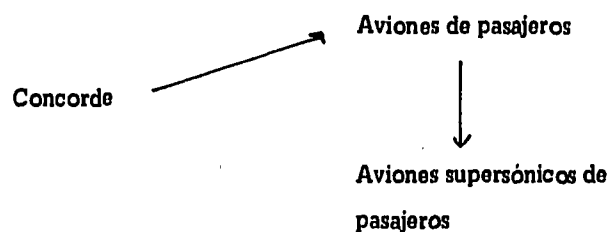
Trasplante
de riñón



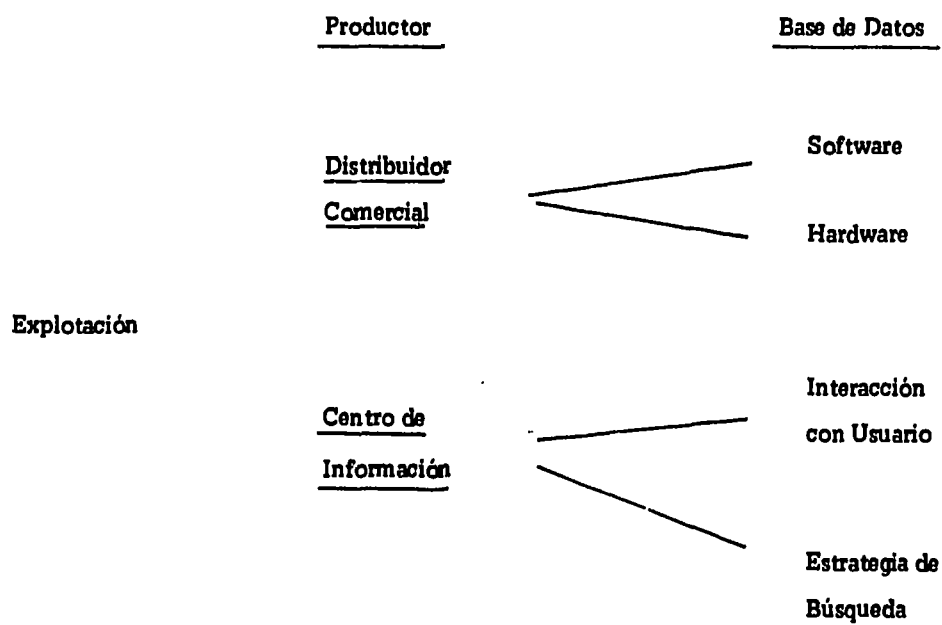
Trasplante
de órganos

TRANSPARENCIA I-38

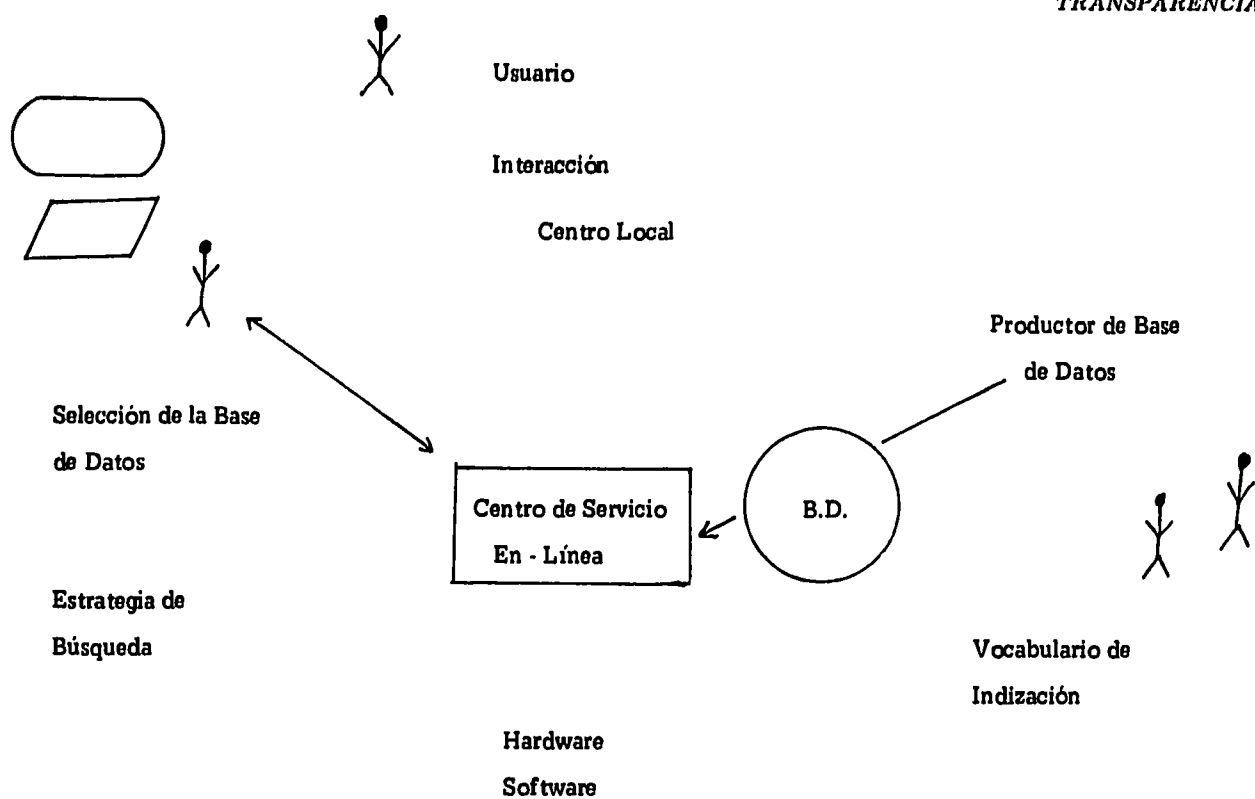
3. Falta de especificidad en la Indización



TRANSPARENCIA I-39

Factores Interrelacionados

TRANSPARENCIA I-40



TRANSPARENCIA I-41

Vocabularios Controlados

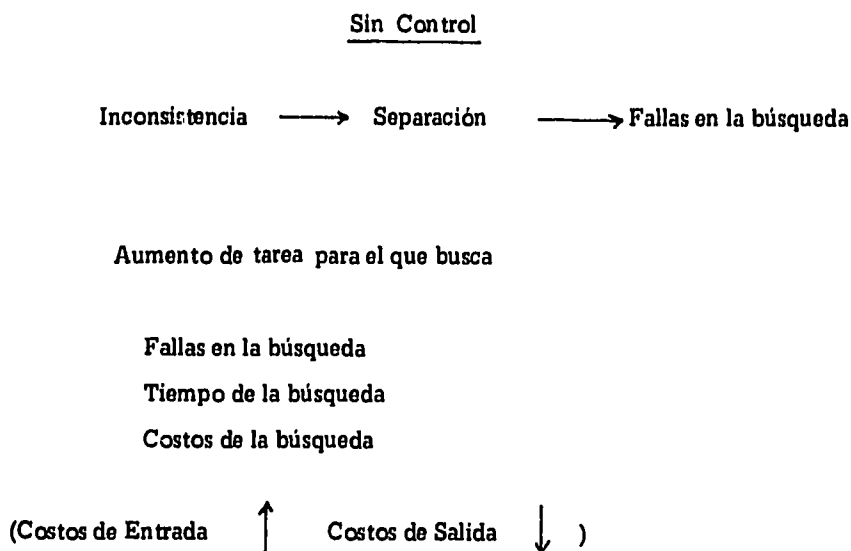
1. Tesouro
2. Encabezamiento de materia
3. Esquema de Clasificación

Conjunto limitado de términos
(Especificadores del tema)

Indización

Búsqueda

TRANSPARENCIA I-42



TRANSPARENCIA I-43



(Se evita la Dispersión)

Sinónimos

Casi - Sinónimos

Antónimos

Homógrafos

2. Facilitar búsquedas amplias (genéricas)

Unen términos relacionados semánticamente

Paradigmática (T.G. - TE)

Sintagmática (Véase además, TR)

TRANSPARENCIA I-44

Sinónimos

Ala triangular	<u>use</u>	Alas Delta
Pie de atleta	<u>use</u>	Tinea pedis
Filatelia	<u>use</u>	Colección de estampillas
Proyectores, películas	<u>use</u>	Proyector de películas
Organización de las Naciones Unidas para la Agricultura y la Alimentación	<u>use</u>	FAO
CA	<u>use</u>	Corriente alterna
DGE	<u>use</u>	Dirección General de Electricidad - México
Light Amplification by the stimulated emission of radiation	<u>use</u>	LASER

TRANSPARENCIA I-45

Casi Sinónimos

Herencia	<u>use</u>	Genética
Exactitud	<u>use</u>	Precisión

Antónimos

Aspereza	<u>use</u>	Lisura
Inestabilidad	<u>use</u>	Estabilidad

Aspero	—————	Liso
Rugoso		Suave

Lenguaje de Indización

1. Falta de especificidad
 2. Relaciones ambiguas
 - Coordinaciones falsas
 - Relaciones incorrectas
 3. Falta de estructura
 - Jerarquía
 - Términos relacionados
-

TRANSPARENCIA I- 47

Tres ratones ciegos

Ratones

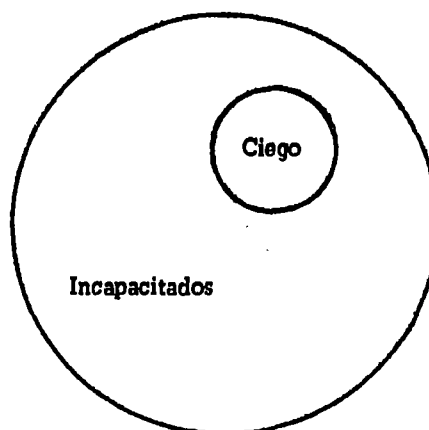
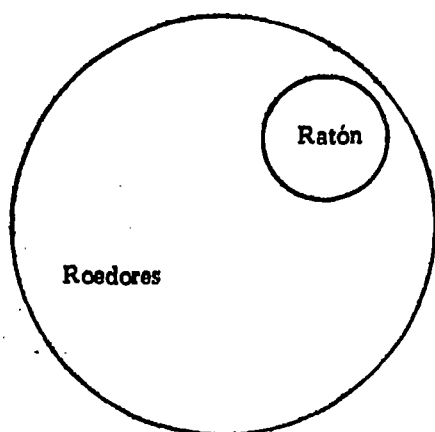
use

Roedores

Ciegos

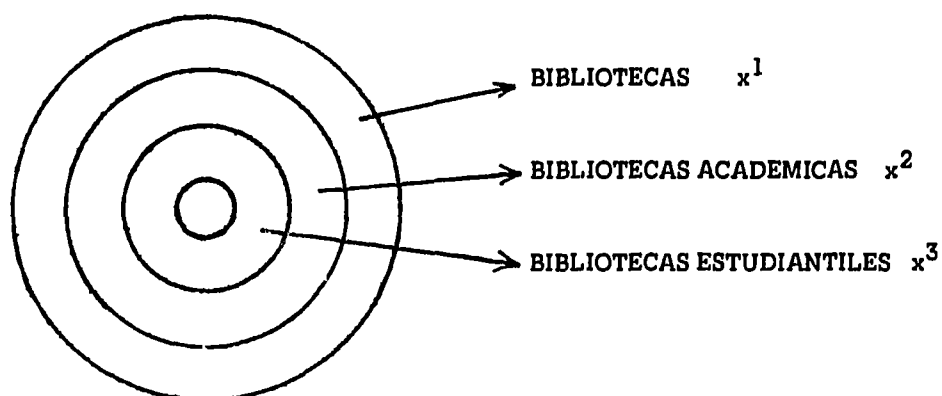
use

Incapacitados



ESPECIFICIDAD

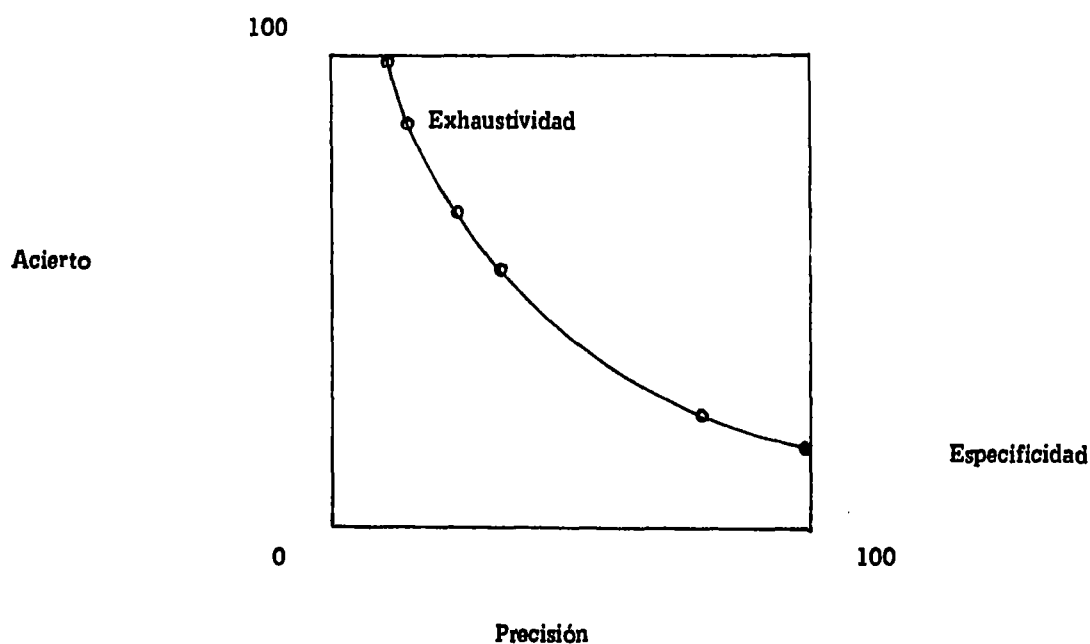
Controla la precisión de la recuperación en forma directa



Pero hace la recuperación exhaustiva (acierto alto) más dificultosa

hace que la indización consistente sea más dificultosa

Vocabulario más amplio, más costoso



TRANSPARENCIA I- 50

Falsa coordinación

Ratones

Ratas

Conejos

Ciego

Sordo

Mudo

TRANSPARENCIA I- 51

Relación incorrecta

Computadora (2)

Lectura

Diseño

Epilepsia

Avión (1)

TRANSPARENCIA I- 52

UnionesPrecoordinación

Ratones

}

3002/A

Ciegos

Ratas

}

3002/B

Sordas

Conejos

}

3002/C

Mudos

Soldadura

Limpieza

Acústica ultrasónica

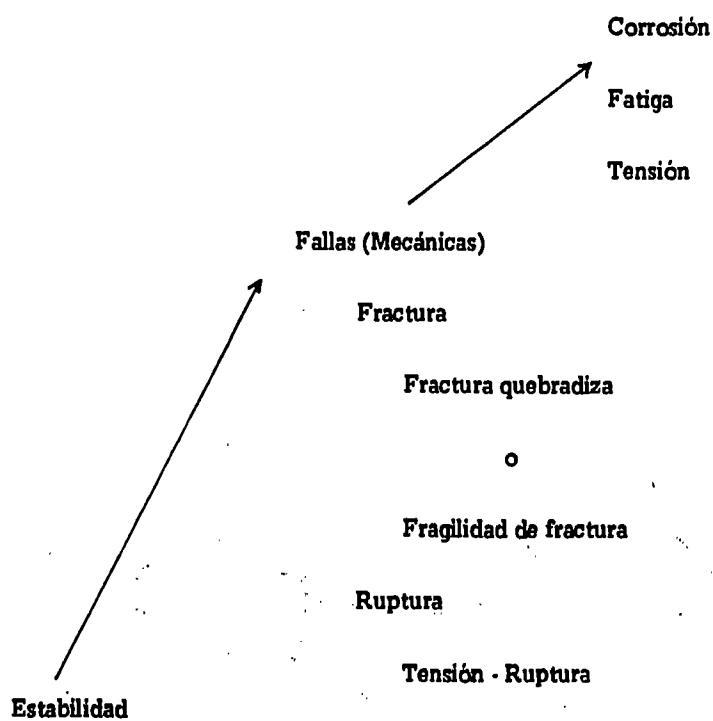
Soldadura ultrasónica

TRANSPARENCIA I- 53

Indicadores de Rol

(2)	Objeto de acción		
(3)	Herramienta	(1)	Modo de realización
	Computadora (3)		
	Diseño		
	Avión (2)		

TRANSPARENCIA I- 54

Estructura del vocabulario

TRANSPARENCIA I- 55

Vocabulario de Entrada

Sinónimos

Casi - Sinónimos

Antónimos

Específico (No utilizado)



General

Gavilán

use

Halcón

use

Aves de rapina

use

Pájaros

(Todos los términos pueden ser considerados términos de entrada)

TRANSPARENCIA I- 56

Vocabulario de Entrada

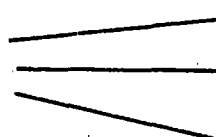
1. Evita inconsistencia
2. Indica a quien busca donde buscar
3. Asegura a quien busca que ese tema está cubierto
4. Reduce la tarea del indizador

TRANSPARENCIA I- 57

Soldadura de Arco Argón

use

Soldadura de arco protegido

Soldadura de
arco
argón

Soldadura de arco protegido

Soldadura de arco

Soldadura por gas

Parámetros del Sistema

1. Alcance

Núcleo

Periferia

2. Tipo de Documento

Libro

Informe

Interno

Externo

Artículo de Revista

3. Volumen de la Literatura

Volumen actual

Tasa de crecimiento

4. Usuarios

Especialistas en campos temáticos

Especialistas en información

5. Tipos de pregunta

Genérico

Específico

6. Entrada

versus

Salida

7. Implementación Física

TRANSPARENCIA II- 1

1951 MOOERS
1951 TAUBE: UNITERMINO

TESAUROS

1959 - Du Pont
1960 - ASTIA
1961 - Chemical Engineering Thesaurus
1964 - E.J.C. Thesaurus of Engineering Terms

ASTIA

E.J.C. Thesaurus of Engineering Terms

Project Lex

1967 - Thesaurus of Engineering and Scientific Terms (TEST)

<p>NORMAS NACIONALES NORMAS INTERNACIONALES</p>

TRANSPARENCIA II- 2

ESTRUCTURA DE UN TESAURO

CLASIFICADA SISTEMATICA

ALFABETICA

CLASIFICACION EXPLICITA

CLASIFICACION IMPLICITA

TRANSPARENCIA II- 3

BIBLIOTECAS

- PUBLICAS
MUNICIPALES
POPULARES

- ACADEMICAS
UNIVERSITARIAS
ESTUDIANTILES
ESCOLARES

- ESPECIALIZADAS
EMPRESARIAS
DE INVESTIGACION

- NACIONALES

MATERIALES BIBLIOTECARIOS

- MICROFORMAS
MICROFICHAS
ULTRAMICROFICHAS
MICROFILMES
MICROTARJETAS

MATERIALES AUDIOVISUALES

- FILMINAS
- PELICULAS
- VIDEOTAPES

CLASIFICACION ENCUBIERTA

BIBLIOTECAS ACADEMICAS
 TG BIBLIOTECAS
 TE BIBLIOTECAS ESTUDIANTILES
 BIBLIOTECAS UNIVERSITARIAS
 TR EDUCACION SUPERIOR
 BIBLIOTECAS ESTUDIANTILES
 TG BIBLIOTECAS ACADEMICAS
 TE BIBLIOTECAS ESCOLARES

RESIDUOS INDUSTRIALES

NA -----

 UP EFLUENTES
 DESECHOS
 TG -----
 TE -----
 TR POLUCION DEL AIRE
 CONTAMINACION
 DISPOSICION
 RECUPERACION
 EFLUENTES use RESIDUOS INDUSTRIALES

EJEMPLO DE UN ASIENTO DE TESAURO

DESCRIPTOR	→	SERVICIOS SOCIALES
NOTA DE ALCANCE	→	NA
USADO POR (REFERENCIA)	→	UP
TERMINO GENERICO	→	TG
TERMINOS ESPECIFICOS	→	TE
TERMINOS RELACIONADOS	→	TR

ADMINISTRACION CORRESPONDIENTE A LAS NECESIDADES DE
 LA COMUNIDAD EN LA ESFERA SOCIAL
 ASISTENCIA SOCIAL
 INSTITUCIONES DE BENEFICENCIA
 ADMINISTRACION SOC'AL
 ASISTENCIA A LOS ANCIANOS
 AYUDA A LA INFANCIA
 AYUDA DE EMERGENCIA
 CUIDADO DE LOS INVALIDOS
 ORFELINATOS
 SERVICIOS DE SALUD
 PROBLEMAS SOCIALES

TRANSPARENCIA II- 7

MOOERS

DESCRIPTOR

1. ETIQUETA

**2. DETERMINACION
DEL ALCANCE**

UNA CLASE DE VERTEBRADOS OVIPAROS DE SANGRE

CALIENTE CON EL CUERPO CUBIERTO DE PLUMAS Y LAS PATAS

DELANTERAS TRANSFORMADAS EN ALAS

↓ ↓ ↓ ↓
PAJAROS BIRDS 598.2 AcK

TRANSPARENCIA II- 8

DIRECTRICES DE LA UNESCO

ENTRADA DIRECTA

FORMACION PROFESIONAL

NO PROFESIONAL, FORMACION

SUSTANTIVO (O EXPRESION SUSTANTIVA)

ASPEREZA NO ASPERO

EJERCICIO NO EJERCITAR

**PERO ALGUNOS ADJETIVOS PUEDEN PERMITIRSE
AUTOMATICO
MANUAL
CONICO**

TRANSPARENCIA II- 11

HOMOGRAFOS

PLANTAS (BOTANICA)

PLANTAS (INDUSTRIA)

TRANSPARENCIA II- 9

SINGULAR / PLURAL

ESTADISTICA

ESTADISTICAS

}

DIFERENTES SIGNIFICADOS

- EL PLURAL ES PREFERIDO, EXCEPTO

a) PROCEDIMIENTOS: SOLDADURA ACIDIFICACION

b) PROPIEDADES: OPACIDAD TEXTURA

- NOMBRES QUE RESPONDEN A CANTIDAD (¿CUANTOS?): PLURAL

LADRILLOS

PERROS

DIARIOS

- NOMBRES QUE RESPONDEN A VOLUMEN (¿CUANTO?): SINGULAR

PAPEL

CONCRETO

ARENA

TRANSPARENCIA II- 10

ABREVIATURAS

DEBEN EVITARSE, EXCEPTO LAS SIGLAS ESTABLECIDAS POR EL USO:

RADAR

LASER

UNESCO

PUNTUACION

DEBE EVITARSE, AUNQUE PUEDEN ACEPTARSE ALGUNOS SIGNOS:

GUION

CUPRO-CLORAMINAS

PARENTESIS

TIEMPO (ATMOSFERICO)

NORMALIZACION DE LOS DESCRIPTORES

Forma directa sobre la forma invertida

Forma sustantiva sobre la forma adjetiva

Forma plural sobre la singular

Sin abreviaturas (salvo excepciones)

Sin puntuación (salvo excepciones)

Aclarando homógrafos

PREFERENCIAS**CANTIDAD DE TERMINOS POR DESCRIPTOR:**

La menor cantidad de términos que expresen un concepto, ya que cada descriptor es una unidad significativa dentro del tesoro.

PARTE ALFABETICA

TRANSPARENCIA II-13

DES NACIONALIZACION <CONT>

RT: EMPRESAS PUBLICAS
NACIONALIZACION
PROPIEDAD PUBLICA

DESPEGUE ECONOMICO

ECONOMIC TAKE OFF / DECOLLAGE ECONOMIQUE -
03.02.03

DESPERDICIOS

WASTES / DECHETS - 12.08.01

UF: DESECHOS

RT: CONTAMINANTES

NT: AGUAS RESIDUALES

DESPERDICIOS AGRICOLAS

DESPERDICIOS DOMICILIARIOS

DESPERDICIOS INDUSTRIALES

DESPERDICIOS RADIACTIVOS

RT: ADMINISTRACION DE DESPERDICIOS

ELIMINACION DE DESPERDICIOS

RECUPERACION

DESPERDICIOS AGRICOLAS

AGRICULTURAL WASTES / DECHETS AGRICOLES -
07.05.05

RT: DESPERDICIOS

DESPERDICIOS DE MADERA

WOOD WASTES / DECHETS DE BOIS - 08.07.01

RT: DESPERDICIOS INDUSTRIALES

RT: MADERA

DESPERDICIOS DOMICILIARIOS

DOMESTIC WASTES / ORDURES MENAGERES - 16.03.04

RT: DESPERDICIOS

RT: SERVICIOS DE SANEAMIENTO

DESPERDICIOS INDUSTRIALES

INDUSTRIAL WASTES / DECHETS INDUSTRIELS -
08.04.00

RT: DESPERDICIOS

NT: CHATARRA

DESPERDICIOS DE MADERA

RT: INDUSTRIA

DESPERDICIOS RADIACTIVOS

RADIOACTIVE WASTES / DECHETS RADIOACTIFS -
08.11.04

RT: DESPERDICIOS

MATERIAS RADIACTIVAS

RT: CONTAMINACION RADIACTIVA

DESPIDO

DISMISSAL / LICENCIEMENT - 13.05.00

SN: FIN DEL EMPLEO CUANDO EL EMPLEADOR LO
DECIDE.

RT: ADMINISTRACION DEL PERSONAL

EXCESO DE PERSONAL

INDEMNIZACION POR CESANTIA

RENUNCIA

DESPLAZAMIENTOS RESIDENCIA-TRABAJO

COMMUTING / MIGRATIONS ALTERNANTES - 14.07.00

RT: MIGRACIONES INTERNAS

DESPOBLACION

POPULATION DECLINE / DEPEUPLEMENT - 14.01.02

RT: DECLINACION DE LA FECUNDIDAD

DINAMICA DE LA POBLACION

POBLACION

DESTILACION

DISTILLING / DISTILLATION - 08.12.03

RT: PROCESOS INDUSTRIALES

DETERGENTES

DETERGENTS / DETERGENTS - 08.12.08

RT: PRODUCTOS DE LIMPIEZA

DETERIORO DE LOS ALIMENTOS

FOOD SPOILAGE / ALTERATION DES ALIMENTS -
15.03.01

RT: ALIMENTOS

ALIMENTOS PERECEDEROS

CONTAMINACION DE LOS ALIMENTOS

DETERIORO DEL MEDIO AMBIENTE

ENVIRONMENTAL DEGRADATION / DEGRADATION DE
L'ENVIRONNEMENT - 16.03.04

RT: DESERTIFICACION

RT: CONTAMINACION

MEDIO AMBIENTE

DETERMINACION DEL SALARIO

WAGE DETERMINATION / FIXATION DU SALAIRE -
13.07.00

SN: DETERMINACION DE LA TASA DE SALARIO O DE LA
ESTRUCTURA DEL SALARIO EN UNA SITUACION
DADA.

RT: IGUALDAD DE REMUNERACION

SALARIOS

DEUDA

DEBT / DETTE - 11.02.02

NT: DEUDA EXTERNA

DEUDA PUBLICA

RT: CONSOLIDACION DE LA DEUDA

CREDITO

EMPRESTITOS

ENDEUDAMIENTO

PRESTAMOS

REBAJA DE LA DEUDA

REEMBOLSO

DEUDA EXTERNA

EXTERNAL DEBT / DETTE EXTERIEURE - 11.03.01

RT: DEUDA

RT: BALANZA DE PAGOS

DEUDA PUBLICA

PUBLIC DEBT / DETTE PUBLIQUE - 11.01.01

RT: DEUDA

HACIENDA PUBLICA

DEVALUACION

DEVALUATION / DEVALUATION - 11.02.01

RT: POLITICA MONETARIA

RT: BALANZA DE PAGOS

MONEDAS

PARTE SISTEMATICA

16.

MEDIO AMBIENTE. RECURSOS NATURALES.

16.01

ECOLOGIA.

16.01.01

ECODESARROLLO
 ECOLOGIA
 ECOLOGIA ANIMAL
 ECOLOGIA HUMANA
 ECOLOGIA VEGETAL
 USE: FITOECOLOGIA
 EQUILIBRIO BIOLOGICO
 USE: EQUILIBRIO ECOLOGICO
 EQUILIBRIO ECOLOGICO
 EQUILIBRIO NATURAL
 USE: EQUILIBRIO ECOLOGICO
 ESTABILIDAD DE LOS ECOSISTEMAS
 USE: EQUILIBRIO ECOLOGICO
 FITOECOLOGIA
 INVESTIGACION ECOLOGICA

16.01.02

BIOMASA
 BIOSFERA
 CALIDAD AMBIENTAL
 ECOSISTEMAS
 MEDIO AMBIENTE
 MEDIO AMBIENTE ACUATICO
 MEDIO AMBIENTE FISICO
 MEDIO AMBIENTE HUMANO
 MEDIO AMBIENTE MARINO
 MEDIO AMBIENTE TERRESTRE
 MEDIO NATURAL
 USE: MEDIO AMBIENTE FISICO

16.02

RECURSOS NATURALES.

16.02.01

ADMINISTRACION DE RECURSOS
 AGOTAMIENTO DE RECURSOS
 EVALUACION DE RECURSOS
 EXPLORACION DE RECURSOS
 USE: INVENTARIO DE RECURSOS
 EXPLOTABILIDAD
 EXPLOTACION DE RECURSOS
 INVENTARIO DE RECURSOS
 SOBREEXPLOTACION

16.02.02

RECURSOS ANIMALES
 RECURSOS DE SUELO
 RECURSOS DEL MAR
 RECURSOS ENERGETICOS
 RECURSOS FORESTALES
 RECURSOS HIDRICOS

16.02.02<CONT>

RECURSOS MINERALES
 RECURSOS NATURALES
 RECURSOS NO-RENOVABLES
 RECURSOS PESQUEROS
 RECURSOS PETROLIFEROS
 RECURSOS RENOVABLES
 RECURSOS VEGETALES

16.03

DESASTRES. CONTAMINACION.

16.03.01

DANOS
 DESASTRES
 DEVASTACION

16.03.02

DESASTRES CAUSADOS POR EL HOMBRE
 DESASTRES NATURALES
 EROSION
 EROSION DEL SUELO
 EXPLOSIONES
 INCENDIOS
 INCENDIOS FORESTALES
 INUNDACIONES

16.03.03

INSECTOS
 INSECTOS PARASITOS
 LANGOSTAS
 PARASITOS DE ANIMALES
 PARASITOS DE PLANTAS
 ROEDORES

16.03.04

AEROSOL
 AGUAS RESIDUALES
 CONTAMINACION
 CONTAMINACION ACUSTICA
 CONTAMINACION ATMOSFERICA
 CONTAMINACION DEL AGUA
 CONTAMINACION DEL LITORAL
 CONTAMINACION DEL MAR
 CONTAMINACION DEL MEDIO FISICO
 USE: CONTAMINACION DEL LITORAL
 CONTAMINACION DEL SUELO
 CONTAMINACION FLUVIAL
 CONTAMINACION QUIMICA
 CONTAMINACION RADIACTIVA
 CONTAMINACION TERMICA
 CONTAMINACION TRANSFRONTERA
 CONTAMINANTES
 CONTAMINANTES ATMOSFERICOS

OCDE. Macrothesaurus

Tesauro por temas. Grupos de descriptores

PARTE SISTEMATICA

IRRADIACION
 . IRRADIACION DE ALIMENTOS

ISOTOPOS
 . RADIOISOTOPOS

JUBILACION
 . JUBILACION ANTICIPADA

LEGISLACION
 . LEGISLACION DEL TRABAJO
 . . CODIGO DEL TRABAJO
 . LEGISLACION ECONOMICA
 . . LEGISLACION ANTITRUST
 . LEGISLACION SOCIAL

LENGUAS
 . LENGUA ARABICA
 . LENGUA CHINA
 . LENGUA ESPANOLA
 . LENGUA FRANCESA
 . LENGUA INGLESA
 . LENGUA RUSA
 . LENGUAS OFICIALES
 . LENGUAS VERNACULAS

LIBERTAD
 . LIBERTADES CIVILES
 . . LIBERTAD DE ASOCIACION
 . . LIBERTAD DE ENSEÑANZA
 . . LIBERTAD DE EXPRESION
 . . LIBERTAD DE INFORMACION
 . . LIBERTAD DE OPINION
 . . LIBERTAD RELIGIOSA

LIDERAZGO
 . LIDERAZGO POLITICO

LIQUIDEZ
 . LIQUIDEZ INTERNACIONAL

MAQUINARIA
 . MAQUINARIA AGRICOLA
 . . TRACTORES
 . MAQUINARIA ELECTRICA
 . . APARATOS ELECTRICOS
 . . . APARATOS ELECTRODOMESTICOS
 . . MOTORES ELECTRICOS
 . MAQUINARIA PARA LA CONSTRUCCION
 . . MAQUINARIA PARA MOVER TIERRAS

MATERIALES DE CONSTRUCCION
 . AGLOMERADOS
 . ARENA
 . ASFALTOS
 . CAL
 . CEMENTO
 . CERAMICAS
 . . LADRILLOS
 . CONTRACHAPADO
 . HORMICON
 . MADERA DE CONSTRUCCION
 . MADERA PENSADA
 . MATERIALES REFRACTARIOS
 . . ASBESTO
 . PIEDRA

MECANIZACION
 . MECANIZACION AGRICOLA

MEDICINA
 . ACUPUNTURA
 . CIRUGIA
 . MEDICINA LABORAL
 . MEDICINA PREVENTIVA
 . . ATENCION PRENATAL
 . . PROFILAXIS
 . . SALUD MATERNO-INFANTIL
 . . VACUNACION
 . MEDICINA TRADICIONAL
 . MEDICINA VETERINARIA
 . PEDIATRIA
 . PSIQUIATRIA

MEDICOS VETERINARIOS
 . ZOOTECNISTAS

MEDIO AMBIENTE
 . MEDIO AMBIENTE FISICO
 . . ATMOSFERA
 . . BIOSFERA
 . . MEDIO AMBIENTE ACUATICO
 . . . MEDIO AMBIENTE MARINO
 . . MEDIO AMBIENTE TERRESTRE
 . MEDIO AMBIENTE HUMANO
 . . MEDIO CULTURAL
 . . MEDIO SOCIAL
 . . . MEDIO ESCOLAR
 RELACIONES DOCENTES ALUMNOS
 RELACIONES ESCUELA-COMUNIDAD
 . . MEDIO FAMILIAR
 . . MEDIO RURAL
 . . MEDIO URBANO

MEDIOS DE COMUNICACION
 . MEDIOS DE COMUNICACION DE MASAS
 . . LIBROS
 . . PELICULAS
 . . . PELICULAS EDUCATIVAS
 . . PRENSA
 . . RADIO
 . . . RADIO EDUCATIVA
 . . TELEVISION
 . . . TELEVISION EDUCATIVA

MEDIOS DE TRANSPORTE
 . TRANSPORTE NO-MOTORIZADO
 . . BICICLETAS
 . VEHICULOS AUTOMOTRICES
 . . AUTOBUSES
 . . AUTOMOVILES
 . . AVIONES
 . . . HELICOPTEROS
 . . BUQUES
 . . . BUQUES CISTERNAS
 . . . BUQUES DE CARGA
 . . . BUQUES PESQUEROS
 BUQUES ARRASTREROS
 BUQUES FACTORIA
 . . CAMIONES
 . . MOTOCICLETAS
 . . TRACTORES
 . . TRENES
 . . . LOCOMOTORAS
 . . . MATERIAL RODANTE

TRANSPARENCIA II-16

PARTE ALFABETICA

ANALYTICAL CHEMISTRY
(CON'T)

AT CHEMICAL ANALYSIS
 CHEMISTRY
 INORGANIC CHEMISTRY
 QUALITATIVE ANALYSIS
 QUANTITATIVE ANALYSIS
 VOLUMETRIC ANALYSIS
 ANALYZERS
 0701 0801 0902 0905
 (EXCLUDES DEVICES FOR PERFORMING
 MATHEMATICAL ANALYSIS)
 BY MEASURING INSTRUMENTS
 AT ENGINE ANALYZERS
 SIGNAL ANALYZERS
 AT CONTROLLERS
 DETECTORS
 MONITORS
 SELECTORS
 TEST EQUIPMENT
 ANALYZING
 0401 0802 1505 3406 3407
 (USE OF A MORE SPECIFIC TERM IS
 RECOMMENDED--CONSULT THE TERMS
 LISTED BELOW)
 UF ANALYSIS
 INSTRUMENTAL ANALYSIS
 AT ALGEBRA
 ANALYSIS (MATHEMATICS)
 APPLICATIONS OF MATHEMATICS
 CHEMICAL ANALYSIS
 CREEP ANALYSIS
 DIAGNOSIS
 DIFFERENTIAL GEOMETRY
 DIFFERENTIAL THERMAL ANALYSIS
 EVALUATION
 EXAMINATION
 FIGURE OF MERIT
 FPCRECASTING
 NETWORK ANALYSIS
 PHOTOINTERPRETATION
 POSTFLIGHT ANALYSIS
 PREFLIGHT ANALYSIS
 SIGNAL ANALYSIS
 SPECTRUM ANALYSIS
 STATISTICAL ANALYSIS
 STRESS ANALYSIS
 STRUCTURAL ANALYSIS
 SYSTEMS ANALYSIS
 TERRAIN ANALYSIS
 TESTS
 TRAJECTORY ANALYSIS
 WEIGHT ANALYSIS
 X RAY ANALYSIS
 ANAPHYLAXIS
 0404 0405 0408
 AT SENSITIZING
 ANASTIGMATISM
 0404 0405 0408
 AT OPTOMETRY
 VISION
 ANATASE
 0403 1804
 UF OCTAHEDRITE
 BY OCHALCOGENIDES
 METAL OXIDES
 OXIDES
 TITANIUM COMPOUNDS
 TITANIUM OXIDES
 AT MINERALS
 PIGMENTS
 RUTILE
 ANATOMY
 0402 0404
 AT ADRENAL GLAND
 AORTA
 ARM (ANATOMY)
 ARTERIES
 BARORECEPTORS
 BLADDER
 BLOOD VESSELS
 BONES
 BRAIN
 BRAIN STEM
 BRONCHI
 BRONCHIAL TUBE
 CAPILLARIES (ANATOMY)
 CARDIAC AURICLES
 CARDIAC VENTRICLES
 CARDIOVASCULAR SYSTEM
 CARTILAGE

NASA THESAURUS (ALPHABETICAL LISTING)

CEREBELLUM
 CEREBRAL CORTEX
 CEREBRUM
 CHEMORECEPTORS
 CHEST
 CHIN
 CHOROID MEMBRANES
 CIRCULATORY SYSTEM
 COCHLEA
 COLLAGENS
 CONGENERS
 CONJUNCTIVA
 CONNECTIVE TISSUE
 CONSTRICTORS
 CORNEA
 CORPUSCLES
 CORTI ORGAN
 CRANIUM
 DIAPHRAGM (ANATOMY)
 DIASTOLE
 EAR
 EARDRUMS
 ELBOW (ANATOMY)
 ENDOCRINE GLANDS
 EPICARDIUM
 ERYTHROCYTES
 ESOPHAGUS
 EUSTACHIAN TUBES
 EYE (ANATOMY)
 FEMUR
 FINGERS
 FLEXORS
 FOREARM
 FOVEA
 GENITOURINARY SYSTEM
 GLANDS (ANATOMY)
 GLOMERULUS
 GONADS
 GRAVIRECEPTORS
 HAND (ANATOMY)
 HEAD (ANATOMY)
 HEART
 HEMATOPOIESIS
 HEMATOPOIETIC SYSTEM
 HIPPOCAMPUS
 HUMAN BODY
 INTRACHANIAL CAVITY
 JOINTS (ANATOMY)
 KIDNEYS
 KNEE (ANATOMY)
 LABYRINTH
 LEG (ANATOMY)
 LEUCOCYTES
 LIMBS (ANATOMY)
 LIVER
 LUNGS
 LYMPHOCYTES
 MAMMARY GLANDS
 MARROW
 MASTOIDS
 MECHANORECEPTORS
 MIDDLE EAR
 MUSCULOSKELETAL SYSTEM
 MYOCARDIUM
 NECK (ANATOMY)
 NOSE (ANATOMY)
 OCCIPITAL LOBES
 OCULOMOTOR NERVES
 ORGANS
 OTOLITH ORGANS
 OVARIES
 PANCREAS
 PARATHYROID GLAND
 PELVIS
 PHARYNX
 PHOTORECEPTORS
 PINEAL GLAND
 PITUITARY GLAND
 PROPRIOCEPTORS
 PROSTATE GLAND
 PUPILS
 RESPIRATORY SYSTEM
 RETINA
 SALIVARY GLANDS
 SCAPULA
 SCIATIC REGION
 SEBACEOUS GLANDS
 SEMICIRCULAR CANALS
 SENSE ORGANS

Ejemplo de una página del Tesoro de la NASA mostrando una jerarquía completa (bajo anatomía) sin discriminación de niveles.

TRANSPARENCIA II- 17

Ej. de una página del Water Resources Thesaurus

PROPERTIES
 SOIL PROPERTIES
 OIL TYPES
 BT SOILS
 NT ACID SOIL
 ACID SOILS
 - LATERITES
 - PODZOLS
 --- GRAY-BROWN PRODZOLIC SOILS
 --- RED PRODZOLIC SOILS
 --- YELLOW PRODZOLIC SOILS
 AEOLIAN SOILS
 - LOESS
 ALKALINE SOILS
 - ALKALI SOILS
 - CALCAREOUS SOILS
 - PRAIRIE SOILS
 AQUATIC SOILS
 AZONAL SOILS
 CALICHE
 CLAY LOAM
 COHESION LESS SOILS
 COHESIVE SOILS
 COMPACTED SOILS
 EXPANSIVE SOILS
 - EXPANSIVE CLAYS
 --- BENTONITE
 FOREST SOILS
 FROZEN SOILS
 - PERMAFROST
 GLACIAL SOILS
 GUMBO
 HYDROSOLS (SOILS)
 IMPERVIOUS SOILS
 - EXPANSIVE CLAYS
 --- BENTONITE
 - HARDPAN
 - QUICK CLAYS
 KADLINITE
 LAKE SEDIMENT
 LOAM
 ORGANIC SOILS
 - HUCK SOILS
 PALEOSOLS
 PEAT
 PEDALFERS
 - PODZOLS
 --- GRAY-BROWN PRODZOLIC SOILS

--- RED PRODZOLIC SOILS
 --- YELLOW PRODZOLIC SOILS
 - PRAIRIE SOILS
 PEDOCALS
 - BROWN SOILS
 - CHESTNUT SOILS
 - SIEROZEMS
 PERVIOUS SOILS
 - SILTS
 RESIDUAL SOILS
 - SAPROLITES
 SALINE SOILS
 SAND-CLAY
 SANDY LOAM
 SATURATED SOILS
 SUB SOIL
 SOIL WATER
 BT MOISTURE CONTENT
 VADOSE WATER
 WATER TYPES
 NT AVAILABLE WATER
 EXCESS WATER
 SOIL MOISTURE
 - CAPILLARY WATER
 --- FRINGE WATER
 - HYGROSCOPIC WATER
 SOILING
 BT MATERIALS DETERIORATION
 SOILING INDEX
 GENERALLY REPORTED AS "RUDS" OR "COMS"
 PER 100 LINEAR FEET OF AIR. UNIT FOR THE
 QUANTITY OF SOLIDS DEPOSITED ON A TEST
 FILTER: "RUD" WHEN MEASURED BY LIGHT
 REFLECTANCE: "COM" WHEN MEASURED BY
 LIGHT TRANSMISSION
 BT AIR QUALITY MEASUREMENTS
 MEASUREMENT
 NT RUD
 RT COM
 SOILS
 RT ALLUVIUM
 SOIL TYPES
 - ACID SOIL
 - ACID SOILS
 --- LATERITES
 --- PODZOLS
 --- GRAY-BROWN PRODZOLIC SOILS

TRANSPARENCIA II- 18

COMPLEX	POLLUTANTS
BT	CONTAMINANTS
	ENVIRONMENTAL POLLUTANTS
NT	ACETIC ACID
	ACIDS
	BENZOIC ACID
	DYES
	FORMIC ACID
	HYDROCHLORIC ACID
	HYDROCYANIC ACID
	HYDROFLUORIC ACID
	INORGANIC ACIDS
	NITRIC ACID
	NITROUS ACID
	ORGANIC ACIDS
	PERCHLORIC ACID
	PERSISTENT HERBICIDES
	PERSISTENT INSECTICIDES
	PHOSPHORIC ACID
	PTHALIC ACID
	SALTS
	SULFONIC ACID
	SULFURIC ACID
RT	POLLUTANTS

ASIENTO ESTANDAR DE TESAURO

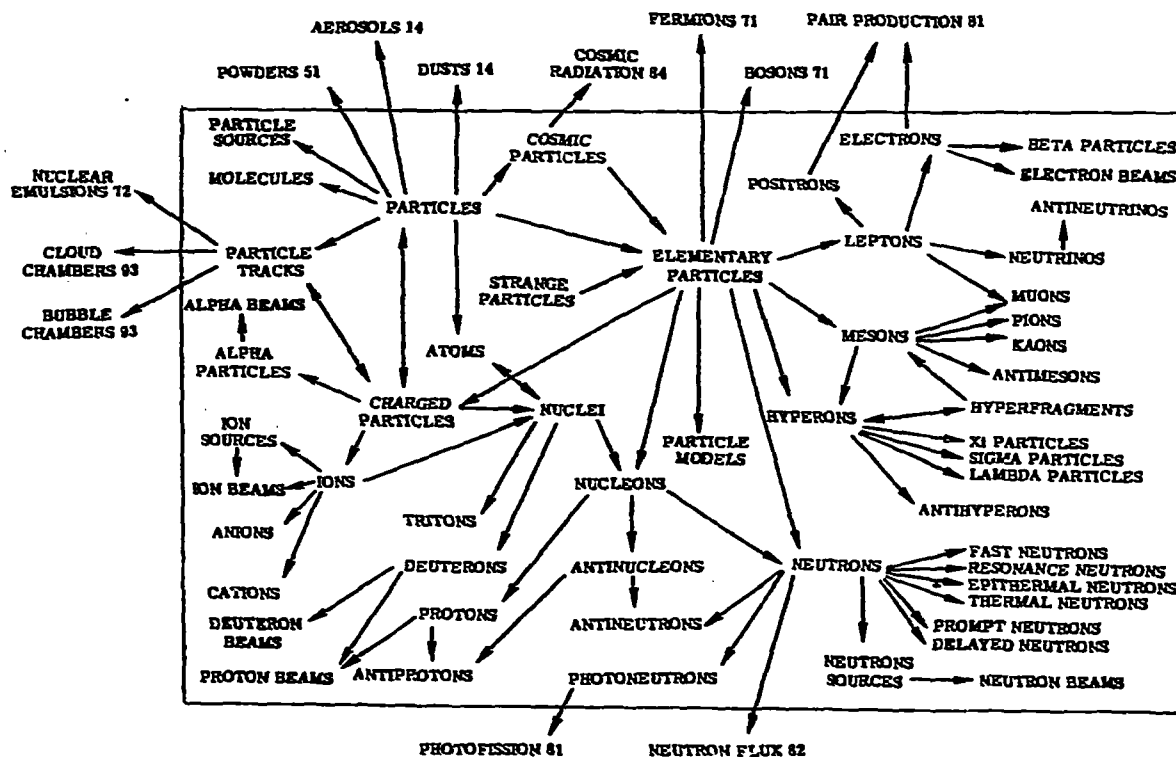
COMPLEX	POLLUTANTS
BT	CONTAMINANTS
	ENVIRONMENTAL POLLUTANTS
NT	ACIDS
	• INORGANIC ACIDS
	• • HYDROCHLORIC ACID
	• • HYDROFLUORIC ACID
	• • NITRIC ACID
	• • NITROUS ACID
	• • PERCHLORIC ACID
	• • PHOSPHORIC ACID
	• • SULFURIC ACID
	• ORGANIC ACIDS
	• • ACETIC ACID
	• • BENZOIC ACID
	• • FORMIC ACID
	• • HYDROCUANIC ACID
	• • PTHALIC ACID
	• • SULFONIC ACID
	DYES
	PERSISTENT HERBICIDES
	PERSISTENT INSECTICIDES
	SALTS
RT	POLLUTANTS

ASIENTO DE TESAURO CON SANGRIAS PARA INDICAR LOS NIVELES JERARQUICOS

PARTE SISTEMATICA**FORMATO CORRIENTE PARA LA PRESENTACION DE LAS CATEGORIAS TEMATICAS**

CHILD ABUSE AND NEGLECT
ACCIDENT PREVENTION
ADOLESCENT ABUSE
CHILD ABUSE
CHILD AT RISK
CHILD NEGLECT
CHILD PROTECTION INVESTIGATIONS
CHILD SAFETY
CHILD WELFARE
COURT REFERRAL
EMERGENCY INVESTIGATIONS
EMOTIONAL ABUSE
EMOTIONAL TREATMENT
ETIOLOGY
FAMILIES AT RISK
HOME ASSESSMENT
IDENTIFICATION
INDEPENDENT INVESTIGATIONS
INCIDENCE
INCIDENT
INCIDENT LOCATION
INCIDENT RECIDIVISM
INSTITUTIONAL CHILD ABUSE AND NEGLECT
INSTITUTIONAL INVESTIGATIONS
INTAKE
INTERVIEWS
INVESTIGATION CONSTRAINTS
INVESTIGATION PROCEDURES
INVESTIGATIONS
LAW ENFORCEMENT INVESTIGATIONS
PHYSICAL ABUSE
PREDICTION
PRELIMINARY INVESTIGATIONS
PREVENTION
REFERRAL
COURT REFERRAL
SELF REFERRAL
SEQUELAE
SEXUAL ABUSE
SITE INSPECTIONS
TREATMENT
TREATMENT CONTINUITY
TREATMENT DURATION
TREATMENT EFFECTIVENESS
TREATMENT GOALS
TREATMENT PLANS
TREATMENT REVIEW
TREATMENT TERMINATION

PARTE SISTEMATICA (CON SANGRIAS)**FORMATO PROPUESTO PARA LA PRESENTACION DE LAS CATEGORIAS TEMATICAS****CHILD ABUSE AND NEGLECT****CHILD ABUSE****ADOLESCENT ABUSE****EMOTIONAL ABUSE****SEXUAL ABUSE****CHILD NEGLECT****CHILD SAFETY****CHILD WELFARE****EMOTIONAL TREATMENT****ETIOLOGY****IDENTIFICATION****INCIDENCE****INCIDENT****INCIDENT LOCATION****INCIDENT RECIDIVISM****INSTITUTIONAL CHILD ABUSE AND NEGLECT****INTAKE****INVESTIGATIONS****CHILD PROTECTION INVESTIGATIONS****EMERGENCY INVESTIGATIONS****INDEPENDENT INVESTIGATIONS****INSTITUTIONAL INVESTIGATIONS****INVESTIGATION CONSTRAINTS****INVESTIGATION PROCEDURES****HOME ASSESSMENT****INTERVIEWS****SITE INSPECTIONS****LAW ENFORCEMENT INVESTIGATIONS****PRELIMINARY INVESTIGATIONS****PREVENTION****ACCIDENT PREVENTION****CHILD AT RISK****FAMILIES AT RISK****PREDICTION****REFERRAL****COURT REFERRAL****SELF REFERRAL****SEQUELAE****TREATMENT****TREATMENT CONTINUITY****TREATMENT DURATION****TREATMENT EFFECTIVENESS****TREATMENT GOALS****TREATMENT PLANS****TREATMENT REVIEW****TREATMENT TERMINATION**



ALPHA BEAMS
ALPHA PARTICLES
ANIONS
ANTHYPERONS
ANTIMESONS
ANTINEUTRONS
ANTINEUTRONS
ANTINUCLEONS
ANTIPROTONS
ATOMS
BETA PARTICLES
CATIONS
CHARGED PARTICLES
COSMIC PARTICLES
DELAYED NEUTRONS
DEUTERON BEAMS
DEUTERONS
ELECTRON BEAMS
ELECTRONS
ELEMENTARY PARTICLES
EPITHERMAL NEUTRONS
FAST NEUTRONS
HYPERFRAGMENTS
HYPERONS
ION BEAMS
IONS
ION SOURCES
KAONS
LAMBDA PARTICLES
LEPTONS
MESONS
MOLECULES
MUONS
NEUTRINOS
NEUTRON BEAMS
NEUTRONS
NEUTRON SOURCES
NUCLEI
NUCLEONS
PARTICLE MODELS
PARTICLE SOURCES
PARTICLE TRACKS
PARTICLES
PHOTONEUTRONS
PIONS
POSITRONS
PROMPT NEUTRONS
PROTON BEAMS
PROTONS
RESONANCE NEUTRONS
STRANGE PARTICLES
SIGMA PARTICLES
THERMAL NEUTRONS
TRITONS
XI PARTICLES

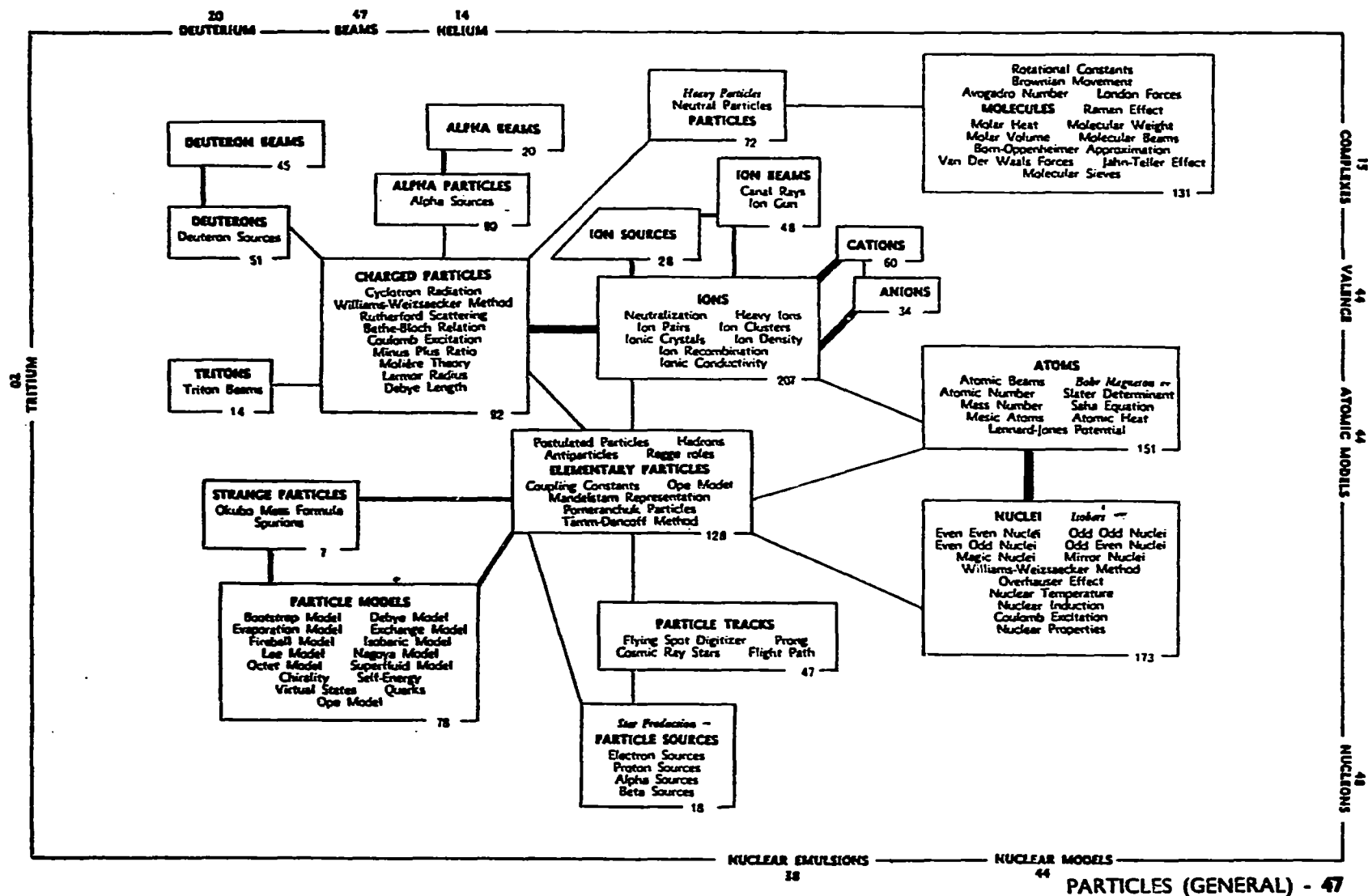
EURATOM/CID
KEYWORD-THESAURUS

ELEMENTARY PARTICLES

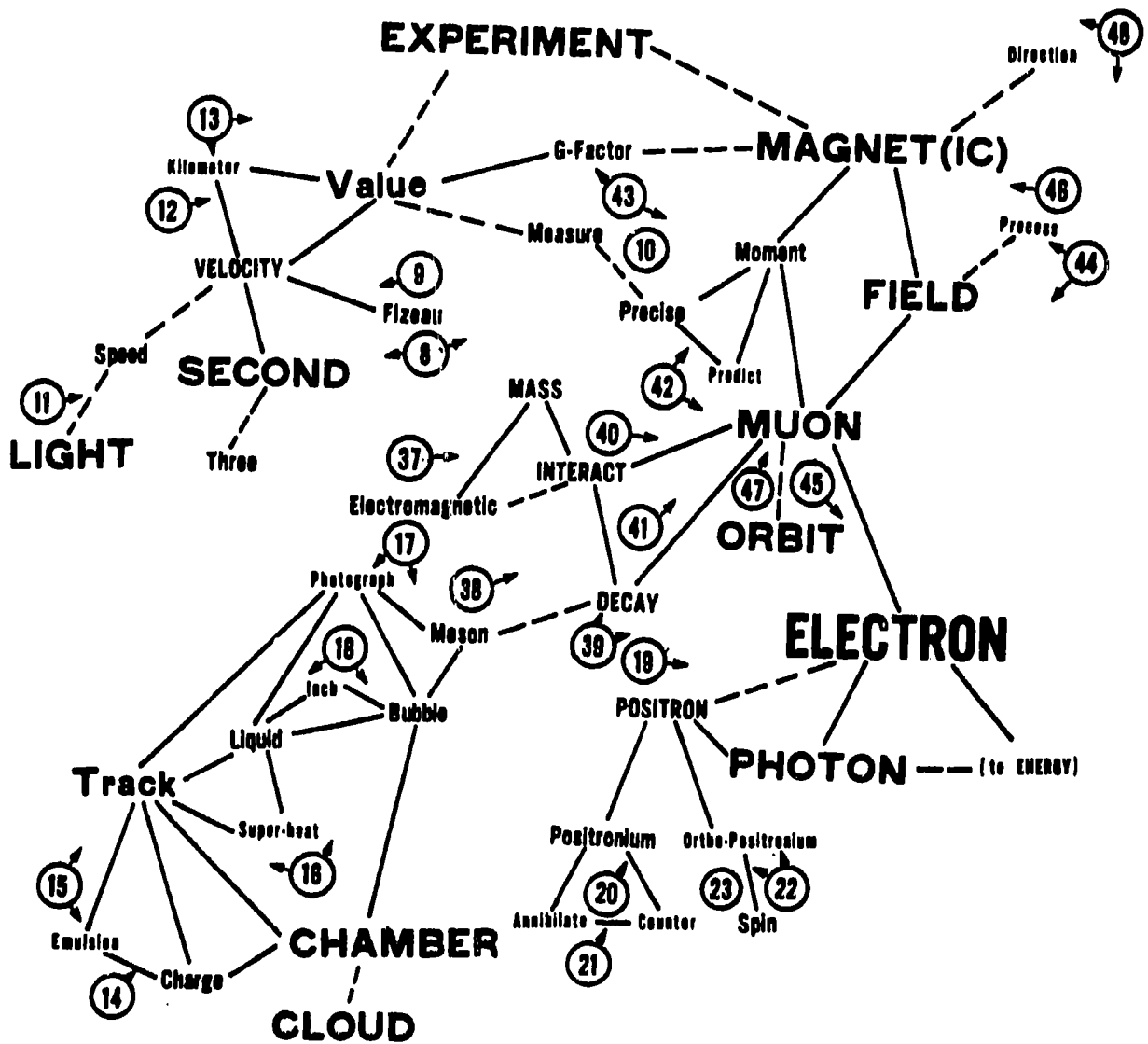
GROUP 85

00	11	12	13	14	15	16	21	22	23	24	31	32	33	34	35	36	37	41	42	43	44
45	51	52	61	62	63	64	71	72	73	74	75	76	81	82	83	84	85	91	92	93	94

Ejemplo de una página del EURATOM THESAURUS, primera edición.
Notar el uso de los gráficos flechados.



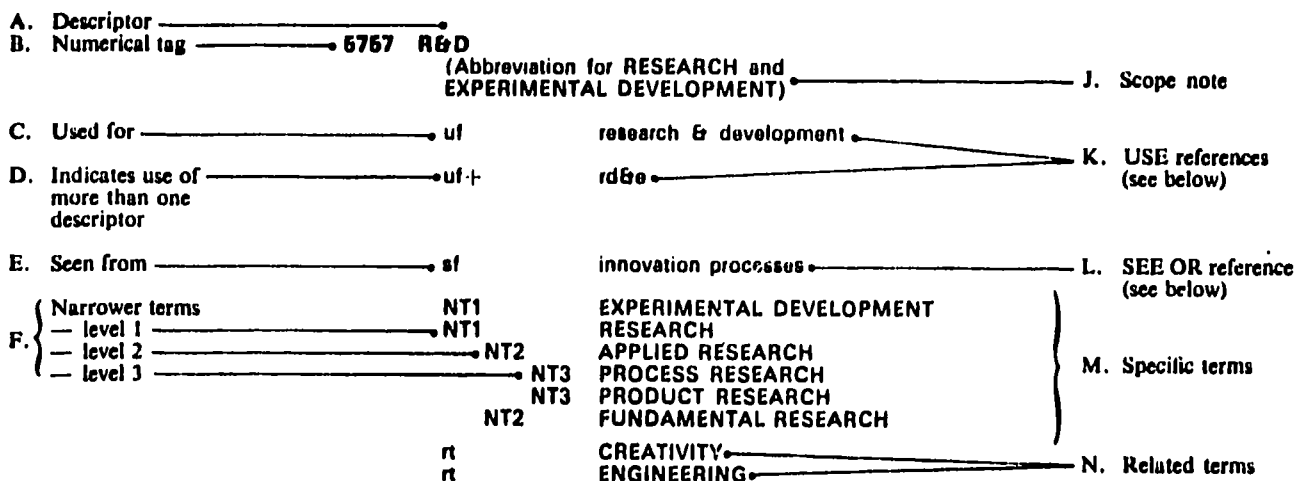
Ejemplo de una página del EURATOM THESAURUS, segunda edición.
Notar las modificaciones en relación con la primera edición.



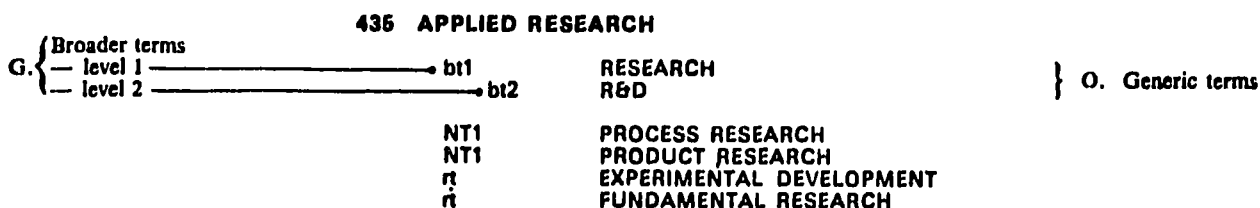
Ejemplo de un "mapa de ruta semántico"

Tesoro SPINES. Ejemplo de asientos y notaciones usados en la lista alfabética estructurada.

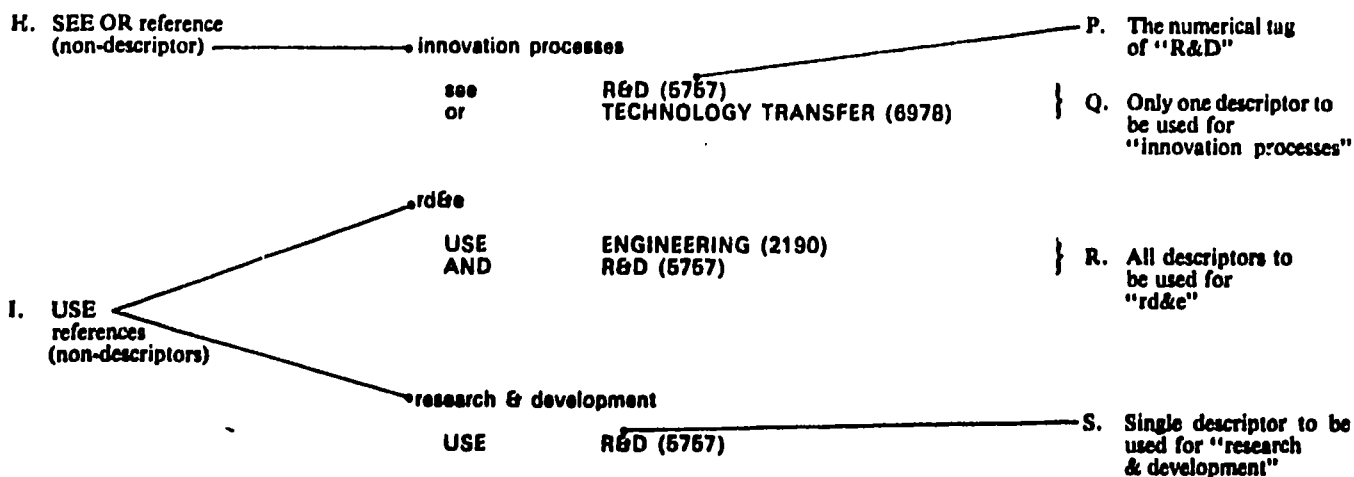
Asiento del descriptor principal



Asiento del descriptor intermedio



Asientos de los no-descriptores



Este ejemplo ficticio muestra todas las situaciones; la correspondiente representación gráfica se encuentra en la Sección C.1.2

SPINES THESAURUS

10 EDUCATION

11 DEMOGRAPHY

14 PHYSICAL ENVIRONMENT

20 MATHEMATICS

19 STATISTICAL ANALYSIS

08 FINANCE

09 PERSONNEL

21 PHYSICS

26 ASTRONOMY

22 CHEMISTRY 1

24 BIOLOGY

23 CHEMISTRY 2

25 MEDICINE

06 PLANNING

04 SOCIOLOGY & PSYCHOLOGY

05 ECONOMICS

07 DECISION MAKING

03 KNOWLEDGE & PHILOSOPHY

15 R&D

13 AGRICULTURE

12 PRODUCTION & DISTRIBUTION

27 TECHNOLOGY 1

28 TECHNOLOGY 2

29 TECHNOLOGY 3

01 POLITICAL SCIENCES

02 JURIDICAL SCIENCES

33 INTERNATIONAL ORGANIZATIONS

16 TECHNOLOGY TRANSFER & INTERNATIONAL RELATIONS

17 INFORMATION

18 DATA PROCESSING

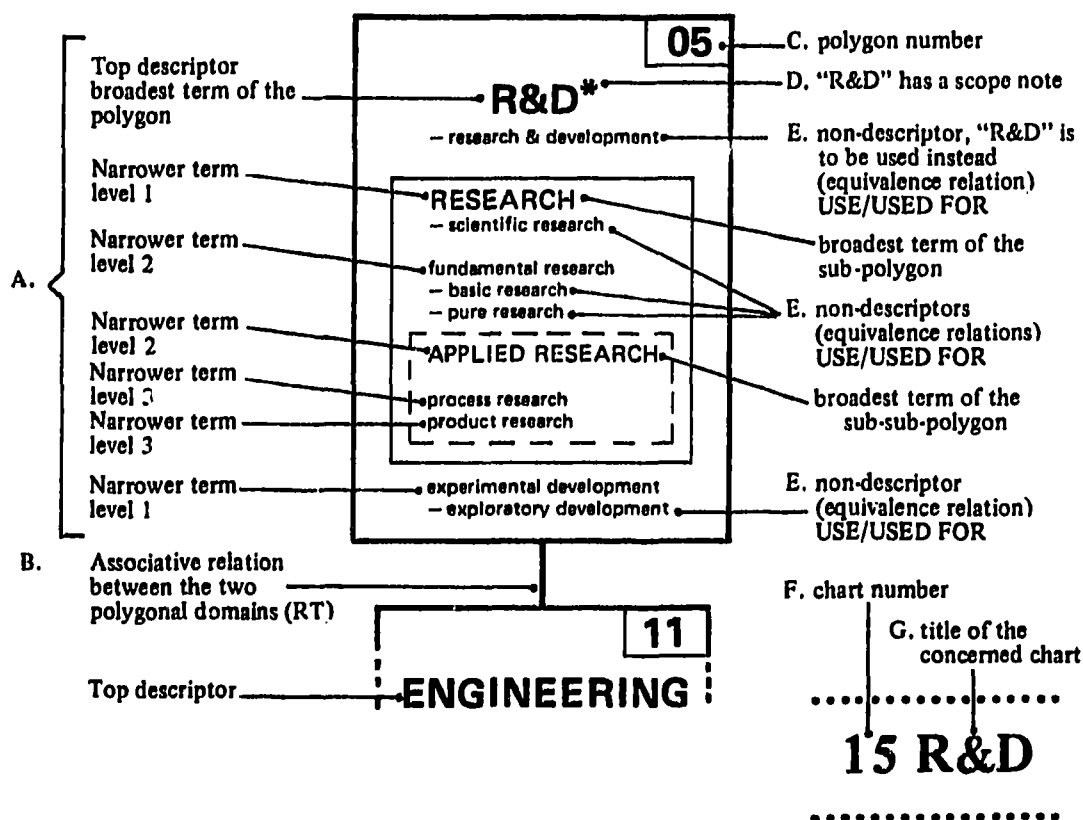
30 HISTORY, TIME & SPACE

31 GEOGRAPHY

32 COUNTRIES

34 WAR & PEACE

GLOBAL GRAPH OF GRAPHIC DISPLAYS

Example² of a polygon and notations

H

ANCILLARY PLANT AND COMPONENTS

ELECTRIC ANCILLARY PLANT AND COMPONENTS (cont'd.)

Hh Switches, Switchgear

By method of connection:

Hhb	Interrupters
Hhc	Contactors
Hhd	Relays
Hhdd	Polarised, hinged, armature
Hhdg	Differential
Hhdi	Slugged
Hhdp	Line
Hhdy	Buchholtz
Hhe	Isolation switches
Hhf	T.R. Switches
Hhg	Selector switches
Hhgb	2-way reversing
Hhgc	Multi-way
Hhgd	Uniselectors, step switches
Hhgm	Combination switches

By operating mechanism:

Hhh	Beam switches
Hhi	Key switches
Hhj	Lever switches, knife switches
Hhl	Drum switches
Hhm	Push button switches
Hhn	Toggle action switches
Hho	Sliding switches
Hhp	Rolling switches
Hhq	Butt switches

Hhr	Circuit Breakers (See F schedule for special components)
Hhrb	Air break (OB, OBX)
Hhrd	Air blast
Hhrg	Oil break (OKF, OKM, OLK, OQF)
Hhrge	S.O.V., L.O.V., low oil breakers
Hhrh	Gas filled (SF6)
Hhrp	Current limiters
Hhrt	Fault make load break oil switches
Hhrv	Fused oil switches
Hhrw	Vacuum circuit breakers

Ejemplo de una página del "Faceted Subject Classification for Engineering" mostrando una clara presentación de facetas y jerarquías

TESAURO CON ASIENTOS COMPLETOS EN LA PARTE JERARQUICA

31.01 WAGE THEORY

EQUITY THEORY

RT: WAGE DIFFERENTIAL 01401

MARGINAL PRODUCTIVITY THEORY

SN: REFERS TO THE THEORY THAT WAGE RATES TEND TO EQUAL
THE VALUE OF THE MARGINAL PRODUCT OF LABOR

RT: PRODUCTIVITY 01402

PHILLIPS CURVE

SN: AN ECONOMETRIC MODEL WHICH EXPRESSES A RELATIONSHIP
BETWEEN INFLATION AND UNEMPLOYMENTRT: INFLATION
UNEMPLOYMENT 01406

SUBSISTENCE THEORY

SN: REFERS TO A THEORY OF WAGE DETERMINATION ACCORDING TO WHICH
THE PRICE OF LABOR EQUALS THE AMOUNT REQUIRED FOR THE
WORKER TO SUBSIST. SHOULD THE PRICE RISE HIGHER THAN THIS
LEVEL, THE SUBSEQUENT INCREASE IN THE SUPPLY OF LABOR WOULD
DRIVE IT DOWN AGAIN 01403PARTE ALFABETICA

WAGE (CONTD.)

WAGE AND SALARY ADMINISTRATIONS = USE: COMPENSATION
ADMINISTRATION 47.01

WAGE BOARD 22.1

WAGE BOARD EMPLOYEE 44.13

WAGE BONUS 47.13

WAGE CHANGE 31.0

WAGE COMPARABILITY = USE: COMPARABLE WAGE 31.1

WAGE CONTROL = USE: WAGE STABILIZATION 22.1

WAGE DECREASE = USE: WAGE ADJUSTMENT 47.1

WAGE FREEZE = USE: WAGE STABILIZATION 22.1

WAGE GARNISHMENT 47.04

WAGE INCENTIVE 47.02

WAGE INCREASE = USE: WAGE ADJUSTMENT 47.1

WAGE MOVEMENTS 31.1

WAGE PARITY 31.1

WAGE PRICE GUIDELINES 22.1

WAGE RATE 31.1

WAGE RESTITUTION 47.12

WAGE STABILIZATION 22.1

WAGE STRUCTURE 31.1

WAGE SUPPLEMENTS = USE: FRINGE BENEFIT 47.3

WAGE SURVEY 31.1

WAGE THEORY 31.01

WAGES

WAGES FUND THEORY 31.01

WAITER

WAITER 42.42

PARTE SISTEMATICA

06.02

POLITICA EDUCACIONAL

06.02

ACCESO A LA EDUCACION / ACCESS TO EDUCATION

RT Oportunidades de educación

COEDUCACION / COEDUCATION

BT Educación

CONDICIONES DE ADMISION / ADMISSION REQUIREMENTS

Condiciones (exámenes, certificados, prueba de calificación, etc.), de entrada a un curso de estudio, de estudio complementario, de capacitación, etcétera.

RT Calificación
Examen

DURACION DE FORMACION / TRAINING DURATION

RT Capacitación

ECONOMIA DE LA EDUCACION / ECONOMICS OF EDUCATION

BT Ciencias económicas

RT Análisis de costos y beneficios
Educación

PARTE ALFABETICA

ACCESO

Acceso a la educación 06.02

Acceso a la información 19.01.1

ACCIDENTE

Accidente de trabajo 13.04

ACCIDENTES

Prevención de accidentes use Seguridad del trabajo 13.04

Seguro de accidentes 11.02.3

ACELERADA

Formación profesional acelerada 06.09.1

ACELERADOS

Cursos acelerados 06.09.2

ACTIVA

Población económicamente activa 13.01.2

ACTIVIDADES

Informe de actividades 19.02.8

ACTIVO

Método activo 06.08.1

ACTUALIZACION

Actualización 06.09.1

ACUERDO

Acuerdo internacional 01.02.4

ADAPTABILIDAD

Adaptabilidad 05.03.3

ADMISION

Condiciones de admisión 06.02

INDICES**GEOLOGY**

- **STRUCTURAL GEOLOGY**
- • **DIASTROPHISM**
- • • **DEFORMATION**
- • • • **BASIN DEFORMATION**
- • • • **ROCK DEFORMATION**
- • • • • **DIAPIRISM**
- • • • • **DILATION**
- • • • • **FAULTING**
- • • • • • **BLOCK FAULTING**
- • • • • • **COMPLEX FAULTING**
- • • • • • **NORMAL FAULTING**
- • • • • • **REVERSE FAULTING**
- • • • • • • **THRUST FAULTING**
- • • • • • • **OVERTHRUST FAULTING**
- • • • • • **WRENCH FAULTING**
- • • • • **FOLDING**
- • • • • **FRACTURING**
- • • • • **GRAVITY SLIDING**
- • • • • **SALT FLOW**
- • • • • **SHRINKAGE**
- • • • • **CRUSTAL SHRINKAGE**
- • • • **STRAIN**
- • • • **SUBSIDENCE**
- • • • **UPLIFTING**
- • • **FORCE (GEOLOGY)**

INDICE JERARQUICO

PROPERTY**GOVERNMENT PROPERTY****REAL PROPERTY****REAL PROPERTY TAX****PROPIONALDEHYDES****PROPOSALS****PROPULSION****ELECTRIC PROPULSION****NUCLEAR PROPULSION****PROPYNES****PROTECTION****ANIMAL PROTECTION****CONSUMER PROTECTION****ENVIRONMENTAL PROTECTION****ENVIRONMENTAL PROTECTION AGENCY****FLOOD PROTECTION****LANDSCAPE PROTECTION****PROTECTION OF ANIMALS****PROTECTION OF NATURE****REGULATIONS (ENVIRONMENTAL PROTECTION)****WIND PROTECTION****PROTECTIVE****PROTECTIVE CLOTHING****PROTECTIVE MASKS****PRESENTACION EN FORMA PERMUTADA (KWOC)**

ASIENTOS DE TESAURO GENERADOS POR AVOCON

ANALYSIS (MATHEMATICS)
 REAL VARIABLES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 BT REAL VARIABLES
 SERIES (MATHEMATICS)
 BT CALCULUS
 CONVERGENT SERIES
 BT SERIES (MATHEMATICS)
 DIVERGENT SERIES
 BT SERIES (MATHEMATICS)
 ASYMPTOTIC SERIES
 BT DIVERGENT SERIES
 POWER SERIES
 BT SERIES (MATHEMATICS)
 TAYLORS SERIES
 BT POWER SERIES
 MACLAURIN SERIES
 BT TAYLORS SERIES

Entrada Inicial Fig. 2.2.A

ANALYSIS (MATHEMATICS)
 NT ASYMPTOTIC SERIES
 CALCULUS
 CONVERGENT SERIES
 DIVERGENT SERIES
 MACLAURIN SERIES
 POWER SERIES
 REAL VARIABLES
 SERIES (MATHEMATICS)
 TAYLORS SERIES
 REAL VARIABLES
 BT ANALYSIS (MATHEMATICS)
 NT ASYMPTOTIC SERIES
 CALCULUS
 CONVERGENT SERIES
 DIVERGENT SERIES
 MACLAURIN SERIES
 POWER SERIES
 SERIES (MATHEMATICS)
 TAYLORS SERIES
 CALCULUS
 BT ANALYSIS (MATHEMATICS)
 REAL VARIABLES
 NT ASYMPTOTIC SERIES
 CONVERGENT SERIES
 DIVERGENT SERIES
 MACLAURIN SERIES
 POWER SERIES
 SERIES (MATHEMATICS)
 TAYLORS SERIES
 SERIES (MATHEMATICS)
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 REAL VARIABLES
 NT ASYMPTOTIC SERIES
 CONVERGENT SERIES
 DIVERGENT SERIES
 MACLAURIN SERIES
 POWER SERIES
 TAYLORS SERIES

CONVERGENT SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 SERIES (MATHEMATICS)
 REAL VARIABLES
 DIVERGENT SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 REAL VARIABLES
 SERIES (MATHEMATICS)
 NT ASYMPTOTIC SERIES
 ASYMPTOTIC SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 DIVERGENT SERIES
 SERIES (MATHEMATICS)
 REAL VARIABLES
 POWER SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 REAL VARIABLES
 SERIES (MATHEMATICS)
 NT MACLAURIN SERIES
 TAYLORS SERIES
 TAYLORS SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 POWER SERIES
 REAL VARIABLES
 SERIES (MATHEMATICS)
 NT MACLAURIN SERIES
 BT ANALYSIS (MATHEMATICS)
 CALCULUS
 POWER SERIES
 REAL VARIABLES
 SERIES (MATHEMATICS)
 TAYLORS SERIES

Entrada del tesauro generado por AVOCON
Fig. 2.2B.

ANALYSIS (MATHEMATICS)
 • • REAL VARIABLES
 • • • CALCULUS
 • • • • SERIES (MATHEMATICS)
 • • • • • CONVERGENT SERIES
 • • • • • DIVERGENT SERIES
 • • • • • ASYMPTOTIC SERIES
 • • • • • POWER SERIES
 • • • • • TAYLORS SERIES
 • • • • • MACLAURIN SERIES

Indice jerárquico de un asiento del tesauro
Fig. 2.2C.

PRESENTACION DEL ASIENTO DE UN SISTEMA EN LINEA

MT		SOLAR CONCENTRATOR
SN		A DEVICE USING LENSES AND REFLECTING SURFACES TO CONCENTRATE SUNLIGHT
UF		SOLAR ENERGY CONCENTRATOR
BT	(1)	CONCENTRATORS
NT	(1)	CONE AND COLUMN CONCENTRATOR
	(2)	DOUBLE MIRROR SOLAR CONCENTRATOR
	(3)	FACETED CONCENTRATOR
	(4)	FIXED-MIRROR CONCENTRATOR
	(5)	LINEAR CONCENTRATOR
	(6)	PARABOLIC CONCENTRATOR
	(7)	PARABOLIC CYLINDER CONCENTRATOR
	(8)	PARABOLIC THROUGH CONCENTRATOR
	(9)	REFLECTING CONE CONCENTRATOR
	(10)	SOLAR FURNACE
RT	(1)	CONCENTRATOR-ABSORBER SYSTEM
	(2)	EXPANDABLE WHIRLING MEMBRANE
	(3)	PARABOLOIDAL WHIRLING MEMBRANE

PARTES CONSTITUTIVAS DE UN TESAURO**PRESENTACION**

PORTADA

SUMARIO

INTRODUCCION

CUERPO DE LA OBRA: parte alfabética (con todas las entradas del tesauo)

parte sistemática (en las diferentes formas en que puede estructurarse)

PARTES AUXILIARES: índices: KWOC
KWIC
GEOGRAFICO
DE SIGLAS
DE FORMULAS
DE EQUIVALENCIAS A OTROS IDIOMAS
ETC.

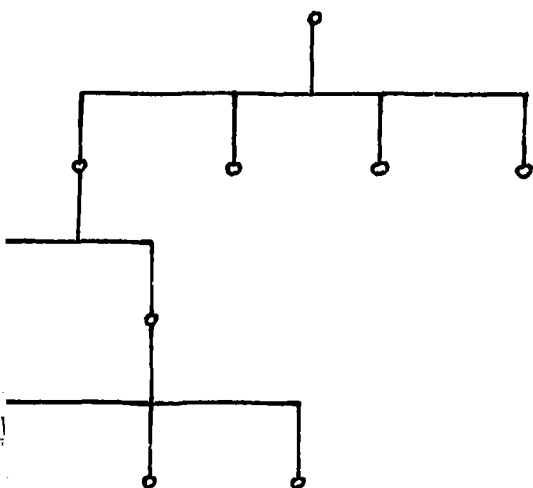
TRANSPARENCIA III- 1

1. ADOPTAR
 2. ADAPTAR
 3. TRADUCIR
 4. HACER UN MICROTESAURO COMPATIBLE
 5. ADOPTAR PARTES (PERIFERIA)
 6. CONSTRUIR UN NUEVO TESAURO DESDE EL COMIENZO
-

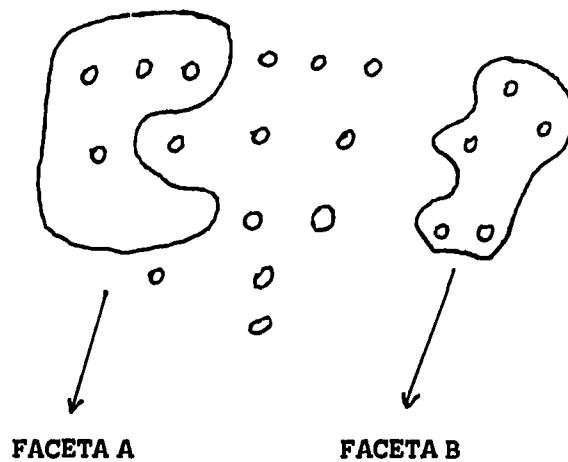
TRANSPARENCIA III- 2

FORMAS DE CONSTRUCCION

DE ARRIBA HACIA ABAJO



DE ABAJO HACIA ARRIBA



TRANSPARENCIA III- 3

JUSTIFICACION LITERARIA

JUSTIFICACION BIBLIOGRAFICA

JUSTIFICACION DEL USUARIO

TRANSPARENCIA III- 4

JUSTIFICACION LITERARIA

DICCIONARIOS

GLOSARIOS

ENCICLOPEDIAS

MANUALES TECNICOS

INDICES

TRATADOS

REVISTAS DE RESUMENES

LISTADOS DE COMPUTADORA (CORRESPONDIENTES A BASES DE DATOS) POR FRECUENCIA DE APARICION

-----	1172
-----	1107
-----	1094
-----	1068

POR FRECUENCIA DE CONCURRENCIA (APARICION CONJUNTA)

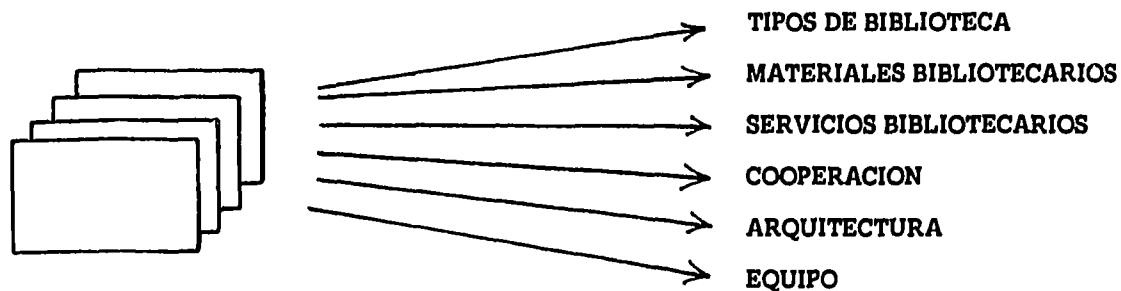
A	-----	Q
	-----	S
	-----	B

JUSTIFICACION DEL USUARIO**TRANSPARENCIA III- 5****1. PEDIDOS DE BUSQUEDAS****2. CUESTIONARIOS****PERFILES DE INTERES****DESCRIPCION DE TAREAS****TITULOS DE PUBLICACIONES****RESUMENES DE PUBLICACIONES****3. METODOS DE DYM Y DE PICKFORD****LOS USUARIOS SELECCIONAN LOS TERMINOS EN LA LITERATURA**/

**JUSTIFICACION
DEL USUARIO**|

**JUSTIFICACION
LITERARIA****TRANSPARENCIA III- 6****REGISTRO**

TERMINO
TG?
TE?
SINONIMOS?
NOTA DE ALCANCE / DEFINICION?
FUENTE?

ORGANIZACIONANALISIS FACETAL

TRANSPARENCIA III- 8

KAISER	RANGANATHAN	VICKERY	CAMPBELL
COSAS CONCRETAS	PERSONALIDAD	SUSTANCIA	COSAS
PROCESOS	MATERIA	ORGANO	PARTES DE COSAS
	ENERGIA	CONSTITUYENTE	SUSTANCIAS
	ESPACIO	ESTRUCTURA	OPERACIONES
	TIEMPO	FORMA	PROCESOS, PROPIEDADES Y FENOMENOS
		PROPIEDAD	ESPACIO
		OBJETO DE ACCION	TIEMPO
		ACCION	SUBDIVISIONES COMUNES
		OPERACION	
		PROCESO	
		AGENTE	
		ESPACIO	
		TIEMPO	

TRANSPARENCIA III- 9

SUBFACETAS Y JERARQUIAS**MATERIALES BIBLIOTECARIOS
(POR FORMA DE PUBLICACION)****LIBROS****PUBLICACIONES PERIODICAS****REVISTAS****ACTAS****FOLLETOS****MICROFORMAS****MICROTARJETAS****MICROFICHAS****ULTRAMICROFICHAS****MICROFILMES****(POR CONTENIDO)****LITERATURA****POESIA****DRAMA****ENSAYO****NOVELA****BIOGRAFIA****AUTOBIOGRAFIA****HISTORIA**

TRANSPARENCIA III- 10

FACETAS DE LONGITUD VARIABLE**MATERIALES****SERVICIOS****ADMINISTRACION****EVALUACION**

ESPECIFICIDADRELACIONADA CON LAS NECESIDADES DE LOS USUARIOS

PERROS ZORREROS IRLANDESES

PERROS ZORREROS

PERROS

UN ANALISIS MATEMATICO DEL FLUJO DE LA SANGRE EN LA ARTERIA CAROTIDA

MEDICOS

ARTERIA CAROTIDA

HEMODINAMICA

ANALISIS MATEMATICO

MATEMATICOS

SERIES DE FOURIER

MEDICINA

PROBLEMAS

¿CUANTA PRE-COORDINACION?

JETS

MOTORES

RUIDO

MOTORES DE LOS JETS

RUIDO

RUIDO DE LOS

MOTORES DE LOS JETS

DE ABAJO HACIA ARRIBA (EMPIRICO)

1. COLECCIONAR TERMINOS
2. ANALISIS FACETAL
3. ORGANIZAR EN JERARQUIAS
4. DETERMINAR RELACIONES UTILES ENTRE LAS JERARQUIAS
5. CONVERTIR EN ORGANIZACION ALFABETICA

UN PASO HACIA ARRIBA, UN PASO HACIA ABAJO**BIBLIOTECAS****ACADEMICAS****ESTUDIANTILES****ESCOLARES****UNIVERSITARIAS****PUBLICAS****ESPECIALIZADAS****BIBLIOTECAS ACADEMICAS****TG****BIBLIOTECAS****TE****BIBLIOTECAS ESTUDIANTILES****BIBLIOTECAS UNIVERSITARIAS****BIBLIOTECAS ESTUDIANTILES****TG****BIBLIOTECAS ACADEMICAS****TE****BIBLIOTECAS ESCOLARES****BIBLIOTECAS ESCOLARES****TG****BIBLIOTECAS ESTUDIANTILES****BIBLIOTECAS****TE****BIBLIOTECAS ACADEMICAS****BIBLIOTECAS ESPECIALIZADAS****BIBLIOTECAS PUBLICAS**

TRANSPARENCIA III- 15

BIBLIOTECAS

TE BIBLIOTECAS ACADEMICAS
 BIBLIOTECAS ESCOLARES
 BIBLIOTECAS ESTUDIANTILES
 BIBLIOTECAS PUBLICAS
 BIBLIOTECAS UNIVERSITARIAS

TRANSPARENCIA III- 16

TG - TE

GENERO / ESPECIE

COSA / CLASE DE

- AUTOMOVILES ✓
TE CAMIONETAS

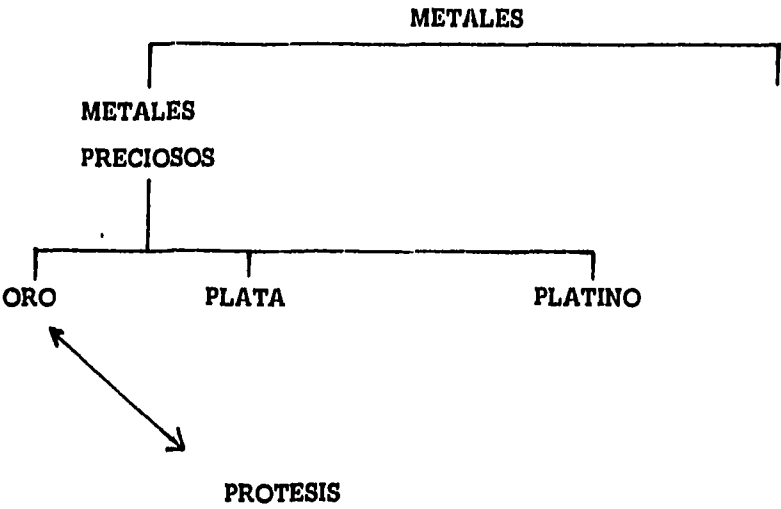
AUTOMOVILES X
TE VOLANTES

Algunos TODO / PARTE

- ARGENTINA
TE BUENOS AIRES

SISTEMA RESPIRATORIO
TE PULMONES

TRANSPARENCIA III - 17



TRANSPARENCIA III- 18

COSA / APLICACION

ABRASIVOS
PULVERIZACION

COLORO
CLORACION

ANTONIMIA

DUREZA
BLANDURA

RECIPROCIDAD

CONDUCTIVIDAD
RESISTIVIDAD

GENETICA

PADRE
HIJO

CAUSA / EFECTO

ENSEÑANZA
APRENDIZAJE

PERMEABILIDAD
INFILTRACION

TRANSPARENCIA III- 19

OPERACION / DISPOSITIVO (INSTRUMENTAL)

INTRUCCION PROGRAMADA

ESCRITURA

MAQUINA DE ENSEÑAR

LAPICES

ENTIDAD / CUALIDAD

NIÑOS HUERFANOS

LASER

ORFANDAD

COHESION

SUELOS

PROPIEDAD DE LOS SUELOS

COSA / PARTE

BICICLETAS

RUEDAS

COSA / PROCESO

BARCO

CONSTRUCCION DE BARCOS

COSA / PRODUCTO

BAUXITA

ALUMINIO

TRANSPARENCIA III - 20

USE

SINONIMOS, CASI-SINONIMOS, ANTONIMOS

AGUA SALADA

USE

AGUA SALINA

HERENCIA

USE

GENETICA

RUGOSIDAD

USE

LISURA

ESPECIFICO A GENERAL

DESTRUCTORES

USE

BUQUES DE GUERRA

CONCORDE

USE

AVION SUPERSONICO

TRANSPARENCIA III- 21

CATALOGACION COOPERATIVA

USECATALOGACION
COOPERACIONCATALOGACION
Y COOPERACIONCATALOGACION
más COOPERACION

CATALOGACION

UP CATALOGACION COOPERATIVA +

COOPERACION

UP CATALOGACION COOPERATIVA +

TRANSPARENCIA III- 24

ACTUALIZACION

1. CONTINUA
2. INDIZACION
3. BUSQUEDA
4. GRUPO DE REVISION DEL VOCABULARIO
5. PROVISIONALES
6. FRECUENCIA

INDIZACION

BUSQUEDA

1. Limitación positiva en el alcance del término:

NUEVA CAPACITACION

Capacitación para un cambio en la ocupación.

(Tal nota limita el alcance del término y lo distingue de otros términos relacionados tales como REHABILITACION)

2. Limitación negativa:

LICENCIA

Excluye títulos y certificados de estudios.

(Tal nota de alcance no sólo excluye sino que también puede dirigir al usuario a conceptos que él pudo pensar cuando buscó en el tesoro LICENCIA. Es preferible una nota de alcance que se refiera inequívocamente a otros descriptores, tal como

Excluye aspectos cubiertos por los descriptores TITULOS ACADÉMICOS, CERTIFICADOS DE ESTUDIO.

3. Definición verdadera:

AGUA DE GRAVEDAD

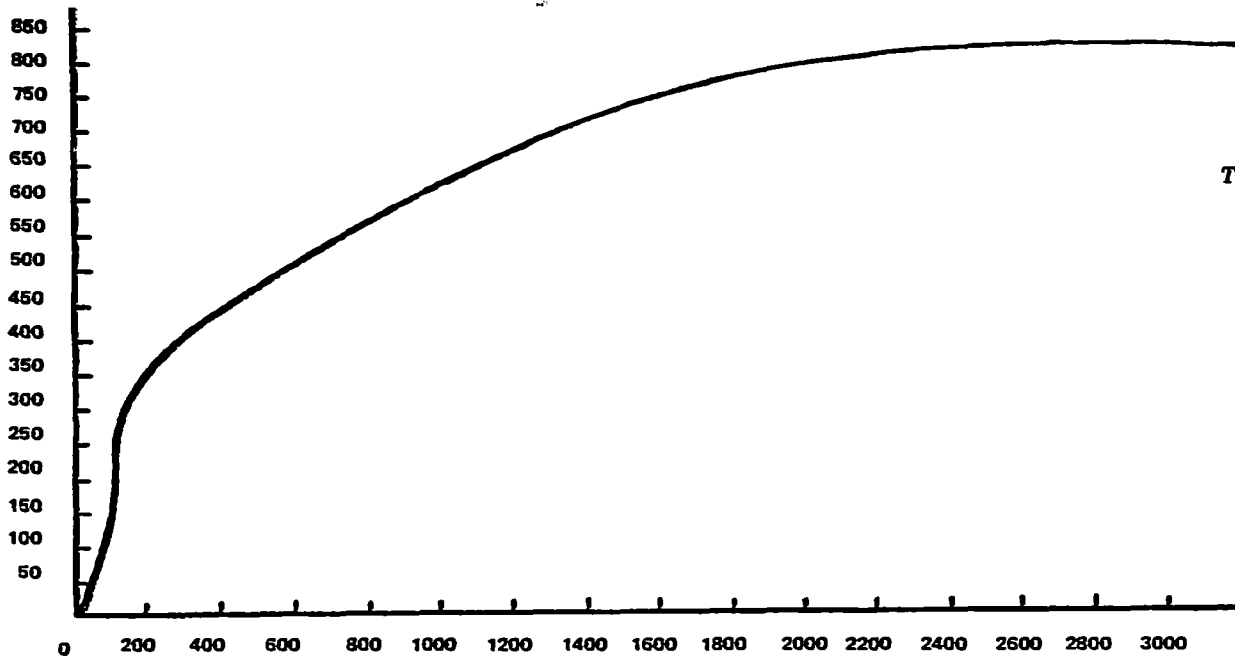
Agua en zonas no saturadas que se mueve libremente bajo la influencia de la fuerza de gravedad.

4. Combinación de definición y limitación positiva:

SUSPENSION LABORAL

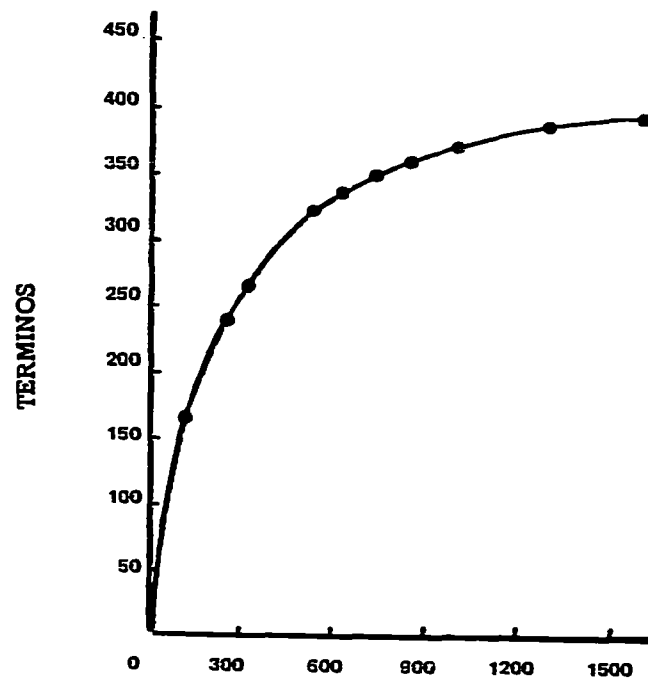
Suspensión de la oferta de trabajo por parte de un empleador con motivo de una penuria de materias primas o de energía, de una avería de la maquinaria, etc. El licenciamiento es normalmente de carácter temporal y sin perjuicio para el trabajador. No se debe confundir con despido.

CANTIDAD DE TERMINOS DIFERENTES
USADOS EN LA INDIZACION



TRANSPARENCIA III-23

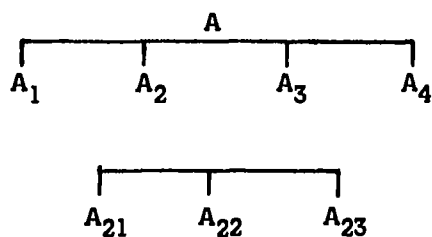
CANTIDAD DE DOCUMENTOS INDIZADOS



TRANSPARENCIA IV- 1

TESAURO EN MAQUINA

1. VERIFICAR LA CONSISTENCIA Y LA ACEPTABILIDAD DE LOS TERMINOS.
2. SUSTITUCION AUTOMATICA DE TERMINOS (ej. TERMINOLOGIA ANTIGUA)
3. ESTADISTICAS SOBRE EL USO
INDIZACION
BUSQUEDA
4. OPTIMIZACION AUTOMATICA DE LA ESTRATEGIA DE BUSQUEDA
5. GENERAR REFERENCIAS VEASE Y VEASE ADEMAS PARA LOS INDICES PUBLICADOS
6. PERMITIR UNA BUSQUEDA GENERICA (EXPLOSION)



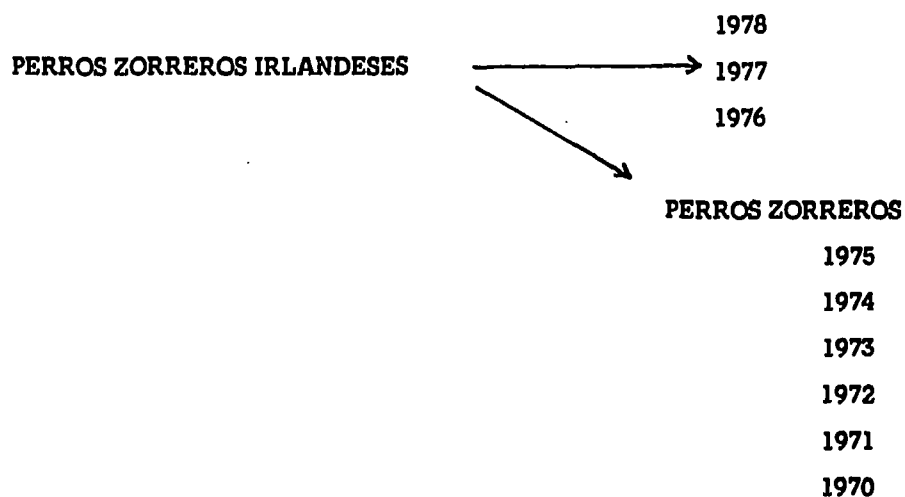
TRANSPARENCIA IV- 2

1970 - 1975

PERROS ZORREROS

1976 - 1978

PERROS ZORREROS IRLANDESES

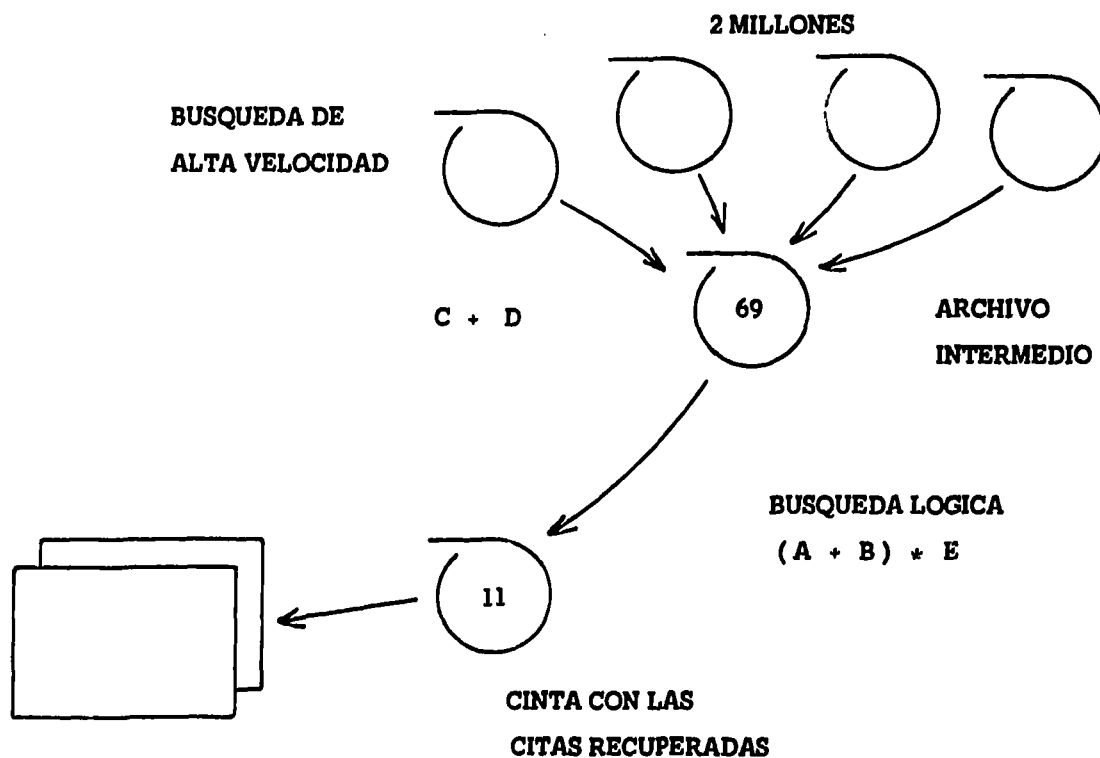


TRANSPARENCIA IV-3

FRECUENCIA			
INDIZACION		BUSQUEDA	
----- 207		----- 84	
----- 202		----- 73	
----- 194		----- 70	
----- 191		----- 62	
----- 178		----- 54	
----- 162		----- 39	
----- 154		----- 36	
----- 148			

TRANSPARENCIA IV-4

$(A + B) * (C + D) * E$
 3029 69 500.000



TRANSPARENCIA IV- 6

RUGOSIDAD

UP LISURA
TE RUGOSIDAD SUPERFICIAL
TR ASPEREZA

1

RUGOSIDAD

véase además

RUGOSIDAD SUPERFICIAL

LISURA véase

RUGOSIDAD

2

RUGOSIDAD

véase además

ASPEREZA

ASPEREZA

véase además

RUGOSIDAD

3

ASPEREZA

TRANSPARENCIA IV- 7

TESAURO FACETADO

L BIBLIOTECAS
 La ACADEMICAS
 Lah ESTUDIANTEILES
 al UNIVERSITARIAS
 Ld PUBLICAS
 Ldc MUNICIPALES
 Ldf POPULARES
 Li EMPRESARIALES
 Lk GUBERNAMENTALES

BIBLIOTECAS MUNICIPALES Ldc
UP BIBLIOTECAS DE CIUDAD
TR INTENDENCIA MUNICIPAL Qp
 BIBLIOTECAS EMPRESARIAS
 TG(A) SERVICIOS DE INFORMACION
 EMPRESARIA Rj
 BIBLIOTECAS DE CIUDAD
USE BIBLIOTECAS MUNICIPALES Ldc

FORMULARIO PARA DESCRIPTORES

DESCRIPTOR:	
CODIGO DE IDENTIFICACION:	
FACETA:	CODIGO DE FACETA:
USADO POR:	NOTA DE ALCANCE
TERMINO GENERICO (TG)	
	Viene del descriptor provisional Usado en fecha: Incorporado al tesauo:
TERMINOS ESPECIFICOS (TE)	Descriptor sugerido por:
	Observaciones:
TERMINOS RELACIONADOS (TR)	
	Fecha de incorporación al tesauo
	Fecha de revisión
	Fecha de descarte del término
Institución:	

Fuente: ICFES. - Norma colombiana 1476 1a. Revisión: Documentación - Directrices para el establecimiento y desarrollo de tesauros monolingües. - Bogotá, 1980.

TRANSPARENCIA IV- 8

M	SERVICIOS TECNICOS	CATALOGACION COOPERATIVA	Maj
Ma	CATALOGACION	UP	CATALOGACION CENTRALIZADA
Mod	CATALOGACION DESCRIPTIVA	TR	CENTRO DE INVESTIGACION
			DOCUMENTARIA – INTI
Mce	CLASIFICACION ALFABETICA		FICHAS IMPRESAS DE CATALOGO
Maj	CATALOGACION COOPERATIVA	TG(A)	PROCESAMIENTO COOPERATIVO

TRANSPARENCIA IV- 9

RECURSOS DE LOS LENGUAJES DE INDIZACIONRECURSOS DE ACIERTO

(AMPLIACION DEL TAMAÑO DE
LAS CLASES)

CONTROL DE SINONIMIA

CONTROL DE LA FORMA DE LAS PALABRAS

AGRUPAMIENTO JERARQUICO

ASOCIACION ESTADISTICA

RECURSOS DE PRECISION

(REDUCCION DEL TAMAÑO DE
LAS CLASES)

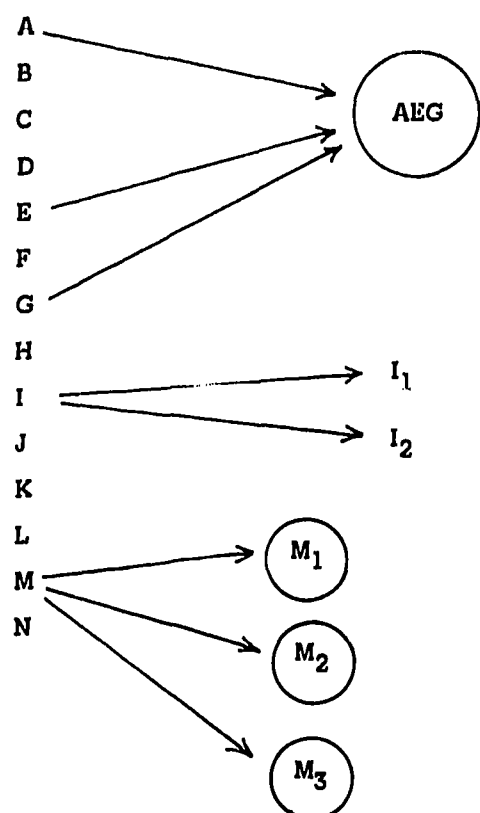
COORDINACION

UNION

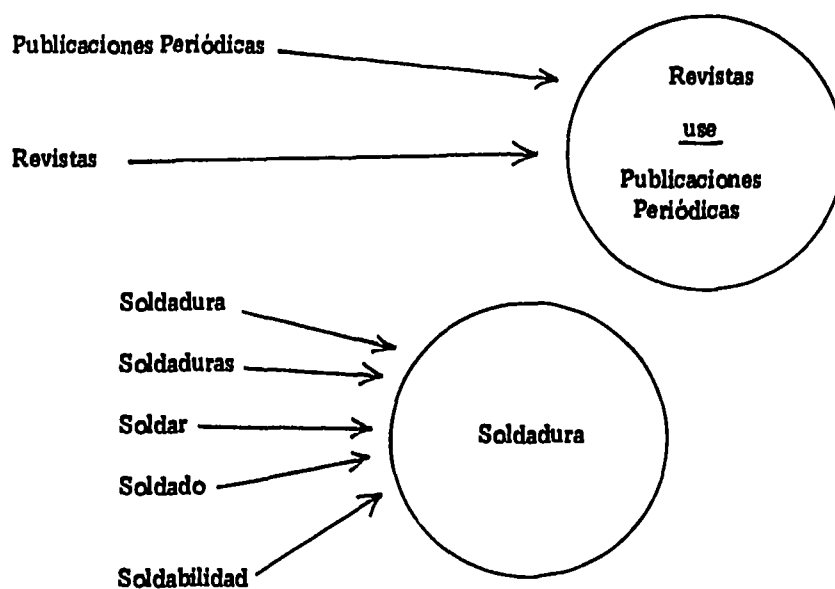
INDICADORES DE RELACIONES

PONDERACION

TRANSPARENCIA IV-10

ACIERTOPRECISION

TRANSPARENCIA IV-11

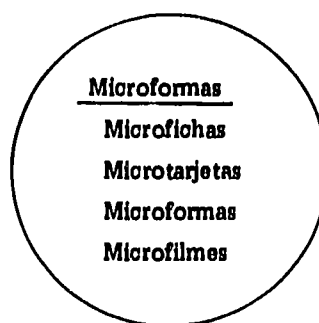
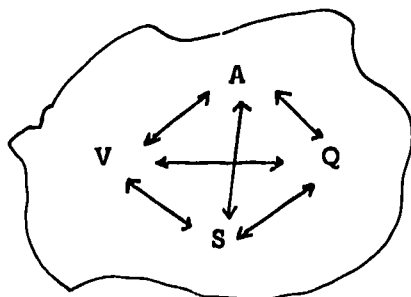


TRANSPARENCIA IV- 12

Microformas

TE

Microfichas
Microfilmes
Microtarjetas

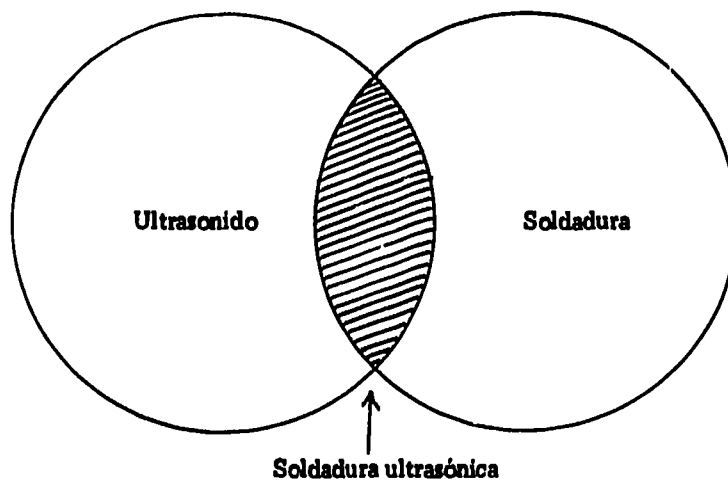


TRANSPARENCIA IV- 13

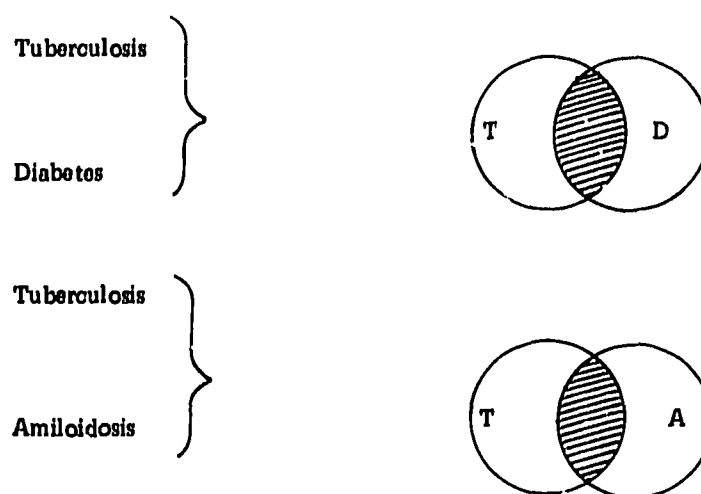
Ultrasonido

Soldadura

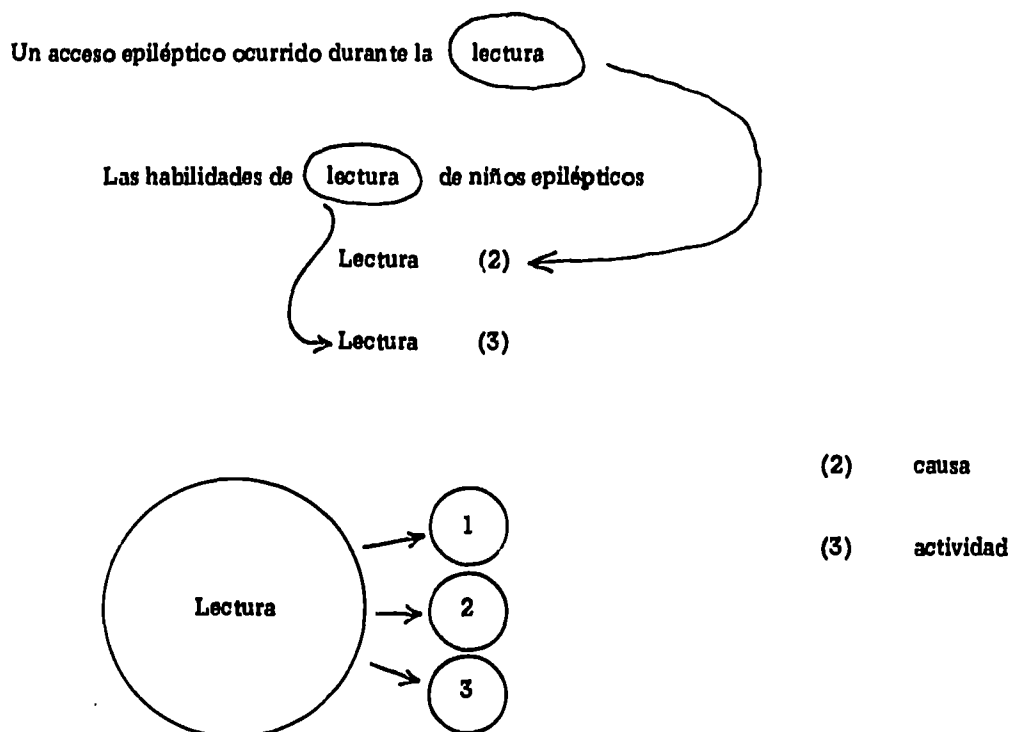
Soldadura ultrasónica



TRANSPARENCIA IV- 14

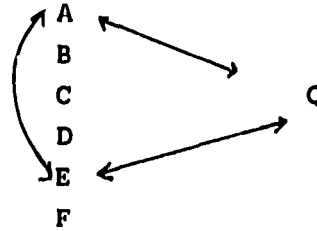


TRANSPARENCIA IV- 15

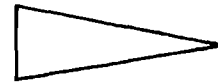


TRANSPARENCIA IV- 16

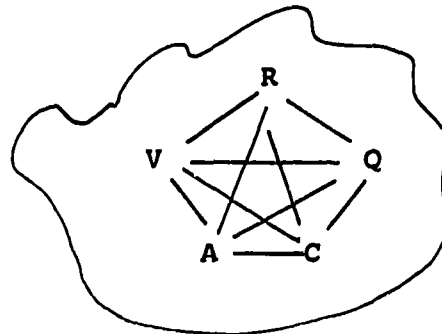
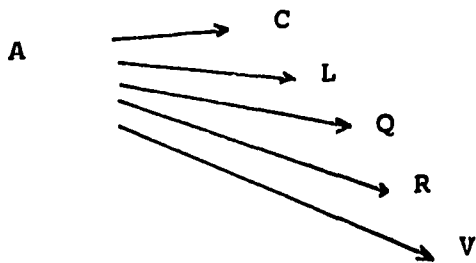
CONCURRENCIA



Q = ALA
A = TRIANGULAR
E = DELTA



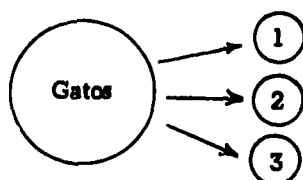
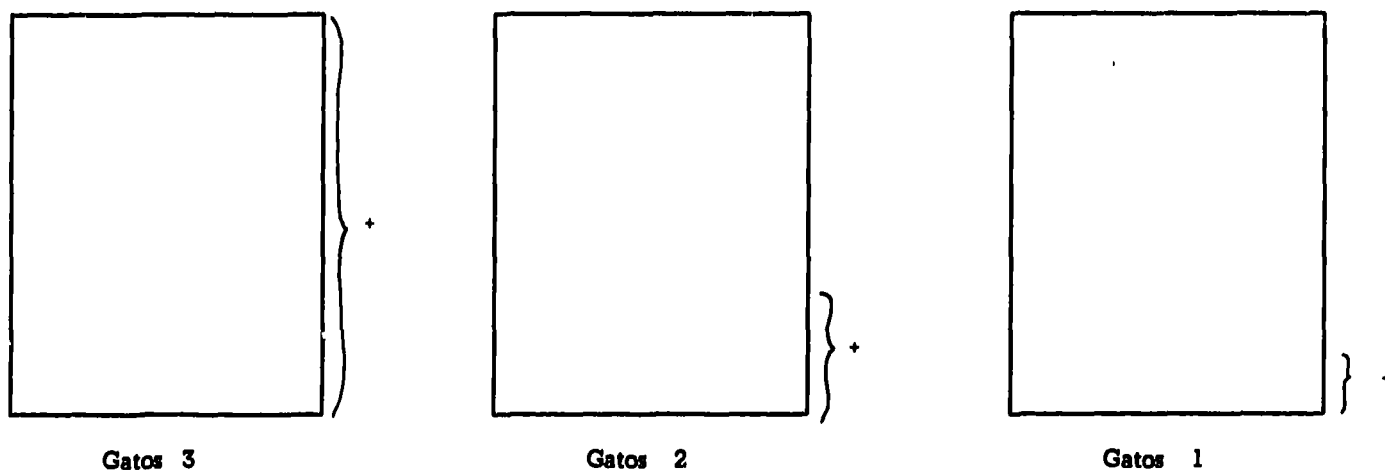
TRANSPARENCIA IV- 17



GRUPOS

AGRUPAMIENTOS

TRANSPARENCIA IV- 18



TRANSPARENCIA IV- 19

LENGUAJE NATURAL

1. COORDINACION FALSA
2. RELACION INCORRECTA ENTRE TERMINOS
3. INDIZACION EXHAUSTIVA
4. BUSQUEDA GENERICA O TOTAL

DISTANCIA DE LAS PALABRAS

LEGISLACION (5 palabras) ESCOLAR

LEGISLACION (1 palabra) ESCOLAR

LEGISLACION (Adyacencia) ESCOLAR

"LEGISLACION ESCOLAR"

(OPERADOR METRICO)

FRAGMENTOS DE PALABRAS

1. TRUNCADO A LA DERECHA. BUSQUEDA POR PREFIJO

COLOR* [COLORES, COLORIDO, COLORANDO, COLORACION]

PERO [COLORADO, COLORETE]

2. TRUNCADO A LA IZQUIERDA. BUSQUEDA POR SUFIJO

* MICINA

3. IZQUIERDA Y DERECHA

* MAGNET *

4. TRUNCADO POR INFIJO

TRI * COBALTATO

17	PROFILE	5	PROJECTIVE	24	PROPAGATED
1	PROFILO	1	PROJEVO	3	PROPAGATING
1	PROFISSIONAL	1	PROJEVY	46	PROPAGATION
2	PROFIT	1	PROKNCROVA	5	PROPANEDIOL
3	PROFOND	1	PROKOP	1	PROPANETRIOL
1	PROFONDA	1	PRUKREVNI	2	PROPANOL
4	PROFONDE	1	PROLER	1	PROPAFHENIN
1	PROFONDITA	3	PROLIFERATED	1	PROPАЗONE
1	PROFONDO	1	PROLIFERATING	1	PROPEASTRIATE
33	PROFOUND	44	PROLIFERATION	10	PROPENSITY
5	PROFOUNDLY	3	PROLIFERATIVE	64	PROPER
1	PROFUNDA	1	PROLIFERING	1	PROPERICIAZIN
1	PROFUNDITA	4	PROLINE	1	PROPERICIAZINA
1	PROFUNDITY	15	PROLONG	1	PROPERICIAZINE
3	PROFUSE	1	PROLONGADA	18	PROPERLY
2	PROFUSION	1	PROLONGADO	248	PROPERTY
1	PROGENITOR	21	PROLONGATION	2	PROPEUNISTRIATE
3	PROGENY	5	PROLONGE	19	PROPHYLACTIC
2	PROGESTATIONAL	374	PROLONGED	2	PROPHYLACTICALLY
10	PROGESTERONE	6	PROLONGEE	2	PROPHYLAXE
2	PROGESTERONE/ESTROGE	5	PROLONGING	27	PROPHYLAXIS
1	PRUGESTIN	1	PROMAZINA	1	PROPIA
1	PROGLOTTID	33	PROMAZINE	1	PROFILAMINA
7	PROGNOSE	37	PROMETHAZINE	1	PROPINQUITY
1	PROGNOSEN	5	PROMINENCE	5	PROPIONATE
2	PROGNOSI	82	PROMINENT	100	PROPO
358	PROGNOSIS	9	PROMISE	106	PROPORTION
71	PROGNOSTIC	18	PROMISING	29	PROPORTIONAL
4	PROGNOSTICALLY	1	PROMJENE	1	PROPORTIONALLY
3	PROGNOSTICATION	2	PROMONOCYTE	2	PROPORTIONATE
1	PROGNOSTICHE	12	PROMOTE	1	PROPORTIONATELY
2	PROGNOSTICO	8	PROMOTED	8	PROPOSAL
8	PROGR	8	PROMOTING	82	PROPOSED
37	PROGRAM	16	PROMPT	1	PROPOSEE
1	PROGRAMMED	7	PROMPTED	1	PROPOSER
51	PROGRESS	1	PROMPTER	2	PROPOSING
13	PROGRESSED	1	PROMPTITUDE	3	PROPOSITI
2	PROGRESSIT	25	PROMPTLY	6	PROPOSITION
8	PROGRESSING	2	PROMYSLOV	20	PROPOSITO
23	PROGRESSION	1	PRONATING	10	PROPOSITUS
1	PROGRESSIVA	1	PRONATIUM	2	PROPOUNDED
336	PROGRESSIVE	29	PRONE	3	PROPOXYPHENE
1	PROGRESSIVEDETERIORA	9	PRONENESS	1	PROPPED
54	PROGRESSIVELY	6	PRONOSTIC	5	PROPANOLOL
7	PROGRESSIVEN	1	PRONOSTICO	4	PROPRIETA
1	PROGRESSIVITY	1	PRONOSTICULUI	1	PROPRIETARY

VOCABULARIOS HIBRIDOS



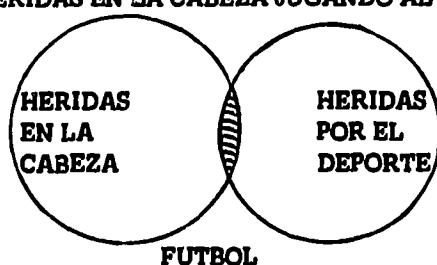
DESCRITORES

PALABRAS DEL TITULO

RESUMENES

BARRIDO

HERIDAS EN LA CABEZA JUGANDO AL FUTBOL



VOCABULARIOS HIBRIDOS

**LOS TERMINOS CONTROLADOS DAN
CONTEXTO A LAS PALABRAS CLAVE
NO CONTROLADAS QUE DAN ESPECIFICIDAD**

EK TRABAJO HUELGA

**WF GUERRA
HUELGA**

CODIGOS TEMATICOS AMPLIOS

~ 250

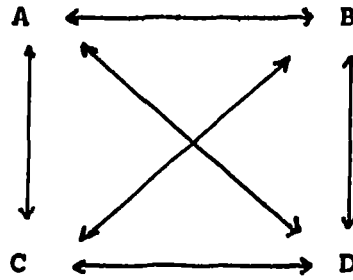
CODIGOS GEOGRAFICOS

DJO = JORDANIA

PALABRAS CLAVE

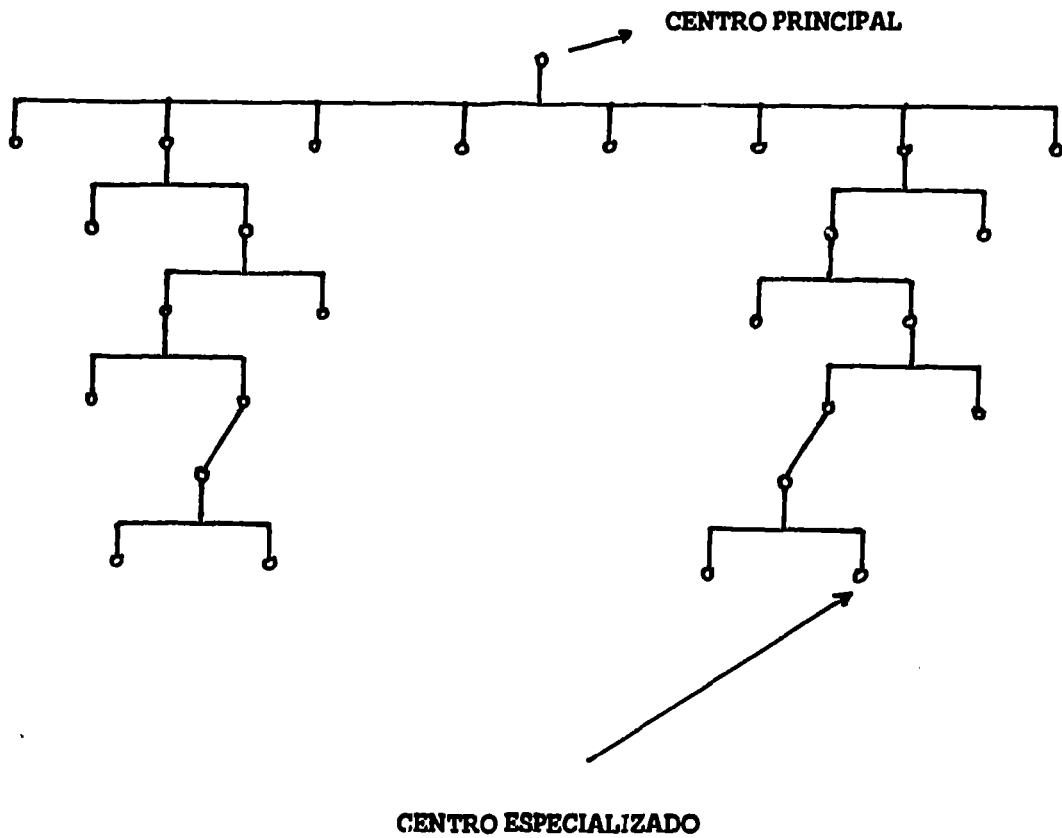
TRANSPARENCIA IV- 27

VOCABULARIO



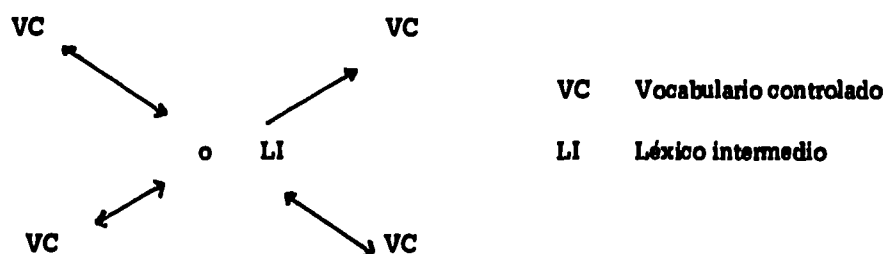
TRANSPARENCIA IV- 28

MICROTESAUROS



TRANSPARENCIA IV- 29

LEXICO INTERMEDIO



TRANSPARENCIA IV-30

CONVERSION POR MAQUINA

EQUIVALENCIA

EPILEPSIA → EPILEPSIA

ESCRITURA VARIABLE

SEPTIEMBRE → SETIEMBRE

FORMA DE LAS PALABRAS

**MEJORA PEDAGOGICA —→ MEJORAMIENTO
PEDAGOGICO**

INVERSIONES

INDUSTRIA TEXTIL → TEXTIL, INDUSTRIA

POR REFERENCIAS CRUZADAS

TNT uso DINAMITA

POR LA JERARQUIA

NIÑO
 ↓
 PARVULO

↗

NIÑO

TRANSPARENCIA IV-31

TABLA 37. CONCORDANCIA DE VOCABULARIOS: Correspondencia exacta

	Source thesaurus	Thesaurus B	Thesaurus C
Original entries	AIRFIELDS	AIRFIELD	FLUGPLATZE
Reconciled entries	AIRFIELDS (0101)	AIRFIELD (0101)	FLUGPLATZE (0101)
Keys code numbers	0101 = AIRFIELDS	0101 = AIRFIELD	0101 = FLUGPLATZE

TRANSPARENCIA IV-32

TABLA 38. CONCORDANCIA DE VOCABULARIOS: Sinonimia

	Source thesaurus	Thesaurus B
Original entries	UNDERGROUND STRUCTURES buried structures, use UNDERGROUND STRUCTURES	SUBSURFACE STRUCTURES underground structures, use SUBSURFACE STRUCTURES
Reconciled entries	UNDERGROUND STRUCTURES (0201) buried structures, use UNDERGROUND STRUCTURES subsurface structures, use UNDERGROUND STRUCTURES	SUBSURFACE STRUCTURES (0201) underground structures, use SUBSURFACE STRUCTURES buried structures, use SUBSURFACE STRUCTURES
Keys to code numbers	0201 = UNDERGROUND STRUCTURES	0201 = SUBSURFACE STRUCTURES

TRANSPARENCIA IV-33

TABLA 39: CONCORDANCIA DE VOCABULARIOS: Descriptor agregado a un Tesauro.

	Source thesaurus	Thesaurus B
Original entries	SNOWDRIFTS	--
Reconciled entries	SNOWDRIFTS (0301)	SNOWDRIFTS (0301)
Keys to code numbers	0301 = SNOWDRIFTS	0301 = SNOWDRIFTS

TRANSPARENCIA IV- 34

TABLA 40. CONCORDANCIA DE VOCABULARIOS: Descriptor relacionado con un término más amplio en un tesoro

	Source thesaurus	Thesaurus C
Original entries	SNOW	SNOW
Reconciled entries	SNOW (0300) SNOWDRIFTS (0301)	SNOW (0300) SNOWDRIFTS (0301), use SNOW
Keys to code numbers	0300 = SNOW 0301 = SNOWDRIFTS	0300 = SNOW 0301 = SNOWDRIFTS

TRANSPARENCIA IV- 35

TABLA 41. CONCORDANCIA DE VOCABULARIOS: Descriptor relacionado con un término más amplio en un tesoro *

	Source thesaurus	Thesaurus B
Original entries	ALUMINUM ALUMINIUM STRUCTURES	ALUMINIUM
Reconciled entries	ALUMINIUM (material) (0701) ALUMINIUM STRUCTURES (0702)	ALUMINIUM (material 0701) (structures 0702)
Keys to code numbers	0701 = ALUMINIUM (material) 0702 = ALUMINIUM STRUCTURES	0701 = ALUMINIUM (material) 0702 = ALUMINIUM (structures)

* El tesoro fuente contiene términos en ambas formas sustantivadas y adjetivadas pero el otro Tesoro no lo tiene.

TRANSPARENCIA IV- 36

TABLA 42. CONCORDANCIA DE VOCABULARIO: El descriptor del tesoro fuente no es deseado en el otro tesoro.

	Source thesaurus	Thesaurus D
Original entries	SNOWDRIFTS	-----
Reconciled entries	SNOWDRIFTS (0301)	snowdrifts (0301)
Keys to code numbers	0301 = SNOWDRIFTS	0301 = snowdrifts

TRANSPARENCIA IV- 37

TABLA 43. CONCORDANCIA DE VOCABULARIOS: Descriptores simples y términos combinados (antónimos)

	Source thesaurus	Thesaurus B
Original entries	CONTRACTION	contraction, use EXPANSION / CONTRACTION
	EXPANSION	EXPANSION / CONTRACTION
Reconciled entries	CONTRACTION (0402)	contraction (0402), use EXPANSION / CONTRACTION
	EXPANSION (0403)	expansion (0403), use EXPANSION / CONTRACTION EXPANSION / CONTRACTION
Keys to code numbers	0402 = CONTRACTION 0403 = EXPANSION	0402 = contraction 0403 = expansion

TRANSPARENCIA IV- 38

TABLA 44. CONCORDANCIA DE VOCABULARIOS: Términos precoordinados y descriptores separados.

	Source thesaurus	Thesaurus B
Original entries	FROST PENETRATION	frost penetration, use FROST + PENETRATION
Reconciled entries	FROST PENETRATION (0501)	frost penetration (0501), use FROST + PENETRATION
Keys to code numbers	0501 = FROST PENETRATION	0501 = frost penetration

TRANSPARENCIA IV- 30

TABLA 45. CONCORDANCIA DE VOCABULARIOS: Términos precoordinados y descriptores separados.

	Source thesaurus	Thesaurus B
Original entries	STIFFNESS METHODS	stiffness methods, use STRUCTURAL ANALYSIS + DISPLACEMENT
Reconciled entries	STIFFNESS METHODS (0502)	stiffness methods (0502), use STRUCTURAL ANALYSIS + DISPLACEMENT
Keys to code numbers	0502 = STIFFNESS METHODS	0502 = stiffness methods

TRANSPARENCIA IV- 40

TABLA 46: CONCORDANCIA DE VOCABULARIOS: Términos desmembrados semánticamente

	Source thesaurus	Thesaurus B
Original entries	THERMOMETER	thermometer, use TEMPERATURE + MEASUREMENT INSTRUMENT
Reconciled entries	THERMOMETER (0503)	thermometer (0503), use TEMPERATURE + MEASUREMENT INSTRUMENT
Keys to code numbers	0503 = THERMOMETER	0503 = thermometer

TABLA 47. CONCORDANCIA DE VOCABULARIOS: Modificadores con equivalencias indirectas.

	Source thesaurus	Thesaurus B
Original entries	HIGH	---
Reconciled entries	HIGH (-) (HIGH + TEMPERATURE 0801) (HIGH + BUILDINGS 0802) (HIGH + STRENGTH 0803) high temperature (0801), use HIGH + TEMPERATURE high buildings (0802), use HIGH + BUILDINGS high strength (0803), use HIGH + STRENGTH multi-storey buildings (0802), use HIGH + BUILDINGS	HIGH TEMPERATURE (0801) MULTI-STOREY BUILDINGS (0802) HIGH-STRENGTH (0803) high buildings, use MULTI-STOREY BUILDINGS
Keys to code	0801 = high temperature 0802 = high buildings 0803 = high strength	0801 = HIGH TEMPERATURE 0802 = MULTI-STOREY BUILDINGS 0803 = HIGH STRENGTH

TESAUROS MULTILINGUES

1. RECOLECTAR TERMINOS DE AMBAS LENGUAS
 LUEGO BUSCAR LAS EQUIVALENCIAS
2. TRADUCIR TESAuros EXISTENTES
3. FUSIONAR DOS O MAS TESAuros

TRANSPARENCIA IV-43

EQUIVALENCIA EXACTA

<u>ESPAÑOL</u>	<u>INGLES</u>	<u>FRANCES</u>	<u>ALEMAN</u>
FISICA	PHYSICS	PHYSIQUE	PHYSIK
MIRLO	BLACK BIRD	MERLE	AMSEL

SIN EQUIVALENCIA EXACTA

<u>ESPAÑOL</u>	<u>INGLES</u>	<u>FRANCES</u>
GANADO MAYOR	CATTLE HORSES	GROS BATAIL
OVEJA	SHEEP MOUTON	MOUTON
PROGRAMA	SOFTWARE	PROGRAMME

TRANSPARENCIA IV-47

CONDICION DE COMPLETO

CARACTERISTICAS DE LOS TERMINOS

PRECOORDINACION	OBSERVABLE	
ESPECIFICIDAD	OBSERVABLE	TIENE QUE COMPROBARSE
TERMINOS DE ENTRADA / DESCRIPTORES		
¿ TG / TE ESTAN USADOS CORRECTAMENTE?		
DEFINICIONES / NOTAS DE ALCANCE		

TESAURO INGLES - FRANCES

AMBASSADOR AT LARGE

AMBASSADEUR ITINERANT

AMBASSADOR DESIGNATE

AMBASSADEUR DESIGNE

A PERSON NAMED BUT NOT YET

ACCREDITED AS AMBASSADOR

BT AMBASSADORS

AMBASSADEURS

CHARGE D'AFFAIRS

CHARGE D'AFFAIRES

DIPLOMATIC TITLES

TITRES DIPLOMATIQUES

BT DIPLOMATIC APPOINTMENTS

NOMINATIONS DIPLOMATIQUES

AMBASSADOR DESIGNATE

AMBASSADEUR DESIGNE

AMBASSADORS

AMBASSADEURS

THE PERSONAL REPRESENTATIVE OF A CHIEF OF STATE. AMBASSADORS NORMALLY
ARE ACCREDITED TO ANOTHER CHIEF OF STATE

UF CHIEFS OF MISSION

CHEFS DE MISSION

DIPLOMATIC MISSION CHIEFS

CHEFS DE MISSION DIPLOMATIQUE

BT CHARGE D'AFFAIRS

CHARGE D'AFFAIRES

DIPLOMATIC TITLES

TITRES DIPLOMATIQUES

NT AMBASSADOR AT LARGE

AMBASSADEUR ITINERANT

AMBASSADOR DESIGNATE

AMBASSADEUR DESIGNE

BT AMBASSADORS WHEREABOUTS

DEPLACEMENTS D'AMBASSADEURS

DIPLOMATIC MISSION CHIEFS

CHEFS DE MISSION DIPLOMATIQUE

PERSONAL REPRESENTATIVE

REPRESENTANT PERSONNEL

TESAURO BILINGUE FRANCES-INGLES

CONFLICTS FRONTALIERS

BORDER CLASHES

- BT CONTESTATIONS DE FRONTIERE
BORDER DISPUTES

CONFRONTATIONS

CONFRONTATIONS

- BT CONNAISSANCE TOUCHANT L'ADVERSAIRE
ADVERSARYS PERCEPTIONS

CONNAISSANCE TOUCHANT L'ADVERSAIRE

ADVERSARYS PERCEPTIONS

- BT THEORIE DES RELATIONS ETRANGERES
FOREIGN RELATIONS THEORY

- BT CONFRONTATIONS
CONFRONTATIONS

CONSEIL EXECUTIF

EXECUTIVE COUNCIL

- BT CHEF D'ETAT
CHIEF OF STATE

CONSUL SUPPLEANT

ACTING COUNSULS

PERSONNES HABILITEES A AGIR EN QUALITE DE CONSUL DURANTE L'ABSENCE DE CE DERNIER

- BT AGENTS COUNSULAIRES
CONSULAR OFFICERS

CONTESTATIONS DE FRONTIERE

BORDER DISPUTES

DISPUTES ENTRE ETATS RESULTANT DE DIFFICULTES FRONTALIERES

- UF CONTESTATIONS FRONTALIERES
BOUNDARY DISPUTES

- NT REVENDICATIONS FRONTALIERES
BOUNDARY CLAIMS

- BT CONFLITS FRONTALIERS
BORDER CLASHES

FRONTIERES DE FACTO
DE FACTO BOUNDARIES

INCIDENTS DE FRONTIERE
BORDER INCIDENTS

PROBLEMS DE FRONTIERE
BORDER PROBLEMS

RECTIFICATIONS DE FRONTIERES
BOUNDARY ADJUSTMENTS

TESAURO MULTILINGUE

SOCIAL COMMUNICATION / COMMUNICATION SOCIALE / COMUNICACION SOCIAL

BT COMMUNICATION / COMMUNICATION / COMUNICACION

RT SOCIAL INTEGRATION / INTEGRATION SOCIALE / INTEGRACION SOCIAL

10200 COMMUNICATION STRUCTURE

10210

COMMUNICATION CHANNEL / CANAL DE COMMUNICATION / CANAL DE COMUNICACION

RT COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

COMMUNICATION STRUCTURE / STRUCTURE DE COMMUNICATION / ESTRUCTURA DE COMUNICACION

USE COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

SN Use in connection with the whole of the elements implied in information transfer (transmitter, receiver, channel, code, etc.) and with types of ordering / Designe l'ensemble des éléments impliqués dans la transmission de l'information (émetteur, récepteur, canal, code, etc.) et les modalités de leur agencement / Indica el conjunto de elementos que necesita la transferencia de información (emisor, receptor, canal, código, etc.) y las modalidades de su arreglo.

UF COMMUNICATION STRUCTURE / STRUCTURE DE COMMUNICATION / ESTRUCTURA DE COMUNICACION

NT ONE-WAY COMMUNICATION / COMMUNICATION UNIDIRECTIONNELLE / COMUNICACION UNILATERAL
POINT-TO-POINT COMMUNICATION / COMMUNICATION DE POINT A POINT / COMUNICACION
DE PUNTO A PUNTO

TWO-WAY COMMUNICATION / COMMUNICATION BIDIRECTIONNELLE / COMUNICACION RECIPROCA
RT COMMUNICATION / COMMUNICATION / COMUNICACION
COMMUNICATION CHANNEL / CANAL DE COMMUNICATION / CANAL DE COMUNICACION

MEDIATION / MEDIATION / MEDIACION

RT VICARIOUS EXPERIENCE / EXPERIENCE VICARIALE / EXPERIENCIA VICARIAL

NETWORK / RESEAU / RED

RT NETWORK ANALYSIS / ANALYSE DE RESEAU / ANALISIS DE REDES

ONE-WAY COMMUNICATION / COMMUNICATION UNIDIRECTIONNELLE / COMUNICACION UNILATERAL

BT COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

POINT-TO-POINT COMMUNICATION / COMMUNICATION DE POINT A POINT / COMUNICACION DE
PUNTO A PUNTO

BT COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

TELECONFERENCING / CONVERSATION A DISTANCE / CONVERSACION A DISTANCIA

RT DISTANCE / DISTANCE / DISTANCIA

TWO-WAY COMMUNICATION / COMMUNICATION BIDIRECTIONNELLE / COMUNICACION RECIPROCA

TWO-WAY COMMUNICATION / COMMUNICATION BIDIRECTIONNELLE / COMUNICACION RECIPROCA

BT COMMUNICATION SYSTEM / SYSTEME DE COMMUNICATION / SISTEMA DE COMUNICACION

RT TELECONFERENCING / CONVERSATION A DISTANCE / CONVERSACION A DISTANCIA

TRANSPARENCIA IV-48

CONSISTENCIA

RECIPROCIDAD

FORMA DE LAS PALABRAS

SINGULAR / PLURAL

ESTRUCTURA

GRADO DE UNION

ESTETICA

PRESENTACION

TIPOGRAFIA

TRANSPARENCIA IV-49

CONEXION

DESCRIPTORES CON CONEXIONES:

DESCRIPTORES

a = DESCRIPTORES SIN CONEXIONES

b = TOTAL DE DESCRIPTORES

$$\frac{b-a}{b} =$$

~ 1 RECOMENDADO

EN UN ESTUDIO REALIZADO SE ENCONTRO QUE $\frac{11}{20}$ TENIAN UNA RELACION
DE CONEXION

DEFINICION

(b - a) / b

a = DESCRIPTORES QUE TENGAN LA POSIBILIDAD DE SER AMBIGUOS
(P.E. UN SIGNIFICADO INSUFICIENTEMENTE DEFINIDO POR LAS
NOTAS DE ALCANCE POR SU CALIFICADOR, DEFINICION, RELACIONES
JERARQUICAS, ETC.)

DEBERIA SER:

16/20 DE 20 TESAuros ESTUDIADOS SE ENCONTRO QUE 16 REUNIAN
ESTA CUALIDAD.

TASA DE EQUIVALENCIA

NO DESCRIPTORES : DESCRIPTORES
SE ENCONTRARON 0.24 DEBERIAN SER > 1

TASA DE RECIPROCIDAD

EN QUE MEDIDA LAS RELACIONES (DE EQUIVALENCIA, JERARQUICA, ASOCIACION
SON RECIPROCAS?)

TASA DE ENRIQUECIMIENTO (ACCESIBILIDAD)

RELACIONES

JERARQUICA / ASOCIATIVA : DESCRIPTORES

11/20 TESAuros DAN VALORES ENTRE 2 y 5 CINCO DAN > 5

DEBERIA SER 2 - 5 , DEMASIADOS ACCESOS SON OBSTACULOS

TRANSPARENCIA IV- 52

FLEXIBILIDAD

RELACION ENTRE LAS PALABRAS COMPONENTES DE DESCRIPTORES QUE,
A SU VEZ, PUEDEN SER DESCRIPTORES, Y LAS QUE NO PUEDEN SER DESCRIPTORES.

Puede ser 0.6 ó mayor

8 / 20 exceso

TAMAÑO DEL GRUPO DE TERMINOS

150 RECOMIENDA 30 / 40 POR GRUPO

TRANSPARENCIA IV- 53

MEDICIONNIVEL DE PRECOORDINACION

INGLES 1.9

FRANCES 1.7

ALEMAN 1.1.

Debería ser 1.5 a 2 para Inglés
Francés

1.1 a 1.2 para Alemán

TRANSPARENCIA IV- 54

HACIA UN LENGUAJE NATURAL (LN)

1. BASES DE DATOS SECUNDARIOS EN FORMA DE LN
COMPLETAMENTE
EL LN COMO COMPONENTE

2. BASES DE DATOS PRIMARIOS EN LN

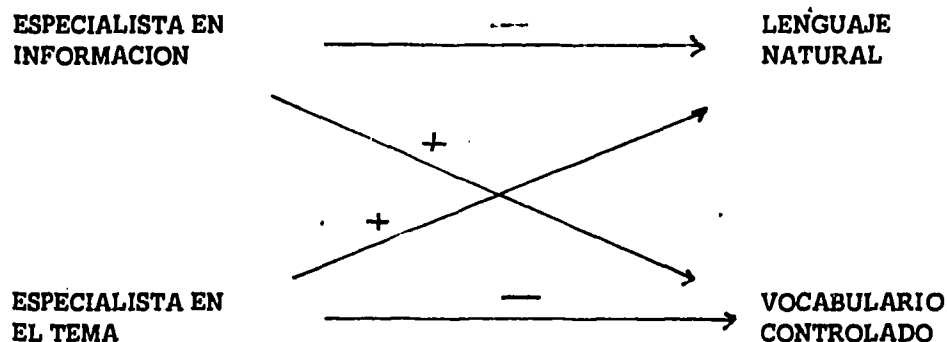
3. LOS LN SON CADA VEZ MAS FACTIBLES

FISICAMENTE

INTELECTUALMENTE

4. BUSQUEDA NO DELEGADA AUMENTARA EN GRAN ESCALA

TRANSPARENCIA IV-55



TRANSPARENCIA IV-56

	<u>INDIZACION</u>	<u>BUSQUEDA</u>	
1.	CONTROLADA	CONTROLADA	<u>CONTROLADO</u>
2.	CONTROLADA	L. NATURAL	PRE-CONTROLADO
3.	LENGUAJE NATURAL	L. NATURAL	NO CONTROLADO
4.	LENGUAJE NATURAL	CONTROLADA	<u>POST-CONTROLADO</u>

TRANSPARENCIA IV-57

TABLAS DE SINONIMOS

METXABITAL = METXILFENCBARBITAL =

GEMONIL = ENDIEMAL

GRUPO CONCEPTUAL

"SICOLOGICO"

SIC x x x x

NEUROSIS

DEPRESION

SUICIDIO

ETC., ETC.

TRANSPARENCIA IV- 58

SULTIAME
SULTHIAME
OSPOLOT

NIVEL DE SANGRE
NIVEL DE SUERO
CONCENTRACION DE SANGRE
CONCENTRACION DE SUERO
CONCENTRACION EN LA SANGRE
CANTIDAD DE . . . EN LA SANGRE

TRANSPARENCIA IV- 59

LINEAS AEREAS

LINEAS AEREAS
TRANSPORTES AEREOS
AEROLINEAS ARGENTINAS
AUSTRAL
LUFTHANSA
SAS
SCANDINAVIAN AIRLINES
SWISSAIR
VARIG

TRANSPARENCIA IV- 60

MEDICINA

QUIMIO *****
***** TERAPIA
***** IATRA
PSI *****
*** PSI *****
ELECTRO ****
RADIO *****
