

Análisis de terminologías de salud para su utilización como ontologías computacionales en los sistemas de información clínicos

Maria Teresa Romá-Ferri^a / Manuel Palomar^b

^aDepartamento de Enfermería, Universidad de Alicante, Alicante, España;

^bDepartamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante, Alicante, España.

(Analysis of health terminologies for use as ontologies in healthcare information systems)

Resumen

Objetivos: Las ontologías son un recurso que permite trabajar informáticamente con la conceptualización del significado y evitar la limitación impuesta por los términos normalizados. El objetivo de este estudio es establecer el grado de usabilidad de las terminologías para el diseño de ontologías, que contribuyan a resolver los problemas de interoperabilidad semántica, y de reutilización de conocimiento en los sistemas de información clínicos.

Métodos: Se han analizado 6 de las terminologías más relevantes para el ámbito clínico, epidemiológico, documental y administrativo-económico. Se valoraron las siguientes cualidades: cobertura conceptual, estructura jerárquica, granularidad conceptual, relaciones conceptuales y grado de formalismo utilizado en la representación conceptual, para establecer el grado de usabilidad.

Resultados: Se consideran como ontologías ligeras los MeSH, los DeCS y el UMLS, aunque con diferencias entre ellas, al explicitar los conceptos, el tipo de relación y las restricciones entre los conceptos asociados. SNOMED y GALEN, con su formalismo declarativo basado en descripciones lógicas, incluyen la explicitación de las cualidades, una mayor restricción para relacionar conceptos y las reglas de combinación entre ellos, por lo que se consideran como ontologías pesadas.

Conclusiones: El análisis de la representación declarada de las terminologías muestra las posibilidades de su reutilización como ontologías. Su grado de usabilidad dependerá de si se pretende que los sistemas de información clínicos resuelvan los problemas de interoperabilidad semántica (ontologías ligeras) o además reutilizar su conocimiento para sistemas de ayuda a la toma de decisiones (ontologías pesadas) y para

tareas de recuperación, extracción y clasificación de información no estructurada.

Palabras clave: Terminologías. Semántica. Aplicaciones de informática médica. Sistemas de gestión de información avanzada integrada.

Abstract

Objectives: Ontologies are a resource that allow the concept of meaning to be represented informatically, thus avoiding the limitations imposed by standardized terms. The objective of this study was to establish the extent to which terminologies could be used for the design of ontologies, which could be serve as an aid to resolve problems such as semantic interoperability and knowledge reusability in healthcare information systems.

Methods: To determine the extent to which terminologies could be used as ontologies, six of the most important terminologies in clinical, epidemiologic, documentation and administrative-economic contexts were analyzed. The following characteristics were verified: conceptual coverage, hierarchical structure, conceptual granularity of the categories, conceptual relations, and the language used for conceptual representation.

Results: MeSH, DeCS and UMLS ontologies were considered lightweight. The main differences among these ontologies concern conceptual specification, the types of relation and the restrictions among the associated concepts. SNOMED and GALEN ontologies have declaratory formalism, based on logical descriptions. These ontologies include explicit qualities and show greater restrictions among associated concepts and rule combinations and were consequently considered as heavyweight.

Conclusions: Analysis of the declared representation of the terminologies shows the extent to which they could be reused as ontologies. Their degree of usability depends on whether the aim is for healthcare information systems to solve problems of semantic interoperability (lightweight ontologies) or to reuse the systems' knowledge as an aid to decision making (heavyweight ontologies) and for non-structured information retrieval, extraction, and classification.

Key words: Terminology. Semantics. Medical informatics applications. Integrated advanced information management systems.

Correspondencia: Maria Teresa Romá Ferri.
Departamento de Enfermería. Universidad de Alicante.
Campus de San Vicente del Raspeig.
Apdo. 99. 03080 Alicante. España.
Correo electrónico: mtr.ferri@ua.es

Recibido: 30 de julio de 2007.

Aceptado: 9 de enero de 2008.

Introducción

El avance progresivo e imparable de las tecnologías de la información y de la comunicación está originando cambios con respecto a la gestión de la información¹⁻⁴. De hecho, el reto actual se sitúa en convertir la abundante información existente en un conocimiento que pueda ser operativo y funcional en el contexto de su aplicación: los cuidados de los pacientes, la evaluación de la calidad asistencial, la investigación y la epidemiología, la planificación y la gestión o la formación, considerando, al mismo tiempo, los factores de riesgo presentes de sobreabundancia, redundancia, infrutilización o inadecuada utilización de la información existente⁴⁻⁸.

La historia clínica es uno de los elementos esenciales del sistema de información asistencial enfocado en el paciente. En el ámbito nacional, la transformación electrónica de la historia (HCE) se está llevando a cabo por la integración de varios subsistemas parciales de información (laboratorio, radiología, anatomía patológica, farmacia...). Es decir, cada subsistema ha pasado de ser un sistema aislado (PC sin conectar a la red) a formar parte de un sistema distribuido y donde la información está repartida en varios ordenadores interconectados (su nueva arquitectura). No obstante, todavía cada subsistema se mantiene como una unidad discreta con su modelo de datos, su lógica, su base de datos y su propia terminología⁹.

Las terminologías en el dominio de la salud han sido el recurso utilizado para la representación de conocimiento, el intercambio de información y la recuperación de documentos profesionales. Las terminologías, indistintamente del criterio de agrupación que se adopte (clasificación, nomenclatura o taxonomía), tienden a modelizar un sistema bajo la forma de un conjunto de términos con una intención fundamentalmente normativa¹⁰. En el actual contexto de la digitalización, las terminologías siguen actuando en los sistemas de información (SI) como una forma de entrada/salida de datos estandarizados, como por ejemplo con las etiquetas para los diagnósticos médicos que se seleccionan desde jerarquías preestablecidas (menús), utilizadas para igualar las opciones de los usuarios¹¹ y para almacenar la cadena de caracteres que corresponden a dicha etiqueta. Los SI sólo aceptan una forma específica de pedir una información y limitan la expresividad natural de cada individuo en su idioma.

Además, los datos que constituyen el SI pueden proceder tanto del ámbito interno como del externo a la organización⁵. En un hospital, los datos complementarios se encuentran en diversos recursos de información externos (bases de datos bibliográficas, guías clínicas, estadísticas, de otros centros asistenciales...). Por otra parte, hay que considerar los desplazamientos de los ciudadanos con sistemas de atención diferenciados (es-

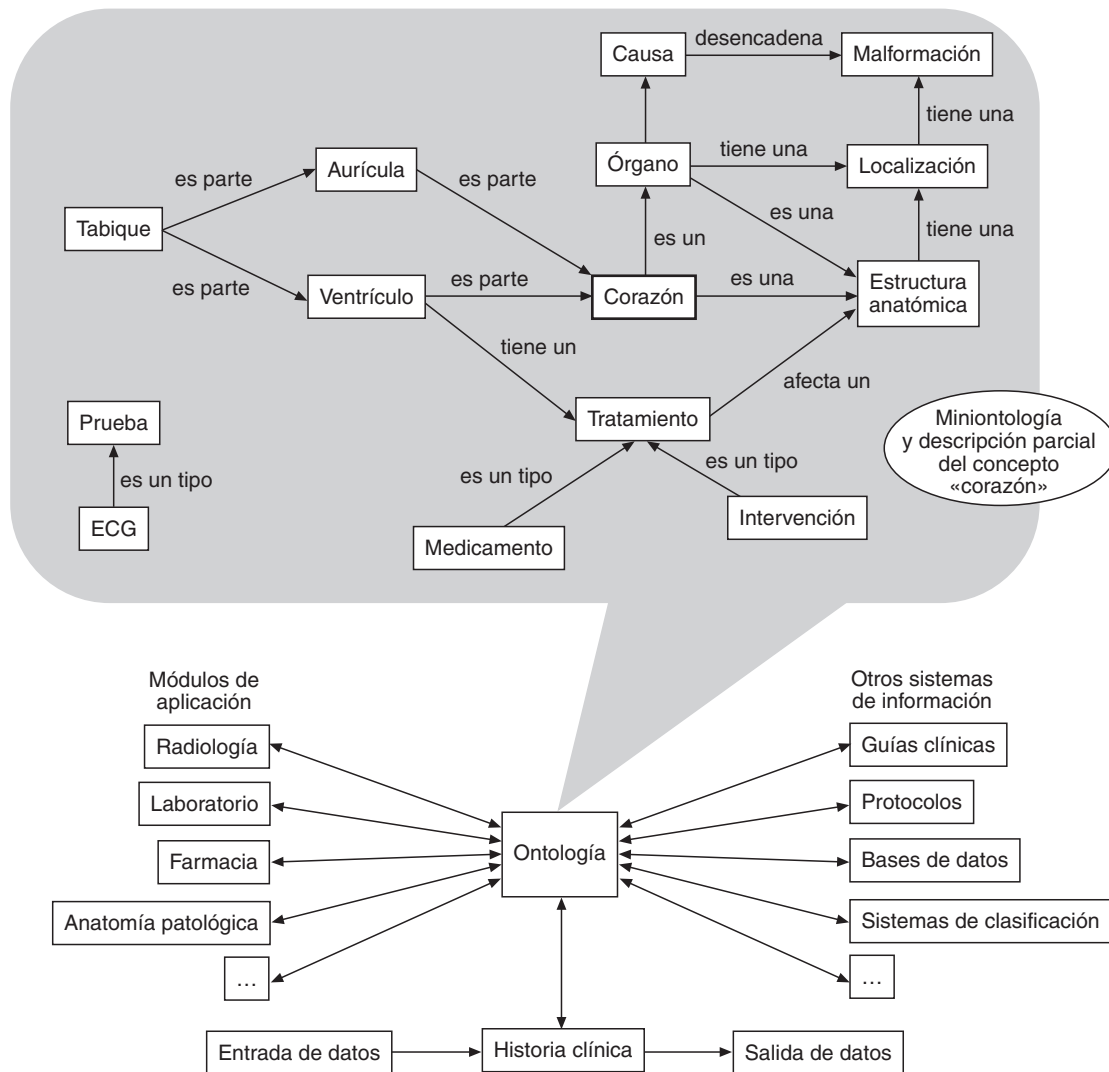
tados miembros de la Unión Europea) y distintos modelos de historia de salud (comunidades autónomas) e idioma.

Para poder utilizar estos datos, ha sido preciso desarrollar protocolos que permitieran la transferencia y la sincronización de los datos entre las diferentes aplicaciones de un sistema o entre los distintos subsistemas (interoperabilidad del sistema). A pesar de ello, persisten problemas para compartirlos, porque los SI, emisor y receptor, necesitan entender de forma común el dato transferido. Ambos SI deberían tener la capacidad de interpretar los datos de forma automática y, consecuentemente, reutilizarlos en aplicaciones o subsistemas que no intervinieron en su creación (a dicha capacidad se la denomina interoperabilidad semántica)^{7,8,12,13}. No obstante, el problema principal es que pueden estar utilizando el mismo término para referirse a conceptos distintos (proteinuria como un trastorno renal o como el resultado de una prueba de laboratorio), o diferentes términos para el mismo concepto (mucoviscidosis, fibrosis quística, disfunción de la proteína RTFQ, mutación de la región 7q31 del cromosoma 7)¹⁴. Esta ambigüedad en los significados tiene consecuencias para los usuarios de los SI. La información localizada puede ser inadecuada tanto por las pérdidas o la redundancia que ocasiona al consultar las bases de datos documentales como por la inexactitud de las traducciones de los términos en los diferentes idiomas¹⁰. En otras palabras, se precisan SI que compartan el significado de los términos y no exclusivamente la normalización de la entrada/salida de los datos. Una estrategia adoptada para superar estas limitaciones es utilizar ontologías computacionales.

Una ontología es un recurso constituido por un conjunto de conceptos organizados por sus relaciones¹⁵. Los conceptos y las relaciones incluidas deben describir el conocimiento consensuado y aceptado de un área o dominio de interés para ser utilizado tanto por los humanos como por las máquinas^{15,16}. Este recurso permite que los SI trabajen automáticamente a partir del significado de los signos lingüísticos (términos) y sus combinaciones. En la figura 1 se incluye una posible descripción del concepto «corazón», en una ontología compartida por un SI.

En la actualidad, las ontologías se están aplicando en áreas heterogéneas¹⁷. Aunque quizá se las conoce más por su papel en el desarrollo de nuevos servicios en la web basados en la descripción del significado de los contenidos de las sedes o portales de internet (web semántica)¹⁸, también se están utilizando para el desarrollo de mecanismos que faciliten la comunicación entre las personas y las máquinas por medio del lenguaje natural (procesamiento del lenguaje natural [PLN]). En el contexto sanitario se están utilizando cada vez más para diversas tareas, como la de recuperación de información, la de búsqueda de respuesta en frag-

Figura 1. Descripción parcial del concepto «corazón», en una ontología compartida por un sistema de información.



mentos de texto que resuelven preguntas y la de extracción de información desde textos narrativos y no desde los campos estructurados de las bases de datos, o bien para la clasificación de documentos de forma automática^{13,19,20}.

En este trabajo se realiza un análisis cualitativo de las terminologías, más referidas en el ámbito de la salud, para determinar su grado de usabilidad a partir de sus características intrínsecas, y establecer los obstáculos que cada una de ellas presenta para la interoperabilidad semántica y la reutilización de su conocimiento como ontologías en los SI clínicos.

Métodos

Para el análisis se seleccionó la Clasificación Internacional de Enfermedades en su novena revisión, Modificación Clínica (CIE-9-MC)²¹, los Medical Subject Headings (MeSH)²², los Descriptores en Ciencias de la Salud (DeCs)²³, la Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT)²⁴, la General Architecture for Languages Encyclopaedias and Nomenclatures in Medicine (GALEN)²⁵ y el Unified Medical Language System (UMLS)²⁶. Los criterios de inclusión

de las terminologías fueron su relevancia tanto para el ámbito clínico, epidemiológico y documental como para el administrativo-económico, y que estuvieran expresadas en español o en alguno de los idiomas de la Unión Europea.

La usabilidad en este trabajo se entendió como la facilidad que presentaba el diseño de las terminologías para poder utilizarse como ontologías computacionales. Se comenzó por equiparar arbitrariamente las etiquetas terminológicas (rúbricas, descriptores, términos) a un concepto. Por medio de este procedimiento se constituyó el conjunto de conceptos que el vocabulario de cada terminología captura^{27,28}. No se diferenció entre conceptos simples (corazón = órgano) o complejos (cardiopatía = enfermedad + músculo + corazón). Las métricas analizadas fueron las siguientes:

1. Cobertura conceptual: finalidad, utilidad y amplitud del conjunto de conceptos representados.

2. Estructura jerárquica. La organización conceptual para representar el conocimiento del dominio puede ser:

– Monoaxial, o una jerarquía compuesta por una lista finita de categorías y subcategorías, elaborada siguiendo un orden lógico determinado por un criterio (una cualidad o un atributo) que aproxima los conceptos más específicos (hijos) y los agrupa dentro de los más generales (padre).

– Multiaxial. El dominio se estructura a través de varias jerarquías. Cada jerarquía o eje se desarrolla a partir de un atributo concreto que le da la naturaleza al eje. Este sistema organizativo permite que un mismo concepto hijo esté asociado a varios padres.

– Composicional. Se caracteriza por permitir la creación de conceptos complejos a partir de la combinación de los conceptos «primitivos» (elementales): clínicos y modificadores (agudo, izquierda, proximal...); no es necesario explicitar por adelantado todos los conceptos precisos de una especialidad o área de atención. Además, para prevenir la creación de conceptos complejos sin significación, se incluyen las reglas y restricciones de combinación de los conceptos.

– Léxico-conceptual. Organiza el léxico (vocabulario) a partir de su significado (tipos semánticos). Los conceptos incluidos y sus relaciones semánticas determinan el modelo de interpretación del dominio.

3. Granularidad conceptual de las categorías o clases representadas. Referida a la profundidad del detalle o especificidad de los conceptos representados.

4. Relaciones conceptuales. El tipo de enlaces existentes, bien por la dependencia jerárquica (implícita o explícita del enlace «es un») o por los enlaces semánticos detallados («es causado», «se localiza»...), que determinan o restringen las cualidades heredadas por los conceptos inferiores de los superiores.

5. Grado de formalización. Referido al lenguaje utilizado para la representación, entendiendo que éste debe ser inteligible y computable por el ordenador. Las posibilidades para su menor o mayor reutilización informática son: a) informal, o expresadas en cualquier lenguaje natural; b) semiinformal, o expresadas en lenguaje natural de forma estructurada y restrictiva; c) semiformal, expresadas en lenguaje artificial formalmente definido, y d) formal, al estar declaradas exclusivamente con lenguajes lógico-matemáticos por la definición metódica de una semántica formal, teoremas y pruebas de validez que suprimen la ambigüedad y aseguran la consistencia de la representación²⁹.

Resultados

Los datos sobre las características cualitativas de cada una de las terminologías analizadas desde una perspectiva ontológica se muestran en la tabla 1. En todas las terminologías hay coincidencia con algunos conceptos, independientemente de la amplitud de la cobertura, lo cual sólo manifiesta la superposición existente en primera instancia. Ahora bien, la interpretación del concepto depende del nivel de granularidad que se le otorga, el cual, a su vez, depende de su localización en la estructura jerárquica a la que pertenece: nivel superior (generalización), intermedio o inferior (especificidad). Así, por ejemplo, en la CIE-9-MC la «tetralogía de Fallot» y el «defecto del tabique ventricular» son conceptos hermanos (igual nivel de especificación); en SNO-MED, «*ventricular septal defect*» está como concepto padre y «tetralogía de Fallot» como concepto hijo. En cambio, en GALEN la «tetralogía de Fallot» se identifica como una instancia (casos concretos) del concepto «*congenital ventricular septal defect*», mientras que en MeSH y DeCS la «tetralogía de Fallot» se considera un hermano de «defectos del septum» (tabla 2), lo cual muestra la heterogeneidad en la representación del conocimiento aceptado.

Discusión

De acuerdo con la literatura médica²⁹⁻³¹, todas las terminologías aquí analizadas pueden considerarse inicialmente como prototipos para diseñar ontologías ligeras, ya que facilitan una conceptualización aceptada del dominio y una organización jerárquica de los conceptos. La CIE-9-MC responde a los criterios de este tipo de ontologías. Sin embargo, esta terminología sólo incluye los conceptos en una estructura taxonómica, donde el enlace «es un» está implícito. De hecho, algunos de los conceptos hijos son «un tipo» o especi-

Tabla 1. Características de las terminologías y de los componentes analizados desde la perspectiva de una ontología

| Terminologías y componentes analizados | Características cualitativas | | | | |
|---|--|--|---|--|---|
| | Cobertura conceptual | Estructura jerárquica | Granularidad conceptual | Relaciones conceptuales | Grado de formalización |
| CIE-9-MC Lista tabular e índice alfabético | Conceptualización de diagnósticos médicos (en función de la etiología y las manifestaciones), así como de procedimientos diagnósticos y terapéuticos Funcionalidad concreta de estandarización de los diagnósticos médicos | Monoaxial: contiene 17 subcategorías mayores (clases principales) para los diagnósticos cuyos códigos numéricos son consecutivos (desde 001 hasta 999.9). Aunque algunas clases son representadas por 5 dígitos (mayor profundidad de especificidad) Los procedimientos se representan por 16 subcategorías (en función de su localización anatómica) con códigos consecutivos (de 00 hasta 99.99) | Cada rúbrica terminológica es equivalente a una clase y se identifica por un código numérico exclusivo. La estructura sintáctica que predomina en las etiquetas terminológicas son los sintagmas nominales y los preposicionales; también destacan las partículas de coordinación y negación (exclusión conceptual). Las etiquetas que representan conceptos simples son casi anecdóticas. Hasta las rúbricas enunciadas por un término simple representan conceptos complejos (apendicitis: concepto de localización más concepto de proceso inflamatorio) | Las relaciones se derivan de la propia dependencia jerárquica en la que se estructura la clasificación. Cualquier categoría de nivel inferior responde a relación de tipo «es un», aunque en unos casos se refiere a ser un tipo y en otros a ser una parte de la clase superior (referencia a localización) | No conlleva ningún nivel de formalización, al ser expresada exclusivamente en lenguaje natural. Desde este punto de vista, se puede entender como un catálogo de términos normalizados. Aunque, si se atiende a la información complementaria de las notas de alcance para su aplicación y las situaciones de inclusión y exclusión, la podemos entender como un catálogo glosado |
| MeSH Listado alfabético anotado y lista jerárquica | Su cobertura y funcionalidad se enfocan a la tarea concreta de recopilar el vocabulario que permite traducir los conceptos incluidos en una publicación y de forma normalizada para la catalogación e indexación, así como para la posterior recuperación de los artículos indexados en la base de datos MEDLINE | Multiaxial: está integrada por 15 ejes o jerarquías que desarrollan clasificaciones independientes e identificadas por una letra distintiva. Cada eje equivale a una categoría principal. Las subcategorías, en cada una de ellas, están asociadas a un código alfanumérico (de lo más genérico a lo más específico) Hay subcategorías que tienen más de un código, cada uno de ellos indica la posición del término en varios de los ejes o bien las diferentes ubicaciones dentro del mismo eje | Los conceptos representados van desde simples a complejos. Las etiquetas terminológicas que los representan tienen estructuras sintácticas constituidas por formas simples (unitérminos, no excluyen la representación de conceptos complejos) o compuestas por más de un término (descriptores precoordinados que son tanto de tipo nominal como preposicional) para dar mayor especificidad del concepto representado | Hay tres tipos especificados de relaciones semánticas. Las relaciones de «equivalencia» señalan los sinónimos o cuasisinónimos por medio del verbo «see» (referencia cruzada que encamina hacia el término aceptado para la representación del concepto). La de «jerarquía» señala la existencia de una relación de orden y, casi exclusiva, del tipo hiperonimia/hiponimia respecto a un descriptor secundario o subordinado, y se identifica con la expresión «see under». La relación de «asociación» indica la existencia de una vinculación temática, | Es un catálogo glosado expresado, de forma estructurada, en lenguaje natural. Es una ontología de tipo semiinformal por el grado de formalización en que se expresa |

(Continúa en pág. siguiente)

Tabla 1. Características de las terminologías y de los componentes analizados desde la perspectiva de una ontología (continuación)

| Terminologías y componentes analizados | Características cualitativas | | | | |
|---|---|--|--|---|---|
| | Cobertura conceptual | Estructura jerárquica | Granularidad conceptual | Relaciones conceptuales | Grado de formalización |
| | | | | tanto en el mismo eje como otro eje diferente, aunque sin relación de equivalencia ni jerárquica, y se señala con « <i>see related</i> » Además, está la relación jerárquica «es un», la cual se deriva de la subordinación. Se puede interpretar según el caso como «es un tipo» o como «es una parte» por la falta de explicitación | |
| DeCS: listado anotado y lista jerárquica | Tiene una acción concreta, delimitada por el conjunto finito de conceptos utilizados en el proceso de catalogación e indexación para la base de datos LILACS y para la recuperación de las publicaciones científicas indexadas en la base de datos MEDLINE y LILACS | Multiaxial: está integrada por 17 ejes. De ellos, 15 son idénticos a los de MeSH (mantienen la misma estructura y codificación) y 2 ejes propios: homeopatía y salud pública. En estos últimos, su sistema de organización y codificación se deriva del utilizado por el MeSH | Al igual que en MeSH, aunque destaca por la inclusión de etiquetas terminológicas trilingües (inglés, español y portugués) | En igualdad de condiciones que MeSH | Catálogo glosado trilingüe y de tipo semiinformal como MeSH |
| SNOMED CT | Contiene conceptos que permiten la captura de los datos clínicos y la recuperación de información médica y de enfermería de las historias clínicas y de las interconsultas entre profesionales (primaria y hospitalización) Se orienta hacia diversas tareas específicas, centradas en el desarrollo de aplicaciones de historias clínicas electrónicas, análisis de resultados clínicos y apoyo a la toma de decisiones | Composicional: los conceptos representados se organizan en 18 ejes (han aumentado con el tiempo y lo pueden seguir haciendo para adaptarse a las demandas y dar solución a los problemas reales). De ellos, 3 son categorías referidas a los tipos de relación entre los conceptos pertenecientes a los otros ejes | Hay un término preferente para denominar a cada concepto, y los posibles términos que se utilizan para su denominación son agrupados (<i>cluster</i> de sinónimos utilizados en la práctica o incluidos en otras terminologías). Los conceptos representados tienen un único significado. Su descripción se realiza a través de la dependencia de los conceptos específicos hasta el concepto raíz o padre (superiores) y se complementa con información textual La versión en inglés se ha traducido al español y al alemán | En el eje « <i>Attribute</i> » se localizan todos los conceptos, que pueden usarse para establecer un tipo de relación semántica («está asociado»). El eje « <i>Qualifier value</i> » contiene los posibles valores de una cualidad; un valor puede usarse en una definición de relación o como un calificador de un concepto («agudo», «remisión»). El tercer eje se denomina « <i>Special concept</i> » y tiene conceptos que comparten una característica común, pero no forma parte de la actual jerarquía lógica de categorías semánticas. En este eje hay dos subcategorías | El lenguaje utilizado para la representación de los conceptos es formal, basado en una descripción lógica que permite hacer una descripción de la semántica de los conceptos Además, en la actualidad está alineado con los estándares más relevantes en el campo de las aplicaciones médicas, como HL7, DICOM, XML, ANSI e ISO para el intercambio de datos |

Tabla 1. Características de las terminologías y de los componentes analizados desde la perspectiva de una ontología (continuación)

| Terminologías y componentes analizados | Características cualitativas | | | | |
|---|---|--|---|--|--|
| | Cobertura conceptual | Estructura jerárquica | Granularidad conceptual | Relaciones conceptuales | Grado de formalización |
| | | | | principales: «concepto no actual» y «concepto de navegación» Además, el resto de jerarquías mantienen, cada una de ellas, una relación jerárquica del tipo « <i>is a</i> ». Es decir, el resto de conceptos se entienden como un subtipo del «concepto raíz» o de un «concepto de nivel superior» | |
| GALEN CRM (Common Referente Model), GRAIL (GALEN Representation And Integration Language) | Su finalidad es la representación formal de terminologías médicas para los distintos usos clínicos y de codificación Se desarrolló para apoyar la construcción de sistemas clínicos que podrían representar y manipular la información clínica sobre los pacientes, en un nivel de detalle suficiente, para apoyar la rutina del cuidado clínico cotidiano (médico y de enfermería) al separar el modelo conceptual del modelo de uso o del idioma. Su utilidad se enfoca a representar conceptos complejos y, posteriormente, generar clasificaciones automáticas según la perspectiva de interés | Composicional: la estructura jerárquica primaria de CRM es una red semántica composicional y generativa de lenguaje de representación de conocimiento Se estructura a partir de 4 jerarquías referidas a conceptos elementales clínicos, a conceptos elementales de las propiedades de los conceptos clínicos y las relaciones En cada jerarquía, las subcategorías sólo tienen una categoría superior. Aunque hay relaciones entre categorías o subcategorías pertenecientes a otras jerarquías para facilitar la asociación entre los conceptos y restricción en la composición de conceptos complejos En su estructura organizativa incorpora jerarquías secundarias superpuestas, cuyo propósito es distinguir entre los conceptos vinculados con enfermedad/trastorno y con función química o con producto químico | Sólo están representados conceptos médicos sensibles, es decir, elementales y de alta especificidad. En el CRM de GALEN se representan los conceptos a través de expresiones formales de GRAIL (<i>Process which hasSpecificOutcome Ulcer) name UlcerationProcess</i> . Estas expresiones están constituidas por dos conceptos primitivos o elementales y su relación semántica; con ellas se definen nuevas categorías (conceptos complejos) Las etiquetas para denominar los conceptos son únicas y con una ambigüedad muy reducida La estructura subyacente interlingua del CRM le permite dar un servicio multilingüe: inglés, francés, italiano, holandés, alemán, finlandés y sueco, sin precisar traducción | Cuenta con una rica jerarquía de relaciones semánticas, especificadas en 592 tipos para expresar conceptos complejos (internamente se denomina «atributos»). Destaca la diferenciación de las relaciones de tipo partitivo desarrolladas y que permiten distinguir entre divisiones o partes de superficies, sólidos, capas, lineales, componentes estructurales específicos y realizados específicamente Los conceptos de las subcategorías presentes tanto en la estructura primaria como en la secundaria de CRM mantienen una relación con las categorías superiores por medio del enlace jerárquico «es un tipo de» Los enlaces jerárquicos establecidos como semánticos han sido verificados para cumplir con la propiedad transitiva. Por tanto, siempre se puede admitir que las propiedades de los conceptos superiores son heredadas por los | Su grado de formalización es máximo, al utilizar la lógica de descripción, que es la base de GRAIL y, a su vez, se fundamenta en el lenguaje de representación de conocimiento KL-ONE Con GRAIL se ha construido el CRM, ya que posee los modeladores que especifican el número de conceptos primitivos, la descripción de sus características y las reglas para la creación de conceptos con términos adicionales o con otros ya presentes en el modelo, así como diferenciar entre conceptos e instancias (casos concretos que representan un concepto) |

(Continúa en pág. siguiente)

Tabla 1. Características de las terminologías y de los componentes analizados desde la perspectiva de una ontología (Continuación)

| Terminologías y componentes analizados | Características cualitativas | | | | |
|---|---|--|--|---|---|
| | Cobertura conceptual | Estructura jerárquica | Granularidad conceptual | Relaciones conceptuales | Grado de formalización |
| UMLS Metathesaurus y Semantic Network (Red Semántica) | <p>Contiene la representación conceptual más amplia del dominio, debido a la combinación y el alcance de sus diversos vocabularios fuente. En total contiene unos 102 vocabularios; aunque muchos de ellos están reiterados por varias revisiones, versiones o traducciones en diversas lenguas</p> <p>Su funcionalidad se orienta a tareas genéricas que se enfocan en la recuperación e integración de información biomédica electrónica desde o para sistemas de información dispares</p> <p>En la actualidad incluye 6 vocabularios fuente en español</p> | <p>En el Metathesaurus se mantiene la estructura jerárquica de cada uno de los vocabularios fuente incluidos</p> <p>La Red Semántica de UMLS se estructura a partir de un conjunto de categorías básicas o tipos semánticos (en total 135). La organización de red está determinada por el tipo de relaciones existente entre las categorías (jerárquicas y semánticas). Todos los conceptos representados en el Metathesaurus son asignados a la categoría semántica más específica de la Red</p> | <p>El Metathesaurus es un repositorio de conceptos, donde se preservan los significados, los atributos, las conexiones jerárquicas y otras relaciones entre los términos presentes en cada uno de los vocabularios fuente que se han incorporado (incluso si tienen una definición textual). En él se localiza la agrupación de los términos con equivalente significado (<i>clusters</i>). El <i>cluster</i> es el que define el concepto y es identificado por una etiqueta preferente para su denominación</p> <p>En la Red Semántica el nivel de granularidad es también variable. Hay zonas de la red con gran número de subcategorías tanto en profundidad (categorías hijo) como en amplitud (subcategorías disjuntas o hermanos). Pero hay otras en que sólo aparecen categorías muy generales. Esta organización tiene implicaciones para la interpretación del significado de los conceptos que le ha sido asignado desde el Metathesaurus. Cada concepto de la Red incluye una descripción o definición textual</p> | <p>conceptos subordinados inferiores, aunque dichas propiedades no estén declaradas explícitamente (por un principio de economía computacional)</p> <p>El enlace primario en la Red es el jerárquico, del tipo «<i>is a</i>» (es un). Además, hay una taxonomía de 54 tipos de relaciones semánticas. Las categorías principales son: '<i>physically_related_to</i>', '<i>spatially_related_to</i>', '<i>temporally_related_to</i>', '<i>functionally_related_to</i>' y '<i>conceptually_related_to</i></p> <p>Las relaciones semánticas son declaradas entre los tipos superiores de la Red cuando es posible y se heredan por todos sus tipos semánticos hijo por medio del enlace «<i>is a</i>». Aunque las relaciones semánticas declaradas pueden mantenerse o rescindirarse entre cualquier par de conceptos tanto de la Red como entre la Red y el Metathesaurus</p> | <p>Es una ontología semiformal al ser expresada en un lenguaje estructurado</p> |

Tabla 2. Representación del concepto «tetralogía de Fallot» en las diversas estructuras jerárquicas

| CIE-9-MC | MeSH |
|---|--|
| 745.2 Tetralogía de Fallot | <i>Definition:</i> A combination of congenital cardiac defects consisting of pulmonary stenosis, interventricular septal defects, dextroposition of the aorta so that it overrides the interventricular septum and receives venous as well as arterial blood, and right ventricular hypertrophy |
| <i>Definición:</i> Defecto del tabique ventricular con estenosis o atresia pulmonar, dextroposición de aorta e hipertrofia del ventrículo derecho | |
| <i>Incluye:</i> pentalogía de Fallot | |
| <i>Excluye:</i> tríada de Fallot [746.09] | |
| Enfermedades y lesiones [no codificado, del 001 al 999.9] | Diseases (MeSH Category) [C] |
| Anomalías congénitas [no codificado, del 740 al 759] | Cardiovascular diseases [C14] |
| Anomalías del bulbo arterioso y anomalías del cierre septal cardíaco [745] | Cardiovascular abnormalities [C14.240] |
| Tetralogía de Fallot [745.2] | Heart defects, congenital [C14.240.400] |
| Defecto de tabique ventricular [745.4] | Tetralogy of Fallot [C14.240.400.849] |
| | Heart septal defects [C14.240.400.560] |
| | Diseases (MeSH Category) [C] |
| | Cardiovascular diseases [C14] |
| | Heart diseases [C14.280] |
| | Heart defects, congenital [C14.280.400] |
| | Tetralogy of Fallot [C14.280.400.849] |
| | Heart septal defects [C14.280.400.560] |
| | Diseases (MeSH Category) [C] |
| | Congenital, hereditary, and neonatal diseases and abnormalities [C16] |
| | Abnormalities [C16.131] |
| | Cardiovascular abnormalities [C16.131.240] |
| | Heart defects, congenital [C16.131.240.400] |
| | Tetralogy of Fallot [C16.131.240.400.849] |
| | Heart septal defects [C16.131.240.400.560] |
| SNOMED ^a | DeCS |
| Tetralogy of Fallot | <i>Descriptor inglés:</i> Tetralogy of Fallot |
| <i>Conceptos padre:</i> | <i>Descriptor español:</i> Tetralogía de Fallot |
| Duplicate concept [] | <i>Descriptor portugués:</i> Tetralogia de Fallot |
| Overriding aorta [] | Categoría: C14.240.400.849 |
| Congenital abnormality of ventricles and ventricular septum [] | C14.280.400.849 |
| Right ventricular hypertrophy [] | C16.131.240.400.849 |
| Pulmonic valve stenosis [] | <i>Definición español:</i> Combinación de defectos cardíacos congénitos constituida por estenosis pulmonar, defectos septales interventriculares, dextroposición de la aorta de forma que cabalga sobre el tabique interventricular y que recibe sangre tanto arterial como venosa, e hipertrofia del ventrículo derecho |
| Overriding aorta [] | |
| Congenital cardiomegaly [] | Enfermedades |
| Ventricular septal defect [] | Enfermedades cardiovasculares |
| <i>Conceptos hijo:</i> | Anomalías cardiovasculares |
| Dextroposition of aorta in Fallot' tetralogy [] | Cardiopatías congénitas |
| Pentalogy of Fallot [] | Tetralogía de Fallot |
| Tetralogy of Fallot NOS [] | Defectos del septum |
| Tetralogy of Fallot with pulmonary atresia [] | Enfermedades |
| Tetralogy of Fallot with pulmonary stenosis [] | Enfermedades cardiovasculares |
| Tetralogy of Fallot, unspecified [] | Cardiopatías |
| Ventricular septal defect in Fallot's tetralogy [] | Cardiopatías congénitas |
| Tetralogy of Fallot with absent pulmonary valve [] | Tetralogía de Fallot |
| | Defectos del septum |
| | Enfermedades |
| | Enfermedades neonatales congénitas y hereditarias, y anomalías |
| | Anomalías |
| | Anomalías cardiovasculares |
| | Cardiopatías congénitas |
| | Tetralogía de Fallot |
| | Defectos del septum |

(Continúa en pág. siguiente)

Tabla 2. Representación del concepto «tetralogía de Fallot» en las diversas estructuras jerárquicas (continuación)

| GALEN | UMLS |
|--|--|
| <p><i>Definition:</i> Tetralogy of Fallot</p> <p>Clinical Situation <i>which</i> shows (presence <i>wich</i> isStateOf (Syndrome <i>which</i> < includes VentricularSeptalDefect includes PulmonaryStenosis includes RightVentricularHypertrophy includes AorticOverriding>)</p> <p>La instancia es asumida por el concepto definido como: ClinicalSituation <i>wich</i> shows (presence <i>wich</i> isStateOfVentricularSeptalDefect)</p> <p>Por lo que en la jerarquía secundaria de enfermedades «tetralogy of Fallot» estaría subordinada a «ventricular septal defects»</p> <p>PathologicalPhenomenon BodySystemPathology GenitalUrinaryPathology CardiovascularPathology HeartDisease HeartAndLungPathology CardiacSeptalDefect AtrialSeptalDefect CongenitalAtrialSeptalDefect VentricularSeptalDefect CongenitalVentricularSeptalDefect</p> | <p>Metathesaurus:</p> <p><i>Concept:</i> Tetralogy of Fallot. CUI: C0039685</p> <p><i>Definition:</i> A combination of congenital cardiac defects consisting of pulmonary stenosis, interventricular septal defects, dextroposition of the aorta so that it overrides the interventricular septum and receives venous as well as arterial blood, and right ventricular hypertrophy. (MeSH) (CRISP Thesaurus)</p> <p><i>Synonyms (cluster):</i> Tetralogy of Fallot Subpulmonic stenosis, ventricular septal defect, overriding aorta, and right ventricular hypertrophy Tetralogy of Fallot, unspecified Tetralogy of Fallot, unspecified (disorder) Tetralogy of Fallot NOS (disorder) Tetralogy of Fallot unspecified TOF-tetralogy of Fallot Ventricular septal defect with pulmonary stenosis or atresia, dextroposition of aorta, and hypertrophy of right ventricle</p> <p>Semantic Network:</p> <p><i>Semantic type</i> congenital abnormality. TUI: T019</p> <p><i>Definition:</i> An abnormal structure, or one that is abnormal in size or location, present at birth or evolving over time as a result of a defect in embryogenesis</p> <p>Entity Physical object Anatomical structure Anatomical abnormality Congenital abnormality</p> <p>El tipo semántico «congenital abnormality» tiene especificadas 133 relaciones con otros tipos semánticos de la Red. A su vez, otros tipos se relacionan con «congenital abnormality» por medio de 122 relaciones especificadas. En concreto: «Congenital abnormality» se relaciona con «Disease or syndrome» por medio de 4 relaciones semánticas: co-occurs_with, location_of, manifestation_of y result_of. En cambio, «Disease or syndrome» se relaciona con «Congenital abnormality» por medio de 3 relaciones: co-occurs_with, complicates y result_of</p> |

^a Información localizada en UMLS Knowledge Source Server, versión 2007AA. Disponible en: http://umisks.nlm.nih.gov/kss/servlet/Turbins/template/admin,user,HSS_login.vm

ficación de los conceptos padres, y en cambio, en otros están asociados por ser «una parte» del concepto superior. En otras palabras, los conceptos se encuentran definidos por dos tipos distintos de relaciones no explicitadas, lo que acarrea ambigüedad para la interpretación de su significado. En la actualidad los conceptos representados en la CIE-9-MC sólo pueden ser procesados informáticamente comparando los signos lingüísticos con un patrón preestablecido sin significación semántica (normalización terminológica). Dicho inconveniente sería superable si se contara con una formalización computacional para su explicitación.

Un modelo de conocimiento con algunas restricciones interpretativas es el que se construye con los tesauros³¹. Tanto los MeSH como los DeCS propor-

cionan una limitación a la interpretación semántica debido a los tipos de relaciones que incluyen en la estructura organizativa de los conceptos (de equivalencia, de jerarquía y de asociación). Sin embargo, la jerarquización de los conceptos a partir de la explicitación de sus relaciones no es suficiente para reducir la ambigüedad de la interpretación. También se precisa establecer las propiedades de los conceptos para delimitarlos. El lenguaje semiinformal que se utiliza para formalizar la estructura de los MeSH y de los DeCS no especifica las propiedades de cada concepto; sólo se incluyen aclaraciones textuales para los catalogadores. Esta debilidad hace imposible resolver aspectos relacionados con la ambigüedad cuando los conceptos no son verdaderos conceptos hijos o no son necesaria-

mente hermanos (tetratología de Fallot y defectos del septum). Ambos recursos sólo pueden ser considerados como ontologías ligeras por su descripción limitada de los conceptos y sus relaciones, lo cual es su inconveniente para la reutilización de su conocimiento³¹.

SNOMED y GALEN han incluido la definición de las propiedades de los conceptos, añadiendo restricciones explícitas y verificando la consistencia de su taxonomía al incluir verdaderos subconceptos y conceptos hermanos. Las restricciones entre los conceptos relacionados incrementan el control de la ambigüedad; con ellas se limita la transmisión de las propiedades de los conceptos padres a los conceptos hijos, y sólo heredan las propiedades adecuadas²⁹⁻³¹. Ahora bien, hay diferencias entre ellas: GALEN es un recurso abierto y SNOMED es un recurso propietario. El acceso limitado a SNOMED sesga el análisis realizado y es un aspecto que hay que valorar por el coste económico que conllevaría su reutilización. No obstante, ambas han sido expresadas con descripciones lógicas que las dotan de mayor expresividad para la representación del dominio; aunque, con la información localizada, sólo GALEN diferencia entre concepto específico e instancia (tabla 2), si bien ambas pueden considerarse como ontologías pesadas³¹, al poder utilizarse también para los procesos de razonamiento (inferencia y descubrimiento de nuevo conocimiento) definidos en la base de conocimiento¹⁵.

La situación de UMLS es intermedia, y su formalismo de representación es semiformal. Este recurso sólo lo podemos entender como una ontología ligera con la representación más amplia del dominio. No obstante, desde el punto de vista de las restricciones semánticas, hay diferencias sustanciales con MeSH y DeCS. El UMLS tiene una taxonomía de relaciones semánticas compuesta por 54 tipos diferentes, que amplían o restringen la descripción de los conceptos. En él no se añaden axiomas para la deducción, pero se puede realizar la inferencia por el seguimiento de las asociaciones semánticas preestablecidas dentro de la Red Semántica, y complementadas por las existentes en el Metathesaurus. Esta limitación le da la versatilidad para adaptarlo a diferentes SI o tareas concretas para el PLN. Por ejemplo, la «tetratología de Fallot» en UMLS es «una enfermedad/síndrome resultado de una anomalía congénita, la cual es un tipo de anomalía anatómica...» (*«tetratology of Fallot» is a «congenital abnormality» is a «anatomical abnormality», «disease or syndrome» result of «congenital abnormality»...*), y dicho concepto puede representarse textualmente por medio de diversas etiquetas o sinónimos, como la abreviatura TOF (tabla 2).

El análisis realizado pone de manifiesto las posibilidades y las limitaciones que deben considerarse para lograr que los SI, emisor y receptor, entiendan de forma común el dato transferido (términos clínicos) a partir de la reutilización del conocimiento representado en las terminologías. Tanto los MeSH y los DeCS como el UMLS

tienen una formalización reducida que permite controlar parcialmente la ambigüedad de los conceptos y, aunque hay diferencias significativas entre ellas, puede reutilizarse para el desarrollo de ontologías ligeras. Una ontología ligera permite la integración y la interoperabilidad semántica de fuentes de información heterogéneas al trabajar con conceptos que pueden asociarse a diversas etiquetas terminológicas y responder a diferentes intereses simultáneamente: clínicos, investigación, epidemiológicos, económicos y formativos. Las nomenclaturas SNOMED y GALEN cuentan con una mayor expresividad para especificar los conceptos, las propiedades, las restricciones y el tipo de relaciones incluidas. En ambos casos podrían reutilizarse para el diseño de ontologías pesadas, debido a que su formalismo declarativo basado en descripciones lógicas les posibilita el descubrimiento por inferencia de nueva información. Las ontologías pesadas son idóneas para su inclusión en los SI que ayudan a la toma de decisiones diagnósticas o de tratamientos^{32,33}, y como recurso para el PLN. La CIE-9-MC es la única de las terminologías analizadas cuya reutilización como ontología conllevaría el mismo esfuerzo e inversión que si se creara una ontología nueva.

Las limitaciones del trabajo presentado son las restricciones propias de los estudios exploratorios; asimismo, el análisis se ha centrado específicamente en el conocimiento declarado y no en la exactitud teórica de la representación que proporciona cada una de las terminologías estudiadas.

Como conclusión, el análisis realizado aconseja replantear la metodología aplicada en España. El mantenimiento de las terminologías de subsistemas integrados plantea inconvenientes para la utilización adecuada del SI clínico. Por una parte, el uso de terminologías diferentes da lugar a pérdidas de datos significativos para el sistema y, por otra, los profesionales precisan un entrenamiento previo¹¹. Además, cuando el profesional consulta información para apoyar sus decisiones, se encuentra con sistemas de clasificación heterogéneos o con la necesidad de aplicar términos diferentes al interrogar a cada sistema de recuperación de información; por ejemplo, en el caso de precisar localizar historias clínicas de casos similares, datos de mortalidad en el Instituto Nacional de Estadística y evidencias en MEDLINE e IME. Trabajar con una ontología compartida por todos los subsistemas que conforman el SI clínico ha mostrado ventajas prácticas para los usuarios³⁴.

Además, las expectativas para el desarrollo de ontologías se incrementan si consideramos la gran cantidad de información textual que se introduce en la HCE y que no es procesable en la actualidad, salvo por la codificación manual en CIE-9-MC. Asimismo, contar con una ontología compartida en el SI clínico facilitaría la realización de diversas tareas automáticas: la clasifi-

cación de documentos, la recuperación de información o la extracción de datos clínicos presentes en las notas de evolución o en los informes al alta (p. ej., para hacerlos anónimos y utilizarlos en ámbitos de docencia e investigación). Estos beneficios para el usuario repercutirían en el propio sistema asistencial, al igual que toda nueva tecnología sanitaria que se incluye, aunque este beneficio se encuentra supeditado a la adaptación de los perfiles profesionales de los documentalistas en los centros hospitalarios.

Por otro lado, hay que tener en cuenta que generar y mantener terminologías es costoso, tanto en recursos humanos como en tiempo¹⁰. Estos costes pueden llegar a ser comparables a los que se producen para el desarrollo de nuevas ontologías, pero se reducen con respecto a su mantenimiento^{14,17}; aunque pueden optimizarse si se opta por la reutilización de los recursos disponibles. Todos estos factores deberían ser valorados por los gestores en la toma de decisiones sobre la innovación y el desarrollo de SI clínicos.

Agradecimientos

Este trabajo forma parte del proyecto «Desarrollo de una ontología multilingüe para el dominio médico-farmacológico» (PI051438), financiado por el Fondo de Investigación Sanitaria en su convocatoria de 2005.

Bibliografía

1. Pons JMV, Castells X. Conocimiento, evaluación y práctica: introducción a una serie de artículos sobre evaluación de tecnologías médicas. *Gac Sanit.* 2003;5:420-1.
2. Castiel DL, Álvarez-Dardet C. Las tecnologías de la información y la comunicación en salud pública: las precariedades del exceso. *Rev Esp Salud Pública.* 2005;79:331-7.
3. Monteagudo Peña JL, Hernández Salvador C, García-López F. Metodología de introducción de servicios de e-Salud para el seguimiento y control de pacientes crónicos. *Rev Esp Salud Pública.* 2004;78:571-81.
4. Perpiñá Tordera M. Conocimiento y práctica. *Med Clin (Barc).* 2005;124:215-6.
5. Itami H. *Mobilizing invisible assets.* Cambridge: Harvard University Press; 1987.
6. Nonaka I. A dynamic theory of organizational knowledge creation. *Organization Science.* 1994;5:14-37.
7. Stead WW, Miller RA, Musen MA, Hersh WR. Integration and beyond: linking information from disparate source and into workflow. *J Am Med Inform Assoc.* 2000;7:135-45.
8. Ciccarese P, Caffi E, Quaglini S, Stefanelli M. Architectures and tools for innovative health information systems: the Guide Project. *Int J Med Inform.* 2005;74:553-62.
9. Escolar Castellón F, Carnicero Giménez de Azcárate J, coordinadores. *El sistema integrado de información clínica. Informe 6 SEIS [monografía electrónica].* Madrid: Sociedad Española de Informática de la Salud; 2004 [citado 15 May 2005]. Disponible en: <http://www.conganat.org/SEIS/informes/2004/PDF/informeseis2004.pdf>
10. Zweigenbaum P. Encoder l'information médicale: des terminologies aux systèmes de représentation des connaissances. *Innovation Stratégique en Information de Santé.* 1999;2-3:27-47.
11. Orueta JF, Urraca J, Berraondo I, Darpón J. ¿Es factible que los médicos de primaria utilicen CIE-9-MC? Calidad de la codificación de diagnósticos en las historias clínicas informatizadas. *Gac Sanit.* 2006;20:194-201.
12. Codina C, Corominas N, Roca M, Tuset M, Del Cacho E, Soy D, et al. Estudio comparativo de la aplicación de un sistema experto en la prescripción de medicamentos. *Med Clin (Barc).* 1997;109:538-41.
13. Arribas P, Cirera E, Tristán Polo M. Buscando una aguja en un pajar: las técnicas de conexión de registros en los sistemas de información sanitaria. *Med Clin (Barc).* 2004;122 Supl 1:16-20.
14. Pisanelli DM, Gangemi A. If ontology is the solution, what is the problem? En: Pisanelli DM, editor. *Ontologies in medicine. Studies in health technology and informatics. Volumen 102.* Amsterdam: IOS Press; 2004. p. 1-19.
15. Swartout B, Patil R, Knight K, Russ T. Toward distributed use of large-scale ontologies. En: Farquhar A, Gruninger M, Gómez-Pérez A, Uschold M, van der Vet P, editores. *AAAI-97: Spring Symposium Series on Ontological Engineering.* California: Stanford University; 1997. p. 138-48.
16. Gruber TR. Toward principles for the design of ontologies used for knowledge sharing. En: Guarino N, Poli R, editores. *Formal ontology in conceptual analysis and knowledge representation.* Dordrecht: Kluwer Academic Publishers; 1993. Stanford University: Stanford Knowledge Systems Laboratory. Technical Report: KSL-93-04 [on line] [citado 10 Dic 2004]. Disponible en: <http://ksl-web.stanford.edu/knowledgesharing/papers/onto-design.rtf>
17. Guarino N, editor. *Formal ontology in information systems. Proceedings of FOIS'98 Amsterdam.* IOS Press; 1998. p. 3.
18. Berners-Lee T, Hendler J, Lassila O. The semantic web. *Scientific American.* 2001;284:34-43.
19. Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports. A study of natural language processing. *Ann Intern Med.* 1995;122:681-8.
20. Recuperación de respuestas en documentos digitalizados (R2D2). Proyecto TIC2003-07158-C04 financiado por el Ministerio de Ciencia y Tecnología 2003-2006. Alicante: Grupo de Investigación en Procesamiento del Lenguaje y Sistemas Informáticos, Universidad de Alicante [actualizado 14 Ene 2005; citado 16 Dic 2005]. Disponible en: <http://gplsi.dlsi.ua.es/r2d2/>
21. Clasificación Internacional de Enfermedades, 9.^a Revisión, Modificación Clínica (CIE-9-CM). 5.^a ed. Madrid: Ministerio de Sanidad y Consumo; 2006 [citado 16 Feb 2007]. Disponible en: <http://www.msc.es/estadEstudios/ecie9mc/webcie9mc/webcie9mc.htm>
22. Medical Subject Headings (MeSH), Introduction 2004. Bethesda: US National Library of Medicine (NLM) [citado 5 Dic 2004]. Disponible en: <http://www.nlm.nih.gov/mesh/introduction2004.html>
23. Descriptores en Ciencias de la Salud (DeCS), edición 2003. Biblioteca Virtual en Salud (bvs), Bireme [citado 5 Dic 2004]. Disponible en: <http://decs.bvs.br/E/decs2003e.htm>
24. What is SNOMED CT (Clinical Terms). SNOMED Internacional [citado 16 Feb 2007]. Disponible en: http://www.snomed.org/snomedct/what_is.html
25. The OpenGALEN Common Reference Model. OpenGALEN [citado 16 Feb 2007]. Disponible en: <http://www.opengalen.org/themodel/structure.html>

26. Unified Medical Language System (UMLS). Bethesda: US National Library of Medicine (NLM) [actualizado 22 Mar 2004; citado 5 Dic 2004]. Disponible en: http://www.nlm.nih.gov/research/umls/about_umls.html
27. Feliu J, Vivaldi J, Cabré MT. Ontologies: a review. Barcelona: Instituto Universitario de Lingüística Aplicada, Universitat Pompeu Fabra; 2002. Serie Informes n.º 34.
28. Chandrasekaran B, Josephson JR, Benjamins VR. What are ontologies, and why do we need them? *IEEE Intell Syst.* 1999;2:20-6.
29. Uschold M, Jasper R. A framework for understanding and classifying ontology applications. En: Benjamins VR, Chandrasekaran B, Gómez-Pérez A, Guarino N, Uschold M, editores. *Proceedings of the IJCAI-99: workshop on Ontologies and Problem-Solving Methods (KRR5)*; August 2, 1999; Stockholm, Sweden [citado 16 Feb 2007]. Disponible en: <http://www.cs.man.ac.uk/~horrocks/Teaching/cs646/Papers/uschold99.pdf>
30. Lassila O, McGuinness D. The role of frame-based representation on the Semantic Web. *Electronic Transactions on Artificial Intelligence [revista electrónica]*. 2001; 6(5) [citado 16 Dic 2005]. Disponible en: <http://www.ep.liu.se/ea/cis/2001/005/cis01005.pdf>
31. Gómez-Pérez A, Fernández-López M, Corcho O. *Ontological engineering*. London: Springer-Verlag; 2004. p. 8, 11.
32. Taboada M, González JL, Argüello M, Des J, Mira J, Martínez D. Adquisición de conocimiento médico para aplicaciones en telemedicina. En: González F, Zamarrón C, editores. *Telemedicina. Aplicaciones y nuevas tecnologías*. Santiago de Compostela: Sociedad Gallega de Telemedicina; 2004 [citado 27 Mar 2006]. Disponible en: <http://aiff.usc.es/~elchus/personal/download/xornadastaboada.pdf>
33. Sobrado FJ, Pikatza JM, Larburu IU, García JJ, López de Ipiña D. Towards a clinical practice guideline implementation for asthma treatment. En: Conejo R, Urretavizcaya M, Pérez de la Cruz JL, editores. *Current topics in artificial intelligence, lecture notes in artificial intelligence*. Vol 3040. Berlin: Springer-Verlag; 2004. p. 587-96.
34. Cimino JJ, Clayton PD, Hripcsak G, Jonson SB. Knowledge-based approaches to the maintenance of large controlled medical terminology. *J Am Med Inform Assoc.* 1994;1:35-50.