

INFRAESTRUCTURAS LINGÜÍSTICAS

Plan de impulso de las Tecnologías del Lenguaje

Maria Cecilia Perez Araujo

Julio/2018



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y EMPRESA

SECRETARÍA DE ESTADO
PARA EL AVANCE DIGITAL

Plan TL

Plan de Impulso de las
Tecnologías del Lenguaje



Este estudio ha sido realizado dentro del ámbito del Plan de Impulso de las Tecnologías del Lenguaje con financiación de la Secretaría de Estado para el Avance Digital, que no comparte necesariamente los contenidos expresados en el mismo. Dichos contenidos son responsabilidad exclusiva de sus autores.

Reservados todos los derechos. Se permite su copia y distribución por cualquier medio siempre que se mantenga el reconocimiento de sus autores, no se haga uso comercial de las obras y no se realice ninguna modificación de estas.

ÍNDICE

1. INTRODUCCIÓN	6
2. UMLS (UNIFIED MEDICAL LANGUAGE SYSTEM)	8
2.1 ¿EN QUÉ CONSISTE UMLS?	8
2.1.1 Metathesaurus	9
2.1.2 Semantic Network (Red semántica)	16
2.1.3 SPECIALIST Lexicon	18
2.1.4 MetamorphoSys	19
2.2 ESTRUCTURA DE BIBLIOTECA NACIONAL DE MEDICINA (NLH)	19
2.3 PRESUPUESTO PARA UMLS.....	24
3. SNOMED (GLOBAL STANDARDS FOR HEALTH TERMS)	26
3.1 COMPONENTES DE SNOMED CT	28
3.1.1 Conceptos.....	29
3.1.2 Descripciones.....	29
3.1.3 Relaciones.....	29
3.1.4 Modelo Lógico de SNOMED CT.....	30
3.1.5 Modelo Conceptual de SNOMED CT.....	34
3.2 LA INTERNATIONAL HEALTH TERMINOLOGY STANDARDS DEVELOPMENT ORGANISATION (IHTSDO) 36	
3.2.1 Estructura de IHTSDO	37
3.2.2 Desarrollo del contenido de SNOMED CT.....	39
4. INFRAESTRUCTURAS DE DATOS ESPACIALES (IDE)	42
4.1 OPEN GEOESPACIAL CONSORTIUM (OGC)	45
4.1.1 Estructura de OGC	46
4.1.2 Membrecía	48
4.1.3 Programas.....	50
4.1.4 Dominios.....	52
4.1.5 Estándares OGC	55
5. CONCLUSIONES.....	56
6. FUENTES Y BIBLIOGRAFÍA.....	59

ÍNDICE DE FIGURAS

FIGURA 1: FUNCIÓN BIOLÓGICA.....	17
FIGURA 2: PORCIÓN DE RED SEMÁNTICA.....	18
FIGURA 3: ESTRUCTURA DE LA BIBLIOTECA NACIONAL DE MEDICINA.....	22
FIGURA 4: ESTRUCTURA DE SNOMED CT.....	30
FIGURA 5: DESCRIPCIONES DE UN CONCEPTO.....	32
FIGURA 6: EJEMPLO DE UTILIZACIÓN DE RELACIONES DE ATRIBUTO.....	34
FIGURA 7: ESTRUCTURA DE IHTSDO.....	39
FIGURA 8: ESTRUCTURA DE OGC.....	48

ÍNDICE DE TABLAS

TABLA 1: CONCEPT, TERM, ATOM, AND STRING IDENTIFIERS..... 14

1. INTRODUCCIÓN

¿Qué es una infraestructura lingüística?

Si bien no se ha encontrado una definición de infraestructura lingüística, en las fuentes de información consultadas, de acuerdo a los diferentes ejemplos de ellas, y algunas definiciones que se han detectado en el ámbito de las infraestructuras de datos espaciales (IDE), por ejemplo, puede definirse una infraestructura lingüística como:

“el conjunto de datos, tecnologías y acuerdos políticos que permiten a ciudadanos e instituciones acceder de forma ubicua a recursos lingüísticos¹, almacenados en servidores institucionales, públicos, privados a través de Internet, con las limitaciones de acceso y uso que el propietario de la información disponga”

Para que este tipo de infraestructuras pueda existir, deben darse necesariamente una serie de requisitos:

- (a) los datos deben estar documentados (con sus metadatos);
- (b) las tecnologías deben estar estandarizadas, de manera que sea cual sea el sistema informático que se utilice, se pueda acceder a los datos, gestionarlos y conjugarlos;
- (c) debe haber una voluntad política que anime y facilite la creación de tales infraestructuras.

En ese sentido, toda infraestructura en el ámbito de las tecnologías lingüísticas debería ser abierta, distribuida, segura e interoperable.

- Abierta, ya que la infraestructura se concibe como una base de recursos escalable y en constante evolución, que incluye recursos libres, recursos de pago y servicios;
- Distribuida, ya que estará formada por repositorios y centros de datos en red, accesible a través de interfaces comunes;
- Segura, ya que garantizará una gestión acorde con la legislación y un acceso seguro a los recursos licenciables.

¹ Datos lingüísticos, como corpus textuales y de voz, datos relacionados con el lenguaje, incluyendo los asociados a otros medios y modalidades en los que el lenguaje escrito y oral juega un papel importante, información geográfica (IG), mapas cartografía, imágenes.

- Interoperable, ya que la base de recursos será compatible con los estándares, tratando de superar las diferencias terminológicas, semánticas y de formato;
- Segura, ya que garantizará una gestión acorde con la legislación y un acceso seguro a los recursos licenciables.

El presente informe pretende hacer una descripción de diferentes infraestructuras lingüísticas existentes, consideradas de alguna manera casos de éxito y analizar sus mecanismos de gobernanza², con el objetivo de identificar principios comunes que permitan establecer mejores prácticas a la hora de desarrollar una infraestructura de este tipo.

Se han tomado como modelos de infraestructuras el UMLS (Unified Medical Language System), SNOMED (Systematized Nomenclature of Medicine – Clinical Terms) y las Infraestructuras de Datos Espaciales (IDE) promovidas por el Open Geoespacial Consortium (OGC).

² Por Gobernanza en el presente informe se entiende el esquema organizativo que coordina la infraestructura, el marco legal que reparte responsabilidades, planifica esfuerzos, regula aspectos vinculados al licenciamiento y acceso a los datos, a las membresías, a los mecanismos internos de toma de decisiones, entre otros.

2. UMLS (UNIFIED MEDICAL LANGUAGE SYSTEM)

UMLS, o Unified Medical Language System, es una infraestructura lingüística, compuesta por un conjunto de archivos, software y recursos lingüísticos de diferentes fuentes, cuyo dominio específico es el lenguaje utilizado en el ámbito de la salud y las ciencias biomédicas, creado con objetivo de permitir la interoperabilidad entre sistemas informáticos con aplicación en este campo de la salud.

Esta infraestructura comenzó como un proyecto de I + D a largo plazo, a principios de 1986, con el objetivo de constituir una herramienta multipropósito para los desarrolladores de sistemas informáticos que permitiera: mejorar la comprensión del significado de los términos médico en todos los sistemas; superar las barreras para la recuperación efectiva de la información mediante el procesamiento de lenguaje por ordenador; y desambiguar la forma en que se expresan los mismos conceptos en lenguaje humano a fin de reducir errores en el procesamiento de lenguaje por ordenador.

Actualmente UMLS es una infraestructura gestionada por la Biblioteca Nacional de Medicina (NLH por sus siglas en inglés), que depende de la Secretaría de Salud y Servicios Humanos, del Gobierno de los Estados Unidos. NLH es la biblioteca biomédica más grande del mundo, que mantiene y pone a disposición una gran recopilación y producción de recursos de información electrónica sobre una amplia gama de temas vinculados a la salud.

Creada por la Ley de Servicios de Salud Pública de Estados Unidos, la NLH tiene por misión principal ayudar al avance de la medicina y ciencias afines, ayudar a la difusión y el intercambio de todo tipo de información (científica o no) a fin de contribuir al progreso de la medicina y la salud pública.

La Biblioteca Nacional de Medicina (NLM) es a su vez, el organismo central de coordinación para la estandarización de la terminología clínica dentro del Departamento de Salud y Servicios Humanos (HHS) de la Secretaría del mismo nombre y trabaja de manera conjunta con la Oficina del Coordinador Nacional de Tecnología de Información de Salud (ONC) para garantizar la implementación a nivel nacional, en Estados Unidos, de una infraestructura de tecnología de información de salud interoperable que permita mejorar la calidad y la eficiencia de los servicios de salud.

2.1 ¿EN QUÉ CONSISTE UMLS?

Como se ha dicho, UMLS en tanto infraestructura lingüística facilita el desarrollo de sistemas de información que utilizan tecnologías del lenguaje comportándose como si entendieran el lenguaje de la biomedicina y la salud. Con ese fin, la NLH produce y distribuye una serie de fuentes de conocimiento

que componen UMLS (bases de datos) y las herramientas de software asociadas a su utilización (programas).

Esta infraestructura permite a los desarrolladores usar esas fuentes de conocimiento y sus herramientas asociadas, para construir o mejorar sistemas de información que crean, procesan, recuperan e integran datos e información biomédica y de salud.

Las fuentes de conocimiento son multiusos y se usan en sistemas que realizan diversas funciones que involucran diferente tipo de información tales como: historia clínica de pacientes, literatura médica científica, pautas y datos de salud pública, etc. Las herramientas de software asociadas ayudan a los desarrolladores a personalizar o usar las fuentes de conocimiento de UMLS para fines particulares.

UMLS está compuesto por:

- Metathesaurus: más de 1 millón conceptos biomédicos de más de 100 fuentes.
- Semantic Network (Red Semántica): consiste en 135 categorías generales y 54 relaciones entre ellas.
- SPECIALIST Lexicon +Tools: información léxica y programas para el procesamiento de lenguaje.
- MetamorphoSys: es el asistente de instalación de UMLS y la herramienta de personalización de Metathesaurus.

2.1.1 Metathesaurus

Metathesaurus es una base de datos de vocabulario de grandes dimensiones, multiusos y multilingüe que contiene información sobre “conceptos biomédicos y relacionados con la salud”, sus diversos nombres y las relaciones entre ellos.

Diseñado para ser utilizado por desarrolladores de sistemas, Metathesaurus se construye a partir de versiones electrónicas de varios tesauros, clasificaciones, conjuntos de códigos y listas de términos controlados utilizados en la atención al paciente, facturación de servicios de salud, estadísticas de salud pública, indexación y catalogación de literatura biomédica y / o investigación básica, clínica y de servicios de salud. Estos recursos se conocen como los "vocabularios fuente" del Metathesaurus.

El término Metathesaurus se basa en la definición del prefijo "meta" más integral, trascendente, queriendo significar que Metathesaurus trasciende los tesauros, vocabularios y clasificaciones específicos que abarca.

Metathesaurus está organizado sobre la base de conceptos o significados. En esencia, vincula nombres alternativos y puntos de vista del mismo concepto e identifica relaciones útiles entre diferentes conceptos.

A su vez, está vinculado a la Semantic Network y a SPECIALIST Lexicon, en tanto todos los conceptos en Metathesaurus están asignados a al menos un Tipo Semántico de Semantic Network. Esto proporciona una categorización consistente de todos los conceptos de un nivel relativamente general. Además, muchas de las palabras o términos que aparecen en los nombres de conceptos o cadenas de Metathesaurus también aparecen en SPECIALIST Lexicon.

Metathesaurus refleja y preserva los conceptos o significados, los nombres conceptuales y las relaciones de sus vocabularios fuente.

Cuando dos vocabularios fuente diferentes usan el mismo nombre de concepto para diferentes conceptos, Metathesaurus representa ambos significados e indica qué significado está presente en qué vocabulario fuente. Por su parte, cuando aparece el mismo concepto en diferentes contextos jerárquicos de los diferentes vocabularios fuente, el Metathesaurus incluye todas las jerarquías. A su vez, cuando las relaciones entre dos conceptos aparecen en diferentes vocabularios fuente de manera conflictiva, ambas vistas se incluyen en Metathesaurus (aunque en algunos casos, estas diferencias pueden obedecer a cuestiones idiosincrásicas y carecer de validez aparente Metathesaurus las incluye).

En resumen, Metathesaurus no representa una ontología de biomedicina completa autorizada por la NLM, si no que conserva la diversidad de los vocabularios fuente, considerando que esas vistas diferentes pueden ser útiles para diferentes tareas.

Aunque conserva todos los significados y el contenido en sus vocabularios fuente, Metathesaurus almacena esta información en un único formato común. El formato nativo de cada vocabulario se estudia cuidadosamente y luego se "invierte" en el formato común de Metathesaurus.

2.1.1.1 Vocabulario Fuente

Metathesaurus contiene conceptos, nombres conceptuales y otros atributos de más de cien terminologías, clasificaciones y tesauros diferentes. Existe un concepto en el Metathesaurus para cada vocabulario fuente en sí, al que se le asigna el Tipo Semántico "Producto Intelectual".

Los vocabularios fuente que componen Metathesaurus, están enumerados en el siguiente link: <https://www.nlm.nih.gov/research/umls/sourcereleasedocs/index.html>

Algunos ejemplos, de vocabularios fuente de mayor relevancia son:

- **LOINC:** “The Logical Observation Identifiers Names and Codes terminology (LOINC)” fue desarrollada y es mantenida por el “Regenstrief Institute, Inc.”, una organización de investigación médica sin fines de lucro asociada con la Universidad de Indiana. El propósito de LOINC es facilitar el intercambio, la agrupación y la gestión de los resultados de laboratorio en la atención clínica, así como la gestión de reclamos y la investigación, proporcionando un conjunto de códigos y nombres universales. Los códigos LOINC permiten a los usuarios fusionar resultados clínicos de muchas fuentes en una base de datos para la atención del paciente, la investigación clínica o la gestión.
- **RxNorm:** proporciona nombres normalizados para medicamentos clínicos y vincula sus nombres a muchos de los vocabularios sobre medicamentos comúnmente utilizados en software de administración de farmacia e interacción de drogas, incluyendo los de First Databank, Micromedex, MediSpan, Gold Standard Drug Database y Multum. Proporciona enlaces entre estos vocabularios y puede mediar mensajes entre sistemas que no utilizan el mismo software y/o vocabulario.
- **SNOMED CT:** es una terminología clínica integral, multilingüe y codificada de mayor amplitud, precisión e importancia desarrollada en el mundo. Permite la optimización de los datos clínicos para una mejor atención de los pacientes. La terminología es propiedad de la Organización Internacional de Desarrollo de Normas de Terminología de Salud (IHTSDO) quien la gestiona y administra.

Metathesaurus incluye además conjuntos de códigos obligatorios para su uso en transacciones administrativas electrónicas en los EE. UU, según las disposiciones de la Ley de Portabilidad y Responsabilidad de Seguros Médicos (HIPAA por sus siglas en inglés), con la excepción de los Códigos Nacionales de Medicamentos (NDC por sus siglas en inglés)). NHL tiene la intención de incorporar todas las terminologías clínicas designadas como estándares ya sea por la HIPAA, como aquella terminología designadas como estándar Informático Consolidado de Salud (CHI por sus siglas en inglés) por el Comité Nacional de Estadísticas Vitales y de Salud.

La estructura de Metathesaurus permite, además, alojar las traducciones de sus vocabularios fuente a otros idiomas además del inglés. Muchas traducciones en muchos idiomas diferentes están presentes en la edición actual del Metathesaurus.

2.1.1.2 Códigos de Identificación

Como se ha dicho Metathesaurus está organizado por concepto o significado. Uno de sus propósitos principales es conectar diferentes nombres de concepto para el mismo concepto o significado, surgido de los diferentes vocabularios fuente. Para ello, Metathesaurus asigna varios tipos de **códigos de**

identificación a los diferentes conceptos y nombres conceptuales. Además, cada concepto y nombre conceptual retiene todos los identificadores que están presentes en los vocabularios fuente.

Entre los diferentes códigos de identificación pueden citarse:

- **CUI** (identificador de concepto único y permanente); código asignado a **cada concepto o significado** de manera exclusiva y permanente. Este código no tiene ningún valor intrínseco, es decir no está vinculado al concepto o significado, no permite inferir nada de él. Cuando los expertos encargados de la gestión de Metathesaurus detectan la presencia de sinonimia (dos nombres de concepto diferentes para un mismo concepto o significado) atribuyen al concepto el CUI correspondiente.
- **SUI** (un identificador de cadena único y permanente); código asignado a cada **nombre de concepto o cadena** en el Metathesaurus. Cualquier variación en el conjunto de caracteres, ya sea una mayúscula o minúscula es un nombre de concepto o cadena separada, con un SUI por separado. El mismo nombre de concepto o cadena en diferentes idiomas (por ejemplo, inglés y español) tendrá un SUI diferente para cada idioma. Si la misma cadena, por ejemplo, **“Cold”**³, tiene más de un significado, el SUI estará vinculado a más de un identificador de concepto (CUI).
- **AUI** (identificadores de átomos e átomos); los bloques de construcción básicos o "átomos" a partir de los cuales se construye el Metathesaurus son los nombres de concepto o las cadenas presentes en cada uno de los vocabularios fuente. A cada aparición de una cadena o nombre de concepto en cada vocabulario fuente se le asigna un identificador de átomo único (AUI). Si exactamente la misma cadena aparece dos veces en el mismo vocabulario, por ejemplo, como el nombre largo y el nombre abreviado para el mismo concepto o como un nombre alternativo para dos conceptos diferentes en la misma fuente de vocabulario, se asigna un AUI único para cada ocurrencia. Cuando aparece la misma cadena en varios vocabularios fuente, tendrá AUI cada vez que aparece como un nombre de concepto en cada una de esas fuentes. Todos estos AUI se vincularán a un único identificador de cadena (SUI), ya que representan las ocurrencias de la misma cadena en los vocabularios fuente. A diferencia de los identificadores de cadena, una sola AUI siempre está vinculada a un único identificador de concepto (CUI), porque cada aparición de una cadena en una fuente solo puede tener un significado.

³ El nombre de concepto “Cold” en inglés puede referirse a frío o resfriado, entonces “Cold” tendrá un único código de cadena (SUI) pero estará vinculado a dos identificadores únicos de concepto (CUI).

- **LUI** (identificador de término común), este código es solo para las entradas de idioma inglés en el Metathesaurus. Cada cadena o nombre de concepto está vinculado a todas sus variantes léxicas o variaciones menores mediante este código LUI. (En el Metathesaurus, por lo tanto, un "término" en inglés es el grupo de todas las cadenas o nombres de concepto, que son variantes léxicas entre sí). Las variantes léxicas en inglés se detectan utilizando el programa Lexical Variant Generator (lvg), una de las herramientas léxicas de UMLS. A medida que las herramientas similares estén disponibles para otros idiomas, se pueden usar para crear grupos variantes léxicas en otros idiomas. Al igual que un identificador de cadena (SUI), el LUI para una cadena o nombre de concepto en inglés puede estar vinculada a más de un concepto o significado (CUI). Esto ocurre cuando las cadenas que son variantes léxicas entre sí tienen diferentes significados. Por el contrario, cada identificador de cadena (SUI) y cada identificador de átomo (AUI) solo se pueden vincular a un solo LUI.

En la tabla de abajo, pueden observarse los usos de los identificadores de concepto, cadena, átomo y término. Así, cada CUI (concepto) está vinculado a al menos un AUI (átomo), SUI (cadena) y LUI (término), pero también se puede vincular a muchos de cada uno de estos. Cada AUI (átomo) está vinculado a un único SUI (cadena), un único LUI (término) y un único CUI (concepto). Cada SUI (cadena) puede vincularse a muchos AUI (átomos), a un único LUI (término) y a más de un CUI (concepto), aunque el caso típico es un CUI. Cada LUI (término) se puede vincular a muchos AUI (átomos), muchos SUI (cadenas) y más de un CUI (concepto), aunque el caso típico es un CUI.

Table 1. Concept, Term, Atom, and String Identifiers.

Concept (CUI)	Terms (LUIs)	Strings (SUIs)	Atoms (AUIs) * RRF Only
C0004238 Atrial Fibrillation (preferred) Atrial Fibrillations Auricular Fibrillation Auricular Fibrillations	L0004238 Atrial Fibrillation (preferred) Atrial Fibrillations	S0016668 Atrial Fibrillation (preferred)	A0027665 Atrial Fibrillation (from MSH) A0027667 Atrial Fibrillation (from PSY)
		S0016669 (plural variant) Atrial Fibrillations	A0027668 Atrial Fibrillations (from MSH)
	L0004327 (synonym) Auricular Fibrillation Auricular Fibrillations	S0016899 Auricular Fibrillation (preferred)	A0027930 Auricular Fibrillation (from PSY)
		S0016900 (plural variant) Auricular Fibrillations	A0027932 Auricular Fibrillations (from MSH)

Tabla 1: Concept, Term, Atom, and String Identifiers

2.1.1.3 Relaciones e identificadores de Relaciones

Metathesaurus incluye muchas relaciones entre diferentes conceptos (además de las relaciones sinónimas en la estructura del concepto Metathesaurus). La mayoría de estas relaciones provienen de vocabularios fuente individuales, algunos son añadidos por NLH durante la construcción del Metathesaurus, otras han sido aportados por usuarios de Metathesaurus.

Las relaciones se expresan en términos de CUI y AUI, en general, el Metathesaurus indica el autor de cada relación, es decir, el vocabulario fuente del cual surge, o el Metathesaurus en sí o el usuario/proveedor.

Entre las relaciones no sinónimas entre los conceptos que recoge Metathesaurus, se encuentran aquellas que provienen del mismo vocabulario fuente (relaciones de vocabulario intra-fuente) y entre conceptos en diferentes vocabularios (relaciones de vocabulario entre fuentes).

Metathesaurus no incluye todas las posibles relaciones no sinónimas entre los conceptos que contiene, si incluye todas las relaciones presentes en sus vocabularios fuente y algunas relaciones adicionales diseñadas para conectar conceptos relacionados. En general, las relaciones afirmadas por los vocabularios fuente conectan conceptos estrechamente relacionados, como aquellos que comparten alguna propiedad común o están relacionados por definición. Por ejemplo, un miembro de una clase

de medicamentos (por ejemplo, penicilina) se conectará al nombre de la clase (por ejemplo, antibióticos); una infección bacteriana se conectará a la bacteria que lo causa, etc.

Cada relación presente en Metathesaurus tiene un identificador de relación único (RUI), el objetivo principal de estos identificadores es permitir la detección fácil de cambios en las relaciones entre las versiones del Metathesaurus. La aparición o desaparición de un identificador de relación indica un cambio en las relaciones presentes en el Metathesaurus.

Algunos vocabularios fuente tienen sus propios identificadores de relación. Donde existen, estos identificadores también están presentes el RUI de Metathesaurus.

2.1.1.4 Atributos e identificadores de atributos

En Metathesaurus, los atributos incluyen cada información discreta sobre un concepto, un átomo o una relación que no es parte de la estructura básica del concepto de Metathesaurus o se distribuye en uno de los archivos de relación.

Metathesaurus incluye atributos de concepto, atributos de átomo y atributos de relación. Los atributos del concepto se agregan durante la construcción del Metathesaurus y se aplican a todos los nombres de un concepto. Por ejemplo, los Tipos Semánticos " Pathologic Function " y "Finding" son atributos del concepto con el nombre preferido "Atrial Fibrillation" y son aplicables a cualquier átomo conectado a ese concepto.

Los atributos de átomo provienen de un vocabulario fuente particular, algunos de ellos son de interés general; otros son relevantes solo para un vocabulario fuente particular. Por ejemplo, la definición "Trastorno del ritmo cardíaco caracterizado por impulsos auriculares rápidos e irregulares y contracciones auriculares ineficaces" es un atributo de la Fibrilación Auricular Atómica que proviene del vocabulario fuente (MeSH). Puede ser una de varias definiciones relacionadas con los nombres de este concepto, ya que el Metathesaurus incluye todas las definiciones proporcionadas por cualquiera de sus vocabularios fuente. Aunque esta definición particular proviene de MeSH, podría ser útil en las aplicaciones de Metathesaurus que no usan MeSH.

Los atributos de relación provienen de un vocabulario fuente particular y describen características especiales de relaciones particulares en esa fuente.

A cada aparición de cada atributo dentro del Metathesaurus se le asigna un identificador de atributo único (ATUI). La aparición o desaparición de las ATUI indica cambios en el contenido del

Metathesaurus, por lo que las ATUI ayudan a la producción eficiente de un conjunto de cambios completo para cada nueva versión del Metathesaurus.

2.1.2 Semantic Network (Red semántica)

La Red Semántica consiste en un conjunto de categorías de temas amplios, o Tipos Semánticos, que proporcionan una categorización consistente de todos los conceptos representados en Metathesaurus y un conjunto de relaciones útiles e importantes, o Relaciones Semánticas, que existir entre tipos semánticos.

El propósito de la Red Semántica es proporcionar una categorización consistente de todos los conceptos representados en el Metathesaurus y proporcionar un conjunto de relaciones útiles entre estos conceptos.

Toda la información sobre conceptos específicos se encuentra en el Metathesaurus. La red semántica en cambio, proporciona información sobre el conjunto de tipos semánticos básicos, o categorías, que pueden asignarse a estos conceptos, y define el conjunto de relaciones que puede contener entre los tipos semánticos.

La Red Semántica contiene 133 tipos semánticos y 54 relaciones, constituye una autoridad para los tipos semánticos que se asignan a los conceptos en el Metathesaurus, define estos tipos, tanto con descripciones textuales como a través de la información inherente a sus jerarquías.

Los tipos semánticos son los nodos en la red, y las relaciones entre ellos son los enlaces. Existen grandes grupos de tipos semánticos para organismos, estructuras anatómicas, funciones biológicas, productos químicos, eventos, objetos físicos y conceptos o ideas. El alcance actual de los tipos semánticos UMLS es bastante amplio, lo que permite la categorización semántica de una amplia gama de terminología en múltiples dominios.

Cada concepto Metathesaurus tiene asignado al menos un tipo semántico, en todos los casos, el tipo semántico más específico disponible en la jerarquía se asigna al concepto.

La figura 1 ilustra una parte de la red. El tipo semántico "Función biológica" tiene dos hijos, "Función fisiológica" y "Función patológica", y cada uno de estos a su vez tiene varios hijos y nietos. Cada "hijo" en la jerarquía está vinculado a su "padre" por el enlace "isa".

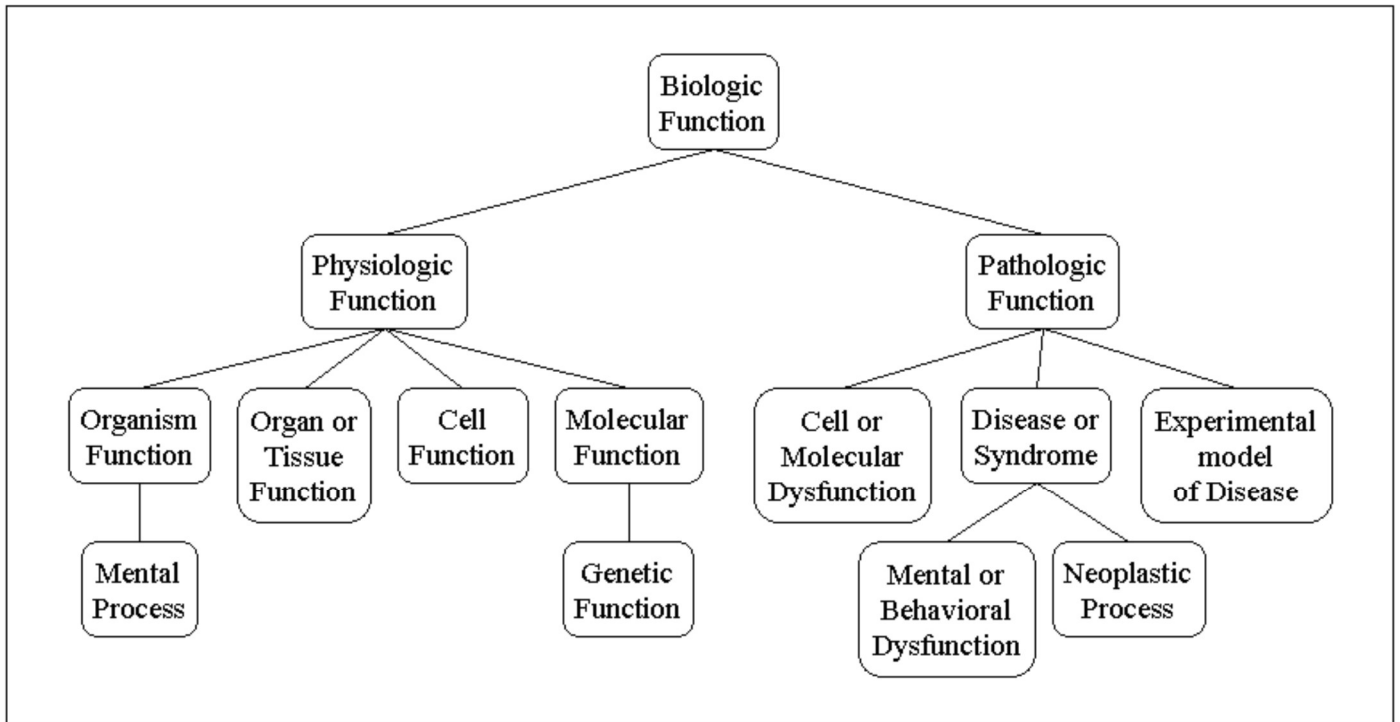


Figura 1: Función Biológica.

El enlace principal en la red es el enlace "is a". Esto establece la jerarquía de tipos dentro de la red y se usa para decidir sobre el tipo semántico más específico disponible para la asignación a un concepto de Metathesaurus.

Además, se identifica un conjunto de relaciones no jerárquicas entre los tipos semánticos, que se agrupan en cinco categorías principales, que a su vez son relaciones: "relacionadas físicamente con", "relacionadas espacialmente", "relacionadas temporalmente con", "relacionadas funcionalmente con" y "relacionadas conceptualmente con".

Las relaciones se establecen entre tipos semánticos de alto nivel en la Red siempre que sea posible y generalmente se heredan a través del enlace "is a" por todos los hijos de esos tipos.

Las relaciones se establecen entre tipos semánticos y no se aplican necesariamente a todas las instancias de conceptos que se han asignado a esos tipos semánticos. Es decir, la relación puede o no mantenerse entre cualquier par particular de conceptos.

La Figura 2 muestra una porción de la Red Semántica, que ilustra las relaciones, ya sean jerárquicas o asociativas, que existen entre los tipos semánticos.

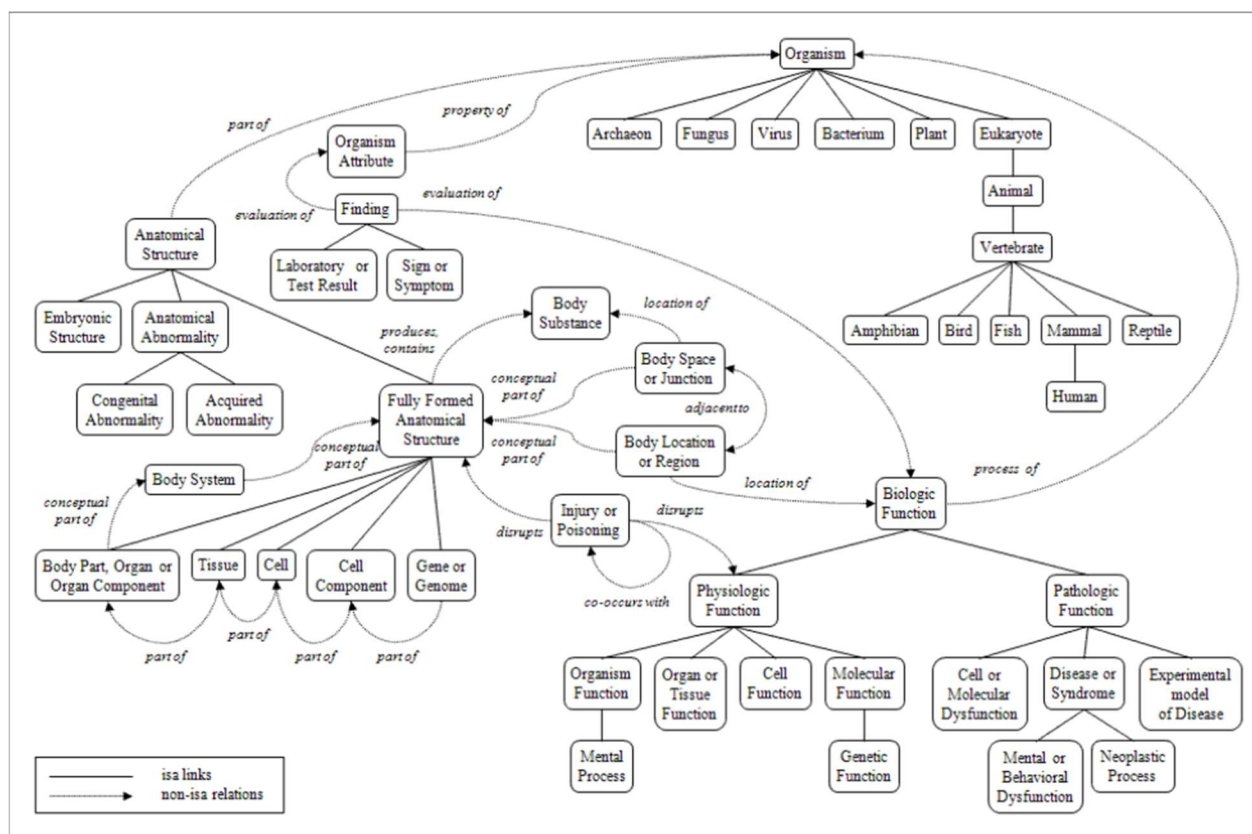


Figura 2: Porción de Red Semántica

Las jerarquías para los tipos semánticos y las relaciones semánticas en la red semántica se pueden encontrar en el sitio web de UMLS, en los enlaces siguientes:

Tipos semánticos:

https://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html

Relaciones semánticas:

https://www.nlm.nih.gov/research/umls/META3_current_relations.html

2.1.3 SPECIALIST Lexicon

“Specialist Lexicon” ha sido desarrollado para proporcionar la información léxica necesaria para el Sistema de Procesamiento de Lenguaje Natural (NLP). Se pretende que sea un léxico general en inglés que incluya muchos términos biomédicos. La cobertura incluye palabras comunes en inglés y vocabulario biomédico. La entrada del léxico para cada palabra o término registra la información sintáctica, morfológica y ortográfica que necesita el “SPECIALIST NLP System”.

Las herramientas léxicas están diseñadas para abordar el alto grado de variabilidad en palabras y términos de lenguaje natural. Las palabras a menudo tienen varias formas flexionadas que se

considerarían correctamente instancias de la misma palabra. El verbo “treat” (tratar en inglés), por ejemplo, tiene tres variantes flexivas: “treats” (trata) referido a la tercera persona del singular en tiempo presente, “treated” (trató) la forma pasada del participio, y “treating” (tratando) la forma presente forma del participio. Los términos de varias palabras en el Metathesaurus y otros vocabularios controlados pueden tener variantes de orden de palabras además de sus variantes de casos inflexionales y alfabéticos. Las herramientas léxicas permiten al usuario abstraerse de este tipo de variación.

2.1.4 MetamorphoSys

MetamorphoSys es el asistente de instalación de UMLS y la herramienta de personalización de Metathesaurus. Los usuarios personalizan sus subconjuntos en Metathesaurus con dos propósitos principales:

1. Para excluir los vocabularios que no son necesarios o no tienen licencia para su uso en una aplicación local. El Metathesaurus consiste en varios archivos, algunos de los cuales son extremadamente grandes; excluir fuentes puede reducir significativamente el tamaño del subconjunto, acotándolo específicamente a la aplicación que se le pretende dar. Dada la cantidad y variedad de vocabularios reflejados en el Metathesaurus, es poco probable que cualquier usuario requiera todos, o incluso la mayoría, de sus más de 100 vocabularios fuente. Además, algunas fuentes requieren acuerdos de licencia por separado para usos específicos, que un usuario de UMLS puede no desear obtener. Estos están claramente indicados en el Acuerdo de licencia.
2. Para personalizar un subconjunto usando una variedad de opciones de salida de datos y filtros.

2.2 ESTRUCTURA DE BIBLIOTECA NACIONAL DE MEDICINA (NLH)

La NLH cuenta con una estructura compleja, integrada por la Oficina del Director, donde coexisten además un Director Adjunto y un Subdirector de Investigación y Educación, configurando el más alto nivel en la estructura.

A su vez, se encuentra la Junta de Regentes de NLH, vinculada directamente con la Oficina del Director, que sirve como cuerpo asesor de este, así como del propio Secretario de Salud y Servicios Humanos, y otros funcionarios de alto nivel dentro de la Secretaría, en materia política y de gestión relativa a NLH, en cuestiones vinculadas subvenciones o ayudas, incluyendo además temas presupuestarios y organizacionales referidos a servicios, publicaciones, instalaciones, etc.

La Junta está compuesta por miembros de oficio y miembros nombrados por el Secretario. Los miembros de oficio son los Cirujanos Generales de la Servicio de Salud Pública del Ejército, la Marina y la Fuerza Aérea; el Director Médico jefe del Departamento de Asuntos de Veteranos; el Decano de la Universidad de Ciencias de la Salud de Servicios Uniformados; el Subdirector de Biología, Conducta y Ciencias Sociales de la Fundación Nacional de Ciencias; el Director de la Biblioteca Nacional de Agricultura, y el Bibliotecario del Congreso.

En cuanto a los miembros designados por el Secretario, serán seleccionados entre los líderes en los diversos campos de las ciencias fundamentales, medicina, odontología, salud pública, administración hospitalaria, farmacología, salud y tecnología de información y las comunicaciones. Cada miembro designado de la Junta desempeñará el cargo por un término de cuatro años.

La NLH cuenta además con tres oficinas dependientes del Director:

1.- Oficina de Administración

2.- **Oficina de Comunicaciones y Enlace Público**

3.- Oficina de Desarrollo de Programas de Información de Salud

Asimismo, **la Oficina de Comunicación y Enlace Público** cuenta con tres Divisiones Principales:

- a) **División de Información Especializada**, de la cual depende la Oficina de Toxicología Clínica, Subdirección de Servicios de Información Biomédica, la Subdirección de Implementación de Archivos Biomédicos; el Centro de Investigación de Gestión de la Información de Desastres y el Área Extensión;
- b) **División de Programas Extramuros**, integrada por un área de revisión científica, un área de programas y un área de gestión de subvenciones. Esta división interactúa directamente con el Comité de Revisión de la Biblioteca Biológica y la Informática (organismo externo);
- c) **División de Operaciones de la Biblioteca**, compuesta por un Comité de Revisión Técnica de Selección de Literatura, una Oficina de Coordinación de la Red Nacional y el Centro Nacional de Información para la Investigación de Servicios de Salud y Tecnología de Cuidado de la Salud. Cuenta además con un área de servicios públicos, otra de servicios técnicos, una división de selección de literatura y otra de historia de la medicina. La División de Operaciones de la NLH trabaja de manera coordinada con el Comité de Revisión Técnica de la Selección de Literatura (organismo externo);

Además de las tres divisiones mencionadas, de la Oficina de Comunicación y Enlace Público depende el **“Centro Nacional ‘Lister Hill’ para Comunicaciones Biomédicas”**.

Dicho centro cuenta con diferentes ramas de investigación y desarrollo, la rama de Ingeniería de Comunicaciones, la rama de Ciencias de la Computación, la Subdivisión de Desarrollo de Programas Audiovisuales; la rama de Ciencia Cognitiva. Asimismo, interactúa directamente con un organismo externo, denominado **Junta LHNCBC de Consejeros Científicos**.

Dentro de este centro de la NLH, es donde de manera coordinada con la Oficina del Coordinador Nacional de Tecnología de Información de Salud (ONC), se desarrolla la implementación, gestión y mantenimiento de UMLS, garantizándose su interoperabilidad, calidad y la eficiencia.

Por último, dependiendo de la misma Oficina se encuentra el Centro Nacional de Información Biotecnológica (NCBI), que al igual que el “Lister hill”, cuenta con tres ramas de investigación y desarrollo, la rama de Biología Computacional, la Rama de Ingeniería de Información, y la rama de recursos de información. Este centro opera de manera coordinada con dos organismos externos, la Junta de Consejeros Científicos del NCBI y el Comité Asesor Nacional PubMed Central⁴.

⁴ PubMed comprende más de 28 millones de citas de literatura biomédica de MEDLINE. Las citas y resúmenes de PubMed incluyen los campos de la biomedicina y la salud, que cubren partes de las ciencias de la vida, las ciencias del comportamiento, las ciencias químicas y la bioingeniería. Es un recurso gratuito que es desarrollado y mantenido por el Centro Nacional de Información Biotecnológica (NCBI), en la Biblioteca Nacional de Medicina de los Estados Unidos (NLM). MEDLINE es la principal base de datos bibliográficas de la Biblioteca Nacional de Medicina, que contiene más de 24 millones de referencias a artículos de revistas en ciencias de la vida con foco en biomedicina. Una característica distintiva de MEDLINE es que los registros están indexados con NLM Medical Subject Headings (MeSH®). MEDLINE es la contraparte en línea de MEDLARS® (Sistema de Recuperación y Análisis de Literatura Médica) que se originó en 1964

En la Figura 3 se presenta un cuadro de la estructura de la NLH.

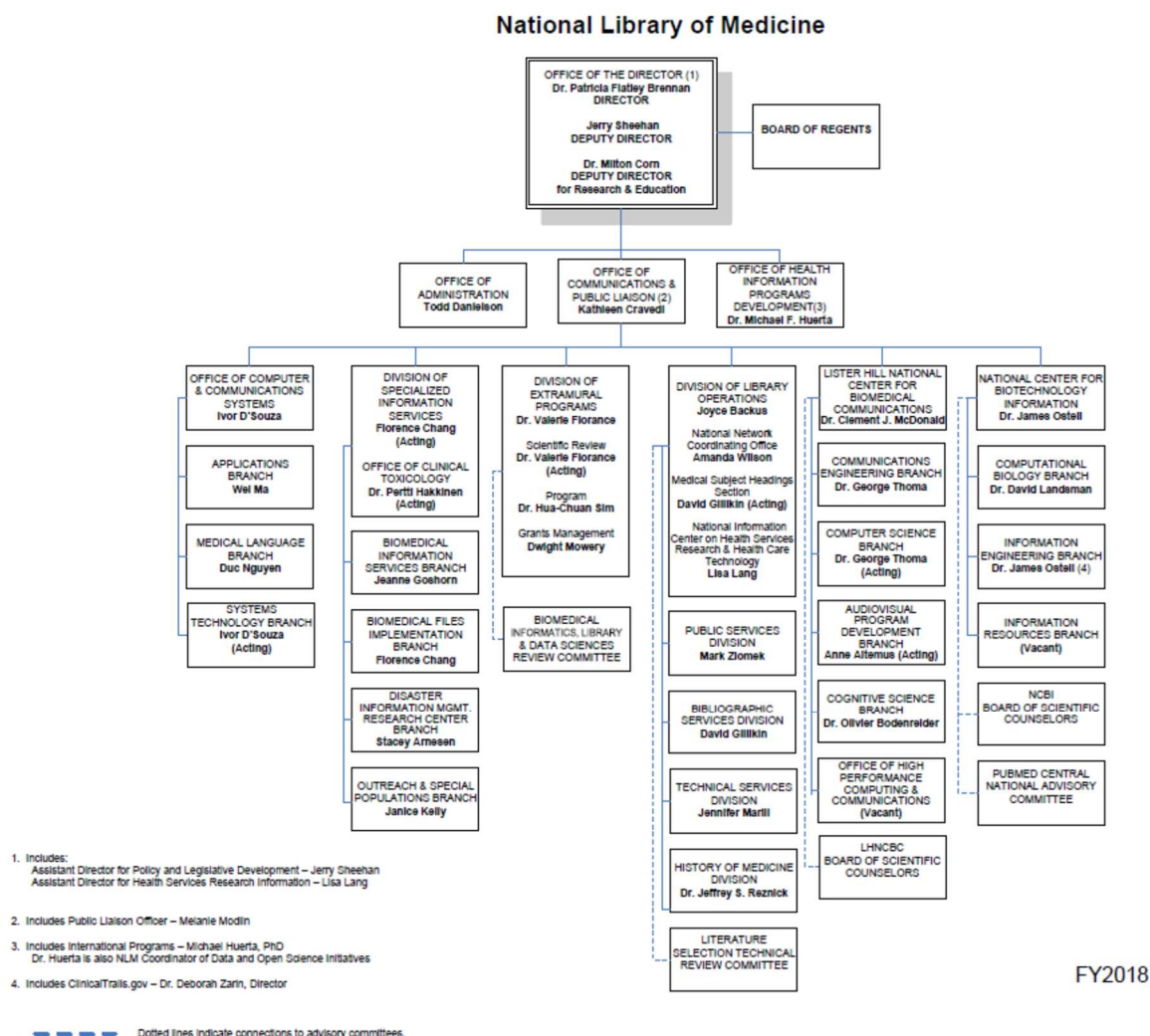


Figura 3: Estructura de la Biblioteca Nacional de Medicina.

Como se mencionó al describir la estructura, la NLH trabaja con Órganos consultivos que prestan asesoramiento externo, en campos como el: medicina, ciencia, academia, bibliotecología y política.

En ese sentido, se hará una breve descripción de estos organismos a fin de comprender como funcionan.

a) Comité de Revisión de la Biblioteca Biológica y la Informática

Establecido por la Ley del Servicio de Salud Pública, está compuesto por 21 miembros, incluidos los Presidentes, designados por el Director de los Institutos Nacionales de Salud (NIH). Se los elige entre las autoridades con conocimientos en los campos de las bibliotecas de ciencia de la salud, de la información, las publicaciones y la informática médica.

Actúa como asesor del Director de los Institutos Nacionales de Salud (NIH por sus siglas en inglés) y del Director de la Biblioteca Nacional de Medicina, y hace recomendaciones a la Junta de Regentes de la Biblioteca Nacional de Medicina sobre las solicitudes de mérito científico y técnico para subvenciones en ayuda y acuerdos de cooperación para investigación, capacitación, apoyo de recursos y publicaciones en las áreas de bibliotecas de ciencias de la salud, gestión de información de ciencias de la salud, publicaciones de ciencias biomédicas e informática médica para incluir informática clínica, bioinformática y temas biomédicos relacionados.

b) Comité de Revisión Técnica de la Selección de Literatura

También ha sido establecido por la Ley del Servicio de Salud Pública, está compuesto de 15 miembros, incluida la presidencia.

Los miembros y la Presidencia son seleccionados por el Director de los Institutos Nacionales de Salud (NIH), de entre las autoridades con conocimientos en los campos de la biomedicina, como médicos, investigadores, educadores, editores, bibliotecarios e historiadores de las ciencias de la salud.

Su función es aconsejar al Director del NIH y al Director de la Biblioteca Nacional de Medicina (NLM) sobre cuestiones de política relacionadas con la evaluación y recomendación de publicaciones biomédicas que se deben considerar para su indexación e inclusión en publicaciones de la NLM, como Index Medicus y MEDLINE.

c) Junta de Consejeros Científicos del NCBI:

Establecida por la Ley del Servicio de Salud Pública la Junta cuenta con 8 miembros, incluido el Presidente, designados por el Director del NIH.

Sus miembros son representantes de diferentes comunidades científicas, con conocimientos en ciencias de la salud, ciencias de la computación, tecnología de la información, biblioteconomía, ciencias del comportamiento, ciencias sociales, tecnología educativa, ingeniería de comunicaciones, biología molecular, bioquímica, genética, química estructural, matemáticas, estadísticas y desarrollo y utilización de medios.

Entre otras tareas de asesoramiento que realiza, en su calidad de junta de expertos, asesora a la Biblioteca Nacional de Medicina (NLM) y el Director del Centro Nacional de Información Biotecnológica (NCBI) sobre los programas de investigación y desarrollo intramuros de la

Biblioteca Nacional de Medicina, mediante la evaluación de dichos programas de investigación y desarrollo. A su vez actúa como supervisor de la productividad y el rendimiento del personal científico de la Biblioteca y del Centro Nacional.

d) Comité Asesor Nacional PUBMED CENTRAL

Establecida de conformidad con la Ley del Servicio de Salud Pública, consta de 15 miembros, incluido el Presidente, designados por el Director, NIH. El Comité estará compuesto por científicos expertos, editores de ciencias biológicas, miembros de sociedades científicas, especialistas en información y bibliotecas médicas y miembros del público en general.

Asesoran al Director del NIH, al Director de la NLH y al Director del Centro Nacional de Información Biotecnológica, NCBI sobre el contenido y el funcionamiento del repositorio PubMed Central. Establece criterios para certificar a los grupos que envían materiales al sistema, supervisa el funcionamiento del sistema y garantiza que PubMed Central evolucione y siga respondiendo a las necesidades de los investigadores, editores, bibliotecarios y público en general.

2.3 PRESUPUESTO PARA UMLS

El desarrollo, gestión, mantenimiento, actualización, distribución de UMLS, a cargo de la NLH y específicamente del Centro Nacional 'Lister Hill' para Comunicaciones Biomédicas se lleva a cabo en el marco de un programa denominado "Enhancing the Interoperability of Electronic Health Records(EHR)" (Mejora la interoperabilidad de los registros electrónicos de salud).

Como se ha dicho anteriormente, la NHL realiza su tarea en estrecha colaboración con la Oficina del Coordinador Nacional de Tecnología de la Información de Salud dentro de la Secretaria de Salud y Servicios Humanos, con el apoyo de los Centros de Servicios de Medicare y Medicaid (CMS) y la Administración de Salud de Veteranos, desarrollando y difundiendo las terminologías clínicas designadas como estándares dentro de EE. UU.

NHL produce una gama de herramientas que ayudan a los desarrolladores y usuarios de los registros electrónicos de salud (EHR) a implementar estándares, poniéndolos a disposición en múltiples formatos, incluso a través de interfaces de programación de aplicaciones (API).

El soporte técnico y financiero de NLM permite que los estándares de terminología clínica se actualicen regularmente para reflejar los nuevos medicamentos, pruebas, dispositivos y cambios en el conocimiento médico y la práctica de la salud. En el año fiscal 2016, por ejemplo, NLM trabajó

rápidamente para desarrollar la terminología necesaria para apoyar la respuesta a la evolución del brote de virus del Zika.

La NHL tiene dos organizaciones que llevan a cabo I + D avanzada en diferentes aspectos de informática biomédica, el Centro Nacional Lister Hill para Comunicaciones Biomédicas (LHC) y el NCBI.

El Centro Nacional Lister Hill para Comunicaciones Biomédicas (LHC), conduce y apoya la I + D en tales áreas como el desarrollo y la difusión de estándares de datos de salud; la captura, procesamiento, disseminación y uso de datos de imágenes de alta calidad; procesamiento del lenguaje médico; alta velocidad acceso a información biomédica; tecnología avanzada para la gestión de emergencias y desastres; y análisis de grandes bases de datos clínicos y administrativos para determinar su utilidad en predecir los resultados de los pacientes y validar los hallazgos de una clínica prospectiva relativamente pequeña Estudios de investigación.

Por su parte, NCBI lleva a cabo I + D sobre la representación, integración y recuperación de datos de biología molecular y literatura biomédica, además de proporcionar un recurso integrado de información genómica que consta de más de 40 bases de datos para investigación biomédica tanto del NIH como de todo el mundo.

Durante el año fiscal 2017, el presupuesto del tesoro del gobierno de los Estados Unidos con el que contó el programa mencionado, dentro del cual se destinan recursos a UMLS, fue de \$ 18.3 millones de dólares estadounidenses, en tanto para 2018, se proyectó una reducción de 1 millón de dólares en el presupuesto, aprobándose un total de 17.3 millones de dólares estadounidenses.

3. SNOMED (GLOBAL STANDARDS FOR HEALTH TERMS)

SNOMED CT es una terminología clínica con alcance global que abarca una amplia gama de especialidades, disciplinas y requerimientos médicos. Como consecuencia de su amplio alcance, uno de los beneficios de SNOMED CT es una disminución de los efectos límite entre las especialidades por el uso de distintas terminologías o sistemas de codificación por diferentes médicos o departamentos. Esto permite compartir y reutilizar la información médica estructurada.

Otro beneficio de SNOMED CT es que permite procesar y presentar los mismos datos de diferentes maneras para distintos fines. Por ejemplo, las historias clínicas representadas con SNOMED CT pueden procesarse y presentarse en diversas formas para facilitar la atención directa de los pacientes, las auditorías clínicas, la investigación, epidemiología, gestión y planificación de servicios.

Además, el alcance global de SNOMED CT disminuye los efectos de las diferencias geográficas como consecuencia del uso de diferentes terminologías o sistemas de codificación en distintas organizaciones o países.

Con SNOMED CT, la información clínica se registra con identificadores que refieren a **conceptos definidos formalmente** como parte de la terminología. SNOMED CT permite el registro de información clínica con niveles apropiados de detalle mediante la utilización de conceptos clínicos relevantes.

Las estructuras de SNOMED CT permiten ingresar la información utilizando sinónimos que se adaptan a las preferencias locales a la vez que se registra a información de manera coherente y comparable.

Por otra parte, debido a su disposición en jerarquías, SNOMED CT permite registrar la información con diferentes niveles de detalle para que se adecue a usos particulares (por ejemplo, |neumonía|, |neumonía bacteriana| o |neumonía neumocócica|).

SNOMED CT permite agregar detalles adicionales mediante la combinación de conceptos cuando los conceptos disponibles no son lo suficientemente precisos (por ejemplo, |neumonía neumocócica| con localización del hallazgo en |lóbulo superior de pulmón derecho|).

SNOMED CT ofrece una amplia gama de opciones diferentes para la recuperación inmediata y posterior reutilización para abordar de inmediato y a más largo plazo los requerimientos médicos y los de otros usuarios.

La naturaleza de las jerarquías de SNOMED CT permiten recuperar y reutilizar selectivamente la información para cumplir con diferentes requerimientos en diversos niveles de generalización (por

ejemplo, tanto la recuperación de subtipos de |trastorno pulmonar| como de |infección bacteriana| incluiría a |neumonía bacteriana|).

El modelo conceptual de SNOMED CT también permite considerar otros detalles al recuperar datos. Por ejemplo, el concepto |neumonía neumocócica| es un subtipo de |neumonía bacteriana| que tiene una relación definitoria que especifica que el |agente causal| es el |Streptococcus pneumoniae| y de esta manera es posible analizar el microorganismo que causa la enfermedad.

SNOMED CT tiene una amplia cobertura de temas relacionados con la salud. Se puede utilizar para describir los antecedentes médicos de un paciente, los detalles de un procedimiento ortopédico, la propagación de una epidemia y mucho más.

Al mismo tiempo, la terminología presenta varios niveles de profundidad, que permiten a los médicos registrar datos en el nivel de granularidad (especificidad) adecuado.

Las aplicaciones específicas tienden a centrarse en conjuntos limitados de SNOMED CT, por ejemplo, conceptos relacionados con la oftalmología. Estos subconjuntos pueden utilizarse para presentar partes relevantes de la terminología, según el contexto clínico y los requerimientos locales. Esto significa, por ejemplo, que una lista desplegable para seleccionar diagnósticos en una historia clínica electrónica de un centro de salud mental se puede adaptar a ese ámbito. En forma similar, es posible definir subconjuntos para listas de problemas para especialidades médicas o brindar listas de medicamentos adecuadas para enfermeras en el ámbito de la atención comunitaria.

Cuando jurisdicciones individuales tienen necesidades que exceden las que pueden verse reflejadas en una terminología global, tal vez debido a requerimientos de la legislación local, pueden crearse extensiones locales o nacionales. Por lo tanto, aunque SNOMED CT tiene un alcance global, se puede adaptar a los requerimientos de cada país o área.

Las referencias cruzadas de SNOMED CT proporcionan vínculos explícitos a clasificaciones relacionadas con la atención de la salud y esquemas de clasificación utilizados en todo el mundo.

También están disponibles o se encuentran en desarrollo por miembros de la Health Standards Development Organisation (IHTSDO), que será desarrollada más abajo, referencias cruzadas hacia o desde varios sistemas nacionales de codificación. Grupos de especialidades con los que la IHTSDO tiene acuerdos de colaboración, también mantienen referencias cruzadas desde o hacia sistemas de codificación de dominios específicos. Las referencias cruzadas facilitan la reutilización de datos clínicos basados en SNOMED CT para otros fines, tales como reembolsos o comunicaciones estadísticas.

Asimismo, SNOMED CT es una terminología multinacional y multilingüe. Tiene un marco integrado para manejar diferentes idiomas y dialectos. La Edición Internacional incluye un conjunto de conceptos y relaciones independientes del idioma.

Actualmente SNOMED CT se encuentra disponible en inglés estadounidense y del Reino Unido, castellano, danés y sueco. Se están llevando a cabo traducciones parciales en francés de Canadá, lituano y otros idiomas y los miembros de la IHTSDO planean nuevas traducciones a otros idiomas.

El objetivo básico de cualquier traducción de SNOMED CT es brindar representaciones exactas de los conceptos de SNOMED CT en forma comprensible, utilizable y segura. Las traducciones deben estar basadas en conceptos. Los traductores deben analizar los conceptos basándose en la descripción completa y teniendo en cuenta su posición en la jerarquía, sus descripciones y sus relaciones con otros conceptos.

Esto permite realizar una traducción de un concepto significativa y basada en frases utilizadas frecuentemente y claramente comprendidas en todos los países. La IHTSDO mantiene lineamientos y otros materiales de apoyo para los países que realizan traducciones.

SNOMED CT es propiedad de la International “Health Standards Development Organisation (IHTSDO)”, la organización responsable del mantenimiento, gestión, actualización y distribución de SNOMED CT. Más adelante se hará una descripción detallada de esta organización y su estructura para comprender su funcionamiento y los procesos de toma de decisiones a su interior.

3.1 COMPONENTES DE SNOMED CT

SNOMED CT es una terminología clínica central que contiene conceptos con significados únicos y definiciones basadas en lógica formal organizadas en jerarquías.

El contenido de SNOMED CT se representa con tres tipos de componentes:

- a) Conceptos, que representan significados clínicos organizados en jerarquías.
- b) Descripciones, que vinculan términos legibles con conceptos.
- c) Relaciones, que vinculan cada concepto con otros.

Estos componentes se complementan con conjuntos de referencia, que proporcionan características adicionales flexibles y permiten configurar la terminología para satisfacer diferentes necesidades.

A su vez, el modo en el que se relaciona y representa cada tipo de componente de la terminología y sus derivados se denomina “Modelo Lógico de SNOMED CT”. Los tipos de componentes centrales en SNOMED CT son esos conceptos, descripciones y relaciones. Por lo tanto, el modelo lógico especifica una representación estructurada de los conceptos que se utilizan para representar significados clínicos, las descripciones utilizadas para referirse a ellos y las relaciones entre conceptos.

Por otra parte, la manera de definir los conceptos de la terminología mediante una combinación de lógica formal y reglas editoriales conforma el “modelo conceptual de SNOMED CT”. Las reglas del modelo conceptual especifican los conceptos del nivel superior debajo de los cuales se disponen los conceptos en la jerarquía de subtipos y los tipos de relaciones permitidas entre conceptos en ramas específicas de la jerarquía.

3.1.1 Conceptos

Los conceptos de SNOMED CT representan pensamientos médicos, tales como |absceso| o |cigoto|. **Cada concepto tiene un identificador numérico único.**

Los conceptos de SNOMED CT están organizados en jerarquías. En una jerarquía, los conceptos se organizan a partir de los más generales a los más detallados. Los conceptos relacionados en la jerarquía se vinculan mediante la relación |es un/una|. Ejemplos de algunas jerarquías incluyen |hallazgo clínico|, |procedimiento|, |entidad observable|, |estructura corporal| y |organismo|.

3.1.2 Descripciones

Las descripciones de SNOMED CT vinculan términos legibles por el ser humano con conceptos. Un concepto puede tener varias descripciones asociadas y cada una de ellas representa un sinónimo que describe el mismo concepto médico. Cada traducción de SNOMED CT incluye un conjunto adicional de descripciones, que vinculan los términos en otros idiomas con los mismos conceptos de SNOMED CT. **Cada descripción tiene un identificador numérico de descripción único.**

3.1.3 Relaciones

Las relaciones de SNOMED CT vinculan conceptos con otros conceptos cuyo significado está relacionado de alguna manera. Estas relaciones proveen las definiciones formales y otras definiciones del concepto. Un tipo de relación es la relación |es un/una| que relaciona un concepto con conceptos más generales. Estas relaciones |es un/una| definen la jerarquía de conceptos de SNOMED CT. Por ejemplo, los conceptos |neumonía bacteriana| y |neumonía viral| tienen una relación |es un/una| con |neumonía infecciosa|, que a su vez tiene una relación |es un/una| con el concepto más general |neumonía|.

Otros tipos de relaciones representan aspectos del significado de un concepto. Por ejemplo, el concepto |neumonía viral| tiene una relación |agente causal| con el concepto |virus| y una relación |sitio del hallazgo| con el concepto |pulmón|. **Cada relación tiene un identificador numérico de relación único.**

3.1.4 Modelo Lógico de SNOMED CT

El modelo lógico de SNOMED CT brinda la estructura fundamental de SNOMED CT y especifica cómo deben manejarse los componentes en un ámbito de implementación para cubrir una variedad de usos primarios y secundarios. Define el modo en el que se relaciona y representa los conceptos, las descripciones y las relaciones, especificando una representación estructurada de los conceptos que se utilizan para representar significados clínicos, las descripciones utilizadas para referirse a ellos y las relaciones entre conceptos.

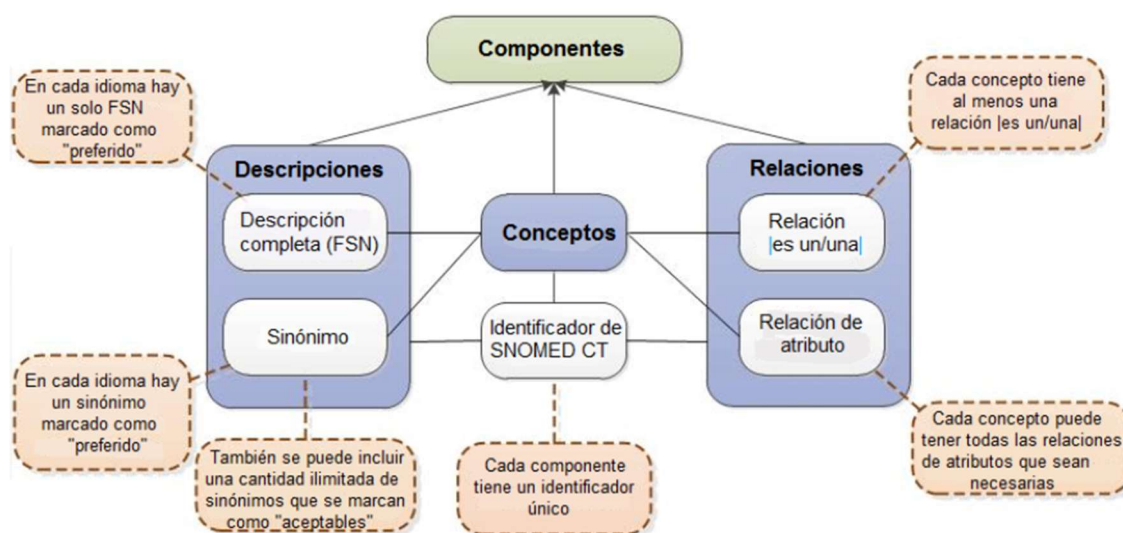


Figura 4: Estructura de SNOMED CT

En ese sentido, cada concepto representa un significado clínico único, al que se referencia con un identificador numérico único de SNOMED CT que puede ser leído por un ordenador. El identificador brinda una referencia única, no ambigua, a cada concepto y carece de cualquier significado interpretable por el ser humano.

Otros tipos de componentes también tienen identificadores únicos --- sin embargo, el identificador del concepto desempeña un papel específico como código utilizado para representar el significado en historias clínicas, documentos, mensajes y datos.

A su vez a cada concepto se asigna un conjunto de descripciones textuales. Éstas constituyen la forma legible de un concepto. Se utilizan dos tipos de descripciones para representar cada concepto --- Descripción completa (FSN, por Fully specified name en inglés) y Sinónimo.

La descripción completa representa una descripción única, no ambigua, del significado de un concepto. La descripción completa no está destinada a ser visualizada en las historias clínicas, sino que se utiliza para desambiguar el significado de cada concepto. Esto es particularmente útil cuando se designan diferentes conceptos con la misma palabra o frase de uso frecuente. Cada concepto puede tener **una sola descripción completa** en cada idioma o dialecto.

Un sinónimo representa un término que puede utilizarse para mostrar o seleccionar un concepto. Un concepto puede tener varios sinónimos. Esto permite que los usuarios de SNOMED CT utilicen los términos que prefieren para referirse a un significado clínico específico. Los conceptos pueden tener numerosos sinónimos y los términos asociados no necesariamente son únicos --- por lo tanto, dos conceptos pueden tener el mismo término sinónimo. Por lo tanto, la interpretación de un término sinónimo depende del identificador del concepto.

Cada concepto tiene un sinónimo que se marca como |preferido| en un idioma, dialecto o contexto de uso. Este se denomina "término preferido" y es una palabra o frase utilizada frecuentemente por los médicos para nombrar a ese concepto. En cada idioma, dialecto o contexto de uso, **sólo un sinónimo puede marcarse como |preferido|**. Otros sinónimos que son válidos en un idioma, dialecto o contexto de uso pueden marcarse como |aceptables|.

En la Figura 5 se recoge un ejemplo de descripciones de un concepto en castellano.

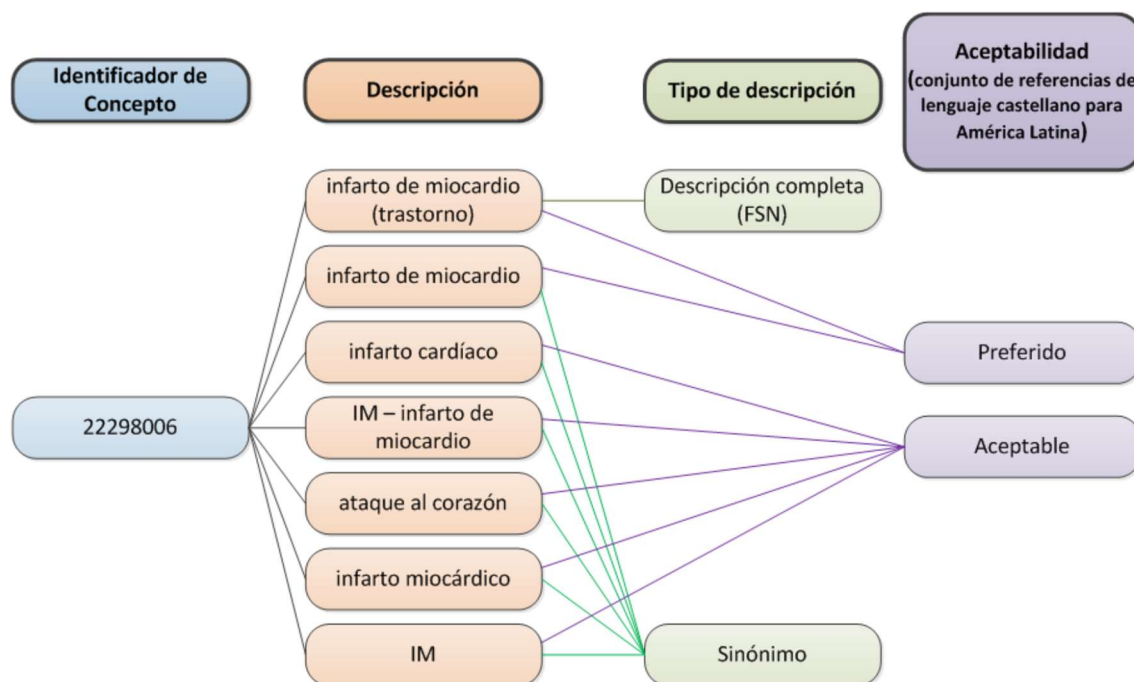


Figura 5: Descripciones de un concepto

Por su parte, como se ha dicho más arriba, una asociación entre dos conceptos representa una relación. Las relaciones se utilizan para definir lógicamente el significado de un concepto de manera que pueda procesarlo un ordenador. Un tercer concepto, denominado el tipo de relación (o atributo), se utiliza para representar el significado de la asociación entre el concepto de origen y el de destino. En SNOMED CT existen diferentes tipos de relaciones.

Las **relaciones de subtipos** son el tipo de relación más ampliamente utilizado. Utilizan el tipo de relación |es un/una| y, por lo tanto, también se conocen como relaciones |es un/una|. Casi todos los conceptos activos de SNOMED CT son los conceptos de origen de al menos una relación |es un/una|. La única excepción es el concepto raíz |concepto de SNOMED CT| que es el más general. La relación |es un/una| especifica que el concepto de origen es un subtipo del concepto de destino. Las relaciones de SNOMED CT son direccionales y la relación |es un/una| leída en sentido inverso señala que el concepto de destino es un supertipo del concepto de origen.

Las relaciones |es un/una| forman las jerarquías de SNOMED CT. Por lo tanto, también se denominan "**relaciones jerárquicas**". El concepto de origen de la relación |es un/una| tiene un significado clínico más específico que el concepto de destino. Esto significa que el nivel de detalle clínico de los conceptos aumenta con la profundidad de las jerarquías.

Si dos conceptos están vinculados directamente por una única relación |es un/una|, el concepto de origen es un "subtipo hijo" del concepto de destino. Éste se denomina "supertipo padre". Cualquier concepto que sea el origen de una secuencia de una o más relaciones |es un/una| que conducen a un concepto de destino especificado es un "subtipo descendiente" de ese concepto.

En forma similar, cualquier concepto que sea el destino de una secuencia de una o más relaciones |es un/una| que conducen a un concepto de origen especificado, es un "supertipo ancestro" de ese concepto. También se dice que el concepto de origen de una relación |es un/una| está "incluido por" el concepto de destino y que el concepto de destino que una relación |es un/una| "incluye" al concepto de origen.

Cada concepto puede tener relaciones |es un/una| con otros conceptos (es decir, un concepto puede tener múltiples conceptos padres supertipos). Como resultado, la jerarquía de SNOMED CT no es un árbol simple, sino que tiene una estructura "polijerárquica".

Existen además las **relaciones de atributo**, que contribuyen a la definición del concepto de origen mediante su asociación con el valor de una característica definitoria. La característica (atributo) se especifica por el tipo de relación y su destino proporciona el valor.

A diferencia de las relaciones |es un/una|, que se utilizan para definir todos los conceptos, la aplicabilidad de cada tipo de relación de atributo se limita a un dominio y un rango definidos.

El dominio se refiere a los conceptos que pueden servir como conceptos de origen para un tipo especificado de relación de atributo. El rango se refiere a los conceptos que pueden servir de destino (valores) de esos atributos. La especificación del dominio y el rango asegura definiciones coherentes que se pueden utilizar para inferir relaciones semánticas adicionales para permitir la recuperación confiable, basada en el significado, de significados compuestos.

Por ejemplo, el concepto |absceso de corazón|, utilizan las relaciones de atributos |morfología asociada| y |sitio del hallazgo| para asociar el concepto de origen |absceso de corazón| con los conceptos de destino |absceso| y |estructura del corazón|.

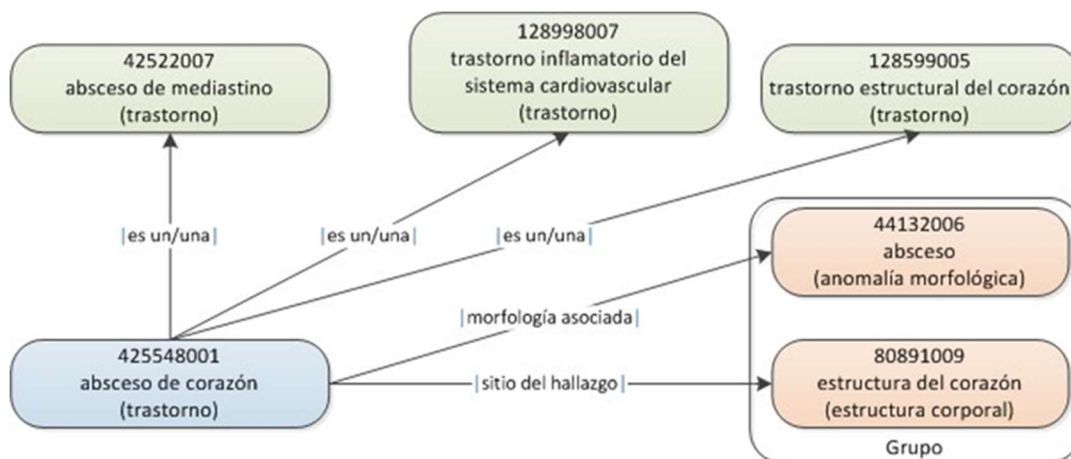


Figura 6: Ejemplo de utilización de relaciones de atributo

3.1.5 Modelo Conceptual de SNOMED CT

El modelo conceptual de SNOMED CT especifica la manera de definir los conceptos de la terminología mediante una combinación de lógica formal y reglas editoriales. Las reglas del modelo conceptual especifican los conceptos del nivel superior, debajo de los cuales se disponen los conceptos en la jerarquía de subtipos y los tipos de relaciones permitidas entre conceptos en ramas específicas de la jerarquía.

En el extremo superior de la jerarquía de SNOMED CT se encuentra **el concepto raíz** (concepto de SNOMED CT). Todos los conceptos descienden de este concepto raíz a través de al menos una secuencia de relaciones **|es un/una|**. Esto significa que el concepto raíz es un **supertipo** de todos los otros conceptos y éstos, a su vez, son subtipos de él.

Los subtipos directos del concepto raíz se denominan **"conceptos del nivel superior"**. Se utilizan para nombrar **las ramas principales de la jerarquía**. Cada uno de estos conceptos del nivel superior, junto con sus numerosos subtipos descendientes, forma una de las ramas principales de la jerarquía de SNOMED CT y contiene tipos de conceptos similares. A medida que se desciende por las jerarquías (es decir, a medida que se agregan más relaciones **|es un/una|** por debajo de los conceptos del nivel superior) los conceptos incluidos en ellas son cada vez más específicos.

Los conceptos del nivel superior son, entre otros:

- **Hallazgo clínico:** representa el resultado de una observación, una evaluación o un juicio clínico e incluye estados clínicos normales y anormales (por ejemplo, **|asma|**, **|cefalea|**, **|ruidos respiratorios normales|**). La jerarquía **|hallazgo clínico|** incluye a los conceptos utilizados para representar diagnósticos.

- **Procedimiento:** representa actividades que se llevan a cabo durante la atención de la salud. No sólo se incluyen procedimientos invasivos sino también administración de medicamentos, diagnóstico por imágenes, educación, terapias y procedimientos administrativos (por ejemplo, apendicectomía |, |fisioterapia |, |inyección subcutánea |).
- **Estructura corporal:** representa estructuras anatómicas normales y anormales (por ejemplo, |estructura de la válvula mitral |, |adenosarcoma |).
- **Organismo:** representa organismos relevantes para la medicina humana y veterinaria (por ejemplo, |Streptococcus pyogenes |, |sabueso |, |raza de ganado bovino texon |).
- **Sustancia:** representa sustancias en general, los constituyentes químicos de los productos farmacéuticos/biológicos, sustancias corporales, sustancias alimenticias y diagnósticas (por ejemplo |metano |, |insulina |, |albúmina |).
- **Producto farmacéutico / biológico:** representa los productos farmacológicos (por ejemplo, |amoxicilina 250 mg, cápsula |, |paracetamol + codeína comprimido |).
- **Espécimen:** representa entidades que se obtienen (por lo general del paciente) para realizar exámenes o análisis (por ejemplo, |espécimen de orina |, |espécimen obtenido de próstata mediante biopsia por punción |).
- **Estadificaciones y escalas:** representa escalas de evaluación y sistemas de estadificación tumoral (por ejemplo, |escala de Coma de Glasgow |, |estadificación de cáncer ginecológico de la Federación Internacional de Ginecología y Obstetricia |).

Por su parte, los atributos de SNOMED CT (o tipos de relación) se utilizan para representar una característica del significado **de un concepto**. La terminología actualmente utiliza más de cincuenta atributos definitorios para definir el significado de los conceptos. Cada atributo de SNOMED CT **puede aplicarse a conceptos en una rama de la jerarquía o más**. El conjunto de conceptos a los que se puede aplicar un atributo se denomina el '**dominio**' del atributo. El conjunto de **valores permitidos para cada atributo se denomina 'rango' del atributo**.

El dominio es la jerarquía a la que se puede aplicar un atributo específico. Por ejemplo: El dominio del atributo |morfología asociada| es la jerarquía |hallazgo clínico|. Por lo tanto, un |procedimiento| no puede tener una |morfología asociada|. En cambio, puede tener una |morfología de procedimiento|.

El rango es el conjunto de conceptos de SNOMED CT permitidos como valores de un atributo especificado. Por ejemplo: El rango para el atributo |morfología asociada| es el concepto |estructura

morfológicamente anormal| y sus subtipos descendientes. El rango para el atributo |sitio del hallazgo| es |estructura corporal anatómica o adquirida| y sus subtipos descendientes en la jerarquía |estructura corporal|.

3.2 LA INTERNATIONAL HEALTH TERMINOLOGY STANDARDS DEVELOPMENT ORGANISATION (IHTSDO)

La IHTSDO es la organización internacional⁵ propietaria y administradora de SNOMED CT, que tiene los derechos sobre SNOMED CT y estándares terminológicos relacionados.

IHTSDO es una asociación gobernada por una Asamblea General que contiene un representante de cada miembro nacional (países o estados miembros), su objetivo principal es intentar mejorar la salud de la humanidad mediante el fomento del desarrollo y el uso de terminologías clínicas estandarizadas adecuadas, en particular SNOMED CT, con el fin de apoyar el intercambio seguro, preciso y eficaz de información médica y relacionada con la salud.

La atención está centrada en permitir la implementación de historias clínicas semánticamente precisas que sean interoperables. El propósito de la IHTSDO es desarrollar, mantener, promover y habilitar la captación y uso correcto de sus productos terminológicos en sistemas, servicios y productos para la atención de la salud en todo el mundo y llevar a cabo actividades tendientes a cumplir con el objetivo de la Asociación para beneficio de sus miembros

La IHTSDO actualmente cuenta con 32 miembros nacionales (<https://www.snomed.org/members>). Los miembros de la IHTSDO pueden ser una agencia de un gobierno nacional u otro organismo respaldado por una autoridad gubernamental del país que representa.

Los miembros de la IHTSDO pagan una cuota, basada en el patrimonio nacional, lo que les da derecho a formar parte de la Asamblea General. La IHTSDO no cobra a los Licenciarios Afiliados por el uso de la Edición Internacional de SNOMED CT en los países miembros. La licencia también es gratuita en los países más pobres y para ciertos usos de bien público en otros países.

En otros casos, se aplican cargos por institución de bajo costo por el uso de SNOMED CT en sistemas operativos involucrados en la creación y/o análisis de datos.

⁵ De acuerdo a lo que surge del contrato de Licencia que se debe suscribir para la utilización de SNOMED IHTSDO es una sociedad limitada por garantía constituida de acuerdo con la legislación inglesa (registrada con el número 09915820), con domicilio social en One Kingdom Street, Paddington Central, Londres

Todo uso de SNOMED CT está sujeto a la aceptación de los términos del acuerdo de Licencia de Afiliado de SNOMED CT y los de las sublicencias emitidas a organizaciones de usuarios finales por los Licenciarios Afiliados.

Los miembros también pueden imponer otras condiciones para el uso de SNOMED CT en sus países. Por ejemplo, un miembro puede requerir que los licenciarios den soporte e incluyan su Extensión Nacional en sus implementaciones.

3.2.1 Estructura de IHTSDO

Como se puede observar en la Figura 7, el máximo órgano de gobierno de la IHTSDO es la Asamblea General, que incluye un representante designado formalmente por cada Estado Miembro de SNOMED.

La Asamblea General nombra un Consejo de Administración para dirigir la Asociación que es responsable de las decisiones clave. El Consejo de Administración nombra un Director Ejecutivo responsable de la conducción cotidiana de la organización.

Por su parte, la organización cuenta con un Foro de miembros, que actúa como un órgano asesor de IHTSDO, y cuenta con la representación de todos los países miembros. El Foro de Miembros busca optimizar la colaboración y la coordinación entre sus miembros, establecer las prioridades en materia de proyectos y productos nuevos, que serán propuestos a la Asamblea General, promoviendo además el intercambio de experiencias y mejores prácticas

Los representantes del Foro de Miembros, deben ser idealmente conocedores de todos los aspectos operacionales de la agencia nacional del estado miembro al que representan. Tienen canales directos de comunicación con la Asamblea General, y aunque estructuralmente, deben asesorar al Consejo de Administración, también tiene fuertes vínculos con los niveles de trabajo de la organización (Comités y Grupos de Trabajo).

Existen dentro de la IHTSDO cuatro comités permanentes que asesoran al Consejo de Administración, y son elegidos por la Asamblea General para cada comité.

Se reúnen en persona dos veces por año y entre esas reuniones se llevan a cabo conferencias telefónicas en intervalos regulares. En general, las reuniones de estos comités son abiertas para observadores.

Los comités permanentes son:

- **Comité de Contenido:** Es responsable de brindar asesoramiento sobre aspectos relacionados con la definición y mantenimiento del contenido médico y la estructura de SNOMED CT y sus estándares relacionados.
- **Comité de implementación y educación:** Es responsable de brindar asesoramiento, ayuda y capacitación para una implementación práctica y efectiva de SNOMED CT de maneras que faciliten la administración de la información médica y la adopción de decisiones.
- **Comité de Aseguramiento de la Calidad:** Es responsable de desarrollar y garantizar la calidad de SNOMED CT y sus estándares relacionados, así como de otros productos terminológicos de la Asociación en sintonía con el debido respeto por estándares externos.
- **Comité Técnico:** Está centrado en aspectos relacionados con la tecnología en lo que respecta al uso y aplicación de SNOMED CT y su aptitud para el uso de marcos y herramientas adoptadas en la aplicación de SNOMED CT.

Además de los comités permanentes existen los denominados Grupos de Trabajo, constituidos para dar tratamiento a cuestiones específicas, cuya duración puede ser más o menos extensa de acuerdo a los términos que demande la concreción de su tarea.

Los grupos de trabajo incluyen:

- **Grupos de Interés Especial** (SIG, por Special Interest Groups en inglés), que son foros abiertos que analizan aspectos de temas particulares o áreas de especialidades relacionadas con la atención de la salud.

Hay dos tipos diferentes de Grupos de Interés Especial, los **basados en tópicos**, que abarcan un área general (por ejemplo, implementación, traducción) y reportan a uno de los comités permanentes y los **de especialidades profesionales**, que se centran en los requerimientos de una disciplina o especialidad médica en particular (por ejemplo, Enfermería, Anestesia) y reportan al **Grupo de Coordinación Profesional de la Atención de la Salud**, que incluye a los presidentes de cada uno de los SIG profesionales.

- **Grupos de Proyectos** que tienen objetivos definidos, de tiempo limitado.

Finalmente, la organización cuenta con un Foro de Enlace de Proveedores de Servicio (Vendor Liaison Forum) que al igual que el Foro de Miembros es un asesor del Consejo de Administración, aportando la opinión de los proveedores sobre el desarrollo, lanzamiento e implementación de SNOMED CT.

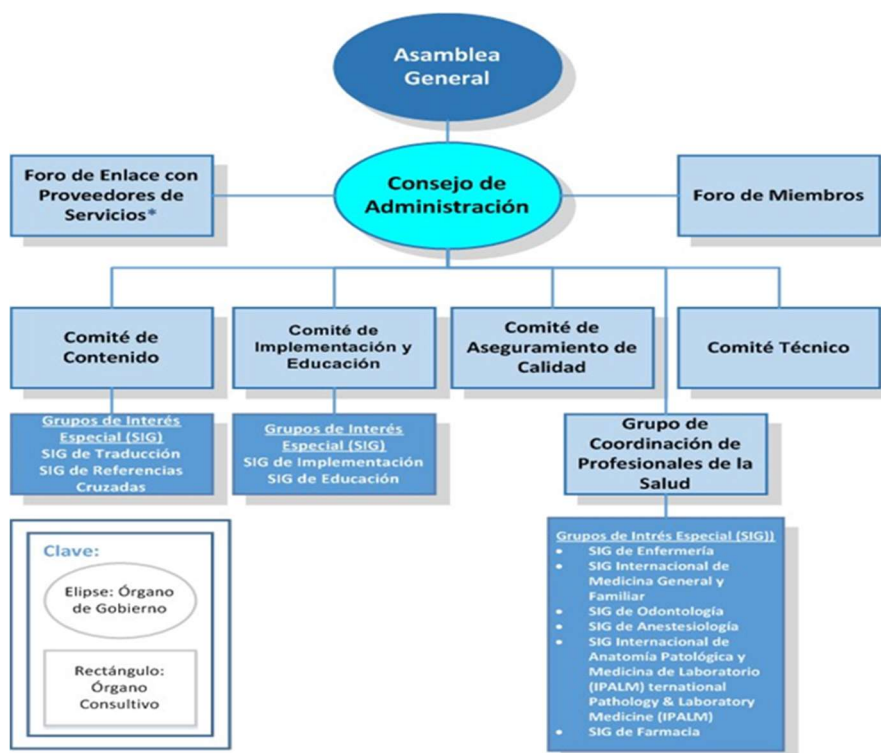


Figura 7: Estructura de IHTSDO

3.2.2 Desarrollo del contenido de SNOMED CT

El desarrollo del contenido de SNOMED CT se basa en cuatro principios básicos que guían el contenido clínico y el diseño técnico. Esos principios son incluyen los siguientes preceptos:

- Los esfuerzos de desarrollo deben abarcar una amplia participación de diversos grupos clínicos y expertos en informática médica,
- El contenido clínico debe centrarse en la calidad y adherir a reglas editoriales estrictas,
- El proceso de mejora de la calidad debe estar abierto a la inspección por el público y a las opiniones de proveedores para garantizar que la terminología sea verdaderamente útil en las aplicaciones relacionadas con la atención de la salud,
- Las barreras para su adopción y utilización deben ser mínimas.

El contenido de SNOMED CT evoluciona y los tipos de cambios realizados pueden incluir nuevos conceptos y descripciones, nuevas relaciones entre conceptos y nuevos conjuntos de referencias, así

como actualizaciones y retiro de cualquiera de estos componentes. Estos cambios se llevan a cabo debido a modificaciones en la comprensión de procesos normales y patológicos, introducción de nuevos fármacos, investigaciones, tratamientos y procedimientos y nuevas amenazas para la salud, así como propuestas y trabajos que presentan los usuarios de SNOMED CT.

Los tres criterios operativos básicos que ayudan a determinar si el nuevo contenido sigue el principio de crear y mantener la interoperabilidad semántica, plantean que SNOMED CT debe ser:

- **Comprensible:** Debe ser posible comunicar el significado, de manera que sea comprensible por el profesional de la salud promedio sin referencia a significados inaccesibles, ocultos o privados.
- **Reproducible:** No es suficiente que una persona piense que comprende un significado. Debe demostrarse que varias personas comprenden un significado de la misma manera.
- **Útil:** El significado debe tener algún uso o aplicabilidad demostrables relacionados con la salud o su cuidado.

Si el objetivo es que SNOMED CT se convierta en un recurso terminológico internacionalmente aceptado para la atención de la salud, debe ser capaz de permitir representaciones terminológicas multilingües de conceptos comunes. Para que la terminología sea aceptable para la más amplia gama de usuarios debe incluir traducciones, variantes léxicas y otras variaciones de dialectos nacionales y regionales. Además, debe ser capaz de representar diferencias entre conceptos surgidos de variaciones culturales, étnicas o lingüísticas.

El desarrollo de una terminología clínica representa un desafío por diversas razones ya que incluso en un mismo país o conjunto de idioma, las personas utilizan las mismas palabras con significados diferentes o distintas palabras para designar una misma cosa.

Los nombres asignados a algunas afecciones médicas en ocasiones se basan en una comprensión incompleta o errónea y a menudo estos nombres engañosos continúan en uso mucho tiempo después de que el conocimiento ha avanzado. Los avances en los conocimientos médicos y la evolución de los organismos patógenos crean una necesidad continua y creciente de agregar nuevo contenido y de revisar las definiciones.

Los esfuerzos por establecer criterios diagnósticos y escalas de estadificación de entidades especializadas también provocan cambios y, en ocasiones, divergencias entre fuentes de autoridad diferentes o superpuestas.

Ante estos desafíos, el desarrollo del contenido está orientado a considerar las prioridades actuales y emergentes identificadas por los miembros de la IHTSDO y otros interesados.

El objetivo de la IHTSDO es la mejora continua de la calidad. Los procesos de control de la calidad se incluyen como parte del trabajo que realiza el equipo de modeladores que participan en el desarrollo de SNOMED CT. Múltiples editores médicos definen y revisan el contenido siguiendo un proceso científico documentado. Los conflictos entre editores se resuelven mediante un proceso iterativo, basado en el logro de acuerdos y consensos, antes de incorporar un concepto a la terminología. Según sea necesario, el equipo de editores consulta con otros expertos para revisar la integridad científica del contenido.

La IHTSDO brinda un servicio para reunir y procesar las solicitudes de agregados y cambios al contenido de la Edición Internacional de SNOMED CT. Este servicio es accesible directamente a través de los Centros Nacionales en los países miembros y de Expertos en Terminología reconocidos en las organizaciones con las cuales colabora activamente la IHTSDO.

Los Centros Nacionales, son las organizaciones o agencias de cada país miembro de la IHTSDO, que llevan a cabo todas las actividades relacionadas con su participación en la IHTSDO y su papel en la distribución, extensión y apoyo del uso de SNOMED CT en su país. Los Centros Nacionales brindan un punto único de contacto para comunicarse con la IHTSDO y otros miembros de la IHTSDO y en sus propios países, manejan el uso de SNOMED CT y se comunican con diversas partes interesadas, incluyendo Licenciarios Afiliados a SNOMED CT, instituciones sanitarias, grupos clínicos y usuarios finales.

En ese sentido, las organizaciones de los países miembros pueden presentar sus pedidos de agregados y cambios a sus respectivos Centros Nacionales. En algunos casos, las solicitudes con relevancia local pueden agregarse a una extensión nacional.

Los Centros Nacionales envían las solicitudes que podrían tener relevancia internacional a la IHTSDO para que decida sobre ellas. Si una solicitud se considera altamente prioritaria debe reflejarse en una acción en el siguiente ciclo de edición. Sin embargo, las solicitudes que requieren cambios significativos que podrían influir sobre otro contenido podrían demorar más tiempo.

4. INFRAESTRUCTURAS DE DATOS ESPACIALES (IDE)

Una infraestructura de datos espaciales, puede definirse como un proyecto colectivo implantado por una comunidad de actores, consistente en un conjunto de recursos orientados a compartir Información Geográfica en la web de manera abierta y estandarizada.

Cuando se dispone de datos georreferenciados, de cierta disponibilidad de recursos informáticos y se quiere o se tiene la necesidad de publicar la información geográfica de la manera más eficaz posible, es necesario contar con una infraestructura que permita compartir, intercambiar, combinar, analizar y acceder a los datos geográficos de forma estándar e interoperable.

Esta infraestructura no es más que el conjunto de recursos cartográficos disponibles en la red, sobre la que los datos mismos serán más útiles al formar parte de un todo más completo.

Como se ha mencionado al principio de este trabajo, la idea de estandarización e interoperabilidad es fundamental a la hora de hablar de una infraestructura lingüística o de una infraestructura de datos espaciales.

Los estándares son reglas generales, que facilitan la adopción de soluciones genéricas y la posibilidad de gestionar todos los componentes del mismo tipo de la misma manera.

Para la gestión eficaz de la información geográfica, los estándares son fundamentales, si un usuario descarga de una institución la descripción geográfica de un río en un formato y sistema de referencia determinados, lo ideal es que no encuentre problemas para cargarla y visualizarla en una aplicación SIG (sistemas de información geográfica) junto con el resto de sus datos.

La interoperabilidad, por su parte, se basa en la idea de aplicar la filosofía de los sistemas abiertos a los SIG. En tal sentido, si se dispone de un conjunto de sistemas que gestionan IG, que mediante protocolos e interfaces estándares permiten acceder a los datos en remoto, es posible generar aplicaciones que integren esos datos y que aparezcan virtualmente frente al usuario como si fuesen un único sistema, sin necesidad de conocerlos en detalle, sólo basándose en información genérica sobre los estándares que cumplen todos y cada uno de ellos.

En conclusión, una IDE es un Sistema de Información Geográfica abierto implementado sobre la Red, con todo lo que ello conlleva: componentes distribuidos, interfaces estándares, interoperabilidad, coordinación, acceso a los datos, capacidad de análisis como objetivo, etcétera.

Por su parte, según el portal de la IDE de España, una IDE *“es un sistema informático integrado por un conjunto de recursos (catálogos, servidores, programas, aplicaciones, páginas web,...) que permite el acceso y la gestión de conjuntos de datos y servicios geográficos (descritos a través de sus metadatos), disponibles en Internet, que cumple una serie normas, estándares y especificaciones que regulan y garantizan la interoperabilidad de la información geográfica. Así mismo es necesario establecer un marco legal que asegure que los datos producidos por las instituciones serán compartidos por toda la administración y que promueva su uso entre los ciudadanos⁶”* .

De esta forma, con la puesta en marcha de las IDE se consigue que un usuario, tanto experto en IG como inexperto, pueda realizar, si se cumplen los estándares, las diferentes tareas que ordenadas por cierto orden lógico se citan a continuación:

- Buscar la IG que hay disponible en una zona geográfica con descripción del formato, la manera de acceder a ella, el año en que se produjo, la calidad que ofrece, quién la ha elaborado, si existe un servicio que la publique y demás características relevantes.
- Visualizar y superponer mapas, ortofotos, MDT y datos geográficos en general de diferentes organismos, con diferentes sistemas de referencia, en distintos formatos y con propiedades heterogéneas.
- Buscar una entidad geográfica por su nombre y ver dónde se ubica sobre una cartografía.
- Acceder a las entidades geográficas en un formato estándar, así como a sus atributos, coordenadas, topología y geometría.
- Realizar operaciones de análisis básicas, como enrutamiento, cálculo de perfiles o análisis de superficies.
- Realizar transformaciones de un modelo de datos a otro diferente, si ambos están descritos de forma normalizada.
- Descargar los datos que se precisen para analizarlos en un SIG, si los servicios ofrecidos por la IDE no satisfacen las necesidades del usuario.

⁶ Ver http://www.idee.es/resources/documentos/Introduccion_IDEE.pdf

Con el establecimiento de las IDE se pretende eliminar los obstáculos que dificultan la disponibilidad y accesibilidad de la IG, que ocasionan problemas y pérdidas de tiempo, y que hacen muy difícil y costosa la reutilización de datos geográficos para un propósito diferente al original.

Hoy en día, gracias al desarrollo de internet, organizaciones de todo el mundo pueden conectarse a una gran red de comunicación y publicar información a la que otros pueden acceder y utilizar. Se han desarrollado varias infraestructuras basadas en Internet, como es el caso de las IDE, que permiten la integración de información geográfica.

Las IDE, al igual que otras infraestructuras, están basadas en acuerdos entre sus miembros que, entre otras cosas, concretan el uso de tecnologías para publicar, acceder, visualizar y procesar los datos por Internet.

Sin la existencia de estándares la utilización de la información sería muy difícil y se consumirían grandes cantidades de recursos en homogeneizarla, por ello la necesidad de disponer de normas y estándares que faciliten la interoperabilidad de la información geográfica.

Esas dificultades para reutilizar la información generada por diferentes sistemas geográficos y la pérdida de información (colores, estilos, grosores, topologías) que ocurría cuando se exportaban ficheros generados en un sistema a otro, impulsó la creación de organizaciones que eliminaran esas barreras y facilitaran compartir la información.

Por un lado, la Asociación de Estandarización Internacional (ISO) mediante su Comité Técnico 211, generó el conjunto de Normas ISO 19100. Por otro, un consorcio compuesto de agencias gubernamentales, universidades, compañías y centros de investigación, se unieron con la misión de promover el uso de estándares y tecnologías abiertas en el área de sistemas y tecnologías de la información geográfica y afines conformando el Open Geospatial Consortium (OGC).

A su vez, ISO/TC211 y OGC han establecido un acuerdo de cooperación orientado al aprovechamiento de los desarrollos, y a la reducción de la duplicidad de esfuerzos. El acuerdo cooperativo entre estas dos organizaciones de estandarización, permite:

- Adaptar estándares y especificaciones comunes,
- Estructurar programas de trabajo complementarios,
- Compartir recursos,
- Prevenir inconsistencias e incompatibilidades de normativa.

En el marco de este acuerdo, el OGC ha adoptado varias normas ISO/TC 211 como especificaciones abstractas en las que basa sus especificaciones de implementación. De forma recíproca, algunos estándares desarrollados originalmente por el OGC se han llevado al ISO/TC 211, para publicarse finalmente como normas ISO internacionales, existiendo otras adicionales en proceso ante el ISO/TC211. Existe un grupo coordinador de ISO / OGC llamado “Joint Advisory Group (JAG) que se reúne durante la mayoría de las reuniones del Comité Técnico de OGC.

A continuación, se hará una descripción del OGC, su estructura y funcionamiento, sus programas fundamentales para el estímulo y desarrollo tanto de las IDE como de otras estrategias de intercambio de IG en sistemas distribuidos.

4.1 OPEN GEOEPACIAL CONSORTIUM (OGC)

El OGC es una organización que tiene como misión promover el uso de estándares y tecnologías abiertas en el área de sistemas y tecnologías de la información geográfica y áreas afines. Agrupa ya más de 400 organizaciones incluyendo agencias gubernamentales, universidades, compañías y centros de investigación que pretenden colaborar con el desarrollo de especificaciones y estándares.

Actualmente el Comité técnico del OGC promueve la creación de grupos de trabajo, constituidos en su gran mayoría por voluntarios de las organizaciones miembros, que se rigen mediante procesos de consenso. Los resultados de los grupos de trabajo se traducen en estándares abiertos y públicos que permiten soluciones interoperables que facilitan el acceso, la manipulación y el intercambio de información geoespacial en la web.

Las actividades del OGC se articulan alrededor de cuatro programas:

- a) Programa de estandarización o especificación,
- b) Programa de innovación,
- c) Programa de cumplimiento,
- d) Programa de Comunicaciones y Extensión.

El consorcio cuenta con el personal y la tecnología necesaria para dinamizar la actividad de todos los involucrados, formando una infraestructura de colaboración que facilita el trabajo en equipo, inclusive cuando sus miembros se hallen dispersos por todo el mundo.

Entre las tecnologías se encuentra un sistema de gestión de contenidos en la web, foros, infraestructura para teleconferencia y listas de correo. El Comité técnico del OGC se reúne físicamente cuatro veces al año en diferentes partes del mundo.

EL OGC se financia con las cuotas de sus miembros, licencias de certificación y con los ingresos generados por la gestión de proyectos, que forman parte del programa de innovación.

4.1.1 Estructura de OGC

Como podrá observarse en la figura que sigue a continuación, el órgano máximo del OGC es la Junta Directiva (Board of Directors). Los miembros de la junta son líderes de diferentes disciplinas, representantes de diferentes regiones en el mundo.

Los Directores de OGC están dedicados a lograr los objetivos establecidos en la visión de OGC, proporcionando su experiencia profesional para el cumplimiento de esos fines. Su cargo se ejerce a título personal y representan las necesidades de los diferentes sectores a los que representan, y en ningún caso representan a sus organizaciones de origen.

De acuerdo a los estatutos de OGC, los negocios y asuntos del OGC se regirán por su Junta Directiva, que será y tendrá todos los poderes que la Ley General de Corporaciones de Delaware, Estados Unidos, otorga al "Consejo de Administración" de una corporación de membresía sin fines de lucro.

Los directores tienen un mandato de 2 años, y son elegidos previa recomendación del Comité Ejecutivo, que eleva a la Junta Directiva una lista completa de candidatos, integrada por líderes respetados de la industria, los gobiernos y la academia.

Para integrar la lista de candidatos, las personas no deben necesariamente integrar organizaciones, empresas o universidades miembros de OGC.

La Junta se expide en relación a la lista elevada por el Comité Ejecutivo, remitiéndola al Comité de Planificación quién tomará la decisión definitiva relativa a los Directores de la Junta. El número total de Directores no podrá ser menor de nueve ni mayor de veinticinco.

La Junta en el primer orden del día, en la primera reunión de su mandato tiene como tarea la elección del Presidente, y de un Comité Ejecutivo, compuesto por entre cinco y siete de los Directores electos. El presidente de la Junta actuará también como presidente del Comité Ejecutivo.

La junta puede, además, mediante resolución adoptada por la mayoría de los Directores, crear otros Comités que estime pertinentes. Estos comités se reunirán por separado y presentarán sus informes a la Junta.

Otro órgano del OGC es el Comité Asesor Estratégico de Miembros (SMAC) compuesto por los representantes de los Miembros Estratégicos del OGC, cuyas principales funciones son:

- Asesorar en relación a la conveniencia y de oportunidad de determinadas estrategias sobre las operaciones del Consorcio a la Junta Directiva, al personal del Consorcio y a la Membresía;
- Recomendar estrategias de recursos para apoyar los programas del Consorcio; y
- Apoyar el Programas de interoperabilidad y estándares con recursos de gestión y operativos.

Por su parte, el Comité de Planificación (Planning Committee), compuesto por los miembros estratégicos y principales, los miembros del Comité Ejecutivo de la Junta Directiva y dos representantes del Comité Técnico, tiene la responsabilidad de:

- Realizar la planificación estratégica tecnológica del OGC referida al desarrollo de estándares y especificaciones de geoprocesamiento, considerando las posibilidades de su adopción por el mercado,
- Ratificar los planes de desarrollo de la Especificación abstracta de OpenGIS®, la publicación de horarios, planes de pruebas de conformidad y cualquier otro documento importante producido por el Comité Técnico,
- Realizar la aceptación final de los estándares de implementación de OpenGIS según se informa a partir del Comité Técnico;
- Realizar la elección de los directores de la Junta Directiva, en base a la lista de nominados remitida por el Comité Ejecutivo;
- Determinar, para someter a aprobación de la Junta, las políticas y procedimientos internos del Comité, lo que puede involucrar la creación de subcomités, mesas de trabajo, grupos de trabajo especiales, etc.

Otro de los comités del OGC es el Comité Técnico (Technical Committee), integrado por los representantes de la miembros técnicos, principales y estratégicos, presididos por un miembro del personal permanente del OGC, designado por el Presidente de la Junta Directiva.

Sus funciones están sujetas a la supervisión final del Comité de Planificación y son:

- El desarrollo de la especificación abstracta de OpenGIS a través de un proceso cooperativo de consenso que involucrara a todos los Miembros;
- El establecimiento de las pautas para su propia organización y procedimientos internos, la que podrá involucrar la creación de mesas de trabajo, grupos de trabajo de interés especial (SIG), Subcomités y cualquier otra figura que el Comité Técnico considere necesarios;
- Elevar al Comité de Planificación para su aprobación los borradores de las versiones de la especificación abstracta de OpenGIS y sus estándares de implementación, así como cualquier otro documento importante relacionado;

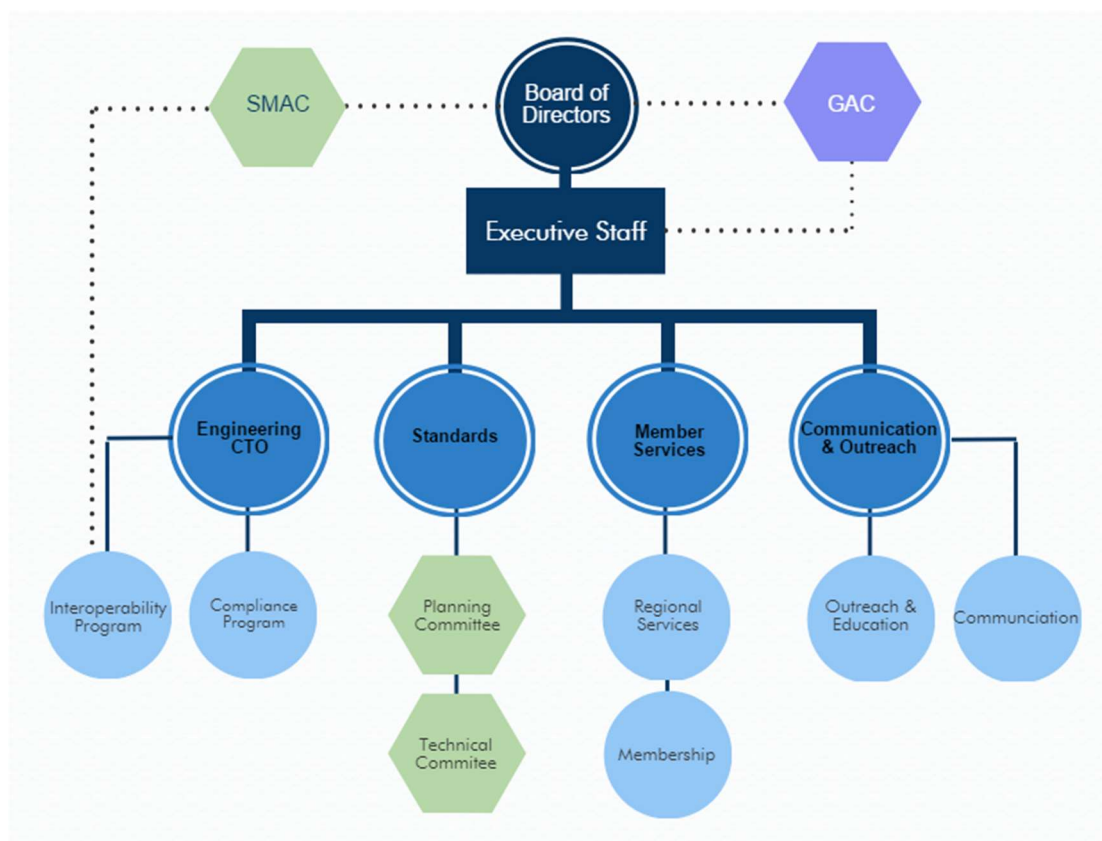


Figura 8: Estructura de OGC

4.1.2 Membrecía

Cualquier asociación, sociedad, organización, compañía o corporación, individuo que tenga un interés en los propósitos del OGC, puede, satisfechas ciertas condiciones, convertirse en Miembro del Consorcio.

Quien tenga interés en ser miembro, realizará una solicitud por escrito, y deberá abonar la membresía, posteriormente a ser admitido.

Existen diferentes niveles y clases de membresía, las que darán derecho, de acuerdo a la estructura descrita en el apartado anterior, a participar en los diferentes comités. El orden en el que se describen las membresías es jerárquico, y la membresía superior goza además de todos los derechos que atribuye la membresía inferior.

- **Nivel de Membresía Estratégica:** mediante la aprobación previa de la Junta Directiva, la membresía se proporciona a las organizaciones que desean proporcionar recursos significativos, además de las cuotas de membresía principal, en apoyo de los programas y objetivos del Consorcio, así como desempeñar un papel activo en el desarrollo de estándares y de especificación técnicas, o que pertenezcan a mercados de valor estratégico para el Consorcio y sus organizaciones miembro.

Cada Miembro Estratégico, mientras esté en regla, tendrá derecho a designar un representante con derecho a voto en el “Comité Asesor de Miembros Estratégicos” (SMAC), trabajar estrechamente con la Membresía del OGC y el personal para crear programas especiales diseñados para mejorar su efectividad en áreas específicas.

- **Nivel de Membresía Principal:** la membresía principal se proporciona a las organizaciones que desean participar en la planificación y gestión de los procesos de desarrollo tecnológico del Consorcio. Cada Miembro Principal, mientras esté en regla, tendrá derecho a designar un representante con derecho a voto en el Comité de Planificación; votar y participar en los subcomités, mesas de trabajo y grupos de trabajo de intereses especiales; votar en el Comité de Planificación para ratificar la lista aprobada de directores candidatos a integrar la Junta Directiva; participar en actividades de planificación y divulgación que el Consorcio organiza por medio del Comité de Planificación con sus Miembros Principales.
- **Nivel de Membresía del Comité Técnico:** la membresía del Comité Técnico es proporcionada a las organizaciones que desean participar en el proceso de desarrollo de la tecnología del OGC. Este nivel de membresía habilita la designación de un representante con derecho a voto en dicho comité, otorgando además el derecho de presentar propuestas tecnológicas para consideración del Consorcio. Además, estos miembros pueden votar y participar en los subcomités, mesas de trabajo, grupos de trabajo de intereses especiales y designar candidatos a la Junta Directiva para someterlos a consideración del Comité de Planificación.
- **Nivel de Membresía Asociada:** cada miembro asociado, mientras esté en regla, tendrá derecho a designar un representante sin voto en el Comité Técnico que podrá participar en las tareas y debates del Comité y en los subcomités, mesas de trabajo y grupos de trabajo de

intereses especiales. Además, tienen acceso a todo el software y la documentación general publicada por el OGC con carácter previo a su lanzamiento, reciben los informes y revisiones de todas las reuniones plenarias de los Comités y de sus grupos de trabajo y todas las publicaciones del Consorcio que están destinadas a la distribución regular. Hay seis categorías de Membresía Asociada, que prevén diferentes beneficios, derechos y privilegios, establecidos por la Junta de Directiva:

- a) **Membresía Asociada Comercial:** se aplica a empresas, agencias gubernamentales y organizaciones no gubernamentales que no son principalmente de naturaleza educativa.
- b) **Membresía Asociada Comercial de Pequeña Empresa:** se aplica a PyMES mediante la documentación que verifique que sus ingresos anuales son menores al umbral establecido regularmente por la Junta Directiva que es y publicado en el sitio web de OGC.
- c) **Membresía Asociada sin fines de lucro:** se aplica a ONGs o asociaciones sin fines de lucro de carácter privado, locales, nacionales o internacionales.
- d) **Membresía Asociada de GovFuture:** aplica a entidades gubernamentales de carácter local (incluyendo alcaldías, provincias, etc.)
- e) **Membresía Asociada Universitaria:** aplica a "instituciones educativas" tales como departamentos universitarios, museos, centros de ciencia, centros de investigación, escuelas e instalaciones de capacitación.
- f) **Membresía Asociada Individual:** individuos y/o contratistas independientes con capacidad para contratar y que no hayan cedido derechos de propiedad a otra persona o entidad legal. Para ser aceptado el individuo debe probar que es un consultor independiente, que no está contratado a tiempo completo a un solo empleador (excepto cuando la organización es miembro del OGC en buen estado), la membrecía es individual y no es transferible.

4.1.3 Programas

El OGC cuenta con cuatro programas básicos que orientados al cumplimiento de sus objetivos, que le facilitan la tarea de innovación en el desarrollo de estándares de alta calidad para la comunidad geoespacial. Esos programas son:

- **Programa de estandarización o especificación:** en el cual el Comité Técnico y el Comité de Planificación trabajan en un proceso formal de consenso para llegar a los estándares OGC aprobados (o "adoptados").

El programa define el proceso mediante el cual se formaliza un estándar OGC. Se caracteriza por abordar el problema de la interoperabilidad desde un punto de vista teórico y deliberativo, donde los resultados deben aprobarse de forma consensuada por parte de los miembros de OGC.

El eje coordinador de este programa es el Comité técnico (Technical Committee, TC), el cual decide qué trabajos se van a emprender y sobre la validez de los resultados.

El punto de partida para la elaboración de un estándar OGC puede ser una propuesta específica llevada a cabo por un miembro, basada en la detección de una carencia o necesidad a partir de los resultados del Programa de interoperabilidad.

El TC forma un grupo de trabajo de estándares (SWG) ad hoc para cada propuesta. El SWG es el encargado de concretar la propuesta, siguiendo las pautas establecidas por el conjunto de documentos de la Especificación Abstracta (Abstract Specification), un modelo para el desarrollo de especificaciones establecido por OGC.

El resultado es una serie de documentos técnicos y propuestas de acciones. El TC es el marco donde los miembros de OGC consideran las propuestas de los SWG y donde, entre otras cosas, se recomienda la adopción de estándares OGC.

El TC también puede promover la creación de grupos de trabajo temáticos (DWG) donde se discuten las cuestiones relacionadas con la interoperabilidad de información geoespacial en sectores profesionales determinados o para temas concretos. Los resultados del trabajo de los DWG también deben formalizarse según la tipología OGC y las propuestas deben elevarse al TC para su consideración.

La última responsabilidad la ostenta el Comité de planificación, quien tiene capacidad de aprobar y adoptar las recomendaciones del TC y hacerlas públicas como estándar OGC.

- **Programa de innovación⁷** : se trata de un programa global, innovador y práctico de prototipos y pruebas diseñado para acelerar el desarrollo y la validación de interfaces a fin de llevar la interoperabilidad al mercado.

⁷ En el siguiente link pueden verse las iniciativas en materia de innovación que el programa tiene en curso <http://www.opengeospatial.org/projects/initiatives>

El Programa permite definir, documentar, mejorar y poner a prueba las especificaciones actuales de OGC. Uno de los propósitos principales de este programa, es conseguir de manera rápida y práctica (basados en esfuerzos de colaboración), el desarrollo de estudios y prototipos de infraestructuras interoperables basados en estándares OGC y especificaciones presentados por los participantes de los proyectos.

- **Programa de cumplimiento:** el Programa de Cumplimiento OGC proporciona los recursos, procedimientos y políticas para mejorar el cumplimiento de las implementaciones de software con los estándares OGC. Brinda un servicio gratuito de pruebas en línea, un proceso para la certificación y desarrollo de marcas de productos que cumplen con los requisitos, y la coordinación de una comunidad que desarrolla y respalda los scripts de prueba. También proporciona servicios de pruebas de conformidad que permite evaluar y certificar la compatibilidad de una aplicación con un estándar determinado y ejecuta eventos a corto plazo para aumentar la interoperabilidad entre los productos de los proveedores.
- **Programa de Comunicaciones y Extensión:** El OGC y sus miembros ofrecen recursos para ayudar a los desarrolladores y usuarios de tecnología a aprovechar los estándares abiertos de OGC. Los documentos técnicos, materiales de capacitación, suites de prueba, implementaciones de referencia y otros recursos de interoperabilidad desarrollados por OGC están disponibles de manera pública. Además, OGC y sus miembros dan apoyo y soporte a publicaciones, talleres, seminarios y conferencias para ayudar a los desarrolladores de tecnología, integradores y gerentes de adquisiciones a introducir los estándares y aplicaciones desarrolladas en el marco de OGC en sus arquitecturas.

4.1.4 Dominios

Como se ha mencionado más arriba, el objetivo de OGC es lograr el consenso de las diferentes comunidades de interés, en busca de alcanzar el objetivo de resolver problemas relacionados con la creación, comunicación y uso de información espacial.

Esas comunidades de interés representan diferentes dominios, cuya actividad es muy diversa, como por ejemplo la observación del océano, el modelado 3D de entornos urbanos, información de ubicación de eventos durante los desastres, etc.

Dentro de OGC, esas comunidades de interés se clasifican en dominios y el Consorcio busca favorecer la cooperación y el diálogo dentro y entre estos dominios, a fin de implementar pruebas, proyectos piloto y experimentación de interoperabilidad para ser examinados en el Programa de Estándares OGC.

Los dominios de OGC son:

- **Aviación:** para mejorar la seguridad operacional y la eficiencia operacional de los viajes en avión, la comunidad aeronáutica mundial avanza en la adopción de un marco internacional de normas que permitan la comunicación en un Sistema de Transporte Aéreo (ATS) globalmente interoperable y centrado en la red. Debido a que la información de ubicación es crítica en prácticamente todas las actividades de aviación, la interfaz de ubicación y los estándares de codificación del OGC juegan un papel importante.
- **Sector de la Construcción y modelado 3D:** los interesados en cada etapa del ciclo de vida de los edificios y otros proyectos de capital dependen de la información espacial interna y externa. Desafortunadamente, comunicar información espacial a menudo es problemático. El OGC, trabajando con otras organizaciones de estándares y asociaciones industriales, desarrolla estándares abiertos para respaldar la productividad en las cadenas de suministro del diseño de edificios, la infraestructura física, los proyectos de capital y las industrias de administración de instalaciones.
- **Inteligencia de negocios:** la ubicación geográfica suele ser un factor crítico en la investigación y los cálculos comerciales sobre clientes, proveedores, distribuidores, recursos naturales, centros de transporte, energía y otras cuestiones comerciales.
- **Defensa e Inteligencia (D&I):** casi todas las aplicaciones de tecnología de la información de Defensa e Inteligencia (D&I) - planificación, inteligencia, logística, etc. - involucran información geoespacial. Casi todos los activos y todas las amenazas, humanas y materiales, tienen una ubicación o un área geográfica. Los sistemas D & I son diversos y están cada vez más interconectados, por lo que las organizaciones y contratistas de D & I han jugado un papel importante en el desarrollo de estándares OGC, y ahora juegan un papel importante en D&I en todo el mundo.
- **Respuesta a emergencias y gestión de desastres:** respuesta a emergencias y gestión de desastres son diferentes dominios de actividad con diferentes requisitos para compartir información. En ambos dominios, sin embargo, existe la necesidad de descubrir, compartir, integrar y aplicar rápidamente la información geoespacial. El desarrollo de estándares de

OGC juega un papel clave para abordar esta necesidad. Acelerar el flujo de información de ubicación significa más vidas y bienes ahorrados y menos riesgos para los primeros en responder.

- **Energía y servicios públicos:** la comunicación sobre la ubicación es esencial en todo lo relativo a la energía y los servicios públicos. Desde la exploración de petróleo y gas hasta la ubicación de parques eólicos, la conservación de energía municipal y el mantenimiento de tuberías y cables. La comunicación por ubicación se vuelve aún más esencial con el rápido aumento en la generación y el almacenamiento de energía distribuida. Por otra parte, las compañías de generadoras de energía y prestadoras servicios públicos necesitan comunicarse con los administradores de desastres y las fuerzas de emergencia, con los sensores que poseen y que no son de su propiedad. La industria de comercio de carbono necesita un estándar abierto para comunicar los parámetros relacionados con la ubicación de los pasivos por emisión de carbono y los activos de compensación de carbono. El OGC tiene un papel clave en la evolución del régimen energético del siglo XXI.
- **Geociencias y medio ambiente:** los sistemas de la tierra están acoplados y, por lo tanto, el descubrimiento y el intercambio de recursos de información geoespacial en todas las disciplinas es fundamental para los investigadores y los administradores ambientales. En el OGC, las universidades, las organizaciones de investigación, las ONG, las empresas y las organizaciones gubernamentales trabajan juntas para eliminar los obstáculos técnicos al intercambio de información institucional.
- **Infraestructura de datos gubernamentales y espaciales:** la gestión de la infraestructura física requiere infraestructura de comunicación e información. Las infraestructuras de datos espaciales (IDE) desempeñan un papel importante, porque la información de ubicación es fundamental en la gestión de todo lo que deben administrar los gobiernos, desde las carreteras y las alcantarillas hasta la educación y la salud pública. Al igual que la mayoría de los tipos de infraestructura, las IDE también proporcionan una plataforma para el desarrollo económico. Las organizaciones gubernamentales participan en el OGC tanto para avanzar en los estándares como para aprender sobre ellos.
- **Servicios móviles de Internet y localización:** OGC trabaja para lograr la interoperabilidad para que la información espacial disponible en la web pueda ser procesada por los dispositivos móviles conectados a internet móvil.

- **Webs de sensores:** hay cientos de millones de sensores conectados a Internet alrededor de la Tierra, y la cantidad crece rápidamente. La estandarización es esencial para comunicar información, incluida la información de ubicación, sobre sensores y fenómenos detectados. Los estándares de OGC cumplen este requisito tanto en aplicaciones complejas como simples, y facilitan la integración de información de sensores en miles de aplicaciones geoespaciales que implementan los otros estándares de OGC.
- **Universidad e investigación:** la membresía de OGC ayuda a universidades y organizaciones de investigación a contribuir al desarrollo de prácticas de recopilación, procesamiento y diseminación de datos científicos. También proporciona un ambiente fértil en el que los departamentos de geomática, informática, geografía y geociencia de las universidades pueden modernizar y avanzar sus planes de estudios, agendas de investigación y planes de transferencia de tecnología, trabajando codo a codo con otras universidades y miembros de la industria y el gobierno.

4.1.5 Estándares OGC

El principal producto o resultado de los programas OGC son los estándares materializados en especificaciones, es decir, en documentos que detallan interfaces informáticas o formas de codificación de datos. Cada estándar se piensa para solucionar aspectos específicos de interoperabilidad, de manera que su implementación en productos y servicios produzca resultados independientes del productor y de la implementación. En la aproximación cliente servidor, los estándares deben permitir que los desarrollos informáticos a cada lado puedan intercambiar datos sin necesidad de adecuar los correspondientes códigos.

OGC trabaja también en términos de estándar abierto, lo que significa que deben estar disponibles de forma pública y libre, sin discriminación alguna, sin costes, independiente de quien lo proporciona o de los datos que maneja y aprobado formalmente mediante consenso.

La elaboración de estándares debe llevarse a cabo de forma coordinada y coherente con el estado de desarrollo ya conseguido dentro de OGC.

El modelo de referencia OGC es un documento que describe las relaciones entre los documentos considerados básicos, es decir, los documentos de **estándares abstractos y de implementación** (interfaz, codificación, perfil o esquema de aplicación), y **las buenas prácticas**. El modelo de referencia se actualiza periódicamente y proporciona una visión general de los resultados del trabajo de los miembros de OGC que han contribuido a esa documentación básica.

5. CONCLUSIONES

No es difícil encontrar casos en los que diferentes organismos, organizaciones, empresas, administraciones, sin contacto entre ellas, generan recursos lingüísticos ya existentes o contratan servicios para obtener un producto que ya está disponible, produciendo gastos innecesarios.

Tampoco es raro el hecho de que estos recursos, producidos por un organismo público, permanezcan ajenos al resto de los organismos de la misma administración, ya sea por falta de información u otras causas.

Estas circunstancias generan una serie de inconvenientes, además del sobregasto ya mencionado, que van desde el desconocimiento de la existencia de estos recursos, a la multiplicidad de formato en los que se encuentra almacenada la información en cada institución productora y desde luego a la manera de compartirla de una forma rápida en caso de necesidad.

Es recomendable que este tipo de recursos no exista de manera duplicada y que los que existen estén homogeneizados.

La necesidad de compartir estos recursos, pone de manifiesto la importancia de crear infraestructuras lingüísticas, que los pongan a disposición de manera abierta, distribuida, interoperable y segura.

Las infraestructuras que se han analizado en el presente informe, han surgido producto de la necesidad de compartir información y estandarizar la utilización y el intercambio de la misma, que los dominios específicos a los cuales responden han hecho surgir.

La gestión de esas infraestructuras es muy diferente, y en algunos casos, como las IDE, existe un Consorcio que de manera colaborativa consensua normas y estándares para promover que el desarrollo de cualquier infraestructura de este tipo, sea interoperable con los desarrollos realizados en otros ámbitos o dominios específicos.

Lo cierto es que a la hora de pensar o imaginar la implementación de una infraestructura de estas características (ya sea lingüística o de datos espaciales) no cabe duda que el componente político es fundamental.

Una infraestructura lingüística es un proyecto colectivo en el que participan una gran diversidad de actores y necesita iniciativas y actuaciones legales que establezcan y regulen su participación.

Intervienen las administraciones, las empresas privadas, las universidades y los usuarios, cada uno con intereses y necesidades diferentes y es muy importante que sus iniciativas y actuaciones, estén

coordinadas, armonizadas e integradas en un contexto más amplio, con pleno reconocimiento ante terceros interesados y con efectos jurídicos y técnicos de conformidad con la normativa aplicable en los diferentes ámbitos, tanto en razón de la materia como del territorio.

Pero, ¿quién o qué organización tiene autoridad para esta labor de regulación? una solución sería que ese papel lo desempeñara un organismo colectivo, en el que todos los actores relevantes estuvieran representados, cuya organización y burocracia no fuera demasiado compleja y en el que todos los implicados se sientan cómodos.

La solución organizativa óptima depende, del ámbito de aplicación, de la cultura los países o regiones que la integren y fundamentalmente de los medios de que se disponga.

De acuerdo a los casos relevados, donde existen diferencias fundamentales de organización, es necesario identificar desde donde se parte, a la hora de proyectar el desarrollo de un proyecto de infraestructura lingüística.

Si existe voluntad política, recursos humanos especializados, recursos financieros establecidos en los presupuestos, y un organismo público con competencia específica, determinada por ley, como en el caso de la Biblioteca Nacional de Medicina (NLH) de la Secretaria de Salud y Servicios Humanos del Gobierno de los Estados Unidos, no cabe duda que ese modelo representa un caso de éxito, aunque en el dominio específico de la salud.

Esta infraestructura, hoy financiada por el gobierno de Estados Unidos, es aprovechada por toda la estructura compleja que integra el sistema de salud de ese país, pero además está disponible para los desarrollos y la utilización del público en general, de otros países y de otras organizaciones, viéndose limitada su utilización en los casos en los que no se posea la licencia para aquellos recursos específicos que son de pago.

Ahora bien, si no contamos con los requisitos mencionados anteriormente, como para que, mediante la inversión pública de un único país, se financie una infraestructura de este tipo, entonces quizás deba recurrirse a un modelo asociativo similar al que posee la International Health Terminology Standards Development Organisation (IHTSDO).

Esta Organización, si bien constituida de acuerdo a la ley inglesa, se sustenta en base a los aportes realizado por los estados miembros y las contribuciones surgidas del pago de las licencias de uso que realizan organismos u organizaciones pertenecientes a estados no miembros.

Claramente, para que este modelo pueda replicarse, debe existir voluntad política de los gobiernos de los países, para conformar este tipo de organización, y recursos disponibles en sus presupuestos para aportar a este organismo, cuya constitución debería de adecuarse a la normativa de uno de los países que la integren, o en su defecto, analizar la posibilidad de crear mediante un tratado internacional, una estructura supranacional de acuerdo al derecho internacional público.

Sin embargo, si la necesidad de crear una infraestructura de este tipo, no ha sido divisada con claridad por todos aquellos actores que se considera importante que participen en su desarrollo, una buena opción es la implementación de un proyecto piloto, con la participación y los aportes, en dinero o especie, que cada uno de los actores interesados pueda realizar, para poner en marcha, aunque sea de manera imperfecta, un primer atisbo de esta infraestructura.

El establecimiento de un marco legal común y la necesidad de coordinación entre los agentes implicados, en caso de un proyecto piloto, podría ser la firma de un acuerdo marco, que defina las líneas principales del proyecto, los objetivos, los recursos necesarios, los aportes que cada parte se compromete a realizar, la conformación de grupos de trabajo tanto a nivel técnico como a nivel político, entre otros aspectos.

Por último, la cuestión de la interoperabilidad y la normalización, a fin de que cualquier desarrollo de una infraestructura de estas características pueda ser compatible con otras ya existentes o a desarrollarse en el futuro, es fundamental, y aquí juegan un papel vital los organismos de normalización existentes a nivel local, regional y global, su comunicación y cooperación, como puede observarse en el ámbito de las IDE, en el que de manera coordinada adoptan estándares y especificaciones técnicas internacionales tendientes a garantizar la interoperabilidad.

6. FUENTES Y BIBLIOGRAFÍA

- [1] Unified Medical Language System (UMLS) <https://www.nlm.nih.gov/research/umls/>
- [2] About the National Library of Medicine <https://www.nlm.nih.gov/about/index.html>
- [3] UMLS® Reference Manual <https://www.ncbi.nlm.nih.gov/books/NBK9676/>
- [4] PUBLIC HEALTH SERVICE ACT <https://legcounsel.house.gov/Comps/PHSA-merged.pdf>
- [5] License Agreement for Use of the UMLS® Metathesaurus
<https://uts.nlm.nih.gov/help/license/LicenseAgreement.pdf>
- [6] Department of Health and Human Services National Institutes of Health National Library of Medicine (NLM) Justification of Budget Request <https://www.nlm.nih.gov/about/2018CJ-NLM.pdf>
- [7] Welcome to SNOMED International <https://www.snomed.org/>
- [8] Governance and Advisory <https://www.snomed.org/about/governance-and-advisory>
- [9] Technical Implementation Guide <https://confluence.ihtsdotools.org/display/DOCTIG>
- [10] OGC Organization <http://www.opengeospatial.org/ogc/organization>
- [11] Welcome to the OGC <http://www.opengeospatial.org/>
- [12] BYLAWS OF OPEN GEOSPATIAL CONSORTIUM
portal.opengeospatial.org/files/?artifact_id=6947
- [13] Guide to the Open Geospatial Consortium
https://portal.opengeospatial.org/files/?artifact_id=68511
- [14] FUNDAMENTOS DE LAS INFRAESTRUCTURAS DE DATOS ESPACIALES Miguel A. Bernabé Poveda, Carlos M. López Vázquez
http://redgeomatca.rediris.es/Libro_Fundamento_IDE_con_pastas.pdf