



E-Ciencias de la Información

ISSN: 1659-4142

Universidad de Costa Rica, Escuela de Bibliotecología y
Ciencias de la Información

Suárez Sánchez, Adriana; López Hernández, Arlin
Taxonomías digitales: fundamentos teóricos y metodología de construcción
E-Ciencias de la Información, vol. 10, núm. 1, 2020, Enero-Junio, pp. 87-108
Universidad de Costa Rica, Escuela de Bibliotecología y Ciencias de la Información

DOI: <https://doi.org/10.15517/eci.v10i1.39881>

Disponible en: <https://www.redalyc.org/articulo.oa?id=476864645005>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

UNER
redalyc.org

Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto



UNIVERSIDAD DE
COSTA RICA

EBCI

Escuela de
Bibliotecología y Ciencias
de la Información

e-Ciencias de la Información

Taxonomías digitales: fundamentos teóricos y metodología de construcción

Adriana Suárez Sánchez
Arlin López Hernández

Recibido: 3/09/2019 | Corregido: 23/10/2019 | Aceptado: 13/11/2019

DOI: [10.15517/eci.v10i1.39881](https://doi.org/10.15517/eci.v10i1.39881)

e-Ciencias de la Información, volumen 10, número 1, Ene- Jun 2020
ISSN: 1649-4142



¿Cómo citar este artículo?

Suárez Sánchez, A. y López Hernández, A. (2020). Taxonomías digitales: fundamentos teóricos y metodología de construcción. *e-Ciencias de la Información*, 10(1). Doi: 10.15517/eci.v10i1.39881

Taxonomías digitales: fundamentos teóricos y metodología de construcción

Digital taxonomies: theoretical fundamentals and construction methodology

Adriana Suárez Sánchez ¹

Arlin López Hernández ²

RESUMEN

El artículo analiza las taxonomías digitales desde la perspectiva bibliotecológica. Su objetivo fue construir una taxonomía digital del área de conocimiento “Sistema solar” que pudiera ser implementada como sistema de acceso temático jerárquico-visual para un conjunto de recursos de información digitales albergados en un sitio web. La metodología empleada en la construcción de la taxonomía consistió en una propuesta particular que se nutre de dos metodologías relevantes sobre el tema. Se encontró que las taxonomías son un Sistema para la Organización del Conocimiento (SOC) que, debido a su potencial representador y organizador, han migrado al espacio de los Sistemas para la Organización del conocimiento en Red (SOCR) cuyas funciones en la web están asociadas, entre muchas aplicaciones, a la indización y clasificación de recursos de información. Se concluye que el conocimiento teórico y aplicado de las taxonomías digitales es prioritario para los profesionales de la información que día a día se enfrentan a más retos relacionados con el ordenamiento de recursos de información insertos en la World Wide Web.

Palabras Clave: *Taxonomías digitales, Lenguajes documentales, Sistemas para la organización del conocimiento, Recuperación de la información.*

Abstract

The article analyzes digital taxonomies from the librarianship perspective. Its objective was to build a digital taxonomy about “Solar System” knowledge area that can be implemented as hierarchical-visual subject access system for a set of digital information resources stored in a website. The methodology used to build the taxonomy was a proposal based on two relevant methodologies about the topic. It was found that taxonomies are a Knowledge Organization Systems (KOS) that, due to its representational and organizer potential, have migrated to the space of Networked Knowledge Organization Systems (NKOS), whose tasks on the web are associated, among many applications, to the indexing and classification of information resources. It is concluded that the theoretical and applied knowledge of digital taxonomies

¹ Instituto de Investigaciones Bibliotecológicas y de la Información, Universidad Nacional Autónoma de México, MÉXICO. asuarez@iibi.unam.mx. ORCID: orcid.org/0000-0003-0215-337X

² Dirección General de Sitios y Monumentos del Patrimonio Cultural, Secretaría de Cultura, MÉXICO. netarlin@gmail.com. ORCID: orcid.org/0000-0001-8455-4884



is a priority for information professionals who every day deals with challenges related to the organization of information sources inserted in the World Wide Web.

Key Words: *Digital taxonomies, Documental Languages, Knowledge Organization Systems, Information Retrieval*

1. Introducción

La actividad taxonómica está presente en casi todos los ámbitos de la actividad humana “-de hecho, las taxonomías son cristales a través de los cuales percibimos y concebimos el mundo en que vivimos” (Lambe, 2007, p.4). En gran medida, nuestro mundo está poblado de estructuras taxonómicas o categorías relativas a la familia, los alimentos, las cosas, los conceptos, las actividades, los sentimientos, los lugares, etc. Su utilización extendida radica en que el ser humano es un ordenador nato (Taylor & Joudrey, 2009, p.1): desde la parte epistemológica organiza para darle sentido a lo existente, desde la parte social para el desarrollo del conocimiento individual y grupal y desde el ámbito pragmático para recuperar algo con mayor facilidad.

Las taxonomías han sido usadas en diversas disciplinas como la filosofía, la biología, las ciencias cognitivas y la bibliotecología (Abbas, 2010). Originalmente, estuvieron asociadas al contexto impreso, pero, como señala Centelles (2005), desde la década de los noventa han estado vinculadas con la web. En ésta, algunos sitios desestiman estructurar la información, pero, en comparación, otros son extremadamente conscientes de que un sitio bien estructurado “puede ser fácilmente entendido por otros sitios web y por los buscadores, así como cumplir mejor sus objetivos de diseminar la información lo más ampliamente posible, además de permitir al usuario encontrar lo que busca” (Fernández-Hernández, 2007, p.2).

Díaz, Joyanes & Medina señalan que “en el ambiente tecnológico, para la gestión eficaz de información y contenidos, la taxonomía es elemento esencial” (2009, p.246). En vista de que organizan información, datos y recursos de información en el ciberespacio y, de igual modo, son aplicables tanto en la categorización de un sitio web como en la creación de un sistema más complejo fundamentado en temas, nombres individuales, nombres corporativos, etc.

Considerando lo anterior, el presente artículo tiene por objetivo: construir una taxonomía digital del área de conocimiento “Sistema solar” que sirva como herramienta de acceso temático a un conjunto reducido de recursos de información digitales (imágenes, videos y recursos textuales) que serán organizados en un sitio web considerando sólo su perspectiva temática. El supuesto del que se parte es que las taxonomías digitales pueden ser aplicadas en la organización temática de recursos digitales en sitios web, bajo nodos conceptuales-categoriales. Tal postulado se estableció a partir de la definición de Hodge (2000, p. 6) quien caracteriza las taxonomías como un “sistema para la organización del conocimiento con características de notación conceptual y estructuración jerárquica que les posibilita representar temas amplios en la red”.

Como antecedente del estudio, se retoma que autores como Centelles (2005), Lambe (2007) y Abbas (2010) han señalado la viabilidad de las taxonomías digitales en la organización de sitios web; no obstante, persisten pocos trabajos documentados y asociados a su utilización en la organización temática de recursos en la red, especialmente en el contexto de la bibliotecología latinoamericana³. En tal sentido, la investigación sobre el tema se vislumbra relevante.

2. Referente teórico

El vocablo taxonomías proviene del griego τάξις *táxis* 'ordenamiento' y νόμος *nómos* 'norma' o 'regla'. Reitz (2004) establece que es

The science of classification, including the general principles by which objects and phenomena are divided into classes, which are subdivided into subclasses, then into sub-subclasses, and so on" [La ciencia de la clasificación, incluidos los principios generales por los cuales los objetos y los fenómenos se dividen en clases, que se subdivide en subclases, luego en sub-subclases, y así sucesivamente] (p. 708)

Martínez de Sousa (2004), coincide con Reitz (2004) cuando señala que es la ciencia de los principios de la clasificación y agrega que ordenan jerárquicamente conceptos que para su presentación son colocados en una estructura arbórea. Graef (2001) destaca que son un sistema que mediante el uso del lenguaje asigna nombres a entidades y las organiza en grupos que muestran características similares. Boeri (2004) también concibe las taxonomías como un sistema que organiza lógicamente categorías mediante palabras clave y estructuras visuales que van de niveles altos a específicos.

Las taxonomías tienen su origen más remoto en las categorizaciones filosóficas (Abbas, 2010; Centelles, 2005). En el campo, Aristóteles (384-332 a.c.), Porfirio (233-305), Isidoro de Sevilla (570-636), Konrad Gesnet (1516-1565), Francis Bacon (1561-1626), Thomas Hobbes (1588-1679), Wilhelm Leibniz (1646-1716) y Auguste Comte (1789-1857) generaron propuestas asociadas a la organización del conocimiento o la realidad que consistían en nomenclaturas terminológicas agrupadas en clases y subclases.

En épocas posteriores, según Centelles (2005), "se desarrolló la taxonomía como un subcampo de la biología" (p. 2). A mediados del siglo XVIII los naturalistas plantearon que si bien los nombres vulgares eran útiles para nombrar a los seres vivos, se observaba necesario un sistema más universal y riguroso: cada especie debía ser nombrada, cada especie debía poseer un único nombre, y cada especie debía ser descrita sin ambigüedades. Abbas (2010) indica que:

³ Al hacer una búsqueda exploratoria en el descubridor de información de la DGB- (<http://dgb.unam.mx/>) que incorpora todas las bases de datos de la Universidad Nacional Autónoma de México- empleando el término "Digital taxonomies and Information sources" se obtienen 647 registros y al buscar "Digital taxonomies and Indexing" se recuperan sólo cinco registros. En la búsqueda en español, al buscar "Taxonomías digitales and Recursos de información" se obtienen sólo 5 registros que incluyen el término en alguna parte del texto y no están vinculados con la implementación de taxonomías en la organización temática de recursos de información en sitios web, mientras que al buscar "Taxonomías digitales and Indización" no se obtienen registros.



The eighteenth-century naturalists Carl Linnaeus, a botanist from Sweden, and Georges-Louis Leclerc, Comte de Buffon, a French nobleman and student of nature, were central figures both in defining natural history as a discipline and for developing structures for organizing knowledge that prevail today in the natural sciences, such as biology, anthropology, and evolutionary biology. [Los naturalistas del siglo XVIII Carl Linnaeus, un botánico de Suecia, y Georges-Louis Leclerc, conde de Buffon, un noble francés y estudioso de la naturaleza, fueron figuras centrales tanto en la definición de la historia natural como una disciplina y el desarrollo de estructuras para la organización de conocimiento que prevalecen hasta hoy en día en las ciencias naturales como la biología, antropología y biología evolutiva] (p.30).

Durante la primera mitad del siglo XX la ciencia cognitiva fue otra área que contribuyó significativamente al desarrollo de las estructuras categoriales en la construcción de sistemas para la organización del conocimiento (Abbas, 2010, p.34). Un vasto grupo de psicólogos, entre ellos Lev Vygotsky (1896-1934), Jerome Seimor Bruner (1915-2016), Gorge Armitage Miller (1920-2012) y Jorge Lakoff (1941-) analizaron cómo funciona la mente humana y descubrieron la existencia de nodos categorizadores que constituyen una estructura a partir de la cual el ser humano concibe la realidad, estructura procesos y realiza deducciones.

Situados en la bibliotecología, las taxonomías son, formalmente, un “sistema para la organización del conocimiento (KOS- Knowledge Organization Systems [SOC-Sistemas para la Organización del Conocimiento])” (Hodge, 2000, p.1). Éstos son todos los tipos de esquemas para “organizar información y promover la gestión del conocimiento” (Hodge, 2000, p.1), tal es el caso de clasificaciones, encabezamientos de materias, tesauros, taxonomías, mapas tópicos y ontologías. Tudhope (2006) coincide con la definición de Hodge y agrega que en el universo de los KOS es posible distinguir tres grupos: listados de temas, clasificaciones y categorizaciones y esquemas de relación. Bajo tal marco, las taxonomías corresponden al segundo grupo y se componen de conceptos estructurados en categorías.

Originalmente las taxonomías eran sistemas enclavados en el ámbito impreso; sin embargo, derivado de las posibilidades tecnológicas, tuvieron un proceso de migración hacia ambientes digitales. Reitz (2004, p. 708) argumenta que “the term has been applied more recently within the information sector to the classification of resources available via the World Wide Web [el término ha sido más recientemente aplicado a la clasificación de recursos disponibles vía World Wide Web]”.

Tal traslado no ha sido exclusivo de las taxonomías, sino de varios KOS entre los que es posible mencionar también los tesauros y las clasificaciones. Sobre el tema, Chowdhury y Chowdhury (2007) señalan que:

With the introduction of the web and digital libraries, the world of information has changed dramatically over the last decade in terms of volumen, variety, format, nature and complexity of electronic information resources. This gave rise to the need for new and improved tools, techniques and standards for organizing information” [Con la introducción de la web y las bibliotecas digitales, el mundo de la información ha cambiado drásticamente en la última década en

términos de volumen, variedad, formato, naturaleza y complejidad de los recursos de información electrónicos. Esto dio lugar a la necesidad de herramientas, técnicas y estándares nuevos y mejorados para organizar la información] (p.18).

En la red las taxonomías son esquemas jerárquicos utilizados para sistematizar el acceso a contenidos o sitios web:

Taxonomies advance information search and retrieval by providing powerful browsing capabilities based on structured content organisation and access via point-and-click directories or menu selections [Las taxonomías permiten avanzar en la búsqueda y recuperación de información al proporcionar potentes capacidades de navegación basadas en la organización de contenido estructurado y el acceso a través de directorios de apuntar y hacer clic o selecciones de menú]. (Corcoran, 2002, p.76).

En la web, las funciones de las taxonomías digitales (Gilrisch, 2002) están asociadas a:

- a) Directorios web
- b) Indización
- c) Categorización
- d) Filtros del Front-End (parte de un sistema que interactúa con los usuarios)
- e) Taxonomías corporativas y de negocios

Como bien señala Gilrisch (2002), una función de las taxonomías digitales en los contextos red radica en la tarea de indización que es:

Analysing the contents of a document (book, pamphlet, audiovisual or machine-readable item, etc.) or collection of documents and translating the results of the analysis into terms for use in an index [el análisis del contenido de un documento (libro, panfleto, audiovisual o ítem legible por máquina, etc.) o colección de documentos y la traducción del resultado del análisis en términos para usar en un índice] (Feather & Sturges, 2003, p. 236).

En tal aplicación, las taxonomías actúan como sistemas para la organización del conocimiento, de forma similar a la función que tienen los encabezamientos de materia o los tesauros (Hodge, 2000) que permiten la representación de entidades o áreas de saber y la correspondiente categorización de recursos de información en nodos contenedores de entidades que comparten contenido temático. Sin embargo, se diferencian de éstos porque se insertan completamente en el ambiente digital donde se visualizan como estructuras organizadas que ayudan a los usuarios en la exploración del tema que están buscando mediante la navegación de temas generales a específicos (Ménard, 2013).

3. Metodología

En comparación con los tesauros donde prevalece una larga historia de estándares para la construcción de los instrumentos -como ISO 2788:1986. *Guidelines for the establishment and Development of Monolingual thesauri* (ISO, 1986), ISO 25964-1:2011. *Thesauri for Information Retrieval* (ISO, 2011)



e ISO-25964-2:2013. *Interoperability with Other Vocabularies* (ISO, 2013)- la construcción de taxonomías no cuenta con estandarización. Cuando un profesional de la información se enfrenta a la tarea de crear una taxonomía, debe acercarse a una metodología particular.

Considerando lo anterior, para la construcción de la taxonomía digital del dominio "Sistema solar" se retomaron dos propuestas metodológicas individuales:

1) Heather Hedden en su obra *The accidental taxonomist*: menciona que no hay escuelas de taxonomía, pero la necesidad de organizar información ha dibujado y atraído a practicantes de diversas disciplinas, como bibliotecólogos, publicistas, arquitectos de la información, indizadores y gestores de la información y el conocimiento (Hedden, 2010, p. xii- xiii). En la construcción de taxonomías considera tres etapas: (1) establecimiento de términos, (2) creación de estructura categorial y (3) estructuración de la taxonomía en un programa.

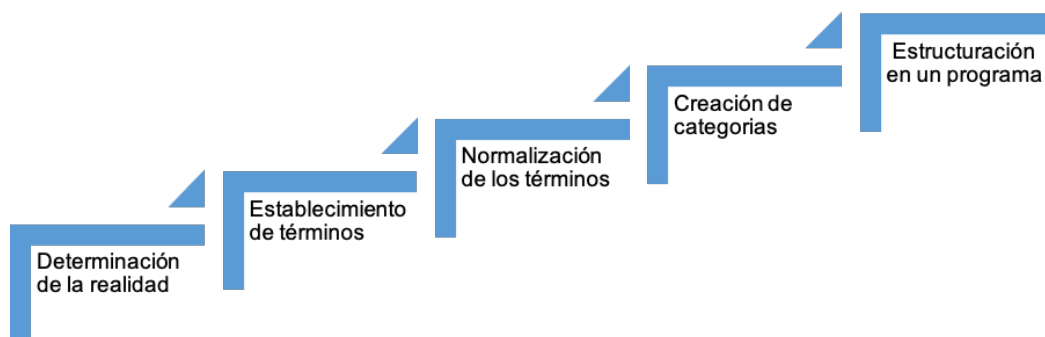
2) Miquel Centelles en *Taxonomías para la categorización y la organización de la información en sitios web*: asume que "las aplicaciones de las taxonomías en el contexto de la web pueden ser diversas" (Centelles, 2005, p. 11). Para la construcción de taxonomías considera cuatro etapas: (1) determinación de la realidad, (2) extracción del conjunto de términos o categorías, (3) normalización o control terminológico de los términos o categorías, (4) establecimiento del esquema y la estructura de organización de los términos o categorías.

Durante la investigación las dos metodologías se observaron útiles para construir la taxonomía digital; no obstante, se observó que su integración producía una metodología más completa en función de las etapas que cada una de ellas incluía, por ejemplo, Centelles (2005) considera la delimitación de la realidad o dominio, mientras que Hedden (2010) incorpora el elemento codificación de la taxonomía en un formato impreso o digital.

De manera que, finalmente, para la construcción de la taxonomía del dominio "Sistema solar" se estableció una metodología particular que integra ambas propuestas y se forma de las siguientes etapas (Figura 1):

FIGURA1

Metodología establecida para la construcción de la taxonomía



Fuente: elaboración propia, 2019

1. *Determinación de la realidad*: es la determinación de entidades, áreas de conocimiento, sectores industriales, etc. que serán representados por la taxonomía (Centelles, 2005).

2. *Establecimiento de términos*: involucra obtener información de cuáles pueden ser incluidos y, más aún, evaluar para decidir cuáles deben ser incluidos (Hedden, 2010).

3. *Normalización de los términos*: consiste en dos tareas: determinar términos preferentes frente a los que no lo son y dar una forma correcta y consistente a todos los términos de la taxonomía (Centelles, 2005).

4. *Creación de categorías*: es la estructuración de la taxonomía, considerando relaciones de equivalencia, jerarquía, asociación y semántica (Hedden, 2010).

5. *Estructuración en un programa*: un vasto grupo de taxonomías son aún creadas y mantenidas en Microsoft Excel pero cuando la taxonomía se hace demasiado compleja o grande, como cuando ésta contiene varios cientos de términos, es momento de moverla a un software para taxonomías (Hedden, 2010).

En los apartados siguientes se retoma la metodología establecida, se desarrollan cada una de las etapas y, a partir de ellas, se presenta como resultado el producto generado en cada etapa que contribuirá a la formación final de la taxonomía.

4. Resultados

4.1. Determinación de la realidad

La realidad sobre la que trata una taxonomía puede ser muy diversa. Como señala Centelles (2005), puede estar asociada a:

- Entidades: dichas entidades pueden ser materiales, "asociadas a elementos físicos" o bien a entidades teóricas "que están presentes en una teoría o hipótesis científica y que denotan una cosa, propiedad, o proceso inaccesible a la observación común" (Bunge, 2001, p. 61).
- Áreas de conocimiento: son ramas o campos de estudio, por ejemplo, física, química, biología, etc. (Bliss, 1939).
- Sectores industriales: son rubros específicos de la industria, por ejemplo, industria química, industria de los alimentos, industria textil, etc.

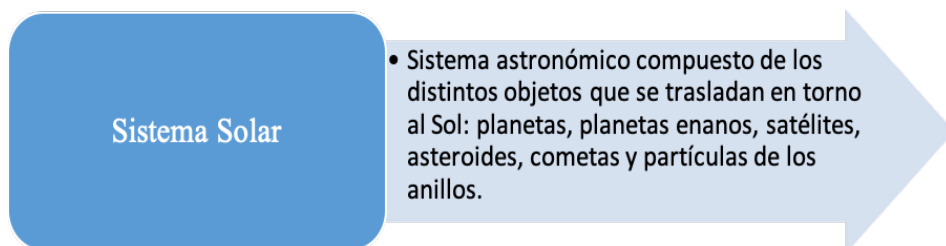
La determinación de la realidad involucra dos momentos (Suárez, 2018, p.65):

- Definición: se precisa con claridad el área que se trabajará.
- Delimitación: radica en el establecimiento de límites y alcances en torno al área, mediante un proceso iterativo.

Para la creación de la taxonomía del "Sistema solar" la determinación de la realidad quedó del modo siguiente (Figura 2):

FIGURA 2.

Determinación de la realidad I



Fuente: elaboración propia, 2019

En lo referente a la delimitación, ésta implica la contextualización de la realidad, y se obtuvo el siguiente producto (Figura 3):

FIGURA 3

Determinación de la realidad II

SISTEMA SOLAR	
Término amplio	Vía Láctea
Término	Sistema Solar
Término específico	Planetas
Término relacionado	Objeto Astronómico

Fuente: elaboración propia, 2019

4.2. Extracción de términos

Los términos constituyen elementos sustanciales en los sistemas para la organización del conocimiento, son su esencia (Dahlberg, 1992). Hjørland (2007) coincide al respecto cuando menciona que básicamente la organización del conocimiento es organización de conceptos y términos que nos permiten asir entidades para hacerlas manipulables.

Para la extracción de los términos que forman cualquier sistema para la organización del conocimiento, se dispone de diversas técnicas, entre las cuales se encuentran las siguientes:

- *Reúso de terminología existente*: se obtiene mediante el reúso de “taxonomías previas o instrumentos de representación del conocimiento ya existentes (desde nomenclaturas de las unidades y recursos existentes en una entidad a los cuadros de clasificación de la documentación administrativa)” (Centelles, 2005, p. 6).
- *Revisión terminológica*: consiste en extraer las estructuras

nominativas que constituyen clases del dominio a partir de un conjunto de documentos (King & Reinold, 2008). Arntz y Picht recomiendan dividir el dominio en líneas fundamentales, adquirir y analizar el material documentario, recopilar los elementos nominales y construir un sistema de conceptos (Arntz & Picht, 1995).

- **Procesamiento de Lenguaje Natural (PNL):** es una rama de la lingüística computacional que se ocupa de la formulación e investigación de mecanismos computacionales para el procesamiento de corpus terminológicos a partir de estructuras textuales (Hedden, 2010).

En la extracción de los términos de la taxonomía del “Sistema solar”, se empleó la técnica de reúso de terminología existente. Se recuperaron términos de la Lista de encabezamientos de materias para bibliotecas (1999), Diccionario de astronomía (1999), Diccionario de astronomía y cosmología (2001) y Nuevo diccionario de astronomía (2002). Derivado de la técnica de reúso, los términos establecidos fueron los siguientes: (Figura 4):

FIGURA 4

Términos del dominio



Fuente: elaboración propia, 2019

4.3. Normalización de términos

Una vez que los términos han sido establecidos es importante llevar a cabo un proceso de normalización (Centelles, 2005). La normatividad disponible para la estandarización terminológica puede ser retomada de ISO 25964-1:2011. *Thesauri for Information Retrieval* (ISO, 2011) e ISO-25964-2:2013. *Interoperability with Other Vocabularies* (ISO, 2013).

- Establecimiento de términos preferidos ante términos no aceptados
- Vinculación unívoca término-concepto

- Asentamiento terminológico:
 - o Unitérminos
 - o Términos compuestos
 - o Términos fraseológicos
- Definición singular/plural
- Presentación terminológica mayúsculas/minúsculas
- Marcas de puntuación (apóstrofes, paréntesis, números, etc.)
- Asentamiento de nombres propios (personas, lugares, ciudades, etc.)
- Fraccionamiento sintáctico y semántico⁴
- Manejo de extranjerismos
- Adecuación de transliteraciones
- Inclusión de neologismos
- Empleo de marcas y copyright
- Uso de abreviaciones
- Homógrafos

4.4. Creación de categorías

Las categorías son agrupamientos de entidades con características comunes o bien:

[...] metaconceptos que nombran agrupamientos de términos de una determinada área. En este sentido, se llama categoría a los agrupamientos que resultan de la división del universo de conocimiento, de acuerdo con las características intrínsecas o fundamentales de cada concepto (Cavalvanti, 1978, p. 15).

Al respecto Glushko (2013) señala que:

Categories are cognitive and linguistic models for applying prior knowledge; creating and using categories are essential human activities. Categories enable us to relate things to each other in terms of similarity and dissimilarity and are involved whenever we perceive, communicate, analyze, predict, or classify. Without categories, we would perceive the world as an unorganized blur of things with no understandable or memorable relation to each other [Las categorías son modelos cognitivos y lingüísticos para la aplicación de conocimiento previo; crear y usar categorías son actividades humanas esenciales. Las categorías nos permiten relacionar las cosas entre sí en términos de similitud y disimilitud y están involucradas siempre que percibimos, comunicamos, analizamos, predecimos o clasificamos. Sin categorías, percibiríamos el mundo como un espacio empañado y desorganizado de cosas, sin relaciones comprensibles o trascendentes entre sí] (p. 237).

En su construcción, la taxonomía puede considerar algunas de las siguientes técnicas categorizadoras (Glushko, 2013, p. 244):

1. Enumeración: consiste en agrupar descriptores bajo una categoría enunciada, por ejemplo: "Animales de granja" > "conejos", "cabras", "vacas", "borregos", etc. Este principio también es conocido como *definición extensional* y los miembros de las categorías son denominados *extensiones*.

⁴ Para mayor información sobre fraccionamiento semántico y sintáctico consúltese Lancaster (1986) e ISO 25964-1 (2011) que establecen reglas para la composición y descomposición terminológica bajo criterios sintácticos y semánticos asociados a organización y recuperación de información.

2. Propiedades únicas: es un proceso intuitivo que considera las propiedades de las instancias, por ejemplo, tamaño, color, forma, peso, material composición, textura, y los agrupa considerando el rasgo. Este principio es conocido como *definición intensional* y las propiedades definidas son *intensiones*.

3. Similitud: se considera una caracterización para los conceptos que se incluirán en la clase, por ejemplo, "Ave" <rasgos: alas, plumaje, dos patas, pico> "canario", "gaviota", "avestruz", etc.

4. Función: se establecen funciones o finalidades que son el eje categorizador, por ejemplo, "Ingredientes para preparar sopa" < "pasta", "jitomate", "condimentos", "aceite", "sal" >.

5. Propiedades múltiples: se establecen diversos rasgos que se van siguiendo en la categorización, por ejemplo, una categorización de vestidos, primero, según su función "vestidos de trabajo" "vestidos de vestir casual" "vestidos para fiestas", luego por propiedad única "vestidos cortos" "vestidos largo medio" "vestidos largos y, luego, por otra propiedad única "vestidos negros" "vestidos rojos" y "vestidos blancos.

Otros aspectos de consideración en la categorización taxonómica son aquellas garantías que validan el instrumento:

1. La garantía del usuario: establece que es importante considerar los requerimientos y expectativas de los usuarios (Lancaster, 1986).

2. La garantía académica: considera el punto de vista de estructuras previas del área de conocimiento, recuperables mediante la consulta de tesauros y clasificaciones, así como la división del dominio en líneas de investigación, especialización académica y producción literaria (Bliss, 1939).

3. La garantía estructural: trata sobre estructuras terminológicas y vocabularios documentales, es un estándar que, debido a la falta de una norma precisa para taxonomías, resulta de utilidad en la construcción conceptual-terminológica de esquemas de dominios (ISO-1087, 2000).

4. La garantía lógica: supervisa la coherencia y lógica de la estructura organizadora (Fraser, 1978).

Retomando el caso ejemplo, luego de aplicar las técnicas categorizadoras y principios antes señalados se logró la siguiente categorización (Figura 5):



FIGURA 5
Categorización

1. Sistema solar
 - 1.1. Anillos
 - 1.2. Asteroides
 - 1.2.1. Asteroides tipo M Ferrosos
 - 1.2.2. Asteroides tipo S Pétreos-Ferrosos
 - 1.3. Cinturón de Kuiper
 - 1.4. Cometas
 - 1.4.1. Cometas de periodo largo
 - 1.4.2. Cometas F-Familia de Júpiter
 - 1.4.3. Cometas HT-Tipo Halley
 - 1.5. Nube de Oort
 - 1.6. Planetas
 - 1.6.1. Planetas gaseosos
 - 1.6.1.1. Júpiter
 - 1.6.1.2. Neptuno
 - 1.6.1.3. Saturno
 - 1.6.1.4. Urano
 - 1.6.2. Planetas rocosos
 - 1.6.2.1. Mercurio
 - 1.6.2.2. Tierra
 - 1.6.2.3. Venus
 - 1.7. Planetas enanos
 - 1.7.1. Planetas enanos determinados
 - 1.7.1.1. Ceres
 - 1.7.1.2. Eris
 - 1.7.1.3. Haumea
 - 1.7.1.4. Makemake
 - 1.7.1.5. Plutón
 - 1.7.2. Planetas enanos indeterminados
 - 1.7.2.1. 1996 TO66
 - 1.7.2.2. 2002 AW
 - 1.7.2.3. 2002 TC
 - 1.7.2.4. 2002 TX300
 - 1.7.2.5. 2002 UX
 - 1.7.2.6. 2007 OR10
 - 1.7.2.7. Caronte
 - 1.7.2.8. Ixion
 - 1.7.2.9. Orcus
 - 1.7.2.10. Quaoar
 - 1.7.2.11. Sedna
 - 1.7.2.12. Varuna
 - 1.8. Satélites
 - 1.8.1. Satélites artificiales
 - 1.8.1.1. Satélites científicos
 - 1.8.1.2. Satélites de comunicación
 - 1.8.1.3. Satélites meteorológicos
 - 1.8.1.4. Satélites militares
 - 1.8.2. Satélites naturales
 - 1.8.2.1. Calisto
 - 1.8.2.2. Europa
 - 1.8.2.3. Ganimedes
 - 1.8.2.4. Luna
 - 1.8.2.5. Mimas
 - 1.8.2.6. Io

Fuente: elaboración propia, 2019

4.5. Estructuración en un programa

Las tecnologías de la información actuales ofrecen diversas herramientas para la construcción de taxonomías digitales. Hedden (2010, p. 135) menciona algunos de los programas más relevantes:

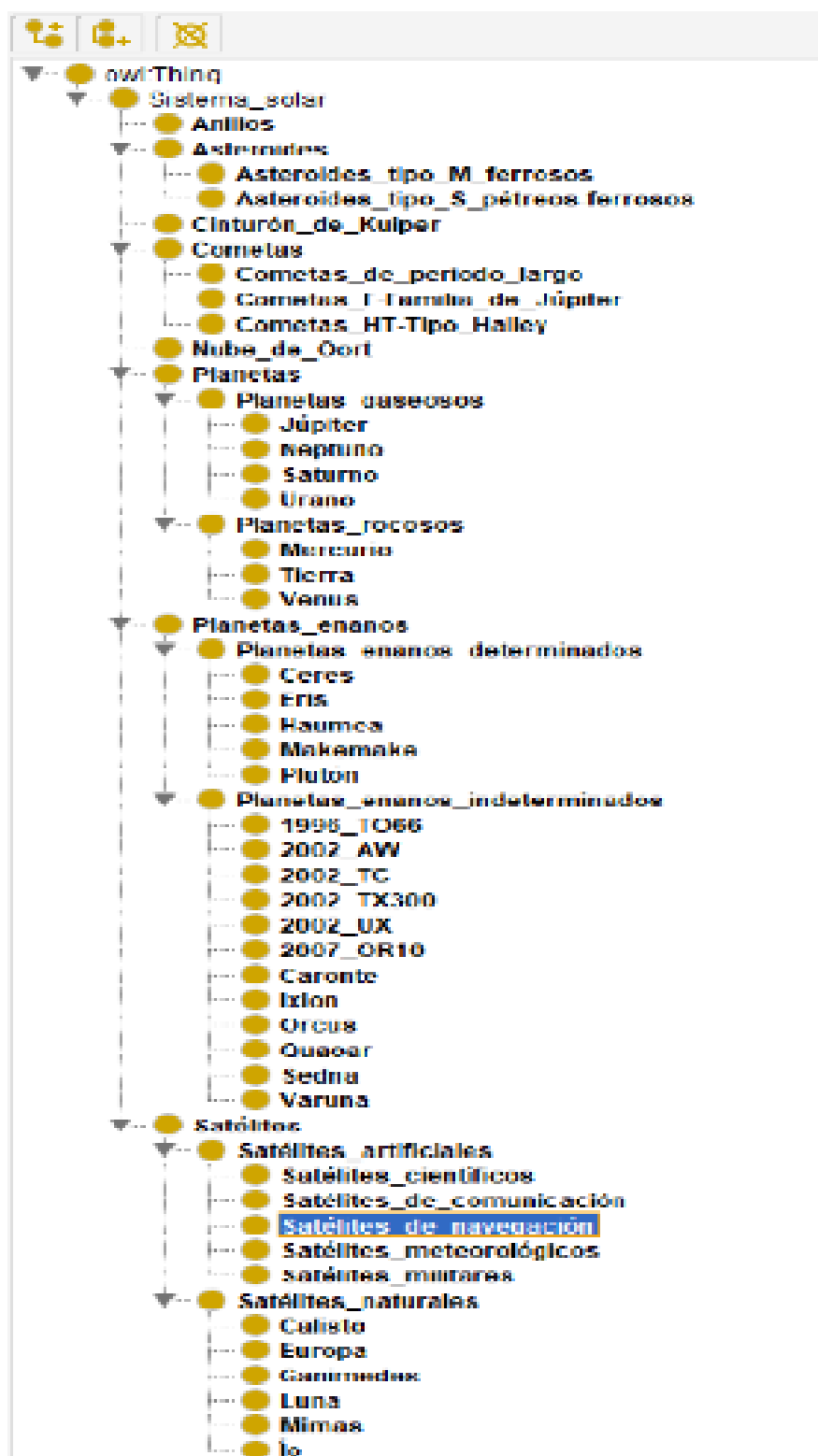
1. *Cognatrix*: es un software enfocado a la construcción de tesauros, pero también permite la construcción de taxonomías. Tiene una versión amplia por un costo alto y una versión más económica que sólo admite hasta 500 términos. Posibilita la adhesión de notas relativas a los términos, crear grupos, permite la importación de bases de datos previas y la importación de datos hacia formatos XML o HTML.
2. *Protégé*: está enfocado a la construcción de esquemas para la web semántica. Permite la creación de categorías bajo etiquetas establecidas. Posibilita la creación de una estructura jerárquica visible en la que se pueden incorporar elementos inferiores (subdivisiones) o al mismo nivel (elementos hermanos). Ofrece la posibilidad de establecer relaciones de equivalencia entre etiquetas para términos no aceptados. Desde la parte visual, la taxonomía puede ser observada como un listado de elementos o como un árbol o nodo, empleando un visualizador (Graphviz) asociado al programa.
3. *Synaptica*: es un programa que cubre diversas necesidades para organizar información en la web, uno de sus componentes está enfocado a taxonomías digitales. Permite la creación de taxonomías sencillas hasta estructuras complejas en diversas lenguas. Posibilita el control terminológico y la vinculación de elementos en la estructura. Ofrece diversos formatos de visualización y altos estándares de interoperabilidad.
4. *MultiTes Pro*: es un programa para la creación y el manejo de vocabularios controlados entre los que se encuentran tesauros y taxonomías. Brinda herramientas para el control del vocabulario, genera autoridades y proporciona métodos para hacer visible el instrumento. Es útil para construir herramientas de diverso alcance y complejidad, desde taxonomías individuales hasta herramientas corporativas. Está activo en el mercado desde 2005, continúa desarrollándose y es usado por muchas instituciones en el ámbito internacional.

Para la estructuración de la taxonomía digital del “Sistema solar” se usó el programa Protégé, 5.5. 0. Beta. La taxonomía final quedó como sigue (Figura 6):



FIGURA 6

Taxonomía en Protégé 5.5.0. Beta.



5. Conclusiones

El estudio cumplió con el objetivo establecido que consistió en la construcción de una taxonomía digital que sirviera como entrada temática para un conjunto reducido de recursos de información sobre el “Sistema solar”, los cuales serán albergados en un sitio web, considerando solamente su perspectiva temática.

La taxonomía resultante se formó de 55 elementos terminológicos desplegados en 8 nodos categoriales distribuidos en 4 niveles de clases. La metodología aplicada para su construcción, que combinó dos metodologías previas: Centelles (2005) y Hedden (2010) fue altamente eficaz para llevar a cabo todo el proceso de construcción del instrumento, desde la determinación de la realidad hasta la estructuración en un programa. Se resume que las metodologías seleccionadas en combinación resultan excelentes para la construcción de taxonomías digitales.

Sobre la diferenciación entre la construcción de una taxonomía para el contexto impreso y una taxonomía para el espacio digital, es decir, la interrogante respecto a si lo digital, requieren un enfoque distinto, la presente investigación ha dejado tal aspecto pendiente -véase como una línea de investigación abierta- ya que las metodologías que siguió consideran el espacio web como objetivo: la metodología de Centelles (2005) está enfocada completamente a taxonomías digitales, mientras que la metodología de Hedden (2010) considera la creación de taxonomías en general, pero incorpora aspectos tecnológicos de las taxonomías en la web.

En comparación con otros estudios, de los cuales hay muy pocos, el presente trabajo sintetiza que las taxonomías son un sistema para la organización del conocimiento viable para la organización temática de recursos de información en la web. Los resultados coinciden con el trabajo de Menárd -que indiza imágenes digitales en un sitio web mediante una taxonomía- especialmente cuando señala que hay un “typical indexing process [proceso de indización típico]” (2013, p. 2) y, a la par, en la web, han emergido nuevas formas de indización como etiquetado social y taxonomías.

Los resultados del trabajo permiten concluir que las taxonomías como sistemas para la organización del conocimiento son herramientas útiles para la indización de recursos en sitios web. No obstante, se cuenta con poca investigación teórica al respecto y casos prácticos implementados. En tal sentido, la investigación sobre el tema es un asunto pendiente, especialmente en los tiempos actuales, caracterizados por una emergente organización de recursos de información en el contexto web en general y también en espacios web asociados directamente a la bibliotecología, como repositorios de información y bibliotecas digitales.



6. Referencias

Abbas, J. (2010). *Structures for Organizing Knowledge*. Nueva York, Estados Unidos de América: Neal Schuman.

Arntz, R., y Picht, H. (1995). *Introducción a la terminología*. Madrid, España: Fundación Germán Sánchez Ruipérez.

Barite, M. *Diccionario de organización y representación del conocimiento*. Recuperado de <http://www.eubca.edu.uy/diccionario/index/htm>.

Bliss, H. (1939). *The Organization of Knowledge in Libraries and the Subject Approach to Books*. Nueva York, Estados Unidos de América: Wilson.

Boeri, R.J. (2004). Playing with Taxonomies. *E-Content*. Recuperado de <http://www.econtentmag.com/Articles/Column/Info-Insider/Playing-with-Taxonomies-7357.htm>.

Bunge, M.A. (2001). *Diccionario de filosofía*. Distrito Federal, México: Siglo XXI.

Cavalvanti, C. (1978). *Indexação y Tesouro: metodología y técnicas*. Brasilia, Brasil: Associação de Bibliotecários do Distrito Federal.

Centelles, Miquel. (2005). Taxonomías para la categorización y la organización de la información en sitios web. *Hipertext* 3, pp.1-16.

Chowdhury, G. y Chowdhury, S. (2007). *Organizing Information: from the Shelf to the Web*. Londres, Inglaterra: Facet Publishing.

Corcoran, M. (2002). Taxonomies: Hope or Hype? *Online*, 26(5), 76-78.

Dahlberg, I. (1992). Knowledge Organization and Terminology: Philosophical and Linguistic bases. *International Classification*, 19(2), 65-71.

Díaz-Piraquive, F. N., Joyanes-Aguilar, L. y Medina-García, V.H. (2009). Taxonomía, Ontología y folksonomía. ¿qué son y qué beneficios u oportunidades presentan para los usuarios de la web? *Universidad & Empresa*, 8(16), 242-261.

Feather, J., & Sturges, P. (2003). *International Encyclopedia of Information and Library Science*. Nueva York, Estados Unidos de América: Routledge.

Ferro Ramos, I. (1999). *Diccionario de astronomía*. Distrito Federal, México: Fondo de Cultura Económica.

Fernández-Hernández, A. (2007). *Organización de los contenidos en los sitios web: las taxonomías*. *ACIMED*, 15(5), 1-11.

Fraser, W. (1978). Literary, user and logical warrants as indexing constraints. En American Society for Information Science. (1978). *The information Age in Perspective: Proceedings of the Annual ASIS Meeting*. Nueva York, Estados Unidos de América: Knowledge Industry Publications.

Gilchrist, A. (2002). Thesauri, Taxonomies and Ontologies: an Ethimological Note. *Journal of Documentation*, 59(1), 7-18.

Glushko, R J (Ed) (2013). *The Discipline of Organizing*. Cambridge, Inglaterra: MIT Press.

Graef, J. (2001). *Introduction to Business Taxonomies*. Massachusetts, Estados Unidos de América: Montague Institute.

Hedden, H. (2010). *The Accidental Taxonomist*. Nueva Jersey, Estados Unidos de América: Information Today.

Hodge, G. (2000). *Systems of Knowledge for Digital Libraries: Beyond Traditional Authority Files*. Washington, Estados Unidos de América: Council on Library and Information Resources.



Hjørland, B. (2007). Semantics and knowledge organization. *Annual Review of Information Science and Technology*, 41(1), 367-405.

Hornak, K.A. (2001). *Diccionario de astronomía y cosmología*. Pennsylvania, Estados Unidos: Castilla La Vieja.

International Organization for Standardization. (1986). *ISO 2788:1986. Guidelines for the establishment and Development of Monolingual thesauri*. Ginebra, Suiza: Autor.

International Organization for Standardization. (2000). *ISO-1087-1:2000. Terminology Work -Vocabulary-Part 1: Theory and Application*. Ginebra, Suiza: Autor.

International Organization for Standardization. (2011). *ISO 25964-1:2011. Thesauri for Information Retrieval*. Ginebra, Suiza: Autor.

International Organization for Standardization. (2013). *ISO-25964-2:2013. Interoperability with Other Vocabularies*. Ginebra, Suiza: Autor.

King, B. y Reinold, K. (2008). *Finding the Concept, Not Just the Word*. Oxford, Inglaterra: Chandos Publishing.

Lambe, P. (2007). *Organising knowledge : Taxonomies, Knowledge and Organisational Effectiveness*. Oxford, Inglaterra: Chandos.

Lancaster, F. W. (1986). *Vocabulary Control in Information Retrieval*. Texas, Estados Unidos de América: Information Resources.

Martínez de Souza, José. (2004). *Diccionario de bibliología y ciencias afines*. Madrid, España: TREA.

Ménard, Elaine. (2013). TIIARA for and IDOL: An Adventure in Indexing. *The indexer*, 31(1), 2-11.

Nuevo diccionario de astronomía. (2002). Barcelona, España: Océano.

Reitz, J. (2004). *Dictionary for Library and Information Science.* Connecticut, Estados Unidos de América: Libraries Unlimited.

Smiraglia, R. (2015). *Domain Analysis for Knowledge Organization.* Nueva York: Chandos.

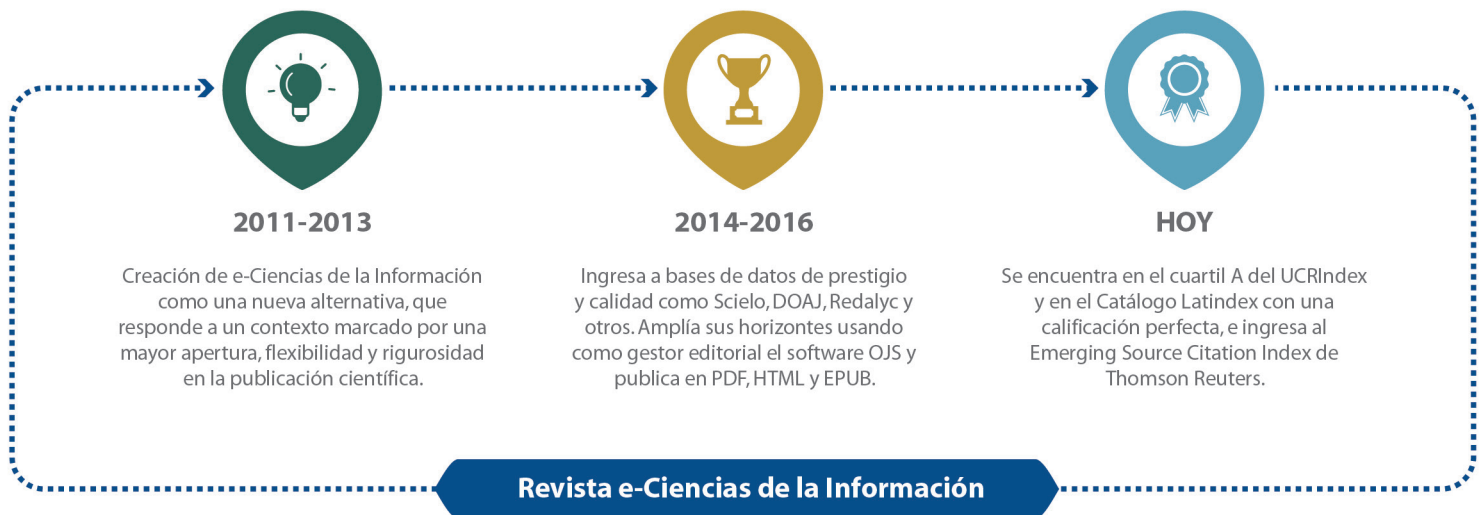
Suárez Sánchez, A. (2018). *Ontologías: fundamentos y aplicaciones en la organización temática de la información. Una perspectiva bibliotecológica.* Recuperado de <http://132.248.9.195/ptd2018/agosto/0778186/Index.html>.

Taylor, A., & Joudrey, D. (2009). *The Organization of Information.* Connecticut, Estados Unidos: Libraries Unlimited.

Tudhope, D. (2006). *A Tentative Typology of KOS: Toward a KOS of KOS?* Recuperado de <http://slideplayer.com/slide/7277501/>.



e-Ciencias de la Información



¿Dónde se encuentra indexada e-Ciencias de la Información?



Para más información ingrese a nuestra [lista completa de indexadores](#)

¿Desea publicar su trabajo?
Ingresa [aquí](#)

O escribanos a la siguiente dirección
revista.ebci@ucr.ac.cr