

**LA PROPIEDAD INTELECTUAL DE LOS CONTENIDOS TERMINOLÓGICOS EN
LA WEB SEMÁNTICA.**

Miguel-Ángel López Alonso,

TU Departamento de Informática de la UEX en la Facultad de Com. y
Documentación,

Licenciado en Derecho UCM y Dr. en Documentación UCIIM.

Boix y Morer, 5 – 8º D (28003 Madrid)

malopalo@alcazaba.unex.es

La Propiedad Intelectual de los contenidos terminológicos en la Web Semántica.

COPYRIGHT / TERMINOLOGY / LAW CONVERGENCE / DIGITAL ENVIRONMENT / SEMANTIC WEB

Survey on the copyright state of the question in the Distribution of the Terminological Contents in the environment of the Internet Semantic Web. A future bet is formulated on the walked road and the one that is to travel for the “Universal Convergence of Copyright Laws”

DERECHOS DE AUTOR / TERMINOLOGÍA / CONVERGENCIA JURISDICCIONAL / ENTORNO DIGITAL / WEB SEMÁNTICA

Revisión del estado de la cuestión de los derechos de autor en la Distribución de los Contenidos Terminológicos en el ámbito de la llamada Web Semántica de Internet. Se evalúa el camino andado y el que queda por recorrer para “una Convergencia Universal de Jurisdicciones”.

Introducción:

Si definimos la terminología como “un conjunto de términos y sus definiciones en determinado campo científico”, y a éstos como “las más pequeñas unidades del conocimiento especializado” (Cabré, 1994), nos estamos refiriendo a la infraestructura del conocimiento especializado y a su aplicación a la gestión terminológica para la creación de terminologías específicas, e incluirá la grabación terminográfica de dicha información como bases de datos terminológicas, diccionarios, léxicos, enciclopedias especializadas, etc.

La diferente dinámica evolutiva de cada campo del conocimiento lleva al discurso científico especializado a una continua diversificación en más y más lenguajes y terminologías especializadas. Estas herramientas de comunicación se vuelven cada día más ambiguas, debido al estricto número de conceptos que deben definirse con los limitados recursos lingüísticos del lenguaje natural, lo que obliga a los especialistas a normalizar su terminología a un nivel multilingüe.

Esta normalización abarca a:

- la verificación sociológica de los neologismos propuestos por los terminólogos (Gaudin, 1994),
- el desarrollo terminológico mediante su armonización y estandarización,
- la investigación y el desarrollo del establecimiento y uso de las bases de datos terminológicas por distintos grupos de usuarios (traductores, tecnólogos, gestores de la información, etc.), y
- la compilación de los vocabularios controlados para fines documentales y de recuperación de la información (tesauros, sistemas de clasificación, encabezamientos de materia, etc.).

Este trabajo revisa los problemas de la *Propiedad Intelectual de las bases de datos terminológicas y otras formas lingüísticas similares* de representación de la información y del conocimiento: lenguajes documentales, bases de datos bibliográficas, etc., aproximándose a los problemas de la distribución electrónica de

las bases de datos digitales (la información multimedia, las bases distribuidas de la Web, etc.). Con su reconducción se derivarán beneficiosas implicaciones en la mejora del comercio electrónico en general y de las industrias de contenidos lingüísticos en particular, en esta Sociedad de la Información todavía pendiente de una normalización jurídica aceptada globalmente al nivel de las redes distribuidas de Internet.

Presupuestos teóricos y desarrollo:

Para el estudio de los aspectos jurídicos de la Propiedad Intelectual en el ámbito terminológico, partiremos de su definición como “protección de los derechos patrimoniales y personales que corresponden al creador de una obra de ingenio” utilizada por la OMPI (Organización Mundial de la Propiedad Intelectual) que engloba tanto los Derechos de autor y afines como la Propiedad Industrial. Pero, siguiendo igualmente a la OMPI en sus dos convenios sobre el Derecho de autor y los Derechos afines, nos referiremos básicamente al “Derecho de autor en el entorno digital”.

Al aplicar dicha definición al campo terminológico, percibimos que es mucho más fácil la protección intelectual de la literatura o del pensamiento científico que la de las “unidades terminológicas” (los pluritérminos). Esto es así porque, fuera de un contexto concreto, la propiedad intelectual de cualquier término es muy difícil de probar. Incluso, se llega al límite de considerar a los términos como un “subproducto” de una actividad cognitiva (diccionario, tesoro, compilación documental, etc.) que se considera “gratuito” o de libre utilización en la creación de cualquier otro tipo de “producto original”, al que sí se considera como titular del Derecho de autor, con protección legal (traducción, documento técnico, libro, etc.).

Dado que la justificación de la protección de la Propiedad Intelectual se apoya en el valor añadido de “innovación”, sus objetos de protección deben superar diversas pruebas de originalidad. Éstas provocan problemas con los recursos terminológicos, ya que los “mono términos” están considerados tradicionalmente como del Dominio Público (aunque no sus definiciones). Además, la identidad y el tamaño de la “unidad a proteger más pequeña” no están determinados satisfactoriamente, pues, ya hemos dicho que cualquier colección terminológica puede contener al mismo tiempo materiales sujetos a la protección del Derecho de autor y otros considerados de libre disposición (mono términos vacíos o de utilización genérica).

Como la necesidad de nueva terminología se va a incrementar con el uso globalizado de la tecnologías de la información en todo tipo de actividades económicas o cognitivas, está claro que los creadores de “unidades terminológicas” con sus correspondientes definiciones innovadoras no van a poder ser identificados individualmente con facilidad y, por tanto, la retribución económica derivada de su meditado y laborioso trabajo intelectual de creación sólo podrá derivarse del reconocimiento de sus “colecciones terminológicas” como *obras de compilación* adaptadas a determinados estándares universalmente aceptados (Graham, 2000).

Este problema es debido a la imposibilidad de escanear datos terminológicos y lexicográficos a partir de sus originales, e introducirlos directamente en una base de datos terminológica (definida como “la colección de trabajos, datos u otros materiales independientes, almacenados de una manera sistemática o metódica, y accesible individualmente por medios electrónicos o no”). Además de la información apreciable a simple vista, se encuentra mucha otra información innovadora oculta

con vínculos entre sus distintos campos, que solamente es posible recuperar mediante la aplicación de meditados procesos cognitivos del cerebro humano (Galinski, 1998).

Constatamos la existencia de tres diferentes corrientes normativas y jurisprudenciales fuertemente arraigadas, tanto al nivel regional como internacional:

- 1) PRIMERO. El concepto internacionalista del derecho a la terminología del Convenio de Berna sobre copyright y del Tratado de la OMPI sobre Derecho de autor (revisados en 1996 y 1999), que *considera a las compilaciones terminológicas como un todo único*, y que con apoyo jurisprudencial defiende el Secretariado TC 37 / SC 1 de la ISO (International Organization for Standardization) en base a la doctrina del Centro Internacional para la Terminología de Viena (INFOTERM) que promueve la adaptación a la “realidad social y tecnológica”.

Como regla general, las contribuciones individuales de los terminólogos se integran en una Propiedad Intelectual conjunta, cuyo “propietario” es la agencia, secretariado o autoridad institucional de la que se recibió el encargo. Las fuentes originales de algunos o de todos los términos (especialmente los normalizados y sus definiciones innovadoras) deben especificarse mediante un conciso código de fuente que en algunos casos remitirá al usuario a otra fuente de información más detallada (registro bibliográfico, lenguaje documental, etc.).

En estos casos, debe utilizarse una metodología consistente tanto por el grupo coordinador como por el gestor de la base de datos, de manera que la mayoría de los datos terminológicos puedan establecerse de acuerdo con determinadas reglas y se permita además el almacenamiento de datos adicionales (fuente, creador, fecha, etc.) hasta la consecución final del proyecto. Siendo a partir de este momento que, todos los registros y sus respectivas relaciones mutuas, se constituirán en objetos del Derecho de autor.

Como CASO práctico de estudio proponemos la “*Guía de acuerdos en terminología*” (*Guide to Terminology Agreements*) de Christian Galinski y Jürgen W. Goebel. En su articulado se propone al creador o al titular legal del Derecho de Propiedad Intelectual un precontrato para aquellos casos de los que pueda derivarse una reclamación basada en la supuesta violación de los Derechos de autor.

Sin embargo, puesto que el contrato no crea ningún tipo de Propiedad Intelectual y únicamente obliga a las partes firmantes, se crea un problema potencial para la diseminación y reutilización a gran escala dado que, cuando se produzca una “fuga ilegal” en la red de datos autorizados, será muy difícil la paralización de la distribución de cualquier información secundaria no autorizada.

- 2) SEGUNDO. El concepto euro-comunitario del Derecho de autor a la Terminología que, a partir de la “Directiva para la protección de las bases de datos” (1996), defiende una protección desde la “posición territorialista del derecho” que *extiende su aplicación de una manera “sui generis”: como protección del derecho del “fabricante o distribuidor de la base de datos” a impedir la extracción y/o la reutilización no autorizada previamente a todos los registros de una determinada base de datos terminológica durante quince años, siempre que haya invertido recursos sustanciales en su compilación, almacenamiento o presentación (art. 7); con independencia de que además*

dicha base de datos esté protegida por el Derecho de autor o por otros derechos sobre su contenido (art. 3).

Este derecho de protección, similar al de las leyes de Defensa de la Competencia, se enfoca más a la protección de los intereses económicos que a los de la "originalidad o innovación" de su contenido y, además, únicamente durante un período medio de vigencia (15 años).

Como CASO práctico de estudio proponemos el informe final del Proyecto europeo POINTER que, constituido por un consorcio de casi cuarenta asociados, representantes de los grandes usuarios de terminología de la industria, el comercio y la administración, agrupa en diez secciones las conclusiones para "una nueva calidad de la organización del conocimiento en la comunidad terminológica de Europa". Sus resultados constituyen un conjunto de proposiciones y de recomendaciones que se han concretado en la formación de asociaciones (EAFT, ELDA, ELRA, ETIS,...) tendentes a *la constitución de una infraestructura terminológica común* que permita la creación, difusión y utilización de una terminología multilingüe accesible a todos los usuarios europeos.

- 3) TERCERO. El concepto angloamericano de Copyright *defiende el principio de "la autorregulación del intercambio terminológico" mediante contraprestación contractual* (entre los creadores o distribuidores y los usuarios o reutilizadores), apoyada en la Copyright Act que protege durante la vida del creador más otros 70 años tanto los trabajos de terminología de los autores individuales como los de las empresas propietarias de los derechos de propiedad intelectual. Dicha ley ha extendido su protección desde 1998 por otros 20 años más con la "Copyright Term Extension Act (S.505)", con una cláusula de excepción de "fair use" para ámbitos de la documentación y la recuperación de la información, siempre que el propietario del copyright no aporte pruebas de que su trabajo se utiliza para "una explotación comercial normal".

Además, la Copyright Act (1976) se ha actualizado al entorno digital con la Digital Millennium Copyright Act (H. R. 2281) desde diciembre de 1998, implementando los dos Convenios sobre Propiedad Intelectual de la OMPI de 1996.

A partir de la publicación de la Directiva europea sobre bases de datos en 1996, para paliar los problemas de colisión jurídica derivados de la exportación de terminologías y otras bases de datos a la Unión Europea por parte de las empresas americanas, el Departamento de Comercio de Estados Unidos y la Comisión Europea han desarrollado un marco común de "puerto seguro" (Safe Harbor agreement), aprobado por la Unión Europea en julio del 2000, que certifica "bajo petición previa" que las compañías norteamericanas proporcionan adecuada protección privada a los datos de acuerdo con lo estipulado por la citada Directiva.

Conclusiones y apuestas de futuro:

La finalidad de cualquier marco jurídico terminológico debe facilitar la efectiva distribución de los recursos y la seguridad de su utilización entre los proveedores y los reutilizadores. No es necesaria la creación de una "solución aislada" específica para el ámbito de la terminología, sino que por el contrario sus afinidades con otras áreas del procesamiento del lenguaje deben ser explotadas al máximo.

La normalización de la jurisdicción terminológica al nivel planetario es muy compleja debido a que se trata de integrar:

- a) No sólo sistemas jurídicos diferentes (de “civil law” o continentales y de “commun law” o anglosajones), sino de:
- b) Asimilar el derecho mercantil internacional de los convenios internacionales de la OMPI, tanto por ambos sistemas jurídicos como por los correspondientes derechos privados (civiles o mercantiles) regionales de áreas geográficas muy dispares: Europa, USA, Japón, Australia, etc.

Todavía deben realizarse numerosos trabajos específicos, mediante grupos de trabajo formados por juristas y demás especialistas, sobre los temas subyacentes para la transición a una Sociedad de la Información Virtual que tenga en cuenta los diferentes modelos jurídicos analizados para una convergencia universal de jurisdicciones. Las soluciones deberán coordinarse entre todas las áreas de las comunidades lingüísticas y de la información y documentación que se enfrenten a problemas similares.

El uso combinado de contratos perfectamente redactados, técnicas automatizadas con sofisticados perfiles de acceso, educación de los grupos de usuarios en un “código de buena conducta” (LÓPEZ ALONSO, 1996), y una eficaz utilización de las acciones legales contra los infractores deberá reducir el problema a una escala factible.

El futuro tardará en completarse, pero, empieza a vislumbrarse prometedor debido a que la globalización de la economía obliga a todos los participantes a estrechar diferencias, y ya hemos visto que extensas áreas geográficas como la Unión Europea o los Estados Unidos acaban de aprobar legislaciones para integrar los recientemente revisados Convenios internacionales de la OMPI, e incluso aprueban procedimientos bilaterales que tiendan puentes para evitar los problemas legislativos derivados de la aplicación a los intercambios de la Directiva europea sobre bases de datos (ej.: Acuerdo “Safe Harbor” entre Estados Unidos y la Unión Europea).

Bibliografía citada:

- Acuerdo UE-EEUU sobre el puerto seguro referente a la protección de datos (2001). In DOCE 18.09.01. Retrieved October 8, 2002 from the World Wide Web: http://www.export.gov/safeharbor/sh_documents.html
- Berne Convention for the Protection of Literacy and Artistic Works (Paris Text 1971). Retrieved October 8, 2002 from the World Wide Web: <http://www.law.cornell.edu/treaties/berne/>
- Cabré, T. (1994). La terminología. Barcelona. Calzolari, N. (1994). European efforts towards standardizing language resources. In P. Steffens, editor, Machine Translation and the Lexicon. Springer-Verlag
- Copyright Act (1976). Retrieved October 8, 2002 from the World Wide Web: <http://www.copyright.gov/title17/>
- Copyright Term Extension Act (1995). Retrieved October 8, 2002 from the World Wide Web: <http://www.copyright.gov/legislation/s505.pdf>
- Digital Millennium Copyright Act (1998). Retrieved October 8, 2002 from the World Wide Web: <http://www.loc.gov/copyright/legislation/hr2281.pdf>

- Directiva 96/9/CE del Parlamento Europeo y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos. Diario Oficial nº L 077 de 27/03/1996 20-28. Retrieved October 1, 2002 from the World Wide Web:
<http://www.legalia.com/dti/derechos%20de%20autor/normativa/eur001.htm>
- Galinski. C. (1998). Copyright on terminological data. Retrieved October 8, 2002 from the World Wide Web: http://www.innovation-showcase.net/01_microsoft/showcase/topics/ipr/pdf/c_g_kr2000_en.pdf
- Gaudin (1994). Socioterminologie. Gauvain, J.-L. and Lamel, L.F. (1993). Sous-corpus BREF 80, disques bref 80-1 et bref 80-2 (CDROM).
- Graham, J.D. (2000). Copyright of terms in terminology: Reflections on practical aspects. In Proceedings of the Conference for a terminology infrastructure in Europe. Paris, France: UNESCO. Retrieved October 8, 2002 from the World Wide Web:
<http://www.unilat.org/dtil/etis/actasTDCnet/graham.htm>
- López Alonso, M.-A. (1996). La protección de los usuarios de Internet: Autorregulación colegiada o leyes supranacionales. Informática y derecho: revista iberoamericana de derecho informático, 1998; (27, 28, 29), 413-422.
- Tratado de la OMPI sobre Derecho de Autor (WCT) (1996). Ginebra (Switzerland): OMPI. Retrieved October 8, 2002 from the World Wide Web:
<http://www.wipo.int/clea/docs/es/wo/wo033es.htm>

Anexo:

CASOS PRÁCTICOS:

- “Guía de acuerdos en terminología” (Guide to Terminology Agreements) de Christian Galinski y Jürgen W. Goebel, editada en Viena (1996) por el INFOTERM (International Information Centre for Terminology) y disponible en el servidor terminológico de TermNet (The International Network for Terminology).
- Informe final del Proyecto europeo POINTER (1996), disponible en el servidor de la Universidad de Surrey (Reino Unido) (<http://www.computing.surrey.ac.uk/ai/pointer/>).