

Lenguajes jurídico-documentales

*Jorge Páez Mañá**

Índice: 1.- Lenguaje jurídico versus lenguajes jurídico-documentales. 1.1.- Características del lenguaje jurídico. 1.2.- Ambigüedad del lenguaje jurídico. 1.3.- Tipología de las fuentes del lenguaje jurídico. 1.4.- Tratamiento documental del lenguaje jurídico. 2.- Terminología jurídico-documental: conceptos y términos. 3.- Tesoros jurídicos. 3.1.- Función documental de los tesauros. 3.2.- Forma de elaboración de los tesauros. 3.3.- Estructura de los tesauros. 4.- Automatización de tratamientos de indización y tesauros. Bibliografía.

1.- Lenguaje jurídico versus lenguajes jurídico-documentales

Como cuestión previa al análisis de los lenguajes jurídico-documentales, conviene prestar atención a la relación entre el lenguaje y el Derecho, a efectos de determinar las peculiaridades propias del lenguaje jurídico, con objeto de facilitar su tratamiento, almacenamiento y recuperación automatizada, en entornos informáticos.

El lenguaje es un elemento vivo y característico de las diferentes comunidades, pudiéndose incluso constatar determinadas variaciones en su utilización según los grupos sociales que las conforman, por lo que, habida cuenta que el lenguaje jurídico tiende a hacer públicos unos preceptos obligacionales de forma tal que sean universalmente entendidos y asumidos por las diferentes personas y grupos sociales que coexisten en la comunidad, evitando interpretaciones arbitrarias o heterogéneas, se exige del mismo un cierto rigor lingüístico que, sin caer en tecnicismos impropios de su finalidad divulgadora, fortalezcan la seguridad jurídica de aquéllos a quienes va dirigido.

La incorporación de la informática a los procesos de almacenamiento, tratamiento y distribución de la información jurídica está generando, por su parte, la creación y consolidación de lenguajes jurídico-documentales adaptados a esta nueva realidad, tal y como pone de manifiesto la incesante aparición de léxicos especializados, aptos para su utilización en los tratamientos automáticos de la información; léxicos que permiten direccionar, obtener o relacionar, en forma automática, los datos almacenados en aquellos documentos jurídicos que se encuentran archivados en bases de datos.

Los lenguajes jurídico-documentales constituyen, en el marco del antedicho fenómeno, los corpora terminológicos creados específicamente para tratar la información jurídica documentada, almacenada en determinados soportes accesibles en forma manual, mecánica o electrónica, con objeto de facilitar su recuperación mediante la utilización de técnicas documentales.

* Doctor en Derecho. Consejero Técnico del Ministerio de Justicia adscrito al Tribunal Supremo

Estos corpora están conformados por conjuntos de términos, estructurados y ordenados en una forma tal que permiten conjugar, mediante la aplicación de las citadas técnicas documentales en las prospecciones retrospectivas de información, los principios de exhaustividad y pertinencia en la recuperación discriminada de información jurídica.

Debe tenerse en cuenta que, mientras el lenguaje jurídico pretende hacer públicos unos preceptos, de forma tal que sean universalmente entendidos y asumidos, los lenguajes jurídico-documentales pretenden normalizar los términos relevantes extraídos de dicho lenguaje, con objeto de facilitar el tratamiento y recuperación de la información almacenada en los textos jurídicos que le sirven de referencia, de forma efectiva.

1.1.- Características del lenguaje jurídico

En el lenguaje jurídico, como lenguaje científico o de especialidad con características propias, se pone de manifiesto la coexistencia de diferentes tipos de términos.

1.- “Términos jurídicos consolidados”, entendiéndose por tales aquellos que se utilizan y aceptan en forma generalizada encontrándose habitualmente en textos y documentos normativos y contractuales, gozando en ciertos casos de una específica delimitación jurisprudencial y doctrinal, lo que les habilita para poner de manifiesto, a lo largo del tiempo, un contenido técnico-jurídico preciso, a veces muy alejado del sentido que a dichos términos le otorga el lenguaje natural (alimentos, mora, servidumbre...).

2.- “Términos del lenguaje común” que, en un momento dado, han servido como elementos de regulación de aspectos y relaciones jurídicas (matrimonio, persona, buque, negocio...).

3.- “Términos científicos”, recopilados de otras lenguas de especialidad, que se utilizan para especificar aquellos aspectos científico-técnicos que adquieren relevancia jurídica (drogas, trastorno mental, fecundación in vitro...).

4.- “Términos ambiguos o en formación” que permiten, sin variar el contenido literal de las normas o documentos en que se incluyen, interpretar los mismos a tenor de las situaciones concretas y del momento en que se aplican (unión de hecho, intereses sociales, orden público). Estos términos, habitualmente utilizados en el discurso jurídico permiten, durante el período previo a su delimitación jurídica concreta y precisa, por un lado solventar conflictos por la vía de su interpretación sistemática, y por otro servir de cauce para sensibilizar a la sociedad sobre aspectos que precisan de una regulación jurídica pormenorizada.

Estos 4 tipos de términos se encuentran habitualmente presentes en los documentos que los profesionales del Derecho utilizan, formando parte esencial de las fuentes del lenguaje jurídico.

La seguridad jurídica exige que, en contra de la práctica habitual del lenguaje literario que se enriquece con su variedad y donde se busca intencionadamente evitar la reiteración de unitérminos utilizando sinónimos remisiones o figuras poéticas, en la formalización del lenguaje jurídico se haya tendido insistentemente a la reiteración de aquellos términos y conceptos que la práctica ha evidenciado como suficientemente

precisos para delimitar determinados aspectos de relevancia jurídica, eludiendo su sustitución por sinónimos o cuasisinónimos más coloquiales pero que pueden propiciar una mayor ambigüedad y por consiguiente incrementar la inseguridad jurídica.

La estructuración del Derecho y los sucesivos intentos de codificación del mismo, han provocado que, junto a la antedicha reiteración terminológica, se hayan producido unas específicas relaciones entre los diferentes conceptos jurídicos, en función de su ubicación sistemática en los textos legales, generándose de hecho unas relaciones jerárquicas entre los mismos, que faciliten su interpretación y delimitación.

En la documentación jurídica automatizada puede observarse la coexistencia junto al lenguaje propio del contenido de los documentos, de otro lenguaje documental, que trata de resolver las ambigüedades producidas tanto en expresiones explícitas como implícitas de dichos documentos. El lenguaje documental, que recoge parte de la terminología jurídica, suele estar conformado por términos documentales con significados precisos, agrupados en campos semánticos, y ordenados mediante una estructura arborescente jerarquizada en forma de términos genéricos a específicos, sobre la que se inserta otra relacional que vincula a los términos que pudieran tener una cierta afinidad.

Las estructuras jerárquica y relacional establecidas sobre los términos documentales, que constituyen los diferentes tesauros utilizados para el procesamiento de la documentación jurídica, permiten una optimización de los procesos de detección, tratamiento y recuperación de la información contenida en la misma por medios automatizados y, asimismo, facilitan el desarrollo de la informática analítica, decisional, registral y documental.

La especial importancia de estas herramientas terminológicas documentales, han originado la elaboración de una serie de normas nacionales e internacionales en las que se estipula la metodología, características y especificaciones concretas que deben utilizarse para su elaboración, con la finalidad de lograr una normalización que facilite su compatibilización, utilización y tratamiento por medio de diferentes aplicaciones informáticas.

Junto al desarrollo de estas herramientas terminológicas documentales se hace preciso seguir avanzando asimismo en la normalización de las recensiones y citas de textos legislativos, jurisprudenciales y doctrinales, de las siglas y acrónimos, de los identificadores, etc., a fin de facilitar su almacenamiento, tratamiento y recuperación precisa por medios informáticos.

1.2.- Ambigüedad del lenguaje jurídico

El lenguaje jurídico plantea diversos problemas de naturaleza semántica y sintáctica. Aquí nos ocuparemos de una de las dificultades más evidentes, las ambigüedades, que pueden observarse en particular en los textos legislativos y en general en una gran parte de los actos jurídicos documentados.

La mayor parte de los textos jurídicos son lingüísticamente ambiguos, porque el lenguaje utilizado es en si mismo ambiguo. Asimismo puede observarse una continua incorporación de nuevos términos, que recepcionan genuinos conceptos, entresacados de la evolución normativa, jurisprudencial o doctrinal de terceros países, con idiomas distintos al

nuestro, que se recogen como neologismos de lenguas extranjeras, originando, en múltiples casos, falsas sinonimias en el lenguaje.

De igual forma puede observarse un incremento de términos polisémicos que surgen de los frecuentes trasvases terminológicos de unos campos científicos a otros y de errores en la utilización, poco ortodoxa, de las técnicas jurídicas de delimitación de conceptos.

El lenguaje natural, por su parte, ofrece igualmente un amplio repertorio de vocablos cuyo significado solo puede ser precisado a tenor del contexto en el que se encuentran, lo que debe tenerse muy presente a efectos de establecer la específica interpretación de los mismos cuando éstos se encuentren integrados en el discurso jurídico.

Las dificultades para hacer del lenguaje jurídico un lenguaje preciso, unívoco y lógicamente riguroso, no se limitan a la indeterminación del significado de los vocablos obtenidos del lenguaje común, ya que en el propio lenguaje científico, desarrollado para la concreción de conceptos estrictamente jurídicos, pueden asimismo observarse términos que, en función de su contexto, expresan un diferente contenido.

Es evidente que todo término jurídico se manifiesta a través de expresiones lingüísticas en las que se puede analizar por separado su formalización y contenido. Este análisis lingüístico que contrapone la forma y el fondo tiene una especial relevancia en aquellas disciplinas que, utilizando terminologías científicas propias, como ocurre en el lenguaje jurídico, precisan dotar a sus términos de un alcance preciso que eluda, en lo posible, ambigüedades no deseadas.

En el lenguaje jurídico existen términos que precisan de complementación con otros vocablos obtenidos de otras disciplinas científicas o del lenguaje natural para su posible comprensión. Esta característica del discurso jurídico provoca una cierta indeterminación de conceptos cuando éstos se delimitan mediante palabras que, en si mismas, presentan una cierta ambigüedad lingüística, lo que vulnera la seguridad jurídica que dicho discurso pretende garantizar.

Como ejemplos de ambigüedad del lenguaje jurídico y sin pretensión alguna de exhaustividad, podemos citar los siguientes:

a) Ejemplos de ambigüedades provocadas por diferir los conceptos referenciados por términos idénticos utilizados en el lenguaje natural y en el lenguaje jurídico.

Fabricante

En el lenguaje natural se considera fabricante al que fabrica; Entendiéndose por fabricar: - El hacer una cosa por medios mecánicos; - El construir un edificio, dique, muro o casa análoga; - Elaborar.

Por contra en el lenguaje jurídico y en el marco regulador de las bases de datos se entiende por "Fabricante de la base de datos" a la persona natural o jurídica que toma

la iniciativa y asume el riesgo de efectuar las inversiones sustanciales orientadas a la obtención, verificación o presentación de su contenido¹.

Tradición

En el lenguaje natural, el término tradición significa “transmisión de noticias, doctrinas, ritos, costumbres, etc., hecha de generación en generación” ó “doctrina, costumbre, etc., conservada en un pueblo por transmisión de padres a hijos”².

En el lenguaje jurídico, la tradición implica la transmisión del dominio sobre una cosa efectuada mediante la entrega de la misma en forma directa o simbólica³.

Acción

En el lenguaje natural el término acción significa: "ejercicio de una potencia", "efecto de hacer", "postura, ademán",...

En el lenguaje del campo científico circunscrito a las leyes de la física significa: "magnitud que define como producto de la energía absorbida durante un proceso por la duración del mismo".

En el lenguaje militar significa: "batalla", "combate o pelea entre fuerzas poco numerosas".

En el lenguaje jurídico mercantil significa: "Cada una de las partes en que se considera dividido el capital social de una sociedad anónima, y también, a veces, el que aportan los socios no colectivos a algunas sociedades comanditarias, que entonces pasan a llamarse comanditarias por acciones". Asimismo significa: "título que acredita y representa el valor de cada una de aquellas partes"⁴.

En el lenguaje jurídico procesal significa: "Derecho que se tiene a pedir alguna cosa en juicio", "Modo legal de ejercitar el mismo derecho, pidiendo en justicia lo que es nuestro o se nos debe"⁵.

¹ Ley de incorporación al Derecho español de la 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.

² Todas la definición referentes al lenguaje natural han sido obtenidas del *Diccionario de la lengua española*, de la Real Academia de la Lengua. Vigésima primera edición. Madrid 1992.

³ Párrafo segundo del artículo 609 del Código Civil “La propiedad y los demás derechos sobre los bienes se adquieren y transmiten por la ley, por donación, por sucesión testada e intestada, y por consecuencia de ciertos contratos mediante la *tradición*”.

⁴ Artículo 160 del Código de Comercio “El capital social de las compañías en comandita, perteneciente a los socios comanditarios, y el de las compañías anónimas, podrá estas representado por acciones u otros títulos equivalentes”; y en el artículo siguiente de dicho Código se añade que “Las acciones podrán ser nominativas o al portador”. Artículo 33 de la Ley de 17 de julio de 1951 “Las acciones representan partes alícuotas del capital social”.

⁵ Entre otros, el artículo 71 de la Ley 1/2000, de 7 de enero, de Enjuiciamiento civil, que indica “1. La acumulación de acciones admitida producirá el efecto de discutirse todas en el mismo procedimiento y resolverse en una sola sentencia. 2. El actor podrá acumular en la demanda cuantas acciones le competan contra el demandado, aunque provengan de diferentes títulos, siempre que aquellas no sean incompatibles entre si...”.

En el lenguaje jurídico penal, el concepto acción constituye uno de los elementos del delito (acción típicamente antijurídica y culpable)⁶.

Mora

En lenguaje natural, el término *mora* obtenido del latín *mora* "fruto del moral, de unos dos centímetros de largo, con figura ovalada, y que está formado por la agregación de globulillos carnosos, blandos, agridulces y de color morado", "fruto de la morera, muy parecido al anterior pero de la mitad de su tamaño, de color blanco amarillento y enteramente dulce", "fruto de la zarzamora", en algunos puntos "fresa silvestre".

Obtenido como femenino de moro, en latín *maurus*, significa: "natural de la parte de África septentrional, frontera a España, donde estaba la antigua provincia de Mauritania", "perteneciente a esa parte de África, o a sus naturales".

Por extensión del anterior concepto significada "que profesa la religión de Mahoma", "dícese del musulmán de Mindanao y de otras islas de Malasia, figurada y familiarmente "aplícase al vino que no está aguado, en contraposición del cristiano o aguado".

En lenguaje jurídico, obtenido del latín "mora", significa: "dilación o tardanza en cumplir una obligación, por lo común de pagar una cantidad líquida y vencida"⁷.

Monte

En el lenguaje común, del latín *mons, montis*, significa "gran elevación natural de terreno".

En el lenguaje jurídico, bajo este término se contemplan "todos los terrenos o propiedades forestales en que vegeten especies arbóreas, arbustivas, de matorral o herbáceas siempre que no sean características del cultivo agrícola o fueren objeto del mismo", exceptuándose de dicho concepto "los terrenos que formando parte de una finca fundamentalmente agrícola y sin estar cubiertos apreciablemente con especies arbóreas o arbustivas de carácter forestal resultaren convenientes para atender al sostenimiento del ganado de la propia explotación agrícola y, asimismo, los prados desprovistos sensiblemente de arbolado de dicha naturaleza y las praderas situadas en las provincias del litoral Cantábrico"⁸.

Mueble e inmueble

En el lenguaje natural los términos inmueble y mueble del latín *immobilis* y *mobile*, respectivamente tienen como significado "tierras, edificios, caminos, construcciones y minas" (a los que asimismo se denominan como bienes inmuebles o

⁶ Art. 10 del Código Penal "son delitos o faltas las acciones y omisiones dolosas o imprudentes penadas en la ley".

⁷ Artículo 1100 del Código Civil "Incurrir en mora los obligados a entregar o a hacer alguna cosa desde que el acreedor les exija judicial o extrajudicialmente el cumplimiento de su obligación".

⁸ Artículo 1º apartados 2º y 3º de la Ley de Montes de 8 de junio de 1957.

bienes raíces, en función de su inmovilidad) y “cada uno de los enseres móviles que sirven para los usos necesarios o decorativos de casas, oficinas o locales” (que reciben igualmente el nombre de bienes muebles en función de su movilidad).

En el lenguaje jurídico se consideran como inmueble además de las tierras, edificios, caminos y construcciones, otra serie de bienes entre los que se incluyen los árboles y plantas y los frutos pendientes, las estatuas, relieves, pinturas colocadas en edificios, las máquinas y utensilios destinados por el propietario de una finca destinados a satisfacer las necesidades de la explotación de la misma, los viveros de animales, palomares, colmenas, etc., los abonos destinados al cultivo, los diques y construcciones que, aun cuando sean flotantes, estén destinados a permanecer en un punto fijo de la costa, las concesiones administrativas de obras públicas y las servidumbres y demás derechos reales sobre bienes inmuebles⁹. Igualmente se consideran inmuebles los buques y las aeronaves¹⁰ a efectos de hipoteca y muebles¹¹ a los demás efectos.

Por su parte en este mismo lenguaje se reputan como mueble los bienes susceptibles de apropiación no considerados como inmuebles y en general todos los que se pueden transportar de un punto a otro sin menoscabo de la cosa inmueble a que estuvieren unidos¹². Asimismo se reputan como mueble las rentas o pensiones, sean vitalicias o hereditarias, afectas a una persona o familia, siempre que no graven con carga real una cosa inmueble, los oficios enajenados, los contratos sobre servicios públicos y las cédulas y títulos representativos de prestamos hipotecarios¹³.

Alimento

En el lenguaje natural, del latín *alimentum*, significa “la comida y bebida que el hombre y los animales toman para subsistir” o “cualquiera de las sustancias que los seres vivos toman o reciben para su nutrición”.

En el lenguaje jurídico se entiende por alimento a “aquellos bienes indispensables para cubrir las necesidades básicas de las personas”¹⁴.

Censo

En el lenguaje natural, del latín *census*, significa “padrón o lista de la población o riqueza de una nación o pueblo”.

⁹ Artículo 334 del Código Civil.

¹⁰ Artículo 1º de la Ley de Hipoteca naval, de 21 de agosto de 1893.

¹¹ Artículo 585 del Código de Comercio.

¹² Artículo 335 del Código Civil.

¹³ Artículo 336 del Código Civil.

¹⁴ Artículo 142 del Código Civil “ Se entiende por alimentos todo lo que es indispensable para el sustento, habitación, vestido y asistencia médica. Los alimentos comprenden también la educación e instrucción del alimentista mientras sea menor de edad y aún después cuando no haya terminado su formación por causa que no le sea imputable. Entre los alimentos se incluirán los gastos de embarazo y parto, en cuanto no estén cubiertos de otro modo”.

En el lenguaje jurídico se entiende por censo al “contrato mediante el cual una persona adquiere el derecho a percibir una asignación anual a cambio de la entrega de una cantidad de dinero o de una cesión de derechos sobre unos determinados inmuebles que quedan sujetos a dicho gravamen¹⁵”.

Depósito

En el lenguaje natural, del latín *depositum*, significa “acción y efecto de depositar”, “cosa depositada”, “lugar o recipiente donde se deposita”, “sedimento de un líquido”.

En el lenguaje jurídico comercial se entiende por depósito, del latín *depositum*, el “contrato mediante el cual se ponen bajo la custodia de una tercera persona determinados bienes con la obligación de esta última de restituirlos al cedente cuando este así se lo pida”¹⁶.

En el lenguaje jurídico administrativo se entiende por depósito legal la “obligación de los autores o editores de entregar cinco o tres ejemplares (según sean libros u otro tipo de obras impresas) de cada obra en forma previa a su distribución”

En el lenguaje jurídico procesal se entiende por depósito judicial a “aquel que a efecto de asegurar los bienes litigiosos se efectúa por orden o con la intervención del juez”

Servidumbre

En el lenguaje natural, del latín *servitudo*, *servitudinis*, significa “trabajo o ejercicio propio del siervo”, “estado o condición de siervo”, “conjunto de criados que sirven a un tiempo o en una casa”.

En el lenguaje jurídico se entiende por servidumbre a la “limitación del dominio sobre un terreno o bien inmueble a favor de terceras personas”¹⁷.

¹⁵ Artículo 1604 del Código Civil “Se constituye el censo cuando se sujetan algunos bienes inmuebles al pago de un canon o rédito anual en retribución de un capital que se recibe en dinero, o del dominio pleno o menos pleno que se transmite de los mismos bienes”.

¹⁶ Artículo 1758 del Código Civil “Se constituye el depósito desde que uno recibe la cosa ajena con la obligación de guardarla y restituirla”. Artículo 303 del Código de Comercio “Para que el depósito sea mercantil se requiere: 1º Que el depositario, al menos, sea comerciante. 2º Que las cosas depositadas sean objeto de comercio. 3º Que el depósito constituya por sí una operación mercantil, o se haga como causa o a consecuencia de operaciones mercantiles” y artículo 305 del mismo Código “El depósito quedará constituido mediante la entrega, al depositario, de la cosa que constituya su objeto”.

¹⁷ Artículo 530 del Código Civil “La servidumbre es un gravamen impuesto sobre un inmueble en beneficio de otro perteneciente a distinto dueño. El inmueble a cuyo favor está constituida la servidumbre, se llama predio dominante; el que la sufre, predio sirviente”; y artículo 531 del mismo Código “También pueden establecerse servidumbres en provecho de una o más personas, o de una comunidad, a quienes no pertenezca la finca gravada”.

b) Ejemplos de ambigüedades provocadas por la indefinición de los términos jurídicos, en si mismos considerados.

Buena fe

El concepto de buena fe, en el lenguaje jurídico, expresa el convencimiento que puede tener una persona de que actúa, o posee algún bien, de forma legítima.

El alcance de este concepto, que configura una regla de conducta a la que deben adaptarse los ciudadanos ante ciertos actos que exigen su cumplimiento para lograr un determinado efecto jurídico, dada su subjetividad, queda supeditado en múltiples casos al arbitrio y subjetividad de las personas que lo esgrimen o valoran¹⁸.

Orden público

En el lenguaje jurídico el orden público suele utilizarse a tenor de dos acepciones de similar ambigüedad. Por un lado se entiende como tal la “situación de normalidad” que existe cuando los ciudadanos respetan las normas a tenor de la “buena fe” generando un estado de sosiego y tranquilidad en la convivencia, y por otro al conjunto de principios imperativos, que no pueden ser violados por la autonomía de la voluntad, impuestas por el legislador, para preservar el “orden social” o el “bien común de la colectividad”¹⁹.

¹⁸ Artículo 7.1 del Código Civil “Los derechos deben ejercitarse conforme a las exigencias de la buena fe”. Artículos 433 y 434 del mismo Código “Se reputa poseedor de buena fe al que ignora que en su título o modo de adquirir exista vicio que lo invalide. Se reputa poseedor de mala fe al que se halla en el caso contrario” y “La buena fe se presume siempre, y al que afirma la mala fe de un poseedor corresponde la prueba”. Artículo 79 del citado Código “La declaración de nulidad del matrimonio no invalidará los efectos ya producidos respecto de los hijos y del contrayente o contrayentes de buena fe. La buena fe se presume”. Artículo 464 del susodicho Código “La posesión de bienes muebles, adquirida de buena fe, equivale al título”. Artículo 1258 del Código Civil “Los contratos se perfeccionan por el mero consentimiento, y desde entonces obligan no solo al cumplimiento de lo expresamente pactado, sino también a todas las consecuencias que, según su naturaleza, sean conformes a la buena fe, al uso y a la ley”.

¹⁹ Artículo 1 de la Ley de Orden Público de 30 de julio de 1959 “El normal funcionamiento de las Instituciones públicas y privadas, el mantenimiento de la paz interior y el libre y pacífico ejercicio de los derechos individuales, políticos y sociales, reconocidos en las leyes, constituyen el fundamento del orden público”. Artículo 1.3 del Código Civil “La costumbre sólo regirá en defecto de ley aplicable, siempre que no sea contraria a la moral o al orden público y que resulte probada”. Artículo 6.2 del mismo Código “La exclusión voluntaria de la ley aplicable y la renuncia a los derechos en ella reconocidos sólo serán válidas cuando no contraríen el interés o el orden público ni perjudiquen a terceros. Artículo 1255 del Código Civil “Los contratantes pueden establecer los pactos, cláusulas y condiciones que tengan por conveniente, siempre que no sean contrarios a las leyes, la moral, ni al orden público”. El Código Penal indica en su artículo 557 que “Serán castigados con la pena de prisión de seis meses a un año los que, actuando en grupo, y con el fin de atentar contra la paz pública, alteren el orden público causando lesiones a las personas, produciendo daños en las propiedades, obstaculizando las vías públicas o los accesos a las mismas de manera peligrosa para los que por ellas circulen, o invadiendo instalaciones o edificios”, en su artículo 559 que “los que perturben gravemente el orden público con objeto de impedir a alguna persona el ejercicio de sus derechos cívicos será castigado con las penas de multa de tres a doce meses” y en su artículo 561 que “el que, con ánimo de atentar contra la paz pública, afirme falsamente la existencia de aparatos explosivos u otros que puedan causar el mismo efecto, será castigado con la pena de prisión de seis meses a un año”.

La ambigüedad del concepto es clara y aún más patente si nos detenemos en el análisis de los términos que se utilizan para su determinación, que son igualmente ambiguos.

Buenas costumbres

En el ámbito del Derecho Civil se entienden como tales a aquellas que rigen la “convivencia armónica y sosegada de la comunidad”²⁰.

En el ámbito del Derecho Penal se entiende como tales a aquellas cuya vulneración atenta contra la “moralidad sexual colectiva”²¹.

Para evitar innecesarias reiteraciones cabe remitir a los comentarios de los términos precedentes, que sirven igualmente para los términos que se indican posteriormente.

Imprudencia

En el lenguaje jurídico, es aquella que se produce con ausencia de malicia causando un determinado daño, considerándose como imprudencia grave o temeraria cuando se produce con olvido de las precauciones más elementales que han de tenerse presentes para evitar daños a terceros²².

Diligencia de un buen padre de familia

En el lenguaje jurídico se entiende por tal al cuidado y precauciones que cualquier ciudadano medio adopta para evitar daños y perjuicios sobre su persona o sobre un tercero por su culpa²³.

²⁰ Artículos del Código Civil 1116 “Las condiciones imposibles, las contrarias a las buenas costumbres y las prohibidas por la ley anularán la obligación que de ellas dependan”; 1328 “Será nula cualquier estipulación contraria a las leyes o a las buenas costumbres o limitativa de la igualdad de derechos que corresponda a cada cónyuge”; y 1271 *in fine* “Pueden ser igualmente objeto de contrato todos los servicios que no sean contrarios a las leyes o a las buenas costumbres.

²¹ Artículos del antiguo Código Penal de 1973 (derogado por el actual publicado como Ley Orgánica 10/95 de 23 de noviembre, que omite dicho vocablo) 431 “El que de cualquier modo ofendiere al pudor o las buenas costumbres con hechos de grave escándalo o transcendencia incurrirá en penas de arresto mayor” y 566.5 “Incurrirán en pena de multa los que de igual manera ofendieren levemente a la moral, a las buenas costumbres o a la decencia pública”.

²² Artículo 12 del Código Penal “Las acciones u omisiones imprudentes solo se castigarán cuando expresamente lo disponga la ley”; Artículo 152.1 y 152.3 del Código Penal “El que por imprudencia grave causare alguna de las lesiones previstas en los artículos anteriores será castigado: ...” y “Cuando las lesiones fueren cometidas por imprudencia profesional se impondrá asimismo la pena de inhabilitación especial para el ejercicio de la profesión”

²³ Artículos del Código Civil 497 “El usufructuario deberá cuidar las cosas dadas en usufructo como un buen padre de familia; 270 “El tutor está obligado a ejercer dicha administración con la diligencia de un buen padre de familia”; 1555.2 “El arrendatario está obligado: A usar de la cosa arrendada como un diligente padre de familia, destinándola al uso pactado”; 1889 “el gestor oficioso deberá desempeñar su encargo con toda la diligencia de un buen padre de familia”; 1903 *in fine* “La responsabilidad de que trata este artículo cesará cuando las personas en él mencionadas prueben que emplearon toda la diligencia de un buen padre de familia para prevenir el daño”.

Utilidad pública

En el lenguaje jurídico se utiliza este término para indicar el “interés que la comunidad puede tener sobre la utilización de un determinado bien”, empleándose este concepto como causa de legitimación de determinados actos jurídicos (expropiaciones, establecimiento de servidumbres, etc.)²⁴.

1.3.- Tipología de las fuentes del lenguaje jurídico

La tipología tradicional de análisis de los diferentes modos de formalización del lenguaje jurídico parte de las fuentes documentales del Derecho, pudiendo establecerse, a efectos metodológicos, los siguientes tipos de documentos:

- Documentos legislativos

Este tipo de documentos, integrado por todas aquellas normas que en su conjunto conforman el derecho positivo, constituyen la principal fuente del lenguaje jurídico, habida cuenta de su primordial relevancia a efectos de establecer el ordenamiento jurídico aplicable en cada momento.

La exacta interpretación del contenido, de cada una de las partes de estos documentos, es la condición esencial para que los mismos puedan ser utilizados para resolver los conflictos que pudieran surgir de la realidad que pretenden regular, debiendo poner de manifiesto, en forma clara y precisa, su finalidad, permitiendo de esta forma que los ciudadanos puedan conocer los límites de sus derechos y libertades y como contrapartida sus obligaciones legales.

Desgraciadamente, más a menudo de lo necesario, el lenguaje utilizado en los textos legislativos no posee esas características de claridad y precisión, habida cuenta de que, la multiplicidad de sucesos y diversidad de situaciones que potencialmente pueden originarse, impide su específica y concreta regulación, razón por la cual el discurso jurídico precisa recurrir, en su expresión formal, a generalidades y vaguedades que lo convierten en un lenguaje necesariamente ambiguo.

De igual forma puede observarse, que dichos documentos, utilizan una diferente terminología en cada una de las partes en que suelen dividirse.

Así en la exposición de motivos de las diferentes leyes, se utiliza un lenguaje más coloquial procurando informar, en una forma muy asequible a los ciudadanos a quienes va dirigida, sobre las razones que han dado lugar a su elaboración así como sobre su finalidad y objetivos más concretos. Con esta exposición se pretende sensibilizar sobre la conveniencia de la regulación que a continuación se realiza, buscando obtener la aceptación generalizada de la ley como medio de fortalecer el contrato social configurado por el ordenamiento jurídico.

²⁴ Artículos del Código Civil 549 “Las servidumbres impuestas por la ley tienen por objeto la utilidad pública o el interés de los particulares”; 550 “Todo lo concerniente a las servidumbres establecidas para utilidad pública o comunal se regirá por las leyes ...”; y 349 “Nadie podrá ser privado de su propiedad sino por la Autoridad competente y por causa justificada de utilidad pública, previa siempre la correspondiente indemnización”.

Dada la finalidad perseguida, el lenguaje utilizado suele ser genérico y a veces florido incurriendo en barbarismos y otras perversiones del lenguaje cuando se estima que así pudiera ser mejor entendido por las personas o sectores afectados por su regulación. En determinados textos legislativos, especialmente en los elaborados en el ámbito municipal, la especial idiosincrasia de los autores de los mismos queda patente al permitir dichos textos, dada la cercanía entre el legislador y los destinatarios de los mismos y su aplicación a temas concretos de ámbito local, una cierta familiaridad en cuanto a la exposición de las causas que los motivan.

En la parte correspondiente al articulado, se busca, por contra, obtener una mayor precisión terminológica, introduciendo, cuando así se estima oportuno, definiciones que establecen el alcance del concepto legal asociado a determinados términos, excluyendo cualquier otra interpretación de los mismos, cambiando incluso, en algunos de ellos, el significado habitual con el que se utilizan en el lenguaje natural.

Sin embargo, aún en esta parte, se suelen observar ambigüedades generadas tanto de la necesidad de obtener el respaldo político de diferentes grupos parlamentarios, para lo cual se acude a menudo a la utilización de frases o términos vagos o generales que puedan permitir diferentes interpretaciones según el grupo que en cada momento tenga el poder político, como por la política de ir avanzando en el desarrollo legislativo, dejando para futuras regulaciones el desarrollo de aquellos aspectos incipientemente estipulados, utilizando para ello remisiones en blanco o preceptos de similar finalidad, tal y como puede observarse en múltiples tratados y convenios internacionales, en diferentes directivas comunitarias²⁵, en las normas constitucionales²⁶, y en algunas de las leyes de mayor enjundia²⁷.

Junto a estas ambigüedades “forzosas” pueden observarse otras derivadas de una incorrecta o inidónea utilización de la técnica legislativa ocasionadas por la utilización de vocablos excesivamente generalistas o polisémicos, que suelen generar una mayor inseguridad jurídica.

- Documentos jurisprudenciales

Este conjunto de documentos, constituido por autos y sentencias, sirve de complemento al Derecho positivo, estableciendo tanto el marco jurídico aplicable a casos concretos, como las consecuencias legales derivadas de los actos encausados, aportando información terminológica sobre el alcance y contenido de algunos conceptos jurídicos²⁸.

²⁵ Así en el art. 9 de la Directiva 95/46/CE de 24 de octubre de 1995 se expone que “En lo referente al tratamiento de datos personales con fines exclusivamente periodísticos o de expresión artística o literaria, los Estados miembros establecerán, respecto a las disposiciones del presente capítulo, del capítulo IV y del capítulo VI, exenciones y excepciones solo en la medida en que resulten necesarias para conciliar el derecho a la intimidad con las normas que rigen la libertad de expresión”.

²⁶ Así en el art. 18.4 de la Constitución “La ley limitará el uso de la informática para garantizar el honor y la intimidad personal y familiar de los ciudadanos y el pleno ejercicio de sus derechos”.

²⁷ Así en el art. 9.3 de la Ley Orgánica 15/1999, de 13 de diciembre, indica que “Reglamentariamente se establecerán los requisitos y condiciones que deban reunir los ficheros automatizados y las personas que intervengan en el tratamiento automatizado de los datos a que se refiere el artículo 7 de esta Ley”.

²⁸ “Considerando” de la Sentencia de 10 de noviembre de 1975 “La circunstancia eximente de la responsabilidad civil, de *legítima defensa*, es una causa de justificación objetiva y jurídica que se basa en el preponderante interés legítimo de la persona que se defiende, sobre el antijurídico del sujeto activo que realiza la agresión, y puede fundarse en la delegación que el poder público hace en aquel, al hallarse imposibilitado para defenderlo ejercitando el monopolio que le corresponde, por lo que le permite su

La posterior reiteración, en esos u otros textos relacionados, de dichos términos permite, familiarizarse con ellos y consecuentemente utilizarlos en forma precisa en posteriores documentos jurídicos, lo que habitualmente se realiza por parte de los operadores del derecho en una forma excesivamente redundante, siendo ésta una de las características del lenguaje jurídico documentado, que lo hacen más incomprensible para los no expertos en la materia.

Al igual que en los textos legislativos, la estructura de estos documentos, integrados por un encabezamiento, una primera parte de exposición de hechos o resultandos, una segunda de fundamentación jurídica o “considerandos” y una tercera y última de resolución del conflicto planteado o fallo, permite observar diferencias terminológicas en su redacción, utilizando un lenguaje más común y generalizado en el encabezamiento, exposición de los hechos y fallo, y más jurídico en la fundamentación jurídica en la cual se producen los desarrollos terminológicos anteriormente indicados.

- Documentos doctrinales

En este tipo de documentos generados por los autores que publican obras de contenido jurídico, se combina la utilización del lenguaje común con la de los lenguajes especializados en las diferentes ramas del derecho (fiscal, administrativo, penal, etc.), según la particular formación de sus creadores y la finalidad perseguida por los mismos en cuanto a su previsible difusión entre especialistas del Derecho o al público en general.

- Documentos declarativos de voluntad y contractuales

Este tipo de documentos presenta en cuanto a su formalización una serie de características propias, derivadas de su finalidad de establecer, en forma precisa, las condiciones particulares de unos determinados actos, delimitando las obligaciones y derechos concretos de los firmantes y las consecuencias derivadas de su incumplimiento.

Habida cuenta que en los contratos, junto a su objeto principal, establecido como conjunto de derechos y obligaciones asumido por cada una de las partes, se incluyen cláusulas específicas que pretenden evitar los posibles futuros conflictos que puedan surgir de las diferencias de interpretación sobre el alcance de los compromisos adquiridos, se suele acudir, para su elaboración, a la utilización de fórmulas ya acuñadas y sistemáticamente utilizadas, con la finalidad de incrementar la seguridad jurídica de los firmantes.

Esta técnica, origina una cada vez más precisa utilización de vocablos jurídicos consolidados en los textos normativos y jurisprudenciales, lo que dificulta su inteligibilidad para personas no expertas en el campo jurídico, que precisan normalmente de asesoramiento para su más completa comprensión.

- Documentos registrales

Las inscripciones registrales, responden al interés social de poder poner de manifiesto determinados aspectos de relevancia jurídica, que quedan reflejados en las

reacción defensiva, ante el acometimiento que padece, para que disminuya o evite la lesión que se efectúa o se cierne sobre sus bienes jurídicos protegidos de condición personal o real

correspondientes anotaciones, estableciendo una presunción iuris tantum sobre su veracidad.

La utilización generalizada de formulas y expresiones favorecen la consolidación de aquellos términos que permiten incrementar la seguridad jurídica propiciada por estas instituciones que pretenden facilitar, mediante la exhibición de las susodichas anotaciones plasmadas en sus libros, pruebas fehacientes sobre situaciones jurídicas concretas.

1.4.- Tratamiento documental del lenguaje jurídico

La irrupción de la informática en el mundo del Derecho y en particular la aparición de las bases de datos jurídicas como instrumentos de gran capacidad de almacenamiento, tratamiento y transmisión de la información, ha obligado a los juristas a habituarse a la utilización de los lenguajes documentales de acceso a la información, mediante los cuales, se formalizan las peticiones de recuperación de aquella información que, sobre los temas de búsqueda delimitados por las expresiones conformadas por términos de dichos lenguajes, existan en sus fondos, con un cierto ánimo de exhaustividad.

En el diálogo usuario-base de datos, se produce una imposición de partida que exige, al primero, adaptarse al lenguaje utilizado en los documentos almacenados en la segunda, de forma tal que si no se emplean, como términos de búsqueda, aquellos que coincidan con los incluidos en el texto o anotaciones documentales complementarias de dichos documentos, no se recuperará la información existente en los mismos.

La progresiva evolución de la industria de la información automatizada ha permitido desarrollar diversos métodos que suavizan esta difícil adaptación, de los usuarios de las bases de datos, a los presumibles términos literales existentes en los documentos cuya información se pretende recuperar.

A medida que la operatividad de los equipos informáticos fue acrecentándose, aumentando su velocidad operativa y ampliando su capacidad de almacenamiento de registros automatizados, la posibilidad de recuperación de la información almacenada en las bases de datos fue asimismo evolucionando en una doble dirección:

Por un lado se incrementó el volumen de información factible de ser depositado, tratado y distribuido, pasando de un almacenamiento meramente referencial que permitía la recuperación de los documentos por título, autor o clasificación a otros que incluían resúmenes y extractos cada vez más extensos, y por último a la incorporación del texto completo de los documentos.

Por otro se fue complementando paulatinamente la información con una serie de metadatos, elaborados mediante la aplicación de técnicas documentales, que asimismo fueron evolucionando de unas primeras clasificaciones temáticas a listados cada vez más precisos de palabras-clave, a veces adquiriendo la forma de resúmenes documentales, cuyos términos se fueron estructurando en los, cada vez más desarrollados y depurados tesauros asociados a las bases de datos.

La evolución de los tratamientos de la información, que obviamente no ha sido tan lineal como en forma simplificada se ha indicado anteriormente, ha ido acompañada de fuertes polémicas sobre sus ventajas e inconvenientes, que han contribuido a al desarrollo

global de las técnicas de recuperación de la información, confluyendo, en el momento actual, en su aplicación a sistemas semiinteligentes diseñados para tratar y recuperar, en forma precisa, aquellos conjuntos de documentos que contienen información jurídica relacionada con los temas determinados seleccionados por los usuarios del sistema.

Este desarrollo ha puesto de manifiesto el interés del tratamiento del texto completo de la información, que permite recuperar, con una gran precisión, documentos de los que tenemos algún tipo de referencia precisa (ej.: recuperación de una sentencia que se sabe trata sobre un "accidente de tráfico" producido al saltarse un "semáforo en rojo" un "autobús" de unas determinadas "líneas regulares").

Asimismo las búsquedas realizadas sobre archivos que almacenan el texto íntegro de los documentos, permite localizar aquellos que hagan referencia a algún concepto jurídico preciso y acuñado por la práctica habitual (ej.: recuperación de la jurisprudencia existente sobre desahucios), con una pérdida de información muy reducida limitada a documentos que sustituyan, a lo largo de todo el texto, el término empleado habitualmente por algún sinónimo.

Por último, este tipo de almacenamiento permite suministrar en forma inmediata, una vez localizados los documentos pertinentes, su texto completo con el consiguiente ahorro del tiempo de obtención de la información completa al no tener que realizar una posterior tarea de solicitud de localización y transmisión del contenido del documento original.

Sin embargo los tratamientos de textos completos, que no incorporan algún tipo de análisis documental, producen una serie de imprecisiones derivadas de las particularidades, anteriormente indicadas, de los términos utilizados en los documentos jurídicos (ambigüedades, polisemias, sinonimias, etc.). Estas imprecisiones cobran una especial relevancia en búsquedas en las que se precisa utilizar términos no muy precisos en documentos de gran extensión (libros, códigos legales, monografías, tesis, ...) en los que el elevado número de palabras incluidas en los mismos acrecienta la posibilidad de producir combinaciones diabólicas, que desvirtúan totalmente la intencionalidad de las ecuaciones de búsqueda, haciendo prácticamente ineficaz la búsqueda en texto libre (ej.: búsquedas sobre "medidas" de "protección" a la "tercera" "edad"; o sobre el "estado" de "derecho").

Los lenguajes documentales, una vez estructurados, presentan asimismo ciertos inconvenientes:

La elaboración del lenguaje jurídico-documental obliga a la previa selección de términos, a su tratamiento lexicográfico y a su ubicación en las correspondientes estructuras del tesoro, en una forma que difiere de la organización sistemática del Derecho.

El análisis de los documentos, depende de la valoración y apreciación que, el documentalista encargado del mismo, haya realizado tanto de la información almacenada en dichos documentos, como de su correspondencia con los términos del tesoro que considere más idóneos para su recuperación. Estas decisiones producen una cierta reducción respecto al grado de exhaustividad que podría lograrse mediante prospecciones directas sobre el texto íntegro de los documentos.

La ineficacia del tratamiento documental como instrumento que pueda servir para sustituir al contenido de los documentos seleccionados, lo que obliga a una posterior recuperación del texto completo de los mismos.

Por contra, los lenguajes documentales presentan ciertas ventajas dignas de ser puestas en consideración.

La posibilidad de facilitar, a los usuarios de la base, el lenguaje utilizado para el análisis de los documentos (sirviendo éste de puente entre el utilizado por los autores en los documentos y el habitualmente empleado por cada usuario concreto) permitiendo obtener una mayor precisión en la recuperación de la información.

La reducción de pérdidas significativas de información al utilizar términos claramente definidos en forma documental²⁹ que evitan polisemias y ambigüedades, permitiendo recuperar documentos pertinentes en cuyo texto aparezcan sinónimos de dichos términos, conceptos implícitos, etc., sin necesidad de tener que ajustarse a las palabras originalmente utilizadas por los autores de dichos documentos.

La ampliación automática de los términos de búsqueda a aquellos otros que previamente se han estructurado como específicos de los primeros en los correspondientes tesauros³⁰.

La posibilidad de incrementar la exhaustividad de la información incorporando a las ecuaciones de búsqueda, los términos estructurados en el tesoro como genéricos de los inicialmente seleccionados³¹, con el inconveniente de incrementar asimismo el volumen de información recuperada no pertinente.

La facilidad de obtener información complementaria sobre un tema determinado, al establecerse, en los correspondientes tesauros, indicaciones sobre los términos relacionados, existentes en el lenguaje documental, lo que permite analizar el tema de la búsqueda desde diferentes ópticas y posibilitar análisis de temas conexos con la misma³².

La evolución y conjunción de los sistemas basados en el tratamiento del texto libre y en el tratamiento documental, permite aventurar que, en un futuro no muy lejano, se producirá un cambio cualitativo que obligue al sistema de interrogación de la base de datos a adecuarse al lenguaje utilizado por cada usuario, elaborando automáticamente las ecuaciones de búsquedas, buscando los términos sinónimos y específicos, localizando de entre sus fondos los documentos que aporten información sobre el tema solicitado, y suministrando dicha información ordenada en función de su presumible interés para el usuario.

²⁹ El término "acciones" se podría definir documentalmente limitado a títulos representativos de una parte del capital social de las empresas, excluyendo por tanto otras acepciones que, para su aplicación en el mismo lenguaje documental, deberían por tanto ir necesariamente apellidadas ("acciones procesales", "acción reivindicatoria", etc.)

³⁰ Solicitándose "Delitos contra la propiedad" se recuperarían tanto los documentos que traten en forma genérica dicho tema, como aquellos otros que contengan información sobre "Robos", "Apropiaciones indebidas", "Hurto", etc.

³¹ Para ampliar una búsqueda sobre "contratación temporal" se podría incluir en las ecuaciones de búsqueda el término estructurado en el tesoro como genérico del anterior "contratación laboral".

³² Como términos relacionados con cualquier tipo de delitos aparecerían los términos "eximentes", "atenuantes", "agravantes", etc.

En la mayoría de bases de datos jurídicas que almacenan textos completos, puede observarse la coexistencia del lenguaje jurídico, en el que se manifiesta su contenido, con un lenguaje documental, que trata de resolver las ambigüedades del primero.

La estructuración y puesta en relación entre los diferentes términos documentales, conformando tesauros, permiten una mejor optimización de los tratamientos y recuperación de la información por medios automatizados y facilitan el desarrollo de la informática analítica, decisional, registral y documental.

Dada la especial importancia de esta herramienta terminológica documental, se han elaborado una serie de normas nacionales e internacionales sobre la materia, en las que se establece la metodología, características y especificaciones concretas que deben utilizarse como medio de homogeneización de los diferentes lenguajes documentales, con la finalidad de lograr su compatibilización, facilitando su utilización y tratamiento por medio de diferentes aplicaciones informáticas.

La incorporación, a los sistemas informáticos que inicialmente gestionaban las bases de datos jurídicas de diferentes metodologías de recuperación de información, específicamente desarrolladas para su aplicación sobre la terminología obtenida de los textos completos de los propios documentos o sobre la terminología jurídico-documental utilizada como complemento de la anterior, provocó un amplio foro de discusión sobre las ventajas e inconvenientes de cada una de ellas.

Los defensores de las metodologías de recuperación diseñadas para su aplicación sobre el lenguaje literal, defendieron la tesis de que, en las bases de datos legislativas, jurisprudenciales y parlamentarias, debería introducirse todo el texto de los documentos, en función de las necesidades de inmediatez, certeza y seguridad jurídica que la posterior utilización de dichos textos imponía, siendo estos criterios mayoritariamente asumidos por la mayoría de productores de este tipo de bases que de esta forma se convirtieron en los grandes impulsores del desarrollo de sistemas de recuperación adaptados a las peculiaridades de la interrogación sobre el texto completo de los documentos almacenados en soportes informáticos.

La inclusión, en las bases de datos, del texto completo de los documentos, en su lenguaje original, si bien tiene como objetivo principal la transmisión del mismo con el lenguaje auténtico utilizado por sus autores, tiene asimismo como objetivos secundarios los de evitar efectuar análisis previos sobre su contenido, reducir los costes que se derivarían de estos trabajos, eliminar la subjetividad de estos análisis que podía desvirtuar el contenido de la información, eludir la necesidad de tener que recurrir al texto completo en un momento posterior y permitir utilizar los soportes magnéticos utilizados para la edición impresa como instrumento de carga de la información en las bases de datos en forma automática.

Los defensores de los sistemas de recuperación basados en la utilización de lenguajes documentales argumentaron, por su parte, que las bases de datos que almacenaban referencias de los documentos mediante tratamientos documentales, sin incorporar el texto completo de los mismos, tenían como ventajas las de reducir ostensiblemente los costes de almacenamiento, obtener una mayor exhaustividad en la recuperación de documentos relevantes gracias a la posibilidad de los tratamientos documentales de poner de manifiesto

ideas implícitas, evitar el ruido y las imprecisiones derivadas de los lenguajes naturales, y suministrar una mayor exhaustividad en la localización de documentos gracias a la utilización de términos documentales debidamente estructurados y relacionados en los correspondientes tesauros.

En la actualidad las mejoras introducidas en las nuevas aplicaciones informáticas diseñadas específicamente para su utilización sobre textos jurídicos, han originado el establecimiento, en la mayoría de bases de datos normativas, jurisprudenciales o parlamentarias, de unos sistemas mixtos de almacenamiento, tratamiento y transmisión de la información, que operan tanto sobre el lenguaje natural obtenido de los textos íntegros de los documentos normativos, jurisprudenciales o parlamentarios, como sobre los lenguajes documentales que complementan dichos textos, manteniéndose los sistemas basados únicamente en lenguajes documentales para su utilización en las bases de datos bibliográficas.

2.- Terminología jurídico-documental: conceptos y términos

La terminología jurídico-documental, entendida desde una perspectiva estricta de utilidad informativa, representa el resultado de la formalización de un lenguaje científico que refleja, en cuanto a su estructura, la organización conceptual de un determinado dominio del ordenamiento jurídico, y en cuanto a su terminografía, la determinación precisa de los términos que representan los conceptos de relevancia jurídica encuadrados en su cobertura temática.

La elaboración de los corpora terminológicos, vistos desde esa perspectiva, ponen de manifiesto una relación biunívoca entre conceptos jurídicos y términos jurídico-documentales, relación que puede establecerse bien mediante procesos onomasiológicos como mediante procesos samasiológicos.

Los primeros, avalados por Wüster y la Escuela de Viena, parten del supuesto de un conocimiento profundo del campo jurídico a cubrir, lo que permite detectar y determinar todos aquellos conceptos que establecen realidades jurídicas dignas de ser puesta en consideración, y una vez puestos de manifiesto, establecer su representación particularizada mediante términos documentales que reflejan en forma precisa su contenido.

Por su parte los segundos, avalados por la Escuela canadiense, parten del análisis de los términos habitualmente utilizados por los especialistas del campo o dominio en estudio para, una vez detectados, establecer sus relaciones con los conceptos que les sirven de referencia y proceder a su ordenación y estructuración documental.

En el campo del Derecho y a diferencia de esas corrientes terminológicas que propugnan metodologías estrictas basadas en una u otra sistemática, debe plantearse una tercera vía, asimismo propugnada por Rey, A.; Rondeau, G.; Thiry, B., etc., consistente en la combinación de ambos métodos.

A tenor de la metodología establecida en esta tercera vía, por un lado debe tenerse en cuenta la forma de organización del conocimiento en el dominio en estudio, prestando especial atención a la ordenación sistemática del ordenamiento jurídico regulador de la materia a la que se va a circunscribir el trabajo terminológico-

documental, y las opiniones y conocimiento de los especialistas en la misma, y por otro analizar los textos que, en forma explícita, recogen los contenidos de lege data, la interpretación de los mismos, y los debates y propuestas de lege ferenda, a fin de recoger y formalizar los términos concretos utilizados en el discurso jurídico y, partiendo de los mismos, establecer las relaciones con sus correspondientes conceptos.

En una segunda fase deberán ajustarse las imprecisiones o contradicciones surgidas de asociaciones asimétricas entre términos y conceptos establecidas en virtud de la utilización de ambas metodologías.

Con la aplicación de esta sistemática puede obtenerse una herramienta válida para el análisis y tratamiento de aquellos documentos legales³³ que constituyen los fondos de las diferentes bases de datos, con la finalidad de proceder a su tratamiento documental en una forma tal que posibilite su posterior recuperación por los potenciales usuarios de la base en atención de sus finalidades personales.

La unidad básica de los léxicos jurídico-documentales es el "término", que representa, mediante una determinada grafología, el contenido de un determinado concepto.

Sin entrar en la casuística de la formación y consolidación de términos ni en los aspectos relacionados con su fonología, morfología o neología, que excederían del objeto de lo aquí pretendido, si cabría indicar que el lenguaje jurídico-documental esta conformado indistintamente por términos simples (o unitérminos) que se representan mediante una única palabra constituida a base de morfemas³⁴, y términos compuestos (o multitérminos) constituidos por una pluralidad de palabras en forma de sintagma³⁵.

Junto a estas unidades básicas coexisten habitualmente, en los léxicos jurídico-documentales, unidades documentales de utilidad eminentemente práctica, tales como las formadas por siglas, acrónimos, abreviaturas, etc. que permiten simplificar la formación de identificadores, contracciones de términos de uso normal, etc.

Los conceptos, a que hacen referencia los términos jurídico-documentales, están constituidos, pro su parte, por aquellas ideas que sirven, en cuanto a elaboración mental, a delimitar y especificar los objetos del mundo subjetivo u objetivo percibido por cada persona que constituyen su base de conocimiento.

Los conceptos representan, por ende, imágenes del conocimiento que, para su comprensión y divulgación a terceros, precisan de la utilización del lenguaje y de la intercomunicación personal. Su delimitación y difusión suele establecerse mediante aproximaciones sistemáticas que permiten, en el marco del diálogo interpersonal, exponer y transmitir ideas en una forma tal que puedan ser entendidas por los partícipes del mismo, sirviendo de cauce a la crítica, debate o enseñanza.

³³ Esta labor se realiza mediante las tareas de indización de los documentos, que implica la complementación de la información textual mediante la incorporación de los descriptores o palabras clave (delimitación de términos "descriptor" y "palabra clave" que sirve únicamente para reflejar la existencia de un vocabulario controlado estructurado o no en forma de tesoro)

³⁴ Términos configurados mediante una única palabra

³⁵ Términos configurados por una combinación de palabras y una estructura sintáctica

Todo concepto esta constituido por propiedades o características propias que lo diferencian del resto, razón por la cual el acercamiento al mismo puede realizarse mediante la descripción del contenido de las mismas, el establecimiento de analogías o diferencias con las presentes en otros conceptos, o la exclusión de aquellas que le son ajenas.

En cada uno de los campos científicos del saber, los conceptos se interrelacionan ineludiblemente unos con otros, estableciendo estructuras de diverso tipo que facilitan su conocimiento y comprensión. Así, los conceptos de cada disciplina, se entrecruzan formando parte de una estructura relacional a partir de la cual adquieren su verdadero significado.

La relación entre términos y conceptos representa el fundamento básico de la terminología jurídico-documental, que como objeto pretende establecer dichas relaciones en forma biunívoca, eliminando polisemias y homonimias, agrupando sinonimias, y relacionando las estructuras relacionales de conceptos con estructuras arbóreas o analógicas de términos.

Mediante el establecimiento de estas relaciones, y una metodología práctica de elaboración de léxicos normalizados que eviten problemas aparentemente inocuos, pero de vital importancia para el tratamiento y recuperación de información (utilización de singulares y plurales, palabras vacías, tratamiento de prefijos, infijos y sufijos...), se elaboran los productos documentales "ad hoc" que sirven para facilitar el flujo y transmisión de información jurídica, en unas condiciones óptimas de utilidad, a tenor de las necesidades de los potenciales usuarios de las bases de datos y de los servicios de información automatizada.

3.- Tesauros jurídicos

Los tesauros jurídicos constituyen aquellos corpora de términos, configurados en forma estructurada, que en forma biunívoca representan, con aspiración de exhaustividad, la totalidad de conceptos del campo del derecho que se pretende cubrir, estableciendo relaciones semánticas, jerárquicas o asociativas e incorporando, cuando así se estima oportuno, definiciones o especificaciones particularizadas sobre aquellos términos que precisen de las mismas para evitar ambigüedades o polisemias.

3.1.- Función documental de los tesauros

La finalidad de los tesauro es la de facilitar el tratamiento y recuperación de la información almacenada en bases de datos, razón por la cual debe acomodarse a su contenido a fin de incrementar la precisión de los términos que los conforman con vistas a lograr el grado de pertinencia pertinencia y exhaustividad en su utilización que exige la seguridad jurídica, partiendo de su concepción como elemento de complementación documental de la información consignada en cada documento y de recuperación discriminada de aquellos que traten sobre un determinado tema.

Esta característica permite establecer una serie de peculiaridades entre las que podemos destacar las siguientes:

En cuanto al conjunto de documentos que le sirven como fuente de información y por ende son susceptibles de ser recuperados en forma individualizada, cabe indicar la dificultad que entraña, como anteriormente se ha indicado, la coexistencia de diferentes fuentes documentales, lo que dificulta el establecimiento de una terminología uniforme que se adapte, con precisión a la pluralidad tipológica de las mismas (documentos legislativos, jurisprudenciales, doctrinales...), mediante un único lenguaje desarrollado con la profundidad de precisión exigida para la recuperación de textos de diferente finalidad, contenido y amplitud informativa (libros, sentencias, leyes, formularios, contratos...).

Lógicamente esa dificultad será sensiblemente menor en aquellos tesauros elaborados para su utilización sobre bases de datos cuyo ámbito de aplicación se circunscriba a un determinado tipo de documentos, con exclusión del resto.

De igual forma, y en lo que respecta al grado de profundidad y desarrollo sintagmático de los términos del tesoro que reflejan conceptos más o menos específicos, con el los consiguientes problemas documentales de "silencio" y "ruido"³⁶, cabe indicar los inconvenientes que, para los objetivos de pertinencia y exhaustividad, en cuanto al conjunto de términos que lo deberían conformar, presenta la extensa cobertura del Derecho; dificultad que igualmente se ve drásticamente reducida en las bases de datos que abarcan, en forma limitada, algún campo del Derecho (Civil, Penal...) y aun más en las bases de datos especializadas en un determinado tema (Propiedad intelectual, Medio ambiente...).

Lógicamente cuando se producen bases de datos que incorporan combinaciones de tipología documental y ámbito temático, las dificultades de elaboración de tesauros, aptos para su aplicación sobre los fondos que almacenan, se multiplican.

Los tesauros jurídicos deben igualmente facilitar la incorporación de nuevos términos, e incluso la variación de algunos de ya existentes en función de las ampliaciones y modificaciones de los ordenamientos jurídicos que inciden en el ámbito de la documentación almacenada, razón por la cual es preciso, en determinados casos, reseñar la fecha de la incorporación o variación, para tenerla presente, a efectos de delimitar su vigencia, en las futuras aplicaciones del término de referencia.

La estructura orgánica del tesoro deberá asimismo permitir, en una forma acorde con la del ordenamiento jurídico que le sirve de base, navegar entre conceptos relacionados y a ser posible detectar no tan solo los documentos que los contienen sino incluso los párrafos del documento en los que se encuentran ubicados los términos de referencia, utilidad especialmente interesante para aquellas bases de datos que almacenen documentos de gran extensión³⁷.

Ya se ha indicado la dificultad añadida impuesta por la conjunción, en los documentos jurídicos, de términos extraídos de los lenguajes natural, jurídico y

³⁶ Se denomina ruido a la obtención de documentos o información no coherente con el tema de búsqueda, y silencio a la no obtención de documentos o información pertinente existentes en la base de datos que no han sido localizados.

³⁷ Considérese la dificultad que entrañaría, una vez localizado un documento pertinente que contiene el concepto buscado, por ejemplo la Ley de Presupuestos Generales del estado, localizar el lugar preciso donde se encuentra el mismo.

provenientes de otras ciencias o áreas de actividad. Pero esa dificultad resulta más acuciante si tenemos en cuenta la pluralidad y peculiaridades de los posibles usuarios de las bases de datos.

Así podemos observar, que las bases de datos jurídicas, y por ende los tesauros incorporados a las mismas, son habitualmente consultados por juristas, documentalistas y no expertos en ninguna de esas dos ciencias, lo que obliga a conjugar los criterios de utilización de términos jurídico-documentales, con términos habituales del lenguaje corriente o de otros campos del saber, a fin de adaptar la utilidad efectiva de los elementos utilizados para la recuperación de la información almacenada en las bases, a los usos habituales de los potenciales usuarios de las mismas.

Sin embargo no todo son inconvenientes a la hora de elaborar o utilizar tesauros jurídicos ya que, por su propia finalidad, gozan de las ventajas de contar con una serie de estructuras previas de conceptos más o menos relacionados, provenientes de la organización del ordenamiento jurídico. Estas estructuras son aún más perceptibles en los Códigos (Penal, Civil, de Comercio...) que a lo largo de los tiempos han ido configurando, en una forma pormenorizada, las instituciones jurídicas básicas y fundamentales del Estado de derecho.

El acercamiento de la estructura de los tesauros jurídicos al derecho positivo establecido en las diferentes leyes no es solo una pretensión apriorística sino, en la mayoría de los casos, una necesidad devenida de la propia metodología de su conformación y ordenación, metodología que debe ser especialmente rigurosa en la elaboración de tesauros multilingües, donde la traducción precisa de los conceptos está, a menudo, supeditada a la sistematización de los diferentes ordenamientos jurídicos de referencia, que pueden diferir, en mayor o menor grado, en cuanto a su propia concepción y contenido.

3.2.- Forma de elaboración de los tesauros

Si bien la elaboración de Tesauros puede realizarse mediante diferentes metodologías y criterios, a efectos didácticos se expone a continuación la utilizada en el CINDOC³⁸ para la producción de diferentes tesauros jurídicos (Tesauro de Derecho, Tesauro de Propiedad intelectual, Tesauro de Propiedad intelectual,...) basada en las directrices de las Normas ISO y UNE de confección de tesauros, adaptadas a las peculiaridades del campo a cubrir, que se desarrollo en las siguientes fases.

- Primera fase

Como elemento básico de partida se prestó atención a la documentación jurídica existente sobre el tema en estudio, realizándose los siguientes trabajos terminológicos sobre la mismo:

- Extracción de la lista de términos asignados a los conceptos obtenidos de la citada documentación.

³⁸ Centro de Información y Documentación Científica del Consejo Superior de Investigaciones Científicas (CSIC).

- Análisis del desarrollo teórico de la cabecera global del Tesouro; comprobación de la cobertura de la misma; y complementación en los casos en que se consideró insuficiente.

- Integración de la totalidad de términos recopilados, en las diferentes segmentaciones temáticas correspondientes a la cabecera del primer nivel jerárquico del Tesouro, asignándoles a cada uno de ellos una clasificación de dos dígitos en el que el primero corresponde a la subcabecera del Tesouro y el segundo al asignado, siguiendo un riguroso orden alfabético, a la primera subdivisión o nivel temático de la cabecera.

En el desarrollo de esta fase, cuando se detectaban sinónimos al agrupar los diferentes términos, se ponía de manifiesto dicha circunstancia, seleccionando como preferente uno de ellos y remitiendo el otro al seleccionado, clasificando el no seleccionado a tal efecto con una clave que indicaba la sinonimia.

Asimismo cuando durante este proceso se detectaban términos que se consideraba habían sido incluidos por error, bien por su indefinición o por entender que correspondían a otra subcabecera del tesouro, se procedía a su retirada de la subcabecera tomando nota de los mismos para su posterior reubicación, corrección o eliminación.

Al finalizar esta fase, como producto informativo, se redactaban los siguientes listados:

- Lista del desarrollo del tesouro organizada según la clasificación asignada a sus términos con dos niveles de profundidad, colocados en orden alfabético aquellos términos configurados con la misma clasificación.

- Lista de sinónimos con indicación del termino, cuya existencia debía previamente contrastarse, en el que quedaban subsumidos.

- Segunda y posteriores fases

La segunda y posteriores fases se iniciaban con la reubicación de los términos mal clasificados integrándolos en las subcabeceras que se consideraban más idóneas.

Una vez realizada dicha reasignación, se reelaboraba y analizaba la conformación completa de las diferentes subcabeceras, armonizando la precisión del conjunto total de términos y su grado de profundidad.

Posteriormente se realizaba, en forma similar a la indicada en la primera fase, la elaboración del siguiente nivel jerárquico en cada una de las subsegmentaciones que así lo precisasen, reelaborando la integración de los términos correspondientes a dichas subsegmentaciones en las nuevas divisiones establecidas en las mismas.

Esta forma de proceder se repetía, para cada subcabecera del Tesouro, cuantas veces se consideraba procedente hasta completar la jerarquización completa de todos los términos.

- Fase final

Una vez finalizados los trabajos de ordenación jerárquica y de reenvíos entre términos sinónimos se procedía a su tratamiento informático generando los Índices Kwoc y conceptual, que permitían hacer una profunda revisión individualizada de cada término en función de su formalización sintagmática y ubicación, posibilitando el establecimiento de relaciones asociativas con otros términos ubicados en diferentes subcabeceras temáticas.

Por último se añadían, a los términos ambiguos, polisémicos, o de difícil conceptualización, notas de alcance definatorias de su contenido o de otras especificaciones documentalmente útiles, tales como datos sobre su vigencia, desarrollo de siglas o abreviaturas, etc.

3.3.- Estructura de los tesauros

La estructura final de los tesauros así realizados quedaba formada por los siguientes Índices:

- Índice jerárquico

Este Índice permite visualizar globalmente, en forma arborescente, la estructura sistemática completa del Tesauro.

El mismo consta de dos columnas, en la primera se pone de manifiesto la clasificación asignada al descriptor que figura en la segunda, permitiendo determinar la ubicación de éste en la organización temática.

Cada parte del índice comienza con un término cabecera de área o dominio, al que se le asigna una letra clasificadora que se mantiene como inicio de la clasificación del resto de descriptores asignados a las sucesivas subcabeceras en las que se subdivide progresivamente.

A continuación, siguiendo un orden alfabético, se desarrollan las diferentes segmentaciones del dominio añadiendo, en la columna de la izquierda, a la letra clasificadora de la subcabecera cuantos dígitos se hubieren considerado precisos para poner de manifiesto los sucesivos niveles jerárquicos de profundidad, o concreción subtemática, pertinentes para establecer la precisa ubicación de los diferentes conceptos reflejados en los descriptores que figuran en la columna derecha del índice.

Cada uno de estos niveles constituye a su vez el inicio de una subclasificación más precisa, que irá acotando el campo temático de ubicación de los diferentes conceptos, englobando en el concepto representado por cada descriptor a todos aquellos otros conceptos más específicos representados por los descriptores incluidos en los sucesivos niveles de concreción establecidos a partir del mismo.

Ejemplo de Índice jerárquico

I	Derecho Penal
II	Delitos
II1	Aborto
II2	Allanamiento de morada

.....
I1j	Delitos contra la libertad
I1k	Delitos contra la libertad sexual
I1k1	Abusos sexuales
I1k11	Estupro
I1k12	Incesto
I1k2	Acoso sexual
I1k3	Agresiones sexuales
I1k31	Violación
I1k4	Escándalo público
I1k41	Exhibicionismo
I1k5	Prostitución
I1k51	Proxenetismo
I1k52	Trata de blancas
I1k6	Provocación sexual
I1k61	Corrupción de menores
I1k62	Pornografía
I1k621	Difusión de material pornográfico
.....

- Índice alfabético

Este Índice permite visualizar, según un orden alfabético, la totalidad de descriptores (términos preferentes que representan conceptos jurídicos determinados) y no descriptores o reenvíos (términos no preferentes o absorbidos por descriptores), indicándose junto a los primeros su clasificación temática y junto a los segundos, tras el vocablo USE entre paréntesis, el descriptor a que han sido reenviados o por el que han sido absorbidos.

Ejemplo de Índice alfabético

.....
Delitos contra la intimidad (I1i)
Delitos contra la inviolabilidad del domicilio (use) Allanamiento de morada
Delitos contra la libertad (I1j)
Delitos contra la libertad de conciencia (I1e52)
Delitos contra la libertad de mercado (I181)
Delitos contra la libertad del trabajo (use) Delitos contra los derechos de los trabajadores
Delitos contra la libertad sexual (I1k)
Delitos contra la libertad sindical (I1r2)
Delitos contra la navegación (I23)
Delitos contra la navegación aérea (I231)
Delitos contra la ordenación del territorio (I1l)
Delitos contra la paz (I1m)
Delitos contra la propiedad (use) Delitos contra el patrimonio
Delitos contra la propiedad industrial (I196)
Delitos contra la propiedad intelectual (I197)
.....

- Índice permutado Kwoc

Este índice permite determinar aquellos descriptores o no descriptores (reenvíos) que contengan una determinada palabra en su representación formal, con independencia del lugar que esta ocupe en el contexto del término.

Tras cada palabra significativa³⁹, ordenadas en forma alfabética, queda agrupados todos aquellos descriptores o reenvíos que la contengan, ordenados asimismo alfabéticamente.

Ejemplo de Índice permutado Kwoc

Delito

Actos preparatorios del delito
Apología del delito
Aseguramiento de la impunidad del delito
Autores del delito
Consumación del delito
Cuerpo del delito
Delito a distancia
Delito calificado por el resultado
Delito complejo
Delito compuesto
Delito consumado
Delito continuado
Delito contra el derecho de huelga
.....
Instrumentos del delito
Móvil del delito
Objeto del delito
Participación en el delito
.....

- Índice conceptual

Este índice, recoge ordenados en forma alfabética, la totalidad de términos integrados en el Tesauro, permitiendo visualizar, por una parte, el entorno conceptual y temático de los diferentes términos preferentes o descriptores y, por otra, los reenvíos efectuados a partir de términos no preferentes o no descriptores.

Para la elaboración de este índice se adoptaron, siguiendo las pautas indicadas en las Normas ISO 5964-1985 y UNE 50-125 sobre Directrices para la creación y desarrollo de Tesauros multilíngües, los siguientes símbolos:

- => Reenvío del "Término reseñado" al término preferente.
- = Término no preferente reenviado al "Término reseñado".
- CL Clasificación del "Término reseñado".

³⁹ Se excluyen del índice las palabras vacías: artículos, preposiciones, conjunciones...

- < Término genérico del "Término reseñado".
- < Término genérico del genérico del "Término reseñado".
- < Término genérico del genérico del genérico del "Término reseñado".
- > Término específico del "Término reseñado".
- > Término específico del específico del "Término reseñado".
- > Término específico del específico del específico del "Termino reseñado".
- Término relacionado con el "Término reseñado"

La reseña de los diferentes tipos de términos, que se incluyen en este Índice, se realiza en la siguiente forma:

a) Términos no preferentes o reenvíos: En este índice se indica, para cada término no preferente, el descriptor al que ha sido reenviado o por el que ha sido absorbido.

Ejemplo de inclusión de términos no preferentes en el Índice conceptual

Delitos sexuales => Delitos contra la libertad sexual

b) Términos preferentes o descriptores: En este índice se pone de manifiesto, para cada uno de los descriptores del tesoro, los siguientes aspectos:

Clasificación

Precedido de las Siglas "CL" se indica, la clasificación temática del descriptor.

Términos reenviados

Precedidos de la sigla "=", se reseñan los términos no preferentes englobados en el término principal.

Términos genéricos

Precedidos de la sigla "<" se reseñan los términos genéricos del término de cabecera. En la delimitación conceptual de cada descriptor se incluyen, según los casos, hasta tres niveles jerárquicos superiores.

De esta forma se facilita la detección de su ubicación temática particularizada, permitiendo determinar la rama arborescente de la que desciende.

Términos específicos

De similar forma a la establecida respecto a los términos genéricos se indican, para cada descriptor, precedidos por la sigla ">" aquellos otros descriptores integrados en los tres primeros niveles de especificidad del término de cabecera.

Esta delimitación de los conceptos subsumidos en el descriptor referenciado se efectúa manteniendo el desarrollo jerárquico de los diferentes subniveles e

incorporando, a cada uno de ellos, los descriptores integrados en el mismo, todo ello según un orden alfabético.

Este grado de desarrollo, según los casos, de hasta tres subniveles permite establecer, con una gran precisión, la cobertura temática del descriptor de referencia y los conceptos específicos más afines englobados en el mismo.

Términos relacionados

Por último, en el desarrollo conceptual de cada descriptor, se hace constar aquellos otros descriptores asociados con él reseñando esta circunstancia mediante la sigla "-", lo que permite mostrar vínculos asociativos establecidos entre diferentes subcampos temáticos.

Ejemplo de Índice conceptual de términos preferentes

Delitos contra la libertad sexual
cl Ik1
na Se entienden por tales a los regulados en el Título VIII
del Libro II del Código Penal
= Delitos contra la honestidad
= Delitos contra la indemnidad sexual
= Delitos sexuales
< Delitos
< < Derecho Penal
< < Derecho
> Abusos sexuales
> Estupro
> Incesto
> Acoso sexual
> Agresiones sexuales
> Violación
> Escándalo público
> Exhibicionismo
> Prostitución
> Proxenetismo
> Trata de blancas
> Provocación sexual
> Corrupción de menores
> Pornografía
> Difusión de material pornográfico
- Familiares
- Lugar de la acción delictiva
- Menores
- Perdón del ofendido
- Responsabilidad penal

4.- Automatización de tratamientos de indización y tesauros

Los primeros estudios sobre las posibilidades de la informática para el tratamiento de la información documental, y especialmente para su utilización como herramienta en los procesos de indización automática, podemos fijarlos en los comienzos de los años cincuenta, como consecuencia de la producción de específicas aplicaciones informáticas elaboradas con la finalidad de coadyuvar a los procesos de clasificación y ordenación de la información científica.

Los primeros procesos se basaron en análisis estadísticos de frecuencias de aparición, en los documentos, de unitérminos, estableciendo G.K. Zipf (1949) una fórmula de aplicación de este criterio basada en los estudios estadísticos sobre el *Ulysses* de Joyce, línea seguida asimismo por Hans Peter Luhn (1957), y posteriormente por Damerou (1965), Carrol y Roeloffs (1969), Rosemberg (1971), Sparck Jones (1972), Salton y Yang (1973), Bookstein y Swanson (1975) y Robertson y Harding (1984).

Los estudios realizados fueron evolucionando progresivamente, dando lugar a diferentes propuestas de sistemas para la indización automática, como: El Sistema Smart (1973) que, junto a los cálculos estadísticos, incorporaba procesamientos lingüísticos basados en la forma de las palabras y estructura de las oraciones; El sistema de Faraj (1996) basado en análisis sintáctico de multitérminos; Los sistemas elaborados como fruto de los trabajos de automatización de la indización y traducción automática, basados en estudios semánticos y sintácticos, realizados en el CINDOC (1983); Y los procesos de localización de estructuras sintagmáticas obtenidas de análisis morfológicos de textos, previamente segmentados en frases, realizados por Simón Grand a y Lema Garzón (1990).

En cuanto a la utilización de tesauros, como elemento integrante del sistema automatizado, o para su confección automática, podemos citar los trabajos realizados por Martínez, Lucey y Linder (1987), Biebricher, et al. (1988), Lovtsov (1990), Coret et al. (1991), Silvester et al. (1994), y por el equipo de creación automática de tesauros de la Universidad Carlos III formado por Llorens, et al. (1994).

Como consecuencia de dichos trabajos se han desarrollados sistemas de indización automática tales como el Spirit, sustentado por fundamentos lingüísticos y apoyado en un corpus de 500.000 palabras, y sistemas como el Sapphire, Passat y Clarit, que integran tesauros en su esquema funcional.

A efectos de, al menos, entrever los aspectos a tener en cuenta para el establecimiento y progresión de los trabajos de investigación en marcha, podemos reseñar las líneas de actuación desarrolladas, como base de los sistemas de ayuda a la indización (o indización automática), por el equipo del CINDOC dirigido por Antonio Valle Bracero y utilizado para el análisis documental automatizado de textos normativos y jurisprudenciales.

- a) Selección de fuentes terminológicas y preparación del tratamiento automatizado.*

La base terminológica de los trabajos se sustentó en la preparación de un corpus terminológico amplio, para cada cobertura temática, que permitiera cubrir, en la forma más exhaustiva posible, la materia.

Fue para ello de vital importancia la selección de las fuentes documentales sobre las que, mediante un proceso semasiológico, se obtendrían los términos comprensivos de los conceptos que podrían apriorísticamente quedar reflejados en los textos jurídicos sobre los que se iba a realizar la indización, y por ende que debían ser detectados por la aplicación informática.

Dichas fuentes estaban constituidas tanto por textos normativos y jurisprudenciales que incidía sobre los temas de aplicación, como por cualesquiera otros productos terminológicos o documentales (diccionarios, glosarios, manuales jurídicos...), editados en forma impresa o electrónica, que pudieran servir de referencia para la elaboración de un corpus jurídico-documental omnicomprendido de la materia.

En forma paralela, y a tenor de la progresiva conformación del corpus de referencia, se preparó la metodología informática que permitía, en base a los términos preseleccionados y a la forma del contenido literal de los documentos a analizar, establecer los desarrollos de aplicaciones informáticas necesarias para cubrir los objetivos propuestos.

b) Desarrollo de los tratamientos informáticos para la codificación de términos de indización y la localización automática de plurales y sinónimos.

Una vez definida la metodología informática a seguir y en consonancia con la progresiva conformación del corpus terminográfico de referencia se desarrollaron los programas informáticos que permitían codificar los términos incluidos en los corpora de referencia, en una forma tal que quedaban excluidos de dicha codificación las palabras vacías, infijos, sufijos y demás formantes, que aportaban dígitos irrelevantes para el posterior tratamiento de los términos, y se confeccionaron útiles de validación y comprobación de la adecuación de los términos seleccionados a la finalidad perseguida.

De forma paralela se desarrolló otra aplicación informática que codificaba, en igual sentido al del tratamiento del corpus jurídico-documental de referencia, los textos sobre los que se tenía previsto realizar el proceso de indización automática o de ayuda a la indización, a fin de que estos adoptasen una forma idónea, para su cruce informático con los términos que conforman el corpus de referencia.

La detección, en el texto original codificado, de términos entresacados del corpus de referencia jurídico-documental, permitía asimismo el marcado de los párrafos en los que se detectaban los presuntos términos de indización, a fin de que, en una posterior supervisión documental, pudiera contrastarse la validez de los mismos y asimismo comprobarse que las partes no marcadas carecían de conceptos jurídicos que pudieran dar lugar a su representación por nuevos términos que debían ser incorporados al corpus jurídico-documental de referencia.

Sobre los dos corpora (de textos y términos) previamente codificados se realizaron una serie de procesos informáticos, con aplicaciones específicas desarrolladas para ello, que permitían establecer puntos de equivalencia entre descriptores obtenidos del corpus jurídico-documental y términos equivalentes incluidos como lenguaje del texto documentado. Esas equivalencias, que eludían las formas gramaticales (género y número) y reenviaban los sinónimos a los términos seleccionados como principales (descriptores), servían por un lado de elementos de indexación y anclaje inter-documentos, y por otro de fundamento para la evaluación del contenido de cada documento, permitiendo ordenar los mismos según criterios de análisis de reiteración de conceptos.

Asimismo el tratamiento, mediante la aplicación informática, de cada uno de los párrafos incluidos en el texto de los documentos, entresacando las palabras significativas del mismo que coincidiesen con algunas de las asimismo obtenidas de la totalidad de términos del corpus jurídico-documental, permite entrever la posible aparición de nuevos descriptores.

c) Estructuración de la información, y determinación de sinónimos y relaciones de homonimia

Sobre el corpus jurídico-documental de referencia se desarrolló por último una labor documental de ordenación y estructuración en forma de tesoro que permitía, para los términos principales o descriptores, determinar los términos genéricos y específicos correspondientes a su estructuración jerárquica, los términos relacionados con el mismo ubicados en diferentes ramas del árbol temático, y los términos secundarios o no descriptores, que se reenvían a un descriptor, por hacer referencia al mismo concepto.

Una vez elaborados los tesoros correspondientes a los campos de referencia se validaron los mismos con los resultados obtenidos de los procesos informáticos de codificación y contrastación con los textos documentados ampliándose y modificándose, cuando así se considero necesario, aquellos términos que se observó dieron lugar a ambigüedades, ruidos o lagunas informativas en relación con los conceptos a que remitían.

La elaboración y progresiva mejora del tesoro permitió, en forma complementaria, mejorar las posibilidades de tratamiento, gestión y recuperación de la información almacenada en bases de datos, coadyuvó a la normalización de la terminología relacionada con la materia, y facilitó la estructuración y presentación por materias del contenido de los documentos a los que se sometió al tratamiento automatizado.

Es previsible que, de continuarse estos trabajos en el futuro, puedan simplificarse las tareas de tratamiento y recuperación de la información jurídica almacenada en bases de datos, en una forma más precisa y amigable, facilitando a su vez la elaboración de elementos y productos de valor añadido aptos para su utilización en la selección y ordenación de la información obtenida de las bases de datos que incorporen esta, u otras similares, metodologías informáticas.

Bibliografía

- Aitchison, J.; Gilchrist, A.: Thesaurus construction. A practical manual. London. Aslib. 1987
- Alvaro, Concepción: Representación del lenguaje jurídico a través de un lenguaje documental: el tesoro jurídico. III Encuentro sobre la Informática en las facultades de Derecho. Universidad Pontificia Comillas. Madrid, mayo de 1989. Págs. 63-77.
- Arzt, R. y Picht, H.: Einführung in die Terminologiarbeit (1991).
- Cabré, M. T.: La terminología. Teoría, metodología, aplicaciones. 1993.
- Ciampi, C. et al.: Thes/BID: Un thesaurus per l'informatica giuridica e il diritto dell'informatica, radato con l'ausilio dell'elaboratore elettronico. III Congreso Internacional sobre Informática jurídica y la comunidad nacional e internacional. Roma 9-14 de mayo de 1983.
- Cohen, J.: Highlights: Language and domain-independent automatic indexing terms for abstracting. Journal of the American Society for Information Science. Vol 46 (3), 1995.
- Currás, Emilia: Tesauros. Manual de construcción y uso. Kaher II, S.A. Madrid, 1998.
- Felber, H.: Terminology Manual. París, UNESCO, Infoterm, 1984. Págs. 426.
- Felber, H. y Picht, H.: Métodos de terminografía y principios de investigación terminológica. CSIC. Madrid, 1984. Págs. 254.
- Fernández Martín, Claudia. Sistema SEIK. Sistema de Elaboración de Índices Kwic/Kwoc. ICYT. Madrid, 1989.
- Gil Leiva, Isidoro: La automatización de la indización. Propuesta teórico-metodológica: aplicación al área de biblioteconomía y documentación. Tesis Doctoral. Universidad de Murcia. 1997.
- Headings for tomorrow: public access display of subject headings. Chicago. American Library Association. 1992.
- INFOTERM: Selected Readings in Terminology, I Terminologi-wissenschaft. 1990.
- Inventaire des travaux de terminologie recents. Québec OLF, 1994
- Irazabal, A.; Álvarez, S.; Zarco, J.; y Abejón, T.: Curso de Introducción a la terminología. ICYT (CSIC). Madrid, 1992. Págs. 180.

- Jones, S.; Gatford, M.; Robertson, S.; Hancock-Beaulieu, M.; Secker, J.: Interactive thesaurus navigation: intelligence rules ok?. of the American Society for Journal Information Science. 1995, 46 (1):52-59
- Laguna Serrano, E, Irazazabal Nerpel, A. y Valle Bracero, A.: Confección automática de tesauros. Revista Española de Documentación Científica. Vol. 12, Núm. 2. Págs. 129-140. 1989.
- Laguna Serrano, Emilio; Valle Bracero, Antonio; y Valle Bracero, Adolfo. Constructor de Tesauros CAT. Confección Automática de Tesauros. CINDOC, Madrid, 1992.
- Lancaster, F.W.: Indexing and abstracting in theory and practice. London. Library Association. 1991.
- López Alonso, Miguel Ángel: Normalización de la terminología jurídica y su aplicación a la construcción de un tesoro en Derecho comercial. Tesis Doctoral. Universidad Carlos III de Madrid. 1997.
- Llorens, Juan: Definición de una metodología y una estructura de repositorio orientadas a la reutilización: el Tesoro de software. Tesis Doctoral. Universidad Carlos III de Madrid. 1996.
- Melby, A., Budin, G. y Wright, S. E.: Terminology interchange format (TIF). TermNet News. 40-1993. págs. 3-64.
- Muñoz, A.: Redes neuronales para la organización automática de la información en bases documentales. Tesis Doctoral. Universidad de Salamanca. 1994.
- Norma UNE 50-106-90 (equivalente a la ISO 2788-1986) de Directrices para el establecimiento y desarrollo de tesauros monolingües. Aenor, 1990.
- Norma UNE 50-121-91 de Métodos para el análisis de documentos, determinación de su contenido y selección de términos de indización. Aenor. 1991
- Norma UNE 50-125 (equivalente a la ISO 5964-1985) de Directrices para la creación y desarrollo de tesauros multilingües. Aenor.
- Rey, A.: La Terminologie. noms et notions. PUF, Coll. "Que sais-je". N° 1780 (1979). Págs. 127.
- Rondeau, G.: Introduction à la terminologie. Centre Educatif et Culturel Inc. Montreal, 1981.
- Rowley, J.: The controlled versus natural indexing languages debate revisited: a perspective on information retrieval practice and research. Journal of Information Science. 1994 20(2):108-119
- Sager J.C.: A practical course in terminology processing. 1990.

- Salton, G.: Automatic text processing: the transformation, analysis, and retrieval of information by computer. Addison-Wesley. 1989.
- Shonfeldt, R.: Matematische eigenschaften fur thesaurusrelationen. Nachrichten fuer Dokumentation. 1994 45 (4):203-212
- Thiry, Bernard: La Terminología bilingüe (español-frances y francés-español) de la responsabilidad civil extracontractual en los derechos español y belga. Análisis crítico y contribución a su tratamiento terminográfico y a la teoría de la terminología. Tesis doctoral presentada en Louvain-la-Neuve (Bélgica), el 1 de junio del 2001.
- Valle Bracero, A. y Fernández García, J.A.: Automatización de la indización y coordinación de descriptores. Revista Española de Documentación Científica. 6, 1 (1983), págs. 9-16.
- Valle Bracero, A.; Fernández, J.A.; y Morales Fernández, R.: Separación automática de lexemas, sufijos y morfemas y su aplicación a la traducción automática. Revista Española de Documentación Científica. 7, 3 (1984), págs. 185-192.
- Valle Bracero, Antonio; Rey Guerrero, Alfredo del; y Gutiérrez Muñoz, Francisco. Gestor de Recuperación SERIOMIC. Sistema Evolucionado de Recuperación de Información por Ordenador, versión Microordenador. CINDOC, Madrid, 1994.
- Van Slype, G,: Les languages d'indexation: Conception, construction et utilisation dans les systèmes documentaires. Les éditions d'organisation, Paris, 1991.
- Velasco, M.: Generación automática de representaciones de dominios. Tesis Doctoral. Universidad Politécnica de Madrid. 1998.
- Velasco, M.; Díaz, I.; Lloréns, J.; Amescua, A. de; y Martínez, V.: Algoritmo de filtrado multi-término para la obtención de relaciones jerárquicas en la construcción automática de un tesoro de descriptores. Documentación científica, vol. 22, nº 1, enero-marzo de 1999, págs. 34 a 49.
- Wüster, E.: Les classifications de notions et de thèmes. Différences essentielles et applications. Infoterm. (1971), 2-71.
- Wüster, E.: Introducción a la teoría general de la terminología y la terminografía (traducción española). universidad pompeu Fabra, Institut Universitari de Lingüística Aplicada. Barcelona, 1988 Págs. 227.