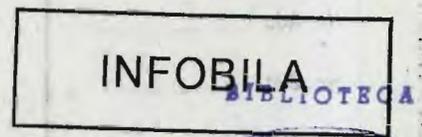


BASES DE DATOS JURIDICAS Y TESAUROS DE DERECHO



CENTRO UNIVERSITARIO DE INVESTIGACIONES BIBLIOTECOLOGICAS

Señoras, señoritas, señores :

Es para mí un alto honor el que se me haya formulado esta invitación de parte de una institución tan prestigiosa como el Instituto de Estudios Peruanos, permitiéndome así dirigir la palabra a un auditorio tan selecto como el presente.

Ante todo debo pedir disculpas, puesto que mis palabras van a constituir no una conferencia, sino tan solo una charla sobre mis experiencias en cuanto a bases de datos jurídicas y una de sus herramientas más importantes, como lo es el tesoro de derecho.

Siendo Jefe de las Bases de Datos de la Junta del Acuerdo de Cartagena y habiendo desarrollado exitosamente la base bibliográfica que prácticamente constituía la memoria de dicho Acuerdo Subregional y a la cual se denominó SIRA (Sistema de Información Bibliográfica Andina) aunque más conocida como JUNINDEX a causa de la publicación a la que dió origen, surgió otro reto, esta vez más complicado;

no se trataba de algo conocido, sino de una base de caracter jurídico y más aun de ámbito subregional en la que se analizarían, uno a uno, todos los instrumentos del Acuerdo de Cartagena, su tratado de creación, el del Tribunal de Justicia, que implicaba hasta sus sentencias, los reglamentos, las decisiones, que en los Países Miembros tienen fuerza de ley y las resoluciones de la Junta que permitían la ejecución de los proyectos. Se trataba pues de una base ^{legal} al servicio de cinco países, que fue denominada LEXA (Legislación Comunitaria Andina)

El proyecto contó con el apoyo económico de la Comunidad Europea y para la implementación informática debía usarse la de la Junta, es decir una computadora IBM "main frame" y el software correspondiente para estos casos, vale decir el SPAIRS.

En lo referente a la política a seguir, se decidió que la unidad de nivel llegaría hasta el inciso e incluso la división en párrafos de los artículos que así lo requirieran. Por otro lado, siendo una base de caracter jurídico, no era suficiente el conocimiento de las técnicas documentarias, ni la experiencia adquirida a través de la constante lectura de los documentos editados por la Junta, se requería también del apoyo jurídico el que encontramos en el abogado que se desempeñaba como coordinador con las entidades internacionales que apoyaban a la Junta y que además de la

2
experiencia en la materia, tenía sumo interés por las tareas documentales.

LEXA (Legislación Comunitaria Andina) si bien tenía carácter jurídico, no presentaba ninguna relación con otras bases de ese tipo, se trataba de una base con características muy especiales, muy distinta de las demás bases de datos jurídicas usualmente dedicadas a la legislación, la doctrina o la jurisprudencia, LEXA debía servir a una comunidad heterogénea de los cinco Países Miembros y representada por un organismo de carácter subregional, por lo tanto su contexto estaba constituido por un ordenamiento jurídico propio.

Las indagaciones que se hicieron por América y Europa para tratar de encontrar algo similar dieron una sola respuesta - CELEX, la base legal de la Comunidad Europea. Se había pues, encontrado el modelo que permitiría componer y enlazar los campos de la hoja de entrada que serviría para la indexación al mismo tiempo que iniciar la elaboración del correspondiente manual de uso, tarea ésta, por cierto bastante ardua, especialmente la parte referente al número de acceso al documento, el cual de acuerdo a la política adoptada, debía llegar a nivel de inciso o párrafo. Los otros campos que requirieron investigación fueron los de firma y vigencia de los tratados, los de modificaciones y concordancias no solo internas dentro de los propios instrumentos sino las externas entre los distintos tipos de documentos.

Contando con una hoja de entrada y un manual tentativo, viajé a Bruselas con el fin de realizar las reuniones de consulta con el Servicio de Documentación de la Comunidad. Se trataba además de conocer lo más ampliamente posible lo relacionado con el tesoro EUROVOC usado en CELEX, puesto que para LEXA tampoco teníamos un tesoro adecuado, ya que el de integración elaborado para SIDA carecía de la terminología jurídica requerida, por lo tanto era necesaria su reestructuración. Fue así como se pudo iniciar el análisis y digitación de unos 9,000 registros con que contaba LEXA al momento de dejar la Junta.

Un par de años después de haber salido de esta institución, tiempo que ocupé realizando asesorías en el exterior, fui llamada por la Facultad de Derecho de la Universidad Femenina del Sagrado Corazón (UNIFE) para el dictado de Informática Jurídica que se consideraba importante que fuera llevado por las alumnas, debido al auge de la informática en las tareas legales.

Como es de suponer el curso tenía un carácter *sui generis*, pues yo sólo podía

ocuparme de la parte documentaria no había la correspondiente a derecho, por lo tanto el curso debió contar con dos profesores. Admitida esta característica excepcional, debimos planearlo, estableciendo su correspondiente sillabus, pues sería la primera vez que se dictaría tal materia en la universidad. El curso tenía como finalidad que las alumnas al dejar la universidad estuvieran capacitadas para crear sus propias bases de datos de gestión, valiéndose de la ayuda de la computadora y así aligerar y mejorar su trabajo en sus respectivos bufetes.

El plan fue concebido en dos grandes lineamientos, una parte teórica que comprendía una introducción a la informática jurídica en sus diversos aspectos y un taller en el que se combinaran, los conocimientos de ^{de} derecho, los de informática y la creación de bases de datos práctica en la que fuimos auxiliados por el centro de cómputo de la universidad, se cumplía pues así la triple acción de diferentes carreras con el objetivo de la creación de bases de datos.

El contrato comprendía además varias horas de investigación que debían emplearse en el proyecto de creación de una base de ^{datos} ~~estrutina~~ de derecho civil, encaminada a servir de sustento a los cursos que se dictan sobre esa materia, lo cual resultaba muy beneficioso desde el punto de vista docente. Sin embargo, un proyecto de esa envergadura no podía ser aceptado sin un estudio previo, por lo menos a nivel de Lima, sobre la existencia de otras bases del mismo tipo. La investigación dio como resultado la inexistencia de una base de datos de derecho civil, aunque sí de algunas de legislación.

La DOCCIV, tal como la bautizamos, tiene como parámetros, que es de carácter referencial y bibliográfica; que los autores analizados ~~peruanos~~ deben ser peruanos, que la unidad de registro es el capítulo de un libro, de preferencia los indicados en los distintos sillabus de los cursos de derecho civil de la universidad - o artículos aparecidos en revistas acreditadas de derecho, en especial aquellas a las que la Biblioteca de la Universidad está suscrita, igualmente se considerarían las ponencias presentadas a congresos de derecho civil. En consecuencia, la base tiene un carácter analítico, con lo cual se complementa en profundidad la base formada por los libros y revistas de derecho civil que forman el acervo de la Biblioteca.

La base cuenta como "hardware" una computadora personal y como "software" el CD/microIsis. La hoja de entrada ha sido elaborada tratando de hacerla lo más compatible

posible con la usada por la Biblioteca, contando naturalmente con otros campos como los referentes a concordancias internas y externas, distintas exigencias en lo relacionado a los resúmenes, etc. Como es lógico la hoja requería del correspondiente manual de uso, el cual debió elaborarse con los detalles suficientes para que las alumnas que conforman el taller del curso, puedan llenar las hojas sin mayor dificultad, como tampoco en los momentos de la digitación para lo cual también debieron ser preparadas especialmente, así como para la debida recuperación.

Posteriormente, la Facultad vió la necesidad de implementar otras tareas de informática jurídica, es decir la creación, por las alumnas de pequeñas bases de datos jurídicas, como una de legislación, mediante el análisis del articulado de la parte del Código Civil Peruano de 1984, correspondiente al Acto Jurídico. Naturalmente esto hizo necesaria la creación de otra hoja de entrada que reuniera los campos necesarios para tal análisis y su correspondiente manual de uso. Para el presente año está en vías de implementarse otra pequeña base ~~taxi~~ sobre puntos referentes a la nueva constitución.

Como en todas las hojas de entrada era imprescindible contar con el campo relativo a los descriptores, con el fin de asignarlos tanto en la indexación como utilizarlos para la recuperación, nos vimos en la ineludible necesidad de contar con un tesoro de derecho. Como es de todos conocido que antes de pensar en la elaboración de un tesoro es preciso comprobar la existencia de algun tesoro o vocabulario controlado del área requerida, no podíamos prescindir de tal norma, naturalmente enfocando la búsqueda hacia los lugares en los que el derecho usado fuera derivado del derecho romano y no del "common law" y sobre todo en español.

Como era lógico la búsqueda se inició por el Perú teniendo como meta el hallar un macrotesoro de derecho y no uno de alguna de sus ramas, lo cual vendría con posterioridad. Aquí solamente encontramos un vocabulario no controlado que usa la base de legislación peruana INFOLEX, de caracter comercial, que contiene toda la legislación a partir de 1985 y que actúa mediante contratos con entidades tales como bancos, etc. Dicho vocabulario, al momento de nuestra investigación, comprendía unos 3000 términos, los que la propia entidad admitía eran ya causa de dificultades en la recuperación a causa del "ruido informático" que generan, lo cual es lógico pues se incluye todos los términos sin relaciones de ninguna índole.

Muestra búsqueda debió extenderse por Sudamérica primero, fué así como supimos que en Argentina, el Sistema de Información Jurídica (SAIJ) *tiene una traducción al español del tesoro de derecho italiano* que los fuera obsequiado por la Corte de Casación de Roma y que corresponde al "software" Italgire. *Los intentos de utilizarlo no han dado los resultados esperados.*

De Chile, tampoco obtuvimos respuesta satisfactoria. En Colombia, logramos saber que existen tesauros para el servicio de determinadas bases de datos jurídicas, aunque en realidad no son propiamente tesauros, sino más bien vocabularios controlados o listas sistemáticas pues carecen de estructura, a la vez que en muchos casos los términos no son tratados como descriptores sino como epígrafes, se diría que tratan de asimilar el formato usado a veces por la Biblioteca del Congreso de Washington en su SUBJECT HEADINGS. *En la mayoría de los casos los proyectos colombianos fueron elaborados por abogados casi sin apoyo de documentalistas.*

El Brasil, poseedor de la base de datos *discussional* más antigua de América Latina, llamada PROBASEN y elaborada inicialmente para satisfacer las necesidades del congreso de la república, cuenta con un vocabulario controlado en el que se encuentran muchos términos propios de derecho y que presenta relaciones de equivalencia y algunas jerárquicas; pero también carece de estructura y tan solo existe una parte alfabética, muy valuminosa por cierto, que se actualiza con relativa frecuencia y un índice permitido.

Con mucha ilusión, nos dirigimos a la búsqueda en Costa Rica, donde conocíamos de la existencia de la entidad conocida como ILANUD, que sí cuenta con un verdadero tesoro, pero dedicado a derecho penal y criminología. Por otro lado, el macrotesauro de derecho usado en la red de bibliotecas jurídicas y el micro tesoro, derivado de aquél, sobre derecho civil, ~~carece~~ carecen de estructura y tan solo tienen una parte alfabética en la que aparecen los términos con jerarquías mínimas y en los que rara vez aparecen términos relacionados.

En México existe el sistema *de información jurídica* llamado UNAM-Jure, basado en la forma usada por la Universidad de Montpellier de Francia, cuyo funcionamiento es bastante diferente del que pensábamos adoptar.

Al investigar por Europa, conseguimos el tesoro llamado THES/BID que es un verdadero tesoro elaborado por el Instituto de Informática Jurídica de Florencia; pero enfocado a la relación derecho-informática. El tesoro fué desarrollado por medio de un programa especial ideado por el propio instituto.

Para nosotros ya era conocido el EUROVOC, ideado para la Comunidad Europea el cual cuenta entre sus diferentes campos llamados microtesoros, con una buena parte relativa al derecho en sus diversas ramas.

En España, nos encontramos con un macrotesoro de derecho, aunque aún en elaboración. Fué iniciado por investigadores del Instituto de Ciencias Sociales (ISOC) del Consejo Superior de Investigaciones Científicas, que en la actualidad ha sido fusionado con el Instituto de Investigaciones Científicas y Técnicas del mismo Consejo, razón por la cual el tesoro se continúa trabajando en dicho instituto. Gracias a mi vinculación con el Consejo, del que una vez también formé parte, pude conocer el desarrollo de este macrotesoro, que en su primera parte, referente a las ramas del derecho, tiene una buena estructura, *cuyo desarrollo es confuso y asistemático.*

Respecto a los microtesoros.

Terminada la investigación creímos llegado el momento de poner manos a la obra. *Como* auxilio, contábamos con los tesoros descritos anteriormente además de otros de carácter general. Examinando estos últimos en su parte relativa al derecho, la cual era *más* o menos amplia según los casos, observamos que el de la Unesco, me refiero a la edición anterior a la última en circulación, que en la parte dedicada a las ciencias sociales, ~~el mismo~~ presentaba un esquema muy interesante que nos *dió* la idea de cómo podíamos estructurar nuestro tesoro, pues solucionaba algunos de los problemas que se nos presentaban. En principio, como Uds. saben, dicho tesoro tiene una nomenclatura alfanumérica lo cual permite acortar los números excesivamente largos y por lo tanto confusos y sobre todo cuenta con una característica especial de relación entre fascetas, mediante el uso de asteriscos, lo cual en derecho permite grandes ventajas para establecer las relaciones entre las diversas ramas.

Encontrado el formato que debía seguir la sección sistemática y de acuerdo a las normas de elaboración de tesauros monolingües de la Unesco y la ISO, empezamos la recopilación de términos, basándonos en las principales fuentes del derecho, comparándolos con la terminología jurídica peruana con el fin de determinar los descriptores y los no descriptores. Para esta tarea contábamos con *el acervo* de la Biblioteca de la UNIFE, ^{en} cuyas obras de referencia, ~~contábamos~~ ^{contamos} todo el apoyo necesario, además de la buena colección de códigos, tratados y legislaciones.

Como habíamos decidido que lo mejor era combinar tanto el método inductivo como el deductivo, ~~optamos~~ ^{optamos} también por una minuciosa revisión de los términos jurídicos que aparecen en los dispositivos legales que se publican en la separata de Normas Legales del Diario Oficial EL Peruano, contando así con el complemento proporcionado por la práctica.

Decidimos llamar a nuestro tesoro, *Macrotesauro de Derecho: JURIVOX*. Dicho tesoro cuenta con una ^{sección} sistemática, otra alfabética y un índice permutado ~~de~~ ^{de} con elaboración.

La sección sistemática comprende tres partes. La primera se refiere a las ramas tradicionales del derecho que han sido ubicadas mediante el uso de las letras que van de la B a la P, habiéndose dejado la letra A para la parte general. La letra Q corresponde a la segunda parte y ha sido asignada a los Derechos Humanos, que como Uds. saben, han sido, en gran parte, incluidos en el Título I de nuestra Constitución actual y por lo tanto están estrechamente vinculados al derecho peruano. Las letras de la R a la Z forman la tercera parte y deben cubrir las áreas reguladas jurídicamente, es decir, aquellas que no constituyen un derecho específico, pero sobre las cuales existe legislación.

han requerido,
 La mayoría de estas áreas ^{han requerido,} por su extensión, de un agrupamiento dentro de la misma letra, por ejemplo bajo la letra ^R fue necesario reunir la Economía y el Desarrollo, lo cual se hizo teniendo en cuenta su obvia vinculación. Lo mismo ocurrió con el caso del Medio Ambiente y los Recursos Naturales, asignados ambos a la letra S. Por otro lado, la Banca, las Finanzas y la Inversión, se agruparon bajo la letra U. Sin embargo y debido a lo amplio de su contenido, a la Educación, la Comunicación y la Información, se les asignó respectivamente las letras, W; Y; y Z.

La Agricultura y la Alimentación; el Comercio y la Industria; y el Transporte fueron consignados bajo la letra T, debiendo ser separados entre sí por la aplicación de una numeración distintiva.

Con la letra V ocurrió algo similar pues comprende la serie temática correspondiente a la Demografía y los Asuntos Sociales; el Bienestar Social y el Medio Humano; así como la Administración, el Trabajo y el Empleo, los cuales también debieron separarse en forma numérica. Por otro lado la Ciencia y la Tecnología; la Cultura y las Humanidades, fueron agrupadas bajo la letra X recurriéndose también a la separación numérica.

El esbozo expuesto no es otra cosa que el sumario de las facetas en las cuales se ha distribuido el primer nivel del macrotesauro. Por otro lado, es conveniente hacer notar que si se observa el desglose de la sección A correspondiente a la parte general del derecho, se verá que las subdivisiones han sido elaboradas recurriendo también a numeraciones que especifican los temas.

En el tercer nivel de esta misma sección A, se manifiesta con mayor claridad las características especiales del tesauro, tal como sucede con el aparente descriptor, "derecho en general", que lleva entre paréntesis la advertencia: "No debe usarse como descriptor", es decir que dicho término no es otra cosa que un clasificador, que naturalmente no puede ser un descriptor ya que por su amplitud y vaguedad, su uso sería totalmente inútil, máxime existiendo el término mucho más concreto de "Derecho". Sin embargo a pesar de su injusticable practicidad, no se le puede considerar así en cuanto a la comprensión de la jerarquía de los términos que incluye. Esta característica también fue tomada del tesauro de la Unesco.

Otra metodología también debida al tesauro mencionado ha sido el uso de asteriscos mediante los cuales se obtiene la relación con distintas facetas del tesauro, tal como sucede con el descriptor "Legislación" y las diferentes legislaciones propias de las áreas reguladas jurídicamente. Asimismo se aplica en el caso de la faceta correspondiente al "Acto jurídico" que evidentemente está relacionada con la parte correspondiente a "derecho civil" ambos situados en el esquema en letras diferentes y por lo tanto la inexistencia de dicho asterístico en ambos términos pudiera llevar al usuario a conclusiones inexactas sobre la no consideración de uno de los términos. Como es lógico, estas relaciones por medio de asteriscos, son como las de todo tesauro de carácter recíproco.

En otro orden de cosas, se debe hacer notar que la estructura del tesoro ha requerido asimismo, como sucede con otros tesauros, en su presentación escrita, el uso de distintos tipos de letra con el fin de resaltar las jerarquías y demás circunstancias de índole similar.

Con respecto a la presentación alfabética, de la que se muestra a quí solo *una de sus páginas* ha de observarse que se ha colocado *debajo* de cada descriptor el correspondiente *etiquetado* que le es propio en el esquema estructural con la finalidad de llevar al usuario a efectuar una consulta más detallada de esta parte proporcionándole así una mejor visión del íntegro de la faceta en la cual se ubica el término objeto de su interés.

En lo referente al "software" utilizado para el desarrollo del tesoro, se ha recurrido a la parte del CD/MicroIisis llamada "Thes", trabajada en Pascal. Por razones técnicas del programa, ningún descriptor puede exceder de los 30 caracteres, lo cual para un tesoro de derecho resulta absolutamente limitante, pues es conocido de todos que la terminología usada en dicha rama del conocimiento sobrepasa, la mayoría de las veces los impuestos 30 caracteres. Asimismo los campos de que se dispone en el "Thes" no fueron considerados suficientes, pues no existía uno correspondiente al número del descriptor y también carecía de la nota de alcance. Sin embargo todo ello fue resuelto con la ayuda de un informático conocer también de las necesidades de la documentación, quien adecuó dicho programa en forma conveniente.

En cuanto a los exponentes usados para indicar las distintas relaciones entre los términos y de acuerdo a las normas actuales, distribuidas para América Latina por la CEPAL, se prefirió el uso de los símbolos en inglés por ser además, de uso más común no sólo en el Perú sino en otros países y con el fin de dar al tesoro una mayor compatibilidad.

En relación al estado actual del tesoro estoy en posición de decir que se encuentran terminadas varias de sus partes tales como la letra F correspondiente al Derecho Civil, cuyo desglose ha sido llevado hasta un cuarto nivel, lo cual hace de dicha sección algo bastante cercano a un microtesoro. Tal decisión fue tomada por ser esta rama del derecho la que *se aplica a casi el 50% de las normas jurídicas relativas al derecho privado, además es la que recibe mayor énfasis en la enseñanza universitaria* y por la necesidad de usar esta parte del macrotesoro para el desarrollo del taller de Informática Jurídica que se realiza con bases de datos referentes a tal derecho.

Asímismo se hayan concluídas las partes correspondientes a las siguientes letras: A, Derecho general y Acto jurídico; B, Derecho Administrativo; D, Derecho Agrario; H, Derecho Constitucional; K, Derecho Internacional Privado; y Q, Derechos Humanos.

Además de lo expuesto es conveniente resaltar que la propiedad de los terminos incluídos como descriptores en el tesauo se evalúa constantemente en los trabajos del Taller de Informática Jurídica.

En otro aspecto y con la finalidad que nuestro trabajo pueda ser de mayor utilidad para otros países de habla española, es nuestra idea tratar de agregar a cada descripto los localismos correspondientes usados sobre todo en España, México y Colombia, países con los que se han tenido conversaciones al respecto. Esto se hará colocando al lado del término aceptado en el otro país, la sigla de acuerdo a la norma ISO para países, señala el nombre del país al cual se adjudica el término.

Asímismo, la Facultad de Derecho de la UNIFE, nos solicitó la elaboración de una base de datos de legislación peruana cuyo contenido está constituido por los dispositivos legales publicados en el fascículo Normas Legales del Diario Oficial El Peruano a partir del 1 de enero de 1996. En dicha base se usa el Jurivox, además de los suplementos consistentes en un vocabulario controlado de aquellas palabras que son indispensables para la descripción del contenido de una determinada disposición; una lista de identificadores, tanto en siglas como desglosadas de entidades del sector público y de instituciones internacionales; y otra relativa a una nomenclatura geográfica, las cuales se elaborarán de acuerdo al avance de los registros de la base. La base proyectada se completará retrospectivamente de manera paulatina. Para todo ello se contará con la Biblioteca de la Universidad, el Taller de Informática Jurídica y el Centro de Computo de la universidad.

Con lo expresado creo haber relatado mis experiencias en cuanto a la informática jurídica desde el punto de vista docente y de investigación.

Muchas gracias

Lima, 29 de marzo de 1996.

había para el Consensatorio CENDOS (Colectivo de Centros de Documentación) en el local del IEP (Instituto de Estudios Peruanos)