

# Metodología para la Creación de Bancos de Datos Personalizados\*

MARTINE FABRE

*Ingeniero de estudios del Centro Nacional de  
Investigaciones Científicas, de Francia*

En el momento en que la tecnología permite a todos conocer informática con micro computadores, se necesitaba una reflexión fundamental sobre la creación de bancos de datos.

El instrumento no es nada sin la mano, así un logicial de tratamiento de datos documentales? tanto perfeccionado que sea dará buenos resultados solamente si la estructura y el contenido del banco de datos han sido pensados previamente.

Ahora bien si el jurista puede crear el mismo su propio banco de datos, importa mucho establecer la manera de tratar la documentación para automatizarla.

La ecuación que resolver es la siguiente:

Finalidades + documentación + sistema = buena información

Para que un banco de datos tenga una eficacia máxima, es necesario que corresponda a la finalidad del servicio jurídico determinado.

---

\* Conferencia impartida el 1 de diciembre de 1990, durante el XVI curso de Informática y Derecho: "La informatización de la Justicia: dos modelos europeos", celebrado en Mérida.

En primer tiempo, cada creador de banco de datos tiene que examinar las necesidades del servicio jurídico: cual información jurídica es necesaria.

Después tiene que hacer una lista de diferentes fuentes documentales que contengan los informaciones jurídicas.

Por ultimo, tiene que conocer todas las posibilidades del logicial elegido, para tratar lo mejor posible estas informaciones.

De la conjugación de estas tres reflexiones van a nacer la Forma y el contenido del banco de datos personalizado .

En nuestra demotración, el logicial elegido es pluri-doc que hicimos especialmente en el IRETIJ para tratar bancos de datos jurídicos.

Esta conferencia se compone de tres partes

I) la estructuración personal de informaciones específicas

II) la reflexión sobre la forma y el contenido documental

III) la presentación del logicial pluri-doc

I) la estructuración personal de informaciones específicas

En primer tiempo, es imprescindible hacer el inventario de lo existente. Se trata de definir los tipos des investigaciones habituales, sus finalidades y las fuentes documentales que sirven par estas investigaciones.

A) las finalidades de la busca de informaciones

Cada servicio jurídico tiene su propia finalidad; un organismo internacional, que proporciona síntesis teóricas, no persigue el mismo objetivo que un bufete de abogado que necesita mas soluciones precisas. Igualmente un servicio jurídica de empresa, que tiene que resolver problemas de reglamentación interna y también cuestiones mas jurídicas, no necesitara la misma información que la que desea un profesor de universidad para defender una tesis.

Así, cuando fueran establecidas las necesidades de servicio, examinaran su funcionamiento y la experiencia que existía antes para las documentalistas.

## **Definición de la finalidad de la información**

Cuando se constituyeron los primeros bancos de datos sobre computadores, se hicieron para el "jurista" en general.

Este se explica muy bien, porque estaban echas para todos y estaban accesibles por terminal sus creadores no podían conocer la especialidad o la función del jurista que podría interrogarlas.

Estos bancos de datos fueron estructurados de manera neutra en su forma y en su contenido. Esta neutralidad en la información estaba indispensable por un banco de datos destinado a todo el público.

Veinte años de experiencia en este sector, nos han permitido someter a tests a los usuarios. Cada jurista, en función de su especialidad, encuentra en estos sistemas generalmente una parte de información? pero no toda la información y se puede sobre una forma que no le satisface.

Cada jurista en su especialidad, tiene necesidades precisas para la información tanto por la origen de la información que por su forma.

Un jurista de empresa, tiene problemas internos tales que derechos sindicales, reivindicaciones salariales y también problemas externos tales que litigios que oponen la empresa a una tercera persona.

Tiene, antes de crear su propio banco de datos que hacer el inventario de tipos de cuestiones que hay que resolver.

Un organismo internacional que centraliza las informaciones tiene, el, que restituir a su consultador, la información bruta, texto integral, y no un simple resumen.

Con tales ejemplos, se ve que la estructura y la forma del banco van a depender de la finalidad de los futuras investigaciones.

Cual son las preguntas que un creador de banco de datos tiene que hacer?

- cuales son las tareas esenciales del servicio jurídico?
- cuales son los tipos de problemas que he tenido que resolver este año?
- con informaciones obtenidas que voy a hacer ? tomar una decisión ? resolver un litigio ? dar consejo ?
- es, la información que quiero automatizar fácilmente accesible ?
- es el banco echo solamente para mi ? o varias personas van a usarla ?
- es una parte de esta información destinada a terceras personas exteriores al servicio ?

-voy yo, con esta información, redactar un documento ? se necesitara una información completa o solamente simples referencias

-se necesitara hacer síntesis sobre puntos del derecho ? se necesitara hacer investigaciones puntuales ? tengo yo que automatizar todos los documentos del servicio o solamente los menos accesibles ?

-puede la automatización de la documentación permitirme tomar en cuenta otros documentos no utilizados ya y que serian muy útiles para mi trabajo de jurista ?

-este banco de datos sera propia al servicio, puedo interrogar comentarios personales y en cual caso ?

La lista no esta exhaustiva, pero cuando el jurista ha respondido a todas estas preguntas, puede tener una idea mas precisa de la función de la información documental en su servicio.

La informática tiene que dar una mejor información, no solo porque sea mas rápida. La creación de un banco de datos da la oportunidad de una reflexión sobre la manera, muchas veces maquinal, de trabajar sobre un problema jurídico preciso. La fijación de los objetivos esenciales del servicio para automatizar los documentos va a conducir a una aclaración de los elementos primordiales para la cualificación de un facto jurídico o para la fijación de los informaciones necesarias por los futuros usuarios del banco ( tercera persona).

## **Funcionamiento del servicio**

La introducción de un micro computador en un servicio es, muchas veces, una pequeña revolución. La automatización de la documentación puede ser mal sentida por las personas que tratan la documentación manualmente y que son muy competentes.

En efecto, el reflejo papel es muy difícil a combatir.

Para que la operación automatización de la documentación sea un suceso, importa mucho implicar todas las personas que trabajan en el servicio documental.

Esta reorganización presupone la estudio completa del funcionamiento actual del servicio. Se necesita conocer todo el circuito de los documentos. En primero, hay que establecer la origen de cada tipo de documentación como llega en le servicio ( revistas, documentos internos) cuando llega...

Hay, después, que notar donde y como todas estas informaciones son clasificadas.

La fase la mas importante de esta reflexión consiste en la descripción por cada persona del servicio de todas las etapas de su trabajo cuando una investigación esta preguntada.

Es muy importante conocer las claves manuales de busca la fecha, el nombre de las personas implicadas en el litigio, la materia de derecho.... Donde se encuentra la información, como se trata ?

La finalidad de estas preguntas es permitir la creación de un producto adaptado a las necesidades del servicio, pero, hay también otro objetivo, mas psicológico: implicar a los documentalistas en la concepción del banco y reproducir, dentro de lo posible, automatismos manuales de busca. Así, si las personas son acostumbradas a buscar un documento con su fecha, crearemos un campo fecha que permitiera conservar este automatismo. El banco de datos tiene que ser un mejoramiento de la información, pero tiene que tomar en cuenta las experiencias anteriores, así los responsables serán menos chocados por este cambio radical, si encuentran en el sistema algunas costumbres que les gustan.

Todos estos informes sobre funcionamiento van a permitir mejoramiento de la gestión de la documentación, pero sobre todo, la implicación de todo el servicio en la operación de automatización, porque nada es peor que una innovación impuesta.

B) tipología de los fuentes documentales y naturaleza de las informaciones

En los bancos de datos general, las fuentes documentales están limitadas a la jurisprudencia y la doctrina, otros bancos tratan de las leyes.

Aparece, hablando con usuarios, que estas informaciones muy útiles, estaban muchas veces insuficientes, además no se encontraban en el mismo banco.

1) tipología de las fuentes documentales: definición de las bases

No hay un jurista pero varios juristas. Así, un jurista de empresa va a necesitar para resolver sus problemas una documentación clásica, pero también una documentación privada

Un servicio jurídico de un organismo internacional tiene que automatizar las leyes nacionales pero también los convenios internacionales y las constituciones de los países concernidos

Un banco de datos en historia del derecho tiene que contener costumbres, cartularios y una base para cada siglo.

Se nota, tantos juristas especializados como fuentes documentales originales.

El objetivo primero, cuando se crea un banco personalizado, es reunir en el mismo instrumento documental todas las fuentes de información necesarias para la decisión del jurista.

En la etapa anterior, todos los documentos que pueden ayudar al jurista en sus investigaciones fueron enumerados. Entre estos documentos, algunas, de origen física diferente, son de misma naturaleza.

Así, en varias revistas muy diferentes, vamos a encontrar el mismo tipo de información.

Se necesita no confundir: la origen de la información, en el sentido del documento mismo ( revistas publicaciones nacionales) con las diferentes fuentes de información de un banco de datos.

La misma fuente de información puede se encontrar en varios documentos. Luego, hay que determinar la naturaleza de los fuentes de información del jurista especializado.

La doctrina es una fuente de información que se encuentra en revistas y en manuales

La jurisprudencia se encuentra en revistas pero puede también provenir de documentos inéditos.

La ley se encuentra en documentos oficiales, pero también en revistas o publicaciones internacionales.

Los convenios internacionales son proporcionados por organizaciones internacionales.

La reglamentación de una profesión, se encuentra en documentos privados tales que convenios colectivos del trabajo.

La fuente de información, en sentido de banco de datos informatizada, se determina como un conjunto de documentos que tienen las mismas características y que dan el mismo tipo de información.

Cuando ha estado determinada la tipología de los fuentes, podemos decidir del numero de base que van a componer el banco.

Es primordial de aislar las informaciones de misma naturaleza porque los elementos esenciales a cada una son diferentes y no se pueden tratar igualmente. Así, para la doctrina, el nombre del autor, el título y la referencia a la revista van a ser esenciales, en cambio, para la jurisprudencia, la jurisdicción, la sala, la fecha son las características principales.

Vamos a crear una base para cada fuente de tipo diferente.

Podremos en el momento de la pregunta, aislar una base o interrogar todas las bases. Juntas.

El principio es: una base para una fuente

Pero este principio no es sistemático, porque, para algunos casos, el recortado del banco en bases puede ser fundado sobre otros criterios más importantes por el servicio.

Así, en el caso de un banco de datos temático con una fuente única, por ejemplo la jurisprudencia, puede ser mejor recortar el banco en bases temáticas; se puede crear una base despedido una base derechos sindicales, una base contrato de trabajo...

Asimismo, para un banco histórico, se puede ser más interesante crear una base para cada siglo, si los elementos de información contenidos en cada fuente no son demasiados diferentes.

La creación de las bases es directamente ligada a la estudio de los finalidades determinadas anteriormente; en el logicial pluri-doc, el número de bases es ilimitado.

Cuando esta determinado el número de bases, hay que darlas un nombre, simple y explícito, porque este nombre va a servir para la captura de los documentos y también para la busca, si desea el usuario que la busca se hace sobre una sola base.

Cuando las tipologías de informaciones fueran establecidas, será imperativo determinar, para cada fuente, los elementos de información y sus naturaleza.

2) naturaleza de los elementos de información para cada base

Cada fuente ley, doctrina, jurisprudencia, convenio colectivo y otras, abarca elementos propios.

Se puede clasificarlos en cuatro categorías

- informaciones constantes
- informaciones aleatorias
- informaciones objetivas
- informaciones relativas al contenido jurídico

## **Las informaciones constantes**

Hay, en cada fuente documental, informaciones que vamos a encontrar sistemáticamente, cual que sea la origen de la información. Es lo que llamamos informaciones constantes

Así, por un documento de doctrina, estamos seguros que conoceremos siempre su referencia, porque es siempre publicado en una revista o en un manual, tendremos siempre un contenido jurídico, pero no estamos seguros conocer el nombre del autor, porque hay algunos artículos que son anónimos.

Esta etapa consiste a hacer el inventario, para cada fuente, de todos los elementos constantes de información.

Para la ley, tenemos siempre una fecha, un título, para la jurisprudencia, una jurisdicción, el nombre de los partes, la fecha. Hay que notar estas diferentes características que van a servir después para la estructuración en campos de cada base.

## **Las informaciones aleatorias**

Se encuentran dos categorías de fuentes documentales: las fuentes estáticas y las fuentes evolutivas.

Una fuente estática contiene en la origen, la totalidad de las informaciones relativas a un documento dado. Un artículo de doctrina es escrito por un autor, un día preciso, en una revista determinada, si este autor escribe otro artículo, sobre el mismo tema, será una información diferente de la primera.

Una fuente evolutiva es formada de una suma de informaciones que constituyen el primer grupo de elementos de información, pero que no están definitivos. En efecto después, otras informaciones sobre este mismo documento pueden venir agregarse. Así, una ley puede ser modificada o derogada, el mismo por convenio colectivo del trabajo. Por la jurisprudencia, cuando hay varios grados de jurisdicción, se necesita hacer una cadena de los diferentes informaciones.

En bancos de datos, existen varias maneras de tratar estas informaciones aleatorias.

Por algunos bancos de datos, se crea tantos documentos que evoluciones relativas al primer documento. Se puede crear, también, cadenas entre los diferentes documentos.

Importe luego, cuando se hace el inventario de las informaciones contenidas en cada fuente imaginar todas las evoluciones de esta documentación en el futuro, para poder preverlas cuando se hará la estructuración de la base.

## **Las informaciones objetivas**

Sobre este tipo de informaciones, el creador del banco de datos no puede intervenir; son datos brutos que hay que transcribir tales que son en el documento de origen por ejemplo la fecha, el nombre del autor, el título de una ley, el nombre de una revista, etc.

## **Las informaciones relativas al contenido jurídico**

Es el elemento fundamental de la información; todas las otras informaciones sirven solamente a precisar aquella.

La forma elegida para este tipo de información va a ser muy importante para su eficacia

Ejemplos contenido de un artículo de doctrina, la parte dispositiva de una decisión judicial, los artículos de la ley el contenido de los convenios colectivos del trabajo, etc...

Veremos en la segunda parte varias posibilidades de tratar esta información

### C) estructuración de las bases

Vamos a establecer la estructura de cada base; es la etapa de creación de los campos.

Un campo es una zona que contiene una o varias informaciones de misma naturaleza por ejemplo campo: autor, campo: título, campo: resumen, etc...

El conjunto de campos constituye el modelo de captura para la fuente documental tratada en esta base.

Hemos, en las reflexiones anteriores, hecho el inventario de la naturaleza de los informaciones contenidas en cada base, tenemos ahora que determinar las funciones de estas informaciones para la finalidad del servicio.

## 1) determinación de las funciones de cada información en el banco de datos

Los documentos contenidos en un banco de datos son, a la vez, sujeto y objeto de información. Esta dualidad es capital. En efecto, en la pantalla, los documentos dan la información jurídica, pero contienen también sus propias claves de acceso puesto que la busca se hace con las palabras incluidas en los diferentes campos de los documentos.

Al extremo, se puede, al momento de la captura, crear documentos que parecen contener una información perfecta, pero que nunca no se podrán recorrer a la interrogación, si no hay ninguna reflexión sobre la manera de poner la información.

Luego, es imprescindible que los informaciones sean organizadas y estructuradas para ejercer estas dos funciones de busca y de información.

Un campo puede tener una simple función de información, o solamente una función de busca o ambas.

Algunos campos tienen un simple valor de información al momento de la consulta, es decir que nunca no se buscará el documento con estos elementos, pero que será muy importante conocerlos cuando se consultara el documento.

Por ejemplo el título exacto de una ley no es una clave de busca, en cambio, para redactar un consejo, es importante conocerlo para ser preciso.

El mismo, un comentario personal después una decisión judicial no servirá a la interrogación, pero el jurista será ayudado por su comentario hecho al momento del litigio.

La cuestión que hay que ponerse es: vamos a buscar este tipo de documento con el contenido de este campo?

Se podría pensar que este no tuviera mucha importancia, pero se necesita saber que los sistemas de banco de datos documental con palabras que son manejadas por léxicos clásicos o automáticos, imponen un rigor en la forma del contenido de los campos que sirven a la busca.

Así, si un campo sirve a la sola información y no a la busca, el usuario será más libre en su redacción y podrá permitirse frases más largas y todos los caracteres que quiere puesto que las palabras de este campo no servirán a la busca y no podrán hacer "ruido".

Algunos campos tienen solamente una función de busca. Están reducidos a la función de clave de busca de los documentos. Es muchas veces el caso de

los campos de tipo index que son una sucesión de palabras que sirven a clasificar el documento y a encontrarlo, la información real se encuentra en otro campo de tipo texto integral o resumen por ejemplo.

Pero, las mas de las veces, el campo tiene a la vez la función de busca y la función de informe, necesita entonces una buena reflexión sobre su contenido que veremos en la segunda parte.

## 2) creación de las estructuras para cada base

En primer tiempo, hay que decidir del numero de campos que va a contener la base; por ejemplo la base doctrina tendrá un campo autor, un campo: titulo, un campo, abstracto o resumen y un campo. Referencia.

El numero de campos es determinado por el numero de informaciones de naturaleza diferente enumeradas anteriormente

Después, vamos a decidir del funcionamiento de cada campo en el banco de datos.

Ahí, para comprender, se necesita presentar las funciones del logicial pluri doc en esta materia.

El logicial pluri doc permite parametrizar cada campo de la manera siguiente

-este campo tiene un nombre:

El mas explicito posible en función de su contenido

-esta campo sera:

Indizable

Parcialmente indizable

No indizable

Indizable todas las palabras del campo sirven para indizar el documento. Todas estas palabras serán claves de acceso al documento.

Parcialmente indizable solo las palabras capturadas entre paréntesis rectangulares son indizables, las otras palabras del texto no podran ser claves de acceso al documento.

No indizable ninguna palabra es indizada. Luego, este campo no sera una clave de acceso al documento, pero solamente un elemento de información a la consulta. Esta función permite la protección de los archivos nominativos.

-este campo sera de tipo:

Texto

Alpha

Numérico

Fecha

Estas características determinan el uso de las teclas del teclado

Texto todas las teclas

Alpha: todas las teclas pero longitud fijada por usuario numérico: ochos cifras máximo

Fecha: ochos letras, dos el día, dos el mes, dos el ano

-este campo va a tener una longitud esta característica es ligada a la precedente: en tipo texto, la longitud es ilimitada, en tipo numérico y fecha, la longitud es fijada a ochos caracteres, en tipo alpha, el logicial da la posibilidad de fijar la longitud en la limita de 250 caracteres.

-este campo sera desplegable o no desplegable: Esta característica permite de decidir que el contenido de un campo sera invisible a la impresión y en la pantalla.

-este campo necesitara una captura obligatoria o opcional

Ejemplos de utilización de estas características:

El nombre de los partes, para los documentos de jurisprudencia, sera declarado no indizable, para la protección de los archivos nominativos, de tipo texto, longitud ilimitada desplegable, captura obligatoria.

Así, no se puede buscar un litigio relativo a una persona particular, pero la información aparece cuando la busca esta manejada sobre un problema de derecho.

Un campo texto integral, sera declarado parcialmente indizable, para evitar el ruido a la interrogación, solo serán indizados los conceptos esenciales del documento, tipo texto, longitud ilimitada, desplegable, captura obligatoria o opcional siguiente las necesidades del servicio.

Un campo comentario sera declarado no indizable, porque no sirve a la busca, es solamente un complemento de información personal, tipo texto, longitud ilimitada, captura opcional, porque el usuario no desea hacer comentario sobre todos los documentos y se puede que desea hacerlo cuando el litigio se acabe o durante una labor de síntesis que suscitara comentarios de su autor.

Las dos características esenciales son la característica indizable, no, parcialmente, y la captura obligatoria o opcional. Vamos a combinarlas para crear un banco de datos mas personal.

La primera permite determinar los blancos exactos de la busca de información, si este campo es declarado indizable; poner el full text parcialmente indizable lo que da las ventajas del texto integral sin los inconvenientes, o eliminar de la busca algunas informaciones personales o nominativas.

Les aconsejamos que tratan todas las notas o comentarios en la forma no indizable, en efecto, para una eficacia máxima del banco, importa que sean indizables solamente las informaciones que se encuentran en el documento mismo, así, si una decisión de jurisprudencia es aberante, hay que ponerla en su aberacion el comentario podra decir en que esta decisión es aberante.

La segunda captura obligatoria o opcional, permite solucionar muchos problemas. Primero, permite tomar en cuenta informaciones no constantes, así en el campo autor para la doctrina elegiremos captura opcional, porque algunos artículos de doctrina están anónimos.

Autoriza también la creación de campos que no serán sistemáticamente es- cribidos en la captura, por ejemplo el campo comentario. Pero, sobre todo, permite manejar informaciones aleatorias y evolutivas. En efecto, cuando un documento puede ser evolutivo, importa prever en la creación de la estructura campos que autorizaran tomar en cuenta la evolución estos campos pueden quedar vacíos mucho tiempo o nunca estar llenos. Por ejemplo una ley puede ser modificada muchos años después su promulgación, puede ser abrogada o no abrogada.

Se necesita notar que esta opción es posible porque el logicial pluri doc permite, con su función modificación, cambiar el contenido de documentos ya indizados.

El numero de campos es limitado a noventa y nueve para un solo documento. Se puede decir que es ilimitado, porque noventa y nueve informaciones de naturaleza diferente para un mismo documento parece imposible.

El creador del banco de datos tiene toda oportunidad de personalizar su banco con campos propios que se añaden a los campos impuestos por la naturaleza de las informaciones.

Por ultimo, la característica desplegable permite dar a algunas informaciones un caracteres confidencial, estas informaciones pueden ser indizables pero no desplegables.

Por ejemplo: para los informaciones nominativas internas de una empresa que tienen un caracteres disciplinario o sindical.

Ademas, puede autorizar el tratamiento de cierto tipo de documentación, creación de un campo index, indizable pero no desplegable a causa de su estilo telegráfico, y después un campo resumen o texto integral, no indizable pero desplegable.

## II) reflexiones sobre la norma y el contenido de las informaciones

La forma y el contenido de cada campo van a depender a la vez de la finalidad del banco de datos y de las funciones del sistema elegido.

Hemos visto que existen varios tipos de informaciones. Vamos primero, a tratar de los datos objetivos, constantes o aleatorios y después de los datos relativos al contenido fundamental.

Tenemos que precisar, antes de hablar de "contenido documental", que un computador no conoce una palabra que en una sucesión de caracteres separados para espacios o puntuación determinada por el programa.

Sus cualidades serán mas o menos grandes en función del sistema lingüístico del logicial elegido: lenguaje cerrado, truncamiento clásico, léxico clásico, léxico automático.

El desarrollo que sigue es fundado sobre el sistema de léxico automático del logicial pluri doc.

### A) forma y contenido de las informaciones objetivas constantes o aleatorias

Son informaciones dadas en el documento y que tenemos que transcribir sin interpretación.

Vamos a ver los datos objetivos los mas frecuentes

#### 1) las fechas

La fecha es un elemento esencial para individualizar un documento.

Tenemos interés por crear un campo fecha, que permitirá hacer selección precisas.

Existe muchas maneras escribir una fecha: 1 1 1990, 1 01 1990, 01 01 1990, 1 1 90, 1 01 90, 01 01 90, 1 enero 1990, 1 enero 90, 01 enero 1990, 01 enero 90, y otras abreviaciones del mes

Existe diez o mas maneras de anotar una fecha. Sabemos que para un computador, 1 no es 01 y 90 no es 1990... Podemos notar la importancia de tener

una escritura uniforme de la fecha sino no podremos recuperar dos documentos de misma fecha, porque la fecha no ha sido escrita de la misma manera.

Importa, luego, imponer, antes de todo, una maneja fija de escribir la fecha.

Aconsejamos poner la fecha en un campo de tipo texto y escribir el día sobre dos caracteres, el mes en letras y el año sobre cuatro cifras.

01 enero 1990

Se puede existir otros problemas sobre la fecha. En efecto, sobre el mismo documento, varias fechas pueden coexistir en campos diferentes. Seguro que si el logicial da la posibilidad de interrogar sobre campo preciso, el riesgo de ruido puede ser reducido.

Tomamos un ejemplo una ley es promulgada el 13 de abril 1988, su decreto de aplicación es del 01 enero 1989, ha sido modificada el 03 de marzo 1990. Vamos a encontrar las tres fechas sobre el mismo documento.

En este caso, aconsejamos de juntar a la fecha en la misma frase, un elemento que permitirá diferenciar esta fecha. Así en el campo aplicación escribimos decreto de aplicación de 01 enero 1990, en el campo modificación escribimos modificación por la ley del 03 marzo 1990. La fecha es, entonces, ligada con un evento preciso y autorizada una busca mas eficaz.

## 2) los nombres propios de personas

Los nombres propios de personas van a estar tratados diferentemente según sus naturaleza. Un nombre de un autor de articulo de doctrina o de comentarios, sera indizable. Va a ser un campo particular campo. Autor. Así, en la busca, se podra buscar precisamente, lo que ha escrito un autor interrogando solamente sobre el campo autor.ademas, hay nombre propios que tienen también un senso preciso como nombre común.

Para los nombres propios de personas particulares, en decisiones de jurisprudencia, se necesita mucha prudencia. Así la ley francesa prohíbe los archivos nominativos. En principio este campo sera declarado no indizable, así estos nombres serán desplegables pero no servirán a la interrogación. De toda manera, los nombres propios no son claves seguras de busca en razón de las variaciones ortográficas.

Sin embargo, si el banco de datos es exclusivamente personal y de uso privado, todos los nombres propios pueden ser indizables.

## 3) el origen del documento

Es muy importante en un banco de datos conocer en la consulta, la fuente exacta del documento desplegado en la pantalla. Estos datos van a ser diferentes para cada fuente documental. Lo que importa es de dar lo máximo de informaciones precisas sobre este sujeto. Por ejemplo el nombre de la revista, el año de edición, la página, el número, se puede también poner en el documento la ubicación física en el servicio para encontrar el documento más rápidamente.

Ej: dalloz 1990 p35 despacho 115 segunda estantería

Ejemplo para la jurisprudencia, el nombre de la jurisdicción, el nombre de la ciudad, la sala y el número de la decisión.

Ej: corte de apelación Montpellier, sala 1, número 30

En este orden, se necesita prohibir las abreviaciones porque puede existir varias maneras de abreviar, y si no se usa las mismas abreviaciones, no se encontraron los documentos en la busca porque serán indizados de manera diferentes.

Consejamos escribir todas las palabras de este campo enteras

4) características precisas del documento

Son, esencialmente, los títulos de artículos de doctrina, de manuales, de las leyes o de los convenios internacionales.

5) los tratamientos posibles de las informaciones objetivas aleatorias

Las informaciones aleatorias son las más de las veces de naturaleza objetiva, pero su particularidad es que son conocidas posteriormente a la creación del documento inicial

Son informaciones que van a completar o modificar un documento que existe ya en el banco.

Varios tratamientos están posibles.

Tomamos varios ejemplos

- modificación de una ley

El documento existe ya en el banco de datos sobre su forma primera y incluye un campo modificación que está vacío.

Primera solución se puede crear un segundo documento por el artículo modificado y escribir en el campo modificación modificación de la ley x de tal fecha.

El inconveniente de esta solución es que si los dos documentos no tienen las mismas palabras (clave de acceso) a la busca o la consulta se puede que sala solo lo mas antiguo y el usuario no conocerá el nuevo texto modificado, lo que es muy grave.

Segunda solución se hace la misma cosa que en la primera solución, pero se llama el primero documento para modificarlo y se escribe en el campo: modificación de la ley x de tal fecha por la ley y de tal fecha.

Tercera solución se crea un nuevo documento con el nuevo artículo modificado, se menciona la modificación en el campo modificación y después se suprime el primero documento.

Esta opción no es satisfactoria, porque la modificación, muchas veces, trata de una parte solamente del artículo, en este caso, el usuario no conocerá la parte que subsiste y la parte que esta modificada.

Cuarta solución: se busca el primero documento y después se llama la función modificación de documentos. Se puede agregar un nuevo párrafo conservando el antiguo, o se puede modificar directamente en el texto la parte que esta cambiada, en este caso, se necesita escribir después cada parte cambiada, entre paréntesis, la nueva ley. En el campo modificación se escribe modificación por la ley y la fecha.

Se concentran así las informaciones sobre el mismo documento y no hay ningún riesgo de silencio en la edición de los documentos.

La elección del mejor tratamiento depende seguramente de las necesidades del servicio, pero sobre todo de las posibilidades del logicial que, para algunas de las soluciones tiene que permitir modificar documentos ya indizados.

Es posible sobre el logicial pluri doc, ese estaba unos de los objetivos esenciales en su concepción.

-Abrogacion de una ley

Es una información aleatoria, pero que tiene consecuencias diferentes de la modificación. Ahí, la ley no es mas aplicable y se necesita capturar nueva ley.

Los documentos relativos a la ley abrogada existen en el banco de datos.

Primera solución : se puede crear nuevos documentos sobre la nueva ley y escribir en el campo : abrogacion de estos documentos : abrogacion de la ley x por la ley y y la fecha.

Los graves inconvenientes de esta solución son que si las palabras del contenido jurídico son diferentes en documentos antiguos y en documentos nuevos, corremos el peligro de salir los documentos sobre la ley abrogada sin saber que la ley es abrogada y que existe una nueva ley. Pero, si los conceptos son los mismos, vamos a salir todos los documentos, antiguos y nuevos, cuando solo los nuevos nos interesan.

Segunda solución : hacemos la misma cosa que en la primera solución, pero nos vamos a llamar en la pantalla todos los documentos sobre la ley abrogada y a escribir en el campo : abrogacion de estos documentos : abrogacion por la ley y del y fecha, este evita el riesgo de tener en la edición solamente la ley abrogada sin saber que exista una nueva ley, pero hay siempre el inconveniente de obtener en repuesta, cada vez, la ley abrogada y la nueva ley sobre una pregunta.

Tercera solución : creamos documentos para la nueva ley y escribimos en el campo : abrogacion que esta ley abroga la ley antigua. Después se suprime todos los documentos relativos a la antigua ley.

Esta solución es mejor, pero, impide toda estudio evolutivo sobre estas leyes. Además, en razón de los regímenes transitorios, se puede tener interés guardar una ley y archivarla.

Cuarta solución : buscamos los números de los documentos de la ley abrogada en el banco de datos. Se puede poner en el campo : abrogacion: abrogacion por la ley y del fecha. Después creamos documentos para la nueva ley y escribimos en el campo : abrogacion de estos documentos : abrogacion de la ley x del y fecha archivada sobre números 30 hasta 50 por ejemplo.

Después , cuando la ley abrogada no se aplica mas, con el mando : suprimir, se puede archivar los documentos de la ley abrogada.

Si se necesita, mas tarde, se puede recuperar esta ley con el mando : restaure que permite recuperar documentos archivados.

Es la mejora solución, la mas fiable, pero es un poco pesada por la gestión del banco. Así, esta solución es preconizada para bancos de datos especializados en el tratamiento de las leyes.

-Tratamiento de los grados de decisiones de justicia

Un litigio puede tener varias etapas, es muy importante conocerlas; así una decisión de un tribunal podra seguir en corte de apelación y después en corte suprema.

Se necesita prever un campo: cadena para los documentos de jurisprudencia.

En este caso, se necesita crear tanto documentos como grados de jurisdicción. No es posible concentrar sobre el mismo documento todas las decisiones porque se pueden que sean muy diferentes o contrarias.

Hay dos soluciones posibles

Primera solución: para cada decisión, se anota la cadena: para el primer grado, no se anota nada.

Para la corte de apelación: se anota el tribunal y la fecha del primer grado.

Para la corte suprema: se anota el tribunal y su fecha y la corte de apelacion y su fecha.

No estamos seguros salir en repuesta los tres documentos, pero si sala el documento de la corte de apelación conoceremos la fecha del primer grado, si sala el documento de la corte suprema conoceremos toda la cadena y podremos buscar los otros documentos , en el banco, con sus fechas.

Segunda solución: para cada decisión, se nota la cadena como en la primera solución. Pero, después, para la corte de apelación o la corte suprema, se puede llamar en la pantalla los documentos anteriores y escribir en sus campos : cadena las otras referencias; así, cualquier documento sale , el usuario conoce toda la cadena jurídica.

Esta solución es un poco pesada para la gestión del banco. Es preconizada para banco de datos especializado en el tratamiento muy preciso de la jurisprudencia.

B) el contenido fundamental de los documentos

En todos los documentos hay una información fundamental, es, para el jurista, la información jurídica incluida en la ley, en la jurisprudencia, en la doctrina y otros.

Tenemos que reglar dos problemas esenciales: la homogeneidad del vocabulario y la naturaleza del contenido documental.

En el banco tenemos documentos de fuentes muy heterogéneas, como vamos a resolver el problema de la homogeneidad del vocabulario, muy importante para la eficacia del banco ?

Para cada fuente, que tipo de informaciones es necesaria para mi ?

En realidad, la distinción entre los dos problemas es artificial porque son ligados. Les hemos tratados distintamente para mas de claridad.

#### 1) homogeniedad documental

El fondo del derecho es el elemento esencial de las investigaciones documentales. El jurista, confrontado a un problema jurídico, desea encontrar informaciones jurídicas que van a ayudarle a resolver este problema. Además, va a interrogar el banco con conceptos jurídicos.

Cuando, por ejemplo, el jurista encuentra un problema de perdida de los equipajes durante un viaje en avión, va a hacer esta investigación con los conceptos: transporte aéreo y perdida de los equipajes.

El acceso principal a la información se hará con conceptos jurídicos.

Es por eso que es capital resolver el problema de la homogeneidad del vocabulario en un banco de datos donde las fuentes son muy diferentes.

Hay que tratar de combinar, al máximo, homogeneidad de las claves de acceso y flexibilidad de la análisis documental.

En un sistema de lenguaje cerrado, el problema es reducido, porque un cierto numero de palabras solamente, sirve a la indización.

En sistemas que permiten un lenguaje natural, la reflexión sobre la forma de la información es capital.

El sistema lingüístico del logicial pluri doc, autoriza una gran flexibilidad: la canonización o truncatura inteligente que trataremos en la tercera parte, facilita el tratamiento lingüístico, pero, para obtener muy buenos resultados, no se puede contentar del texto integral y de las informaciones objetivas, al menos, si el banco de datos no es considerado como un simple instrumento de clasificación pero como un instrumento de investigación sobre el fondo del derecho.

El texto integral, declarado: parcialmente indizable, es ya una mejor solución, porque permite aislar la parte esencial de la información en el texto, dejando todas las frases que podrían ser una fuente de ruido.

Pero este sistema no resuelve totalmente el problema esencial de las formulaciones diferentes para expresar el mismo concepto.

Así, en las varias legislaciones sobre el derecho de autor, para expresar la idea de duración de la protección del derecho de autor, varias formulaciones

se encuentran : el derecho de autor será protegido su vida durante y cincuenta años después su muerte.

La protección del derecho de autor dura toda la vida del autor y cincuenta años después su defunción.

El derecho de autor es protegido del viviente del autor y expira cincuenta años después su muerte.

La duración de la protección del derecho de autor comprende la vida del autor y se prolonga cincuenta años después su muerte

Hay todavía otras formulas para expresar la misma idea, la palabras usadas varían de un texto a un otro; el riesgo existe entonces, pedir ciertos documentos a la pregunta, si las palabras usadas en la pregunta no se encuentran en el texto integral.

Los sistemas de la canonización y de la sinonimia, permiten recuperar familias de palabras con sus raices, pero el riesgo subsiste.

El creador del banco de datos confrontado a una documentación heterogénea, tiene que reflexionar sobre los conceptos esenciales del campo jurídico tratado. Se necesita establecer un equilibrio entre: homogeneidad del vocabulario y especificidad documental.

Para conciliar los dos imperativos, eficacia y flexibilidad, aconsejamos establecer modelos de análisis bastante breve para cuestiones de derecho esenciales y dejar, después, el analista libre de notar los hechos jurídicos como quiere.

Se puede entonces hacer una investigación muy puntual por ejemplo: colorización de una película en negro y blanco, pero se puede también hacer un investigación mas larga sobre la duración de la protección, si ha decidido que cada vez que este problema sea encontrado, el mismo concepto sera usado para caracterizarlo.

En la jurisprudencia, en la doctrina y en la ley, el uso de perífrasis es frecuente para expresar un concepto jurídico. Es por eso que importa tener modelos de análisis para cada problema jurídico fundamental.

Este es mas fácil a hacer para bancos de datos temáticos. No son estructuras rígidas y imperativas, es imposible prever modelos de análisis para todos los problemas jurídicos; pero se puede organizar el vocabulario para las cuestiones esenciales de una materia, optando, por ejemplo, por el concepto en lado de una frases con un verbo. Así serán creados modelos de análisis que van a dar una cierta uniformidad al banco de datos y sobre todo una grande fiabili-

dad, porque, si , para una pregunta puntual y precisa, el logicial no da ninguna repuesta, sera siempre posible, con una pregunta mas larga sobre el problema jurídico, obtener repuestas que permitirán encontrar una solución razonando por analogía.

Algunos ejemplos de modelos de análisis para un banco de datos tratando de los diferentes legislaciones sobre el derecho de autor.

Aquí, el problema de la homogeneidad de la análisis documental estaba muy difícil porque las legislaciones nacionales tenían sistemas jurídicos fundamentalmente diferentes según los países.

Luego, se necesitaba pensar a partir de los elementos invariantes de este derecho: cualquier sea el sistema, hay siempre: un autor, una protección, una obra, etc

### Ejemplos

Derecho de autor/ duración de la protección/ ( escribir : viviente de autor y numero de años) régimen general/ punto de partido de la protección/ primer año civil siguiente al de la muerte o fecha de defunción o fecha de la primera publicación

Derecho de autor/ titularidad del derecho de autor/ derechos protegidos/ derechos morales (si existen en la ley tratada) / derechos patrimoniales/ aquí retranscribir los derechos de la ley tratada por ejemplo:

derecho de reproducción  
derecho de representación  
derecho de suite  
derecho de publicación  
etc....

El derecho de autor fue dividido en ocho capítulos. Ustedes notaran que los modelos de análisis son muy breves y dejan entera libertad reproducir la especificidad de cada ley. Estos modelos eran muy necesarios porque no podíamos crear un solo documento por una ley (demasiado largo) se necesitaba cortar cada ley en unidades informativas.

Cada creador de banco de datos personalizada va a adaptar estos principios al finalidad de su propio banco.

Estos no son que algunos ejemplos para suscitar una reflexión de cada usuario enfrente de la información que va a tratar.

El contenido documental puede ser organizado también en frases y párrafos, si el logicial elegido permita una pregunta con operadores de distancia. Se necesita tomar en cuenta estas construcciones en la organización del contenido documental.

Capturando una documentación, importa pensar la manera como el usuario podrá hacer una pregunta sobre este problema preciso.

Así, escribir : transporte internacional / mercancías/ es una mala opción porque el reflejo a la pregunta sera escribir : transporte de mercancías (en la misma frase) ahora bien en nuestro ejemplo, transporte es en una frase y mercancías es en una otra frases, la pregunta obtendrá cero repuesta cuando había realmente repuesta en el banco.

El principio que tenemos que repetir es de escribir un concepto por una frase y un problema de derecho por párrafo.

Todas estas recomendaciones son destinadas a evitar el silencio que es la peor de cosas para un banco de datos, los problemas de ruido son menos graves porque es mejor encontrar su información en medio de otros documentos que no son pertinentes, que obtener cero repuesta.

Para mejorar los resultados del logicial y no encerrar su usuario en reglas estrictas de captura, los instrumentos lingüísticos tienen una gran importancia.

La canonización propia a pluri doc de que hablaremos mas tarde, resuelve muchos problemas porque agrupa familias de palabras sobre la misma raíz.

Pero cierto logiciales ofrece también posibilidades crear vínculos de sinonimia y declarar cierto palabras como nulas

Las palabras nulas son artículos o preposiciones que sirven a la comprensión del texto pero no tienen ninguna valor informativa. Son en el banco de datos, pero no son indizados.

Los vínculos de sinonimia son muy importantes para mejorar los resultados del banco de datos. Cada usuario podrá crear vínculos entre diferentes palabras que pueden, en una materia particular, tener el mismo significado. Excepto la sinonimia perfecta jurídica, un logicial no puede tener a priori, vínculos de sinonimia. Solo el creador del banco podrá decidir que en la materia que trata hay un interés decir que dos palabras son sinónimas. Así, en materia de derecho de autor, se puede decidir que titular y titularidad son sinónimos, en un banco de datos en derecho de transporte, coche y cochero.

En pluri doc, estos vínculos pueden ser creados posteriormente cuando a-

parece en la pantalla que algunas palabras son usadas por el mismo sentido. El vínculo de sinonimia funciona entonces para el pasado y para el futuro.

2) ejemplos de tratamiento de las fuentes en funciones de la finalidad

## La ley

En función del objetivo buscado, la ley va a ser tratada con minucia, parcialmente tratada o excluida del banco.

Si el objetivo buscado es conocer la posición de cada país sobre un punto de derecho preciso ( UNESCO derecho de autor ) la ley va a ser la fuente principal de información. Tendremos la estructura siguiente.

1) país: país de origen de la ley y continente Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

2) naturaleza: informaciones objetivas sobre la ley: artículo, número, fecha Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

3) título : título exacto de la ley

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

4) abstracto: contenido siguiente los modelos de análisis establecidos antes

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

5) texto integral: reproducción del texto de los artículos tratados en este documento y concerniendo un problema preciso de derecho.

Características :

indizable = parcialmente

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

6) modificación: modificación de la ley x por la ley y

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

7) abrogación: abrogación de la ley x por la ley y

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

8) aplicación: decreto de aplicación del fecha

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

9) lengua: lengua de origen del texto de ley

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

10) comentario: posibilidad de añadir comentario relativos a la interpretación jurisprudencial, por ejemplo

Características :

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Si el banco de datos es temático, por ejemplo sobre el derecho de transporte, tenemos interés a tratar las leyes relativas al derecho de transporte y los convenios internacionales, pero de manera mas sucinta, a titulo de información:

1) naturaleza: informaciones objetivas sobre la ley: articulo, numero, fecha

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

2) abstracto: contenido siguiente los modelos de análisis establecidos antes

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

3) seguido: modificación ,abrogacion, aplicación

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

4) referencia: Referencia exacta del lugar donde se puede encontrar el texto integral

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Si el banco es constituido para un bufete de abogado, sera mas orientada sobre la jurisprudencia y la doctrina; no hay ningún interés capturar el código civil porque es accesible muy fácilmente manualmente. Cierta texto de ley podran ser indizado si el abogado es especializado en una parte precisa del derecho.

## La jurisprudencia

Puede ser esencial para un bufete de abogado, accesoria para un servicio jurídico o anecdótica para un organismo internacional, siguiente las necesidades del servicio.

Para un abogado

1) jurisdicción: jurisdicción ciudad sala numero

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

2) fecha: 03 febrero 1990

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

3) partes: nombre de los partes

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

4) abstracto: contenido siguiente los modelos de análisis establecidos antes

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

5) considerando principal: copiar el motivo principal de la decisión

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

6) cadena: anotar las decisiones anteriores o posteriores

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

7) comentario: posibilidad de añadir comentario relativos a la interpretación jurisprudencial, por ejemplo

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Para un servicio de empresa:

1) jurisdicción: jurisdicción ciudad sala numero

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

2) fecha: 03 febrero 1990

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

3) partes: nombre de los partes

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

4) abstracto: contenido siguiente los modelos de análisis establecidos antes

5) cadena: anotar las decisiones anteriores o posteriores

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Para un banco de datos legislativo.

Se necesita solamente mencionar en el campo comentario del documento relativo al artículo concernido la interpretación jurisprudencial. Se puede presentar interés notar, por ejemplo, que la jurisprudencia añade ciertas condiciones a la aplicación de la ley o que interpreta un concepto en un cierto sentido.

## Los convenios internacionales

Las informaciones contenidas en convenios son a la vez informaciones jurídicas y datos objetivos muy importantes tales que los adhesiones, las reservas o las ratificaciones

Tres soluciones posibles:

1º) crear un solo documento por convenio

naturaleza : nombre del convenio y fecha

Ratificación: ratificación del convenio por (nombre del país) y fecha; tantos párrafos como países

Adhesión: adhesión al convenio por (nombre del país) y fecha tantos párrafos que países

Excepción: mención de los países que han preguntado ser favorecidos excepción; tantos párrafos como países

Declaración: declaración al convenio (nombre del país) y fecha; tantos párrafos como países

Denunciación : mención de los países denunciando el convenio y fecha; tantos párrafos como países

Reservas: mención de las reservas y fechas; tantos párrafos como países.

Abstracto: sucesión de abstractos resumiendo todos los artículos del convenio.

Comentario:

Origen: organismo a la origen del convenio

Esta solución presente graves inconvenientes. En el momento de la consulta, el documento sera muy largo, luego fastidioso para la lectura; ademas, las informaciones relativas a las ratificaciones y adhesiones etc... Serán reagrupadas en el mismo campo, sera luego mas difícil conocer la posición de un país.

2) crear una ficha por país y tantas fichas que países que han firmado el convenio

País: país signatario del convenio y continente

Naturaleza: nombre del convenio y fecha

Ratificación: ratificación del convenio por (nombre del país y fecha;

Adhesión: adhesión al convenio por (nombre del país) y fecha

Excepción: mención si el país ha preguntado ser favorecido excepción

Declaración: declaración al convenio (nombre del país) y fecha

Denunciación: mención si el país ha denunciado el convenio y fecha

Reservas: mención de las reservas y fecha.

Abstracto: sucesión de abstractos resumiendo todos los artículos del convenio.

Comentario:

Origen: organismo a la origen del convenio

Esta solución , ya mas seductora que la precedente, presente todavía algunos inconvenientes a la consulta en razón de la longitud del abstracto.

Además, como el abstracto reproduce el contenido del convenio, al momento de una interrogación sobre conceptos jurídicos, tendremos tantas salidas de documentos como países signatarios del convenio.

En función de las finalidades y del conocimiento del servicio de la UNESCO, parece mejor privilegiar informaciones objetivas, porque el texto de los principales convenios es accesible fácilmente por todos los juristas del servicio. Hemos, luego elegido una tercera solución que toma en cuenta la segunda pero con un abstracto más sucinto que no es más la principal zona de investigación, pero que es aquí a título de breva información.

Aquí está la estructura elegida con las características de cada campo:

País: país signatario del convenio y continente

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

Naturaleza: nombre del convenio y fecha

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

Ratificación: ratificación del convenio por (nombre del país) y fecha

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Adhesión: Adhesión al convenio por (nombre del país) y fecha

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Excepción: mención si el país ha preguntado ser favorecidos excepción

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Declaración: declaración al convenio (nombre del país) y fecha

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Denunciación: mención si el país ha denunciado el convenio y fecha

Características

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Reservas: mención de las reservas y fecha.

Características

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Abstracto: abstracto breve sobre temas esenciales del convenio.

Características:

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatorio

Comentario:

Características:

indizable = no

tipo = texto

longitud = ilimitada

desplegable = si

captura = opcional

Organismo: organismo a la origen del convenio

Características

indizable = si

tipo = texto

longitud = ilimitada

desplegable = si

captura = obligatoria

Estos ejemplos demuestran las múltiples posibilidades de organización de un banco de datos jurídico en función de las necesidades del servicio concertado.

Voy ahora presentar el logicial pluri doc que hicimos a IRETIJ especialmente para tratar bancos de datos jurídicos .

III) un sistema para crear bancos de datos personalizados sobre micro computadores

Los bancos de datos estaban considerados, hasta hoy día, como muestros por muchos juristas, tanto por su potencia como por su técnica.

La creacion de tales bancos es reservada a los especialistas: juristas, lingüístas y especialistas en informática. Las dificultades tecnicas y el coste elevado de realizacion hacen estos instrumentos documentales inaccesibles a los particulares.

Veinte años de experiencia en el campo de la informática jurídica nos permiten pensar que hoy, los resultados descontados en materia de uso de bancos de datos jurídicos no están logrados. La reacción de los juristas no es muy buena en razón de las dificultades de los métodos de interrogación y del numero demasiado largo de documentos que son la consecuencia de una mala pregunta.

La reticencia del publico enfrente del instrumento informático de que hemos mucho tiempo pensado que constituía un obstáculo dirimente, tiende a desaparecer con la llegada de nuevas generación de juristas mejor preparados y mas concernados.

La trivialidad de la micro informática induce interrogarnos sobre las posibilidades de creación de un nuevo instrumento documental para el jurista.

La innovación técnica que constituye el micro computador, accesible a todo particular, induce nos volver a determinar toda la manera de ver el tratamiento documental automático.

Poner a la disposición del jurista, no especialista en informática, un logicial que permite crear un banco de datos adaptado a sus necesidades documentales, tales estaba el reto.

Así nuestra reflexión primera trata de la descomposición del razonamiento del jurista cualquier que fuese su campo de actividad, enfrente un problema de derecho. Después determinamos todos los tipos de necesidades documentales, la forma de las informaciones jurídicas tratadas y la facilidades de uso necesaria a la eficacia del sistema.

De este estudio, resultado de una estrecha cooperación entre jurista y especialista de informática, nació el logicial pluri doc que quiere satisfacer todos estos objetivos.

#### A) de la pasividad a la creación

Un logicial de banco de datos sobre micro computador tiene que dejar la entera maestría del contenido documental a su creador

Los bancos de datos jurídicos, sobre largos sistemas, proponen solamente un acceso consultivo. Los documentos obtenidos en repuesta, han estados, anteriormente, seleccionados, analizados y pensados por otras personas que el usuario. Han estados concebidos, con muchas razones, para satisfacer el mas largo numero, tanto el especialista que los otros.

Estos elementos, variables según la especificidad de las fuentes elegidas y las necesidades del creador, no son impuestos arbitrariamente por un logicial pero dejados a la entera apreciación del usuario.

Y por eso, la información, objeto de una selección y una análisis propia de las necesidades precisas, da una total eficacia al sistema documental perfectamente personalizado

Cuando es determinada, la información tiene que ser organizada racionalmente para permitir su recuperación en las mejoras condiciones.

## 2) libertad de selección del contenido de la documentación

El usuario puede crear tanto archivos como fuentes documentales. La información contenida en el banco es así estructurada en bases distintas. Esta estructuración permite al momento de la consulta, hacer investigaciones globales o sobre una sola base.

A dentro de cada base, la estructuración de la información obedece al mismos principios de flexibilidad y de libertad

Es determinada por el usuario quien, después que sean catalogados los elementos informativos de los datos, puede atribuir a cada uno de entre ellos un campo propio.

Estos campos constituyen las claves de acceso al documento. Se necesita determinarlos de manera precisa. El sistema propone para cada campo las características siguientes:

Un nombre

Indizable:

-Si todas las palabras contenidas en el campo son claves de acceso al documento.

-Parcialmente solo las palabras puestas entre corchetes son claves de acceso al documento.

-No ninguna palabra no son indizadas

Tipo de caracteres

-texto

-numérico

-alpha

-fecha

Longitud

-fija

-ilimitada

Desplegable

-si

-no

Esta posibilidad permite decidir si un campo aparece a la consulta o a la impresión y también la confidencialidad.

Captura  
-obligatoria  
-facultativa

Jugando sobre estas diferentes características ofrecidas por el logicial, el creador del banco va a poder tratar cada fuente documental según sus necesidades en texto integral (parcialmente indizable) en índices, en simple referencias.

### 3) libertad de selección de la forma de la información

Según la naturaleza de la fuente y las necesidades del servicio concernido, la forma de la información va a cambiar de una base a otra. Importaba luego que el logicial ofrezca una flexibilidad máxima en este campo. Autoriza tanto un lenguaje natural como un lenguaje convencional (palabras claves índices referencias texto integral).

Los métodos de tratamiento tradicionalmente opuestas full tex abstracto pueden coexistir en la misma base o en bases diferentes.

Así según las necesidades del servicio, la jurisprudencia va a ser tratada sobre la forma de un simple resumen, de un abstracto y texto integral o de un simple abstracto.

El logicial pluri doc ha previsto una nueva técnica que autoriza la captura del texto integral eliminando los problemas de ruido que se ponen al momento de la investigación. Una función permite declarar el texto integral parcialmente indizable solas las palabras o las frases entre dos paréntesis rectangulares serán claves de acceso al documentos y podran ser interrogadas. Este permite extraer los conceptos esenciales de un texto y ganar al mismo tiempo el texto integral en la edición. Se elimina así todas las palabras que no son esenciales para la recuperación del documento.

El usuario tiene luego una total autonomía en la determinación del banco de datos.

### B) de la dificultad a la facilidad

Los bancos de datos sobre grande sistemas suponen la asistencia de un especialista en informática (porque todo almacenamiento de datos nuevos necesita números tratamientos informáticos) y de un lingüista (puesta el día del léxico o de un glosario).

Estaba imperativo resolver estos dos problemas cuando creamos el logicial sobre micro computador porque no se podía pensar que un jurista fuese

dispuesto para una formación informática y lingüística para crear su propio banco de datos.

### 1) la informática al alcance de todos

El acceso a diferentes funciones del logicial ha estado simplificado al máximo. Se hace con menús en cuales el usuario elige la función con números desplegados en la pantalla. Cada función tiene una asistencia con mensajes desplegados abajo de la pantalla o con ventanas que el usuario hace aparecer si lo quiere y que dan informes mas completos.

Solas las claves de acceso que protejan el banco de datos tienen que ser memorizadas. Todas las etapas que pueden tener efecto sobre el contenido documental son protegidas: creación de las estructuras, captura, modificación, supresión. La consulta es en libre acceso.

Cuando la estructuración de las bases esta echa, el jurista no tiene mas que capturar los documentos en las bases concernadas

Cuando los documentos han estados capturados, sus indización (las palabras toman dirección y vuelven interrogables) puede se hacer inmediatamente o en diferido. Cuando esta operación esta acabada, los documentos son interrogables. Así todas las operaciones de almacenamiento son automatizadas.

Esta automatización fue autorizada por la creación de un instrumento lingüístico original.

### 2) Un instrumento lingüístico original

La elección del lenguaje abierto permite el tratamiento de fuentes documentales las mas variadas (técnicas literarias jurídicas científicas). En cambio, supone una organización lingüística del vocabulario particularmente rigurosa.

Se necesita saber que todas las palabras escritas en un campo declarado indizable son claves de acceso al documento. En la mayor parte de los sistemas documentales, el tratamiento de los variaciones ortográficas de las palabras necesita la intervención de especialistas.

La originalidad de pluri-doc se pone en la elección de un léxico automático que puede tomar en cuenta ciertas flexiones lingüísticas.

Las posibilidades de recurso a un léxico clásico o a una truncatura en momento de la pregunta han estadas sucesivamente abandonadas en provecho de un modo mas adaptado y mas eficaz la canonización.

Se podrá usar las palabras canonización o lematización para hablar de este modo de truncatura inteligente.

#### El sistema del léxico manual clásico

En la mayor parte de bancos de datos tradicionales, la comunicación entre el usuario y los archivos, la recuperación de los documentos se hace por un léxico. Aquel organiza rigurosamente la totalidad de las palabras incluidas en los ficheros, eliminando el obstáculo constituido por la sintaxis y la semántica, para la función investigación.

El léxico vela por la concordancia entre los términos de la pregunta y los contenidos en el corpus tratado. Esta concordancia se hace con ayuda de un archivo constituido por números de identificación de los términos del léxico y de los diferentes referencias de los documentos que los contienen. De esta manera, las variaciones gramaticales y allotaxicas de una misma palabra, cuando no dan problema para la busca, son reagrupadas en una misma noción: cesión es cesionario cesionaria cedente cesionista.

Sin embargo, el enriquecimiento del vocabulario ligado al crecimiento del corpus pone el problema del tratamiento de las nuevas en el archivo

Aquellas tiene que ser incorporadas al léxico y esta operación de actualización supone un labor permanente que no se puede, en el caso de un banco de datos personal sobre micro computador, imponer al usuario.

La mantención de un léxico, característica de un labor largo y delicado (reagrupamiento de los variables, constitución de las nociones lexicales, aislamiento de la rupturas semánticas) es incompatible con la micro informática fundada sobre la flexibilidad la ergonomía y la rapidez de ejecución.

Ahora, la idea es aparecida de establecer un léxico automático fundado sobre una truncatura sistemática en función de las variaciones gramaticales o ortográficas de las palabras.

El logicial pluri doc consagra un método inspirado de la truncatura clásica pero sin embargo mas eficaz en la lucha con el silencio.

#### El sistema de léxico automático

- Esta técnica interviene a priori sobre las palabras capturadas y sin intervención del usuario. Funciona de la izquierda a la derecha y reduce las palabras a una forma normalizada forma canónica. Sin embargo, a la inversa de la truncatura clásica que produce sus efectos de manera ciega, la canonización que se puede

calificar de truncatura inteligente, toma en cuenta las variaciones gramaticales de la palabra para reagrupar la totalidad de sus formas sobre la misma raíz.

Este método es fundado sobre una estudio previa de la gramática española. Las terminaciones regulares o irregulares, que parecen las mas incidentes en la formación del vocabulario, presiden a la elaboración de las normas de canonización.

Ejemplo

organiz/ ar

organiz/ ado

organiz/ ada

organiz/ ados

organiz/ adas

organiz/ ando

organiz/ acion

organiz/ acciones

etc....

Hacen: organiz

En este ejemplo, se puede observar que el plural es reducido al singular, el femenino al masculino, la conjugación del verbo a su radical, el adverbio a su sustantivo y el adjetivo verbal al radical del verbo.

Las normas establecidas permiten tratar la casi totalidad de la lengua, las excepciones son tratadas aparte ejemplo rey reina reyes reinas.

Pero si la canonización reduce el riesgo de silencio, no resuelve sola el riesgo de ruido. En muchos casos, palabras semánticamente distintas van ser reducidas, con la truncatura automática, a una forma canónica idéntica.

Ejemplo: sol/ ido sol/ ida y sol/ sol/ es y sol/ o sol/ a adulter/ o a y adulter/ acion acciones

Este problema ha constituido el principal obstáculo del desarrollo de indización automático de los documentos.

Para resolverlo, una estudio ha estado hecha que tenia para objetivo de enumerar todas las palabras fuente de ambigüedad debido a la truncatura automática. Para hacer lo, una confrontación de las normas de canonización al totalidad del vocabulario de la lengua española estaba necesaria. Ha sido posible con el uso de un diccionario inverso que clasifica las palabras por sus terminaciones según una lectura a tergo. La canonización interviene sobre las terminaciones, luego un diccionario clásico estaba en este caso insuficiente y inadaptado.

Cuando todas las palabras fueran enumeradas, estas palabras fueran disociadas, en el sistema, del vocabulario general y tratadas en un archivo distinto para poder sustraerlas, tanto al momento de la indización, como al momento de la pregunta, de la fase de truncatura automática.

Las normas de truncatura inteligente y el tratamiento de los excepciones (palabras que no tienen que ser truncadas) resuelven totalmente los problemas de las variaciones gramaticales de los substantivos adjetivos o adverbios y algunos verbos.

El tratamiento de los verbos irregulares, para la versión española, ha sido dejado en espera por el momento. En efecto, la interrogación del corpus documental se hace generalmente con conceptos y no con verbos conjugados por cuales el tratamiento lingüístico sería muy pesado.

La canonización no pretende resolver todos los problemas lingüísticos de todos los sistemas documentales. Es una respuesta precisa a un problema preciso: el tratamiento del lenguaje natural por un no especialista. Ha estado concebida y realizada en función del sistema todo entero. así, combinada con las técnicas de gestión de los archivos y la eficacia de los métodos de interrogación, autoriza la creación de bancos de datos en lenguaje natural. Así el usuario no hace ninguno tratamiento lingüístico.

El usuario no tiene que preguntarse a la interrogación si hay que aplicar una truncatura. Todo el vocabulario es manejado a priori, con normas precisas y con un tratamiento de los excepciones. Este trabajo de base, muy importante, permite liberar el usuario de un problema delicado y dirimente para el uso personal de bancos de datos jurídicos.

Pluri-doc reduce al mínimo las tareas que incumben al creador del banco de datos. Hemos deseado ir mas lejos y crear un instrumento documental que no sea solamente un medio de almacenamiento de la información.

### C) de la rigidez a la flexibilidad

Los bancos de datos están mucho tiempo estáticos, cuando un documento esta memorizado, tiene su forma definitiva. Ahora bien, ciertas fuentes documentales pueden ser evolutivas, así la ley puede ser abrogada o modificada. El jurista tiene que poder intervenir sobre el documento de origen para notar estas modificaciones.

Tiene que poder archivar documentos antiguos, tiene que poder agregar comentarios cuando conoce la solución de un litigio.

La gestión del corpus documental debe presentar una larga flexibilidad. La consulta de esta documentación hay también que ofrecer un máximo de posibilidades.

### 1) banco de datos de geometría variable

Cuando un banco de datos es albateada después de muchos años el número de respuestas es abundante, este puede ser molesto; todos los documentos en repuesta son pertinentes pero ciertos son antiguos y no presenten que poco interés para el servicio jurídico. Lo mismo, en un banco de datos tratando de la legislación, es superfluo, incluso molesto guardar una ley abrogada. Así hemos previsto la posibilidad de suprimir de la función de busca ciertos documentos que van a dormir en el banco.

Sin embargo, si para un problema particular, el jurista desea consultar documentos archivados, el logicial da la posibilidad de restaurarlos de ponerlos en línea en el momento de la pregunta y volver a suprimirlos después.

Esta solución permite disminuir la función investigación del sistema y no pedir definitivamente los documentos antiguos, autoriza una buena gestión de la masa documental que es muy importante sobre un micro computador.

Ciertas fuentes documental, como la ley, la jurisprudencia son evolutivas. Hemos pensado que estaba capital poder intervenir sobre documentos ya capturados para tomar en cuenta todas las modificaciones sobre el mismo documento.

Estas modificaciones pueden ser una simple corrección ortográfica o una modificación sobre el contenido jurídico: supresión de un párrafo y reemplazo por otro en el caso de una modificación de una ley por ejemplo. Añadido de un comentario personal sobre un documento de jurisprudencia, añadido de referencias sobre un documento de corte de apelación, etc.

Hemos creado luego una función modificación que permite intervenir sobre el documento cual que sea el momento después la captura o cuando el documento fue ya indizado.

Con estas dos funciones, el banco de datos se vuelve un instrumento documental dinámico que se adapta constantemente a las necesidades del jurista y a las variaciones de cada fuente documental.

Un banco de datos perfectamente concebido eficacia al momento de la consulta.

### 2) una función de investigación y de edición de una gran flexibilidad

Los resultados de un sistema documental dependen de la multiplicación de los modos de acceso a la información.

El acceso a los documentos en pluri doc se hace por intermedio de palabras que contienen (en campos declarados indizables o parcialmente indizables).

Hemos visto ya (instrumento lingüístico: la canonización) que las variaciones gramaticales son tratadas automáticamente.

Así, si en los termos de la pregunta una palabra es al femenino plural y si en el documento exista al masculino singular, el documento saldrá en respuesta.

El usuario no tiene que preocuparse de la forma precisa de la palabra que escribe para formular su pregunta

Puede escribir todos los conceptos jurídicos (grupo de palabras teniendo valor semántica) el logicial combina automáticamente las palabras al nivel de proximidad constituido por la frase.

Ejemplo: contrato de trabajo a duración determinado

Todos las palabras o conceptos de la pregunta pueden ser después combinados entre ellos con los operadores booleanos y los operadores de distancia (and or but anp bup buf la combinación al nivel frase es implícita).

A estos modos de interrogación bastante clásicos vienen se agregar posibilidades mas originales la interrogación sobre una sola base la interrogación sobre un solo campo y ambas combinadas.

La presencia de varias fuentes documentales (bases) en el mismo banco de datos necesita un instrumento que puede orientar las investigaciones sobre una sola de estas bases si el problema lo permite.

Así, anteriormente a la formulación de la pregunta, el sistema propone la opción de interrogar todas las bases o una sola. Esta opción no esta definitiva; a todo momento durante el periodo de investigación, el usuario puede programar la ejecución de su pregunta sobre una otra base. Puede tomar en cuenta sus ficheros base después base, sobre un punto de derecho dato y volver después a una interrogación global si lo desea.

Esta función es muy importante para un micro computador porque aumenta la rapidez del tiempo de respuesta, los accesos al documentos estando selectivos. Permite también de encontrar muy rápidamente una información precisa concernando una fuente documental especifica.

A esta función se agrega la investigación sobre un campo, se puede exigir que los términos de la pregunta sean buscados en un campo preciso. Esta es muy útil en el caso de informaciones objetivas tales las fechas. En efecto, para una fuente documental puede existir varias fechas muy interesantes así para la ley, su fecha, la fecha de su decreto de aplicación, la fecha de sus modificaciones o de sus abrogación; si estas fechas son en campos diferentes será muy fácil, a la pregunta de poner la cuestión sobre un solo campo (ex campo modificación) y no existirá el riesgo de confusión con otras fechas figurando en el mismo documento o en otro.

La interrogación sobre un campo permite individualizar la clave de busca y dirigirla.

La combinación de estas dos opciones da a la investigación su pertinencia máxima. Así si pedimos base doctrina y campo autor para una pregunta sobre un nombre propio, estamos seguros que tendremos en respuesta todos los artículos escritos por este autor.

Asimismo, si elegimos base ley y campo promulgación para la pregunta 24 julio 1966 no tendremos en respuesta que las leyes promulgadas en esta fecha y no los decretos o las modificaciones hechas a la misma fecha.

A todo momento del periodo de interrogación, el usuario puede anular las opciones tomadas para su pregunta para volver a una investigación clásica sobre la totalidad del corpus. El interés de esta manipulación es poder comparar los resultados obtenidos para una misma investigación sobre bases diferentes.

Además, durante todo el periodo las diferentes preguntas quedan desplegadas en la pantalla; esta permite formulaciones diferentes para la misma pregunta y una comparación de los resultados. Esta función ofrece posibilidad a todo momento de volver a una pregunta anterior para ver los documentos o para modificarla.

Por último, el logicial autoriza una edición al momento de la visualización o una edición directa sin ver a los documentos.

Pluri doc es un logicial de creación de bancos de datos sobre micro computador concebido para autorizar bancos de datos jurídicos a la medida; elimina todos los problemas informáticos y lingüísticos. Presida al nacimiento de nuevos instrumentos documentales personalizados y dinámicos.

En la perspectiva europea de 1992, IRETIJ a desarrollado una versión inglesa del logicial con la elaboración de un instrumento lingüístico perfectamente adaptado a la lengua inglesa (canonización en inglés) y esta desarrollando una versión española de pluri doc; las normas de truncatura son

establecidas y testadas las excepciones son enumeradas, se necesita escribirlas en el logicial.

Se puede imaginar la coexistencia en el mismo servicio de varios bancos de datos sobre el mismo tema jurídico en lenguas diferentes. Este tipo de experiencia es actualmente realizado para el servicio del derecho de autor a UNESCO. El mismo banco de datos sobre las legislaciones nacionales los convenios internacionales y las constituciones existe en francés en París en ingles en Washington y va a ser creada en español en Colombia.

"El derecho es la mas poderosa escuela de la imaginación (dice Giraudoux un autor francés ) porqué no dar a esta imaginación un instrumento a su medida".