

CLAUDE BELAIR | SERGIO MATUTE

LAT 385 GUILLERMO AGUILAR

AUTOMATIZACION  
DE LA  
LEGISLACION  
MEXICANA

EL SISTEMA

UNAM-JURE



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

INFOBILA

CLAUDE BELAIR SERGIO MATUTE  
GUILLERMO AGUILAR

AUTOMATIZACIÓN  
DE LA  
LEGISLACIÓN  
MEXICANA

EL SISTEMA

**UNAM-JURE**

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

INFOBILA

No. Lat. 060385  
No. Adq. \_\_\_\_\_  
No. Sist. \_\_\_\_\_  
Tipo de Adq. Donación  
Fecha 23 Sept 2011

INSTITUTO DE INVESTIGACIONES JURÍDICAS

Serie E. Varios. Núm. 22

Primera edición: 1983

D.R. © 1983, Universidad Nacional Autónoma de México  
Ciudad Universitaria. Delegación Coyoacán. C.P. 04510.  
México, D.F.

INSTITUTO DE INVESTIGACIONES JURÍDICAS

## PRÓLOGO

El Instituto de Derecho Comparado, hoy Instituto de Investigaciones Jurídicas de la Universidad Nacional Autónoma de México, fue creado en 1940 y desde 1942 se integró una sección de legislación que ha venido recibiendo, sin interrupción, los diarios oficiales tanto de la federación como los de las entidades federativas de México, de los diversos países latinoamericanos y de algunos europeos. Todo ese material se clasifica y del mismo se ha dado y se da noticia por medio de diversas publicaciones de este Instituto como son: de 1958 a 1967 a través del *Repertorio Anual de Legislación Nacional y Extranjera*; en 1968 a través del *Repertorio Anual de Legislación Nacional*; a partir de 1972 a la fecha a través de la *Gaceta Informativa de Legislación y Jurisprudencia* y también a través del "Alerta de Legislación", desde 1980.

Hoy, el Instituto de Investigaciones Jurídicas cuenta con un fichero sobre legislación mexicana, latinoamericana y de algunos países europeos, que contiene más de un millón de fichas, y que constituye parte muy valiosa del acervo del Centro de Legislación y Jurisprudencia de esta dependencia universitaria.

Fácil es comprender los problemas de almacenaje y fichaje de toda esa documentación; por ello, se deci-

dió que ya era hora de comenzar un programa para automatizar la documentación mencionada, con el ánimo de que el manejo de ese material sea más fácil, y su conservación esté asegurada para el futuro.

En consecuencia, se solicitó al personal técnico del Centro de Documentación del Instituto, la preparación de un proyecto de automatización de la legislación mexicana. Dicho personal, con gran entusiasmo, se dio a la tarea de preparar el proyecto, el cual fue discutido con la Dirección y con la Secretaría Académica, con la que acuerda el jefe del citado Centro. Cuando se tuvo una idea clara del proyecto, se contempló que lo primero que era necesario, era la preparación del personal académico que tendría la responsabilidad del programa. Dos técnicos académicos del Instituto, el licenciado Claude Belair, jefe del mencionado Centro y responsable del programa de automatización, y el señor Guillermo Aguilar Álvarez, viajaron a Montpellier, Francia, en donde se perfeccionaron en técnicas de recuperación automática de información jurídica.

Con el "Institut de Recherches et d'Etudes pour le Traitement de l'Information Juridique" (IRETIJ) se firmó un convenio de cooperación que ha sido muy útil para el diseño del mencionado sistema de automatización.

En las etapas de diseño y desarrollo de este proyecto de automatización, hemos contado con la valiosa colaboración del Centro de Servicios de Cómputo de la UNAM. El personal académico, especialmente el ingeniero Sergio Matute, de dicho Centro, se ha encargado de la realización de UNAM-JURE desde el

punto de vista técnico y ha trabajado estrechamente con el personal del Instituto de Investigaciones Jurídicas.

En la actualidad, el programa es ya una realidad y constituye un éxito; se han automatizado aproximadamente 5000 documentos de la legislación federal pertenecientes al período 1977-1982, y durante 1983 se realizará la automatización de la legislación de las entidades federativas correspondiente al mismo período.

La ejecución de un proyecto de esta naturaleza tiene problemas y complicaciones; sin embargo, nadie discute que estos proyectos son indispensables en los centros de investigación y documentación jurídica. La investigación del presente y más aún la del futuro, se basa en mecanismos automatizados como un apoyo invaluable para la propia investigación.

Con ánimo fraternal y con la finalidad de intercambiar experiencias, los técnicos académicos, Claude Belair, Guillermo Aguilar Álvarez y Sergio Matute, han redactado este folleto para hacer público lo que es hoy el programa UNAM-JURE del Instituto de Investigaciones Jurídicas. A ellos se debe primordialmente la existencia de dicho programa, por lo que les manifiesto el agradecimiento de esta dependencia universitaria y el mío propio.

Hago votos porque la finalidad que se busca con este folleto, para la divulgación de nuestro programa de automatización de la legislación mexicana, se consiga y que al conocerse los esfuerzos que en este sentido está llevando a cabo el Instituto de Investigaciones Jurídicas de la UNAM, los resultados del proyecto

puedan ser útiles no sólo al personal académico del Instituto, sino a todos los estudiosos del derecho mexicano.

JORGE CARPIZO  
Director del Instituto de Investigaciones  
Jurídicas de la UNAM

## *Introducción*

El proyecto de automatización al que haremos referencia se ha desarrollado en su totalidad dentro de la Universidad Nacional Autónoma de México con la participación conjunta del Instituto de Investigaciones Jurídicas y del Centro de Servicios de Cómputo.

Los antecedentes de UNAM-JURE se remontan al año de 1980, fecha del primer proyecto sometido a la consideración de las autoridades competentes del Instituto de Investigaciones Jurídicas. En esta época el Centro de Información de Legislación y Jurisprudencia de nuestro Instituto manejaba en forma manual un acervo de legislación integrado por los diarios o publicaciones oficiales de la federación mexicana así como de todas las entidades que la integran, además de recibir las publicaciones oficiales de 21 países de distintas partes del mundo (Europa, América Latina, Estados Unidos, etcétera). Lo anterior representa un total aproximado de 9,000 ejemplares al año que, a su vez, significan, sólo en relación con la legislación mexicana, una cantidad de más de 100,000 documentos jurídicamente relevantes para el período 1918-1982. La función primordial del Centro de Información de Legislación y Jurisprudencia dentro del Instituto de

Investigaciones Jurídicas es la de servir de apoyo a la investigación que se realiza en el propio seno del Instituto, así como para la libre consulta de universitarios y del público en general. En tal virtud toda la información legislativa que ingresa al acervo del Centro es objeto de selección, clasificación y análisis de contenido por parte de los técnicos académicos de nuestra institución. Es importante señalar que el Instituto de Investigaciones Jurídicas viene realizando esta labor, de manera ininterrumpida desde el año de 1942, y gracias a un reciente proyecto, se logró incorporar al material analizado y clasificado el período 1918-1942 de la legislación federal mexicana. Si tomamos la legislación mexicana (estatal y federal) y la extranjera, observamos que a partir de 1942 se han acumulado en los tarjeteros del Centro alrededor de 1'000,000 de tarjetas de referencia elaboradas en forma manual. Si consideramos la masa de información generada, clasificada y consultada diariamente, el procedimiento para la investigación toma una nueva dimensión. Aun en los ámbitos más específicos, una persona nunca puede tener los elementos indispensables para la investigación a la mano. Ya no se trata de tener una respuesta sino de comenzar por saber dónde encontrarla. Así, los métodos de investigación manual ya no son suficientes en la recuperación precisa de la información legislativa que producen tanto nuestro poder legislativo federal como las legislaturas de los 31 estados de la República mexicana y el Distrito Federal. En virtud de la imposibilidad física de realizar el trabajo anterior en forma eficiente y sin pérdida de información pertinente, presentamos un segundo pro-

yecto que fue aprobado en el año de 1981, y en el que se planteó la necesidad de establecer relaciones de cooperación con el *Institut de Recherches et d'Etudes pour le Traitement de l'Information Juridique* (IRETIJ) de Montpellier, Francia. A raíz del contacto con dicho Instituto francés se realizaron 2 viajes a Francia para conocer el sistema del I.R.E.T.I.J. y trabajar de manera más estrecha con sus investigadores. Posteriormente, dos investigadores y el fundador del IRETIJ viajaron a México para ultimar los detalles de lo que habrá de ser un convenio de cooperación científica a partir de 1983.

Aunque todas ellas de tipo académico, varias son las razones que acercaron al Instituto de Investigaciones Jurídicas al IRETIJ. Además de la disponibilidad mutua para cooperar, coincidimos en ser formaciones de tipo académico inmersas en la estructura jurídica de universidades, tenemos la ventaja de hablar lenguas latinas, nuestros sistemas jurídicos provienen de una misma familia y compartimos un gran interés en la informática jurídica documental.

Por otra parte, la problemática que presenta la automatización de la legislación mexicana, que sólo puede resolverse con la utilización de un lenguaje de análisis libre, nos impide trabajar con el sistema conocido como *full text* (que prescinde de la explicitación) y nos obliga al diseño de ciertos instrumentos lingüísticos que permitan la organización de un lenguaje difícil de clasificar como lo es el jurídico. Estas son otras razones por las cuales Investigaciones Jurídicas decidió acercarse al IRETIJ.

Para efectos de claridad, hemos dividido este traba-

jo en dos partes; la primera se ocupa de los problemas relativos al análisis de contenido y de lenguaje de los documentos legislativos y jurisprudenciales, la segunda hace una breve exposición de los elementos más importantes dentro del *software* de UNAM-JURE.

## I. Problemas de análisis de contenido y de lenguaje

A pesar de las similitudes e intereses comunes apuntadas, el caso de la legislación mexicana presenta algunas particularidades que alejan a UNAM-JURE del sistema del IRETIJ. El instituto francés se ocupa principalmente de la jurisprudencia francesa (Corte de Casación o diversas cortes de apelación) mientras que UNAM-JURE arranca con la legislación federal y estatal mexicana a partir de 1917 hasta la fecha. Resulta evidente que los tipos de documentos a analizar y recuperar son completamente distintos, lo que nos ha llevado a modificar y adaptar ciertas partes del sistema de la institución francesa. En primer lugar analizaremos los problemas específicos que plantea la realidad jurídica mexicana y que, de alguna manera, nos obligan a crear un sistema original, el cual presentaremos al final.

### A. Problemas específicos de la realidad jurídica mexicana

Por diversas razones que apuntaremos más adelante, resulta evidente que una sentencia de la Corte de Ca-

61060  
400282

sación francesa o de la Suprema Corte de Justicia mexicana y la Ley General de Sociedades Mercantiles de México tendrán que ser tratadas de manera diferente. Podemos afirmar que, tanto en la forma como en el fondo, las disposiciones que producen un juez y un legislador son muy distintas. En cuanto a la forma, mientras las leyes o reglamentos contienen invariablemente títulos, libros, capítulos, secciones u otras formas de división que para efectos de análisis y recuperación es importante conservar, la jurisprudencia, sólo excepcionalmente, contiene encabezamientos y éstos, por lo menos en el caso mexicano, nunca son asignados por el propio juez que dicta la sentencia. Además, la estructura de ambos tipos de documento es diferente, la de las decisiones de justicia es constante, el juez considera los elementos de hecho, los analiza a la luz de las disposiciones jurídicas aplicables y produce una resolución. Esto no sucede en el caso de la legislación debido a la diversidad de disposiciones que puede generar el poder legislativo. Así, encontramos que existen leyes, reglamentos, decretos, acuerdos, circulares, oficios, etcétera, y todos ellos, en mayor o menor medida, gozan de fuerza obligatoria. En este orden de ideas, resulta difícil encontrar una estructura constante en la construcción legislativa. En relación con la extensión de los textos, no son menos evidentes las diferencias, pues las ejecutorias, en promedio, nunca rebasan un máximo y la extensión de los textos legislativos depende en gran parte de la naturaleza del documento en cuestión. En relación con el fondo, encontramos que las diferencias son aún más graves. El lenguaje del legislador

prescinde de los elementos de hecho a que tiene que referirse un juez y, por definición, es más abstracto. Esto significa que, si bien ambos tipos de documento contienen consideraciones de derecho, la aparición de cuestiones fácticas es propia de la jurisprudencia.

Si consideramos que la labor del juez es primordialmente interpretativa, entenderemos por qué los conceptos jurídicos son más explícitos en la jurisprudencia que en la legislación. Tomemos un ejemplo típico; el artículo 1910 del Código Civil para el Distrito Federal contiene el concepto de responsabilidad civil y dice lo siguiente: "El que obrando ilícitamente o contra las buenas costumbres cause daño a otro, está obligado a repararlo, a menos que demuestre que el daño se produjo como consecuencia de culpa o negligencia inexcusable de la víctima". Si revisamos cuidadosamente el texto del Código Civil, observamos que de él no se desprende expresamente el concepto de la responsabilidad civil. Sin embargo, el juez que califica la conducta de una persona como ilícita o contra las buenas costumbres y que establece el nexo causal entre tal conducta y el daño causado a otra persona sin que medie culpa o negligencia de ésta, necesariamente tendrá que fincar la responsabilidad civil del sujeto haciendo referencia expresa al citado concepto.

Admitido el carácter interpretativo de la jurisprudencia, no es difícil establecer la diferencia primordial entre ésta y la legislación. A las abstracciones de la ley se opone la concretación de la jurisprudencia. Mientras ésta se ocupa de un problema jurídico en particular, aquélla tiene carácter heterogéneo. Esto es de gran importancia para efectos de la obligatoriedad de ambas.

En el caso mexicano la obligatoriedad de la jurisprudencia se desprende de la Constitución Política federal y de la Ley de Amparo (Reglamentaria de los artículos 103 y 107 constitucionales). Sin entrar en detalles, podemos afirmar que el artículo 94 de la Constitución Política federal y los artículos 192 a 197 de la Ley de Amparo establecen que la jurisprudencia de los tribunales federales superiores será obligatoria para los inferiores. Además de esto, para poder hablar de jurisprudencia obligatoria es necesario que existan cinco ejecutorias en el mismo sentido, sin existir una que interrumpa la acumulación de las anteriores con sentido diverso. La legislación, por otra parte, deriva su obligatoriedad de la observancia de un procedimiento legislativo y de su publicación en el *Diario Oficial* correspondiente. Las tesis relacionadas o ejecutorias aisladas de los tribunales sólo tienen efectos vinculatorios para las personas que fueron parte en el juicio de que se trate. La jurisprudencia, como hemos visto, será obligatoria en los términos de los preceptos legales citados. El ámbito de aplicación personal de la legislación depende, una vez más, del documento base. Las leyes federales son obligatorias en toda la república, las estatales sólo dentro del estado del que emanan; existen circulares destinadas a ciertos funcionarios para su cumplimiento, la observancia de algunos reglamentos particulares se exige exclusivamente a miembros de una dependencia de gobierno, y así podemos mencionar muchos otros ejemplos donde la obligatoriedad de la legislación se restringe o se amplía.

Por último, vamos a referirnos al problema de la

temporalidad de la jurisprudencia y de la legislación. Mientras la jurisprudencia no está sujeta a reformas (existe una excepción tratándose de contradicciones entre diversas tesis de jurisprudencia a cuyo análisis no entraremos), es éste precisamente el instrumento propio del cambio legislativo. Tanto las leyes como los reglamentos y algunos otros tipos de documentos legislativos (*i.e.*, decretos o acuerdos) están sujetos a reformas que al publicarse en el *Diario Oficial* pasan a formar parte del texto original y que, por tanto, es necesario conservar y recuperar como tal.

Aun entre la jurisprudencia mexicana y la francesa existen importantes diferencias. En el caso mexicano no siempre se publican las ejecutorias mismas de los ministros sino resúmenes o tesis elaborados por empleados del *Semanario Judicial de la Federación*, órgano encargado de la publicación de la jurisprudencia. Estas tesis pueden, y frecuentemente así sucede, cambiar el sentido u omitir elementos importantes de la ejecutoria que les da origen. Si a este problema agregamos el de la extemporaneidad en la publicación de las tesis (hasta ahora este retraso es de dos años y medio) y el de su gran extensión, podremos comprender las dificultades que se plantean para la jurisprudencia a un sistema originalmente diseñado para acoger legislación.

Una vez analizados todos estos problemas y particularidades llegamos a la conclusión de que no era posible aplicar de manera absoluta el sistema IRETIJ a la realidad jurídica mexicana. Tomando el sistema francés como base de reflexión hemos introducido importantes adaptaciones y modificaciones, que hacen del

nuestro un sistema original en cuanto al *software* y diferente en lo que se refiere al método de análisis y los instrumentos lingüísticos. Es decir, se pensó en un *software* original que pudiera operar no sólo con legislación sino también, en una etapa posterior, con jurisprudencia mexicana y doctrina.

## B. Descripción general de UNAM-JURE

### 1. Análisis de contenido

El análisis de contenido, tanto de textos legislativos como jurisprudenciales o doctrinales, es un proceso indispensable de todo sistema de información. Sin embargo, cuando la informática jurídica penetra en los sistemas de información tal exigencia adquiere aún mayor relevancia. Los textos susceptibles de recuperación automática deben ajustarse a ciertas reglas de análisis que a continuación apuntamos.

Como punto de partida hemos diseñado un formato de captura de datos propio de UNAM-JURE, cuyas principales características y técnicas se describen más adelante.

Los problemas propios del análisis de contenido varían según se trate de textos legislativos, o jurisprudenciales. En los primeros hemos encontrado que el respeto al lenguaje textual del legislador es necesario, pues los títulos o conceptos utilizados por él son generalmente aquellos que recoge el mundo jurídico (estudiantes, maestros, jueces, investigadores, abogados y funcionarios) y de ahí la utilidad de conservarlos. Sin

embargo, para efectos de una recuperación exhaustiva de los textos pertinentes deseados, el respeto al lenguaje del legislador no resulta suficiente. Por esta razón, la explicitación de conceptos en materia de legislación debe ser aguda. Esto significa que si nos encontramos analizando el artículo 29 de la Constitución mexicana, tendremos que hacer referencia a conceptos como estado de emergencia o facultades extraordinarias del presidente de la República para legislar, aunque no se desprendan expresamente del texto de tal precepto.

La zona más importante del formato de captura está destinada al *abstract* o análisis de contenido del texto correspondiente. Dentro del *abstract* encontramos la información ordenada de acuerdo con ciertos criterios. El paso de una fase a otra debe hacerse respetando un orden lógico (Por ejemplo: /MONEDA/PESO MEXICANO, EQUIVALENCIA CON MONEDAS EXTRANJERAS/) y el cambio de un párrafo a otro debe obedecer a la aparición de un problema o concepto jurídico diferente. Con lo anterior abandonamos la idea tradicional de que debe existir un concepto jurídico por frase y simplemente imponemos ciertas restricciones a la longitud de frases y párrafos. Fuera de estas limitaciones mínimas, el analista está en completa libertad para redactar y utilizar aquellas palabras o conceptos que le parezcan más apropiados.

Hemos tratado, por otra parte, de uniformar el análisis en aquellas materias que lo permitan. Es así como hemos diseñado estructuras fijas de análisis para las disposiciones legislativas que se repiten idénti-

cas en períodos cortos de tiempo (*i.e.*, presupuestos de egresos federal y estatales).

Para terminar con el análisis de contenido, diremos que hay una indicación muy sencilla, pero de la cual depende el éxito del sistema. Siempre debe indicarse el *status* de la disposición objeto de análisis. Es decir, si nos encontramos con leyes o reglamentos nuevos, la frase que los describe debe expresar tal condición (*i.e.*, Ley de Fomento Agropecuario, publicación). Si, por el contrario, estamos frente a un decreto de reformas, la frase correspondiente del *abstract* debe advertirlo (*i.e.*, /Ley de Fomento Agropecuario, artículo 10, reforma/). Esto nos permite obtener, al momento de la interrogación, el texto original y sus reformas, solamente aquél o exclusivamente éstas.

## 2. Instrumentos lingüísticos

Hemos encontrado que los principales problemas lingüísticos en la investigación documental son los de analogía y polisemia. La analogía supone la existencia de un significado que puede ser descrito por varias palabras (significados) diferentes y sin tratarse de equivalencia total (sinonimia). Tal es el caso de palabras como penitenciaría, cárcel y prisión. A la inversa, con la polisemia se agrupan varios significados en un mismo significando como sucede con la palabra acción (tiene significaciones diferentes en lenguaje coloquial, en materia procesal y en derecho mercantil).

Otro problema es el que produce la construcción morfológica de las palabras. Así, resulta que si deci-

mos "delegados sindicales" o "sindicato, delegados" o "delegados del sindicato", las palabras sindicales y sindicato se utilizan en el mismo sentido, de tal suerte que es indispensable recuperar los documentos en que aparezcan "delegados" y "sindicato" aunque la interrogación se haga por "delegados" y "sindicales".

La solución de estos problemas lingüísticos la hemos intentado con el diseño y empleo de dos instrumentos lingüísticos. Estos instrumentos son el léxico y el *thesaurus*. En términos generales, aquél se ocupa de organizar las palabras siguiendo criterios semánticos (significado) y morfológicos (forma ortográfica). Es decir, encontramos que en el léxico se organizan familias de palabras tales como; "administración-administrativa-administrativos-administrativamente".

En el *thesaurus* se vuelen a tomar esas palabras y se desarrollan semánticamente (es decir, el *thesaurus* nos indicará que administración tiene como análoga la palabra gestión).

UNAM-JURE se aleja del sistema del IRETIJ desde que otorga más importancia a la semántica que a la morfología en la construcción de sus instrumentos lingüísticos. Nos interesa más saber que "abono" es análoga de "fertilizante" y no aislar completamente ambas en razón de su procedencia morfológica diversa. En este orden de ideas, UNAM-JURE introduce la idea de un instrumento lingüístico integral, en el que se construyan relaciones circulares entre las palabras sin atención a su tratamiento morfológico.

En su primera versión, el *thesaurus* de UNAM-JURE es una herramienta de auxilio a la interrogación de consulta opcional. Sólo informa al usuario sobre el

estado semántico (polisemia, analogía) de las palabras que utiliza al interrogar y corresponde a aquél decidir si lo consulta o no.

### 3. La interrogación

El objetivo primordial de UNAM-JURE ha sido el de facilitar a los usuarios la interrogación. Esto significa que la persona que lo desee podrá interrogar en lenguaje natural (*ie.*, utilizando cualquier forma verbal), no sin antes haber sido objeto de un perfil de interés en el que precisará las materias o los temas que le interesen.

Además, el proceso de interrogación de UNAM-JURE es completamente automático y permite la simple aparición de respuestas en pantalla o también su impresión en el formato que para tal efecto elija el usuario.

## II. Aspectos técnicos

El sistema UNAM-JURE tal como se ha planteado, queda definido como un sistema de recuperación de información, al cual le son planteadas solicitudes en forma de líneas de interrogación, tecleadas en una terminal remota. A las anteriores solicitudes, el sistema reacciona formando subconjuntos de fichas relevantes a la interrogación planteada. Como tal, el sistema puede ser calificado como recuperador de datos, es decir, que responde a una interrogación con el con-

junto de descriptores de los documentos donde se encuentra la información.

El sistema se ha desarrollado en forma progresiva con lo que ha sido posible lograr resultados parciales que permitan una mejor elaboración de las fichas descriptivas, así como la explotación parcial de la información disponible. El empleo de técnicas de programación estructurada y de documentación metódica, nos permitirá contar con un sistema de fácil mantenimiento.

Siendo un sistema desarrollado por y para la Universidad Nacional, está llevándose a cabo en el equipo Burroughs con el que cuenta el Centro de Servicios de Cómputo de la U.N.A.M., se ha hecho uso del lenguaje ALGOL cuya eficiencia es la mayor en este tipo de equipos.

En cuanto al diseño de los archivos que forman la base de datos, se ha hecho uso sustancial de las estructuras de datos que permitan un comportamiento veloz y confiable del sistema, sin hacerlo demasiado complejo.

En lo referente al diseño funcional del sistema, se ha intentado aplicar lo más avanzado en cuanto a técnicas y algoritmos, de forma que el desarrollo logrado sea flexible y capaz de brindar un comportamiento adecuado al manejo de la información de una manera rápida y económica en cuanto a memoria, tiempo de procesador y de entrada/salida.

#### A. Documento básico

El documento básico del sistema está constituido por la ficha de información, la cual a su vez está compuesta por dos zonas, la primera de ellas, es la zona de campos fijos, la segunda variable, es la zona de texto.

La zona de campos fijos, contiene la información descriptiva general de la ficha y comprende los siguientes campos.

- 1) Número de referencia: identificación que se asigna a cada una de las fichas realizadas.
- 2) Área geográfica: define de qué estado de la república procede el documento, o bien su condición de ser federal.
- 3) Tipo de documento: describe la calidad del documento analizado (reglamento, decreto, ley, convenio, bando de policía, etcétera).
- 4) Nombre de la publicación: publicación en la que apareció el documento analizado (*Diario Oficial, Periódico Oficial, etcétera*).
- 5) Alcance: condición de aparecido como alcance, extraordinario o suplemento.
- 6) Fecha de publicación: día en que apareció el documento en la publicación correspondiente.
- 7) Clave de microfilme: rollo y ubicación del documento en el archivo de microfilmes de que dispondrá el Instituto.

La zona de textos, a su vez, describe el contenido del documento fuente, de forma verbal. Se compone de dos elementos: resumen y *abstract*.

El resumen hace una descripción breve e informal del contenido del documento, de manera que el lector encuentre en él una guía al contenido general de documento.

El *abstract* describe de forma precisa los puntos de los que consta el documento fuente y los estructura en párrafos y frases que permiten una búsqueda exhaustiva y pertinente al momento de interrogación. El *abstract* está constituido por una serie de párrafos separados por puntos, los cuales a su vez están formados por conjuntos de frases, separadas por diagonales.

### B. Módulo de altas

El módulo de altas tiene por función realizar el ingreso de nuevas fichas de información al sistema, para lo cual dispone de submódulos que le permiten realizar la verificación de cada una de las fichas que componen la nueva carga, así como mantener control especial de las fichas rechazadas y generar las formas descriptivas pertinentes en cada uno de los archivos que forman la base de datos, además de realizar reportes sustanciales de fichas aceptadas, de fichas rechazadas y sus errores, y de estadística global y por ficha.

En cuanto al funcionamiento del módulo de altas, la parte medular del proceso es llevada a cabo por un autómatas de Mealy que realiza el análisis sintáctico de las fichas y cuyo alfabeto de salida controla las acciones necesarias para facilitar el resto del proceso.

A pesar de ser posible realizar la entrada de una o

varias fichas en forma interactiva a través de una terminal remota, la forma normal de ingreso de una carga es en forma de lote, fuera de línea, por lo que se han pretendido realizar aquí todos los procesos que faciliten y aceleren los procedimientos de búsqueda, a despecho del incremento en el tiempo de procesador utilizado, ya que éste puede ser consumido fuera de las horas de mayor carga de trabajo en el equipo.

### C. Módulo de edición y control

El módulo de edición y control permite realizar las tareas de modificación de las fichas existentes en la base de datos, su eliminación, la corrección e ingreso de fichas rechazadas por el módulo de altas, la entrada en forma interactiva de nuevas fichas o el inicio diferido del ingreso en forma de lote de una carga de fichas, la edición general de palabras en el léxico y todas y cada una de las acciones de edición y control necesarias en el sistema.

En cuanto a la edición general de palabras en el léxico, el editor cuenta con submódulos apropiados a permitir las acciones siguientes:

- entrada de nuevas familias de palabras al léxico.
- inserción de nuevas palabras en las familias existentes en el léxico.
- modificación o corrección de palabras.
- eliminación, tanto de palabras como de familias de palabras existentes en el léxico.
- emisión de listados de *kwic* (*key words in con-*

*text*), de una o varias palabras, tanto en la pantalla como en papel.

- creación de *thesauri* de las palabras.
- inserción en el *thesaurus* de analogías, antonimias, y en general toda variante pertinente.
- generación de los reportes necesarios (palabras en orden alfabético, por familias, etcétera).

En cuanto a acciones de control podemos mencionar como ejemplos, la creación de claves de acceso al sistema, otorgando diferentes privilegios a los distintos usuarios, la recuperación de fallas por factores externos, la protección general del sistema, la verificación del estado que guarda cada uno de los archivos del sistema, la emisión de estadísticas, etcétera.

#### D. Módulo de interrogación

El módulo de interrogación se ocupará de todo lo relacionado con la explotación del banco de datos por parte de un usuario; para tal efecto, consta de dos submódulos que cumplen las funciones de: control de terminal y búsqueda.

El submódulo de control de terminal realiza todo el proceso de análisis e interpretación de la petición formulada por el usuario a través del teclado de la terminal, así como las acciones encaminadas a permitir las siguientes facilidades:

- inicio y fin de una sesión de consulta.
- control de formato de despliegue en pantalla de

las fichas recuperadas como respuesta a una interrogación.

- control de un cierto número de conjuntos de fichas elaborados durante la sesión.
- clasificación de las fichas recuperadas en el orden que se desee.
- impresión de reportes al arbitrio del usuario.
- desplegado de puntos de interés encontrados en determinado documento con relación a una interrogación dada.
- conservación en disco de conjuntos elaborados durante una sesión para su uso en sesiones de consulta posteriores.
- despliegue de instrucciones para una situación dada a petición del usuario.
- acceso a la información contenida en el *thesaurus* de una palabra, etcétera.

Todas las facilidades enumeradas, se observan desde el punto de vista del usuario, en cuanto al sistema; este submódulo realiza todas las tareas de control de una red de terminales de consulta.

El submódulo de búsqueda se encarga de todo lo relacionado con la formación de un conjunto de fichas adecuado a partir de la consulta realizada por el usuario. La selección de las fichas se lleva a cabo de dos formas: la selección propiamente dicha, que realiza la búsqueda a partir del contenido del documento, representado en el *abstract* de la ficha, y la postselección basada en las características del documento, codificadas en la zona de campos fijos del documento de captura.

La selección es efectuada por medio de operaciones de conjunción, disyunción y diferencia sobre las palabras contenidas en el *abstract* de la ficha, estas operaciones son realizadas a dos niveles: frase y párrafo.

La división en frases y párrafos de una ficha de información permite ofrecer al usuario una mayor exactitud en la recuperación de documentos al momento de interrogación, situando las palabras en conjuntos con coherencia semántica bien definida. Por lo anterior se ha eliminado de la interrogación directa el nivel documento, a pesar de lo cual el usuario puede lograrlo a través de los conjuntos de trabajo que el sistema le ofrece.

Los instrumentos lingüísticos del sistema permiten realizar una selección basada en las palabras de modo literal, o extender la misma, vía léxico o *thesaurus*, a las distintas variantes gramaticales de la palabra, y aun a sus análogos a petición del usuario.

Una forma de recuperación de fichas por medio del número de referencia es también ofrecida al usuario.

### E. Futuro

La realización del sistema UNAM-JURE, aun cuando no está concluida, genera inquietudes y abre nuevas direcciones de desarrollo e investigación cuyos resultados se irán incorporando a futuras versiones del sistema.

El problema del lenguaje, que resulta central en la elaboración y uso de un sistema de este tipo, sobre todo en cuanto se refiere a la identificación de la rela-

ción semántica entre palabras y conjuntos de éstas, debe ser tratado a la luz de nuevas técnicas y desarrollos científicos. La diferencia básica existente entre los procesos del pensamiento humano y aquellos que son realizables por una máquina, en lo que a este problema se refiere, parecen radicar en la capacidad de la mente humana de manejar conceptos y procesos imprecisos, contra la incapacidad de los métodos usados actualmente en la computadora para hacer lo mismo. En 1965 un investigador americano, Lotfi Zadeh, plantea lo que ahora se conoce como teoría de conjuntos borrosos (*fuzzy sets*) que constituye un intento por manejar matemáticamente la referida imprecisión, esta teoría debe ser contemplada para futuros desarrollos.

Las características de un documento, del tipo al que nos hemos estado refiriendo, están plasmadas en las palabras que aparecen en su descripción, esto es, en el *abstract*, en tanto un *abstract* esté mejor realizado, tanto mayor será su probabilidad de ser recuperado por interrogaciones a las cuales es relevante, de aquí la importancia de una formalización y desarrollo de la teoría del *abstract*.

El desarrollo e investigación de cada una de las disciplinas que engloba la teoría de sistemas de información ofrece nuevas herramientas para la realización de sistemas más eficientes y confiables.

La aplicación de modernos métodos estadísticos que permitan la recuperación probabilística, así como la posible estimación de la exactitud lograda al realizar una consulta antes y después de efectuada la búsqueda.

Igualmente la aplicación de análisis de cúmulos (*cluster analysis*), así como la inteligencia artificial, debe ser tenida en mente.

Siendo jefe del Departamento de Publicaciones del Instituto de Investigaciones Jurídicas José Luis Soberanes, se terminó de imprimir *Automatización de la Legislación Mexicana. El sistema UNAM-JURE*, el día 24 de Marzo de 1983 en los talleres de Editorial Tierra Firme, S.A. de C.V.

Se imprimieron 2 000 ejemplares.