

Problemas para el manejo de la información legal en el Grupo Andino

JAVIER ALCALDE CARDOZA

1. LA PROBLEMÁTICA DE LA INFORMACIÓN LEGAL EN EL GRUPO ANDINO (*)

1. *Recuperación computerizada de la información legal*

Desde el punto de vista de la información, hay una distinción fundamental que hacer entre la literatura científica y las leyes. Mientras que la primera es en su mayor parte descriptiva de fenómenos o procesos y se usa para obtener información acerca de éstos, las leyes, siendo prescriptivas son apreciadas por sí mismas y no por la luz que arrojan sobre fenómenos externos; en este último caso el medio es realmente el mensaje (Tapper 1973).

Este hecho ayuda a comprender la importancia que se ha dado últimamente a la información legal, especialmente en relación con el uso de la computadora. Sin embargo, curiosamente, hasta la fecha el nuevo campo "ha prometido mucho, pero brindado poco" (Myers 1973). El optimismo estuvo basado no sólo en el valor "autónomo" de la información legal sino también en la apariencia de los textos legales como una base de datos ideal, principalmente por dos razones:

- a) Su lenguaje está en cierto grado ya formalizado, debido al esfuerzo hecho, cuando se redacta legislación, para decidir las palabras a emplear y la construcción de las oraciones.
- b) La práctica establecida, de ubicar los distintos temas bajo encabezamientos en los textos, que ha evolucionado a través de siglos de análisis e interpretación.

Lamentablemente, la apariencia de lenguaje bien definido de los textos legales ha probado ser superficial, cuando se ha tratado de usarlo como len-

(*) (Traducción del Capítulo 4 de: Alcalde, J. G. "Considerations for an Information Strategy for the Andean Group". 1975).

guaje para indizar, y la clasificación por encabezamientos en las leyes muestra, ante un análisis de mediana profundidad, que no es muy precisa ni realmente sistemática.

No obstante, luego de fracasos iniciales y creando técnicas que aún requieren refinamiento, ha sido posible montar bancos computerizados de textos legales y a base de ellos ofrecer servicios a abogados y al público en general en varios países, entre ellos Estados Unidos, Canadá, Francia y Bélgica. La mayoría de estos servicios se mantiene en una etapa experimental y subvencionada debido a que aún no ha sido diseñado un modelo de sistema que pueda dar a un servicio computerizado de información legal una perspectiva comercial (Myers 1973).

Las razones que recomiendan el uso de la computadora para textos legales y que hacen persistente la tendencia a desarrollar servicios, no obstante su falta de éxito económico, son las siguientes:

- a) Deficiencias de las publicaciones legales para satisfacer las necesidades de los usuarios, a despecho de la larga experiencia de empresas especializadas.
- b) El tamaño de la base de datos, que en todos los casos es suficientemente grande como para crear problemas para su manejo manual o mecanizado.
- c) La durabilidad del contenido informativo, pues la información legal tiene una extensa vida, lo cual significa estabilidad para una buena proporción del material almacenado en la computadora.
- d) En relación con b) y c), un tamaño comparativamente reducido de la información que mantiene la base de datos al día (nuevos textos legales).

Por otro lado, hay dos factores económicos que usualmente atentan contra la viabilidad de los servicios computerizados de información legal :

- a) La proporción de uso por parte del público, que normalmente debe ser alta.
- b) El tamaño de la base de datos debe estar por encima de un mínimo determinado, de manera que el servicio signifique una solución también para los usuarios. De otra manera la aceptación de los usuarios debe obtenerse ofreciendo servicios conexos.

Estos factores, sin embargo, pueden contrastarse a la luz del ejemplo de sistema CREDOC de Bélgica (Tapper 1973), que fue introducido teniendo en cuenta las dificultades especiales experimentadas en la recuperación de la información legal por medios convencionales en un país pequeño. En los países grandes, el problema es el enorme volumen del material existente y las dificultades surgen en el manejo de la información. En los países pequeños, aun

cuando el volumen de la información podría manejarse con métodos convencionales, hay el problema del tamaño reducido del mercado para servicios como enciclopedias, índices, revistas, etc. Estos servicios, son costosos y sería necesario un número mayor de suscriptores de los que el mercado puede ofrecer para volverlos económicamente factibles. En casos como éste, los costos de un servicio computerizado parecen ser una elección más prudente que la de publicar y actualizar, debido precisamente a un poco uso por parte del público y a pesar del modesto tamaño de la base de datos. La aparente contradicción que surge aquí con los factores anteriormente expuestos puede aclararse haciendo explícitas las variables que intervienen en el caso :

- a) Con una base de datos relativamente pequeña y para una demanda limitada, los servicios computerizados parecen ser más económicos que los servicios de publicación, para la información legal y no existe una diferencia importante en cuanto a efectividad entre ambos tipos de servicios.
- b) En el caso de una base de datos y una demanda de dimensiones considerables, los servicios computerizados son mucho más efectivos que las publicaciones, pero estos últimos servicios en la práctica han alcanzado ciertos costos y una aceptación que impiden el crecimiento de la demanda para los servicios computerizados, manteniendo los costos de éstos relativamente altos.

Con el propósito de relacionar estas consideraciones generales con el Componente de Información Legal de la Red Central del Grupo Andino, en las próximas páginas examinaremos con algún detalle la fuerza de las cuatro razones que sustentan un servicio computerizado en el contexto del Grupo Andino, incluyendo observaciones prácticas respecto al funcionamiento del Componente y sugiriendo, finalmente, el perfil que tendría el servicio, analizando también los factores económicos relacionados con su viabilidad.

2. *Deficiencias de las publicaciones legales*

Existen tres categorías de publicaciones que intentan cubrir los más recientes textos legales en los países andinos *:

Servicios de hojas sueltas, compilaciones y publicaciones periódicas. Las más útiles de todas estas publicaciones son producto de empresas comerciales, algunas de las cuales han surgido o tomado fuerza en virtud de las circunstancias favorable creadas por el proceso de integración económica.

Los servicios de hojas sueltas son del tipo texto o enciclopedia, diseñados para proveer información al día en campos como comercio exterior, legislación tributaria, legislación laboral. El sistema que emplean pone una carga sobre el usuario por lo laborioso que resulta el manejo de las hojas de actualización. Además, estos servicios proveen los textos originales incompletos

(*) En textos legales incluimos tratados internacionales, así como disposiciones de la Comisión y de la Junta del Acuerdo de Cartagena.

y fragmentados al adecuarlos a sus esquemas de clasificación. Las compilaciones, que tienen un campo similar al de los servicios de hojas sueltas, son escasas y en cuanto a su actualización, muy pocas veces ésta se realiza, ocurriendo muy tardíamente.

El período mínimo para una edición actualizada es de tres años y cuando el libro aparece ya se encuentra seis meses atrasado. Parte de las razones para esto son económicas, debido al pequeño tamaño de los mercados nacionales. Existen sí, numerosas compilaciones para tópicos populares, como seguridad social, por ejemplo, pero éstas no son en absoluto confiables.

Las publicaciones periódicas son compilaciones de leyes o resúmenes, que cubren campos mucho más amplios que los de las dos categorías anteriormente mencionadas (con referencia a nuestro interés, por ejemplo, las publicaciones más especializadas cubren el campo entero de la legislación económica).

El inconveniente de las publicaciones periódicas es que las referencias se hallan diseminadas a través de los números semanales, quincenales y mensuales y los índices acumulativos, lo cual constituye un obstáculo para la búsqueda.

Ninguna de estas publicaciones intenta cubrir adecuadamente las normas de integración económica (de la Comisión y de la Junta) así como las disposiciones económicas de los Países Miembros. Una publicación periódica de resúmenes (Legislación Económica y Social de los Países Miembros), preparada por la Junta, cubre solamente las normas de integración económica, además de las principales leyes económicas de los Países Miembros.

Otra publicación similar, del Instituto para la Integración de América Latina (Boletín de Información Legal), cubre de la misma manera las principales leyes económicas de los países latinoamericanos. Sólo estas dos publicaciones periódicas más dos o tres compilaciones a la fecha bastante trasnochadas, tienen una cobertura internacional para los países andinos. La razón para esta situación puede atribuirse básicamente a la dificultad para intercambiar textos legales, compilaciones o cualquier clase de publicaciones entre los Países Miembros, debido a restricciones cambiarias y a la falta de arreglos cooperativos entre las organizaciones interesadas. El reducido tamaño de los mercados determina la onerosidad de todas las publicaciones comerciales, que en muchos casos sólo algunas empresas y organismos pueden adquirir. Un último factor común a todas estas categorías de publicaciones (con excepción de la publicación de la Junta) es la ausencia o pobre calidad de sus índices, hechos empíricamente, con muy limitados puntos de acceso y sin control del vocabulario (incluso para sinónimos y palabras afines).

Todos los Países Miembros tienen diarios oficiales (excepto Bolivia cuya *Gaceta* aparece semanalmente). Los organismos a cargo de su publicación no cuentan con mecanismos especiales o no tienen mecanismos en absoluto (Colombia) para el envío de suscripciones al extranjero. El resultado es que en cualquier país miembro el diario oficial de otro país miembro es virtual-

mente imposible de obtener. Aún para la Junta es normalmente muy difícil obtener estas publicaciones.

3. *Dimensión de la base de datos*

Alrededor de veinte mil disposiciones (leyes y decretos) relacionados con las actividades económicas y asuntos exteriores puede estimarse que son promulgadas anualmente en los seis países andinos y por lo menos treinta mil normas menores (resoluciones ministeriales y de otros organismos). Como puede apreciarse en la Tabla N° 7, los tipos de disposiciones que existen en los países miembros varían considerablemente en cantidad, así como también en el promedio anual de promulgación. Las disposiciones que aparecen en la Tabla N° 7 son las que suelen contener los cambios y decisiones que influyen significativamente las economías nacionales. La situación es relativamente simple en Bolivia (*Gaceta Oficial*), Ecuador (*Registro Oficial*), Venezuela (*Gaceta Oficial de Venezuela*), y Perú (*El Peruano*), este último país donde muy pocas decisiones importantes aparecen en decretos supremos, los cuales se originan en los Ministerios, pero es más complicada en Chile (*Diario Oficial*) y Colombia (*Diario Oficial*), donde importantes decisiones aparecen no sólo en los decretos de los Ministerios sino también en disposiciones de otros organismos.

Las disposiciones menores aparecen en la Tabla N° 8. Mientras que en Bolivia, Perú y Venezuela, las resoluciones ministeriales (o supremas) son el instrumento principal, en Chile los Acuerdos del Comité Ejecutivo del Banco Central tienen gran importancia. En Ecuador los Acuerdos Ministeriales e Interministeriales son equivalentes a las resoluciones de los otros países y las Resoluciones de la Junta Monetaria tienen trascendencia (Carta Económica). Colombia es el país con la más amplia gama de disposiciones y organismos reguladores; cabe destacar las Resoluciones de la Junta Monetaria y las del Instituto de Comercio Exterior, con su Consejo Directivo de Comercio Exterior (Legislación Económica). Debe anotarse que la Tabla N° 8 no pretende ser exhaustiva sino puramente reflejar los tipos de disposiciones que han incluido campos económicos y aspectos de interés para el proceso de integración entre 1970 y 1974 (de acuerdo con una revisión hecha de compilaciones); por lo tanto el número de tipos de disposiciones relevantes podría crecer en los próximos años.

TABLA N° 7

Estimación de la producción anual de Leyes relacionadas con las actividades económicas en los Países Miembros

(Leyes y Decretos)

1972 — 1973

BOLIVIA 500 Decretos (Decretos Leyes y Decretos Supremos)

CHILE	4,000	Decretos de los Ministerios de : Finanzas Economía Relaciones Exteriores
COLOMBIA	15	Leyes
	2,400	Decretos Autónomos
	7,000	Decretos de los Ministerios de : Desarrollo Económico Finanzas y Crédito Trabajo
ECUADOR	3,000	Decretos (Decretos y Decretos Supremos)
PERU	600	Decretos Leyes
	500	Decretos Supremos de los Ministerios de: Economía y Finanzas Relaciones Exteriores Industria Comercio
VENEZUELA		Leyes
	1,600	Decretos

TABLA N° 8

Disposiciones menores relacionadas con las actividades económicas en los Países Miembros

(Los estimados de producción anual van en algunos casos entre paréntesis)

BOLIVIA	Resoluciones Ministeriales
CHILE	Resoluciones Ministeriales Acuerdos del Comité Ejecutivo del Banco Central (60 sesiones)
COLOMBIA	Resoluciones Ministeriales Resoluciones del INCOMEX (1,000) Resoluciones del Consejo Directivo de Comercio Exterior (23) Resoluciones del Consejo Nacional de Política Económica y Social Resoluciones del Consejo Nacional de Política Aduanera Resoluciones de la Junta Monetaria (100) Resoluciones de la Superintendencia de Bancos Resoluciones de la Superintendencia de Industria y Comercio

	Resoluciones de la Superintendencia de Precios (500)
	Resoluciones de la Dirección General Marítima y de Puertos
	Acuerdos de la Corporación de Turismo
	Circulares de Ministerios
	Circulares del INCOMEX
	Circulares de la Oficina de Cambio
ECUADOR	Acuerdos Ministeriales
	Acuerdos Interministeriales
	Resoluciones de la Junta Monetaria
	Resoluciones de la Superintendencia de Bancos (246)
	Resoluciones de la Superintendencia de Compañías (3,200)
PERU	Resoluciones Supremas, por Ministerios
	Resoluciones Ministeriales, por Ministerios y Direcciones
	Disposiciones de la Junta de Transacciones Externas del Sector Privado
	Circulares del Ministerio de Economía y Finanzas
	Circulares de la Dirección de Aduanas
	Circulares del Banco Central
VENEZUELA	Resoluciones Ministeriales

A la producción nacional de leyes debe agregarse las normas de integración emitidas por la Comisión y la Junta. Hasta la fecha la Comisión ha emitido 90 Decisiones (desde fines de 1969 hasta 1974) y la Junta 48 Resoluciones (desde 1971 hasta julio de 1975). Estas cifras dan promedios anuales de 16 Decisiones y 11 Resoluciones, las cuales, especialmente en el segundo caso, tenderán a aumentar.

De acuerdo con sus fines, el Componente de información legal debería procesar disposiciones en los siguientes campos:

- a) Aplicación de políticas por los gobiernos en áreas de interés para el proceso de integración (incluyendo las relaciones económicas internacionales).
- b) Formulación y adopción de normas de integración por los órganos del Acuerdo y los Países Miembros, respectivamente.
- c) Comercio exterior.

De estas categorías, en relación con a) podemos especificar campos de interés de dos clases para el proceso: de interés general, como aspectos de administración pública, finanzas, impuestos y relaciones económicas internacionales, para obtener una idea del desarrollo que siguen las economías nacionales; de interés específico, como inversión extranjera, compras estatales,

turismo, seguridad social, promoción de exportaciones, para evaluar la implicación de las disposiciones para ciertos mecanismos o programas de acción conjunta de los Países Miembros. Los campos de interés general varían muy poco (por ejemplo en 1972 el petróleo y la energía resultaron un nuevo campo de interés para seguir la evolución de las economías de los Países Miembros). Pero aquellos de interés especial son modificados (mayormente ampliados) por lo menos anualmente. Según el resultado de las reuniones de la Comisión y el correspondiente contenido del programa de trabajo de la Junta. En la práctica, mucho más disposiciones pertinentes al proceso se encuentran en los campos de interés general.

La adopción de las normas de integración por los Países Miembros es importante para la Junta debido a su función de velar por la aplicación del Acuerdo y por el cumplimiento de las Decisiones en el proceso; estas leyes son valiosas para conocer las características con las cuales cada país implementa las normas de la Comisión. Las normas de comercio exterior c) son un caso distinto. Estas deben ser difundidas ampliamente para contribuir a la expansión del comercio intrasubregional (como mecanismos del Sistema para la Promoción de Exportaciones) y las más importantes son aquellas relacionadas con procedimientos específicos del comercio exterior, que pueden ser de utilidad para fines prácticos.

Para los propósitos del Componente de Información Legal no parece necesario manejar la producción total de disposiciones económicas de los Países Miembros. En efecto, procesar alrededor de cincuenta mil disposiciones cada año resultaría un ideal muy costoso. Sería un ideal, pues los campos de interés del proceso de integración se expanden rápidamente y es imposible predecir al momento de seleccionar disposiciones en qué dirección seguirá la expansión; además, las diversas disposiciones nacionales se hallan bastante interrelacionadas. Pero el conjunto de disposiciones muestra una gran proporción de normas puramente administrativas así como otras de tópicos muy específicos (por ejemplo relacionadas con personas, compañías, eventos formales), lo cual desaconseja intentar la empresa. Por otro lado, podemos tener una idea acerca de la proporción de disposiciones pertinentes al proceso de integración que anualmente se promulga, examinando la publicación periódica de la Junta, "Legislación Económica y Social de los Países Miembros", en la cual se ha resumido disposiciones desde 1970 hasta 1973. Esta publicación, cubriendo las categorías temáticas a) y b) (excluyendo solamente las normas específicas de comercio exterior), además de las normas de integración, ha venido seleccionando un promedio de 250 disposiciones anualmente. Por supuesto, debe tenerse en cuenta que esta publicación es altamente selectiva, que no todos los tipos de disposiciones han sido incluidos (se basa solamente en los diarios oficiales; véase la Tabla N° 9) y que las disposiciones de comercio exterior no eran especialmente consideradas. Pero aun si dudamos acerca de la validez del 0.5 por ciento de disposiciones pertinentes al proceso de integración que las cifras de esta publicación sugieren, no podemos pensar que siendo una publicación oficial, originalmente preparada para satisfacer las necesidades del organismo técnico del proceso de integración, estas cifras se encuentren muy distantes de la verdadera proporción. En este caso, en primer lugar podemos desechar la idea de procesar cincuenta mil disposiciones para no esperar de

ninguna manera que más del 5 por ciento resulten pertinentes. En segundo lugar, podemos intentar llegar a una estimación del ingreso anual de disposiciones que tendrá el Componente, basándose en las cifras que conocemos. Si suponemos que en el caso del Componente un aumento en extensión y exhaustividad en cuanto a los aspectos de interés general, más el incremento normal de los campos de interés especial, así como el de las normas de integración, determinaría tener que duplicar el número de disposiciones, tendríamos un ingreso estimado de 500 disposiciones para las categorías a) y b). Si queremos sumar a este estimado las normas de comercio exterior como categoría diferenciada debemos previamente tener en cuenta dos hechos:

- a) Una buena proporción de normas de comercio exterior está normalmente incluida en las categorías a) y b) (por lo menos representan un 25 por ciento del total de disposiciones seleccionadas).
- b) Pensando en términos de costo-beneficio, no resultaría sensato cubrir todas las normas de comercio exterior de los seis países, sino sólo aquellas que introducen cambios importantes en los procedimientos establecidos.

Si aceptamos estos dos hechos, podemos considerar que la incorporación de c) como una categoría diferenciada traería un incremento del 30 por ciento en el número de disposiciones a ingresarse en el Componente. Esto nos daría un total de 650 disposiciones como estimado de ingreso anual al Componente.

4. *Fuentes y procesamiento retrospectivo*

La revisión de las cincuenta mil disposiciones que anualmente producen los Países Miembros y el procesamiento de una parte de ellas requiere por lo menos de dos condiciones fundamentales: fuentes que provean las disposiciones oportunamente y antecedentes legales que asistan los criterios para seleccionarlas y procesarlas.

Con referencia a las fuentes ya hemos mencionado los diarios oficiales. El problema es que éstos cubren mayormente legislación (leyes y decretos), muestran demora en la publicación y deficiencias para la distribución internacional. Una demora de una o dos semanas entre la fecha de las leyes y la fecha de su publicación es normal, pero se dan casos, por ejemplo en resoluciones del INCOMEX en Colombia, en Acuerdos Ministeriales en el Ecuador, en Decretos en Chile y Resoluciones en el Perú, en los cuales se presentan demoras de uno a tres meses. En el caso de Bolivia, la Gaceta Oficial, semanal, se publica muchas veces atrasada; en el Ecuador algunas leyes nunca aparecen en el Diario Oficial. En cuanto a la distribución internacional de los diarios, como fuentes para la mayor parte de los textos legales que deben estar a la mano en todos los núcleos del Componente, parecería que un arreglo directo entre los Núcleos Nacionales, para intercambiar puntualmente una determinada cantidad de diarios oficiales, sería la mejor solución.

En la Tabla N° 9 podemos apreciar los tipos de disposiciones que no aparecen en los diarios oficiales. Es de destacarse que en Colombia las Resoluciones de Ministerios como los de Agricultura y Desarrollo Económico no son publicadas en el Diario Oficial y que las resoluciones del INCOMEX no siempre aparecen en él. El Ministerio de Agricultura, por ejemplo, emitió en 1972 resoluciones sobre restricciones a la importación y el Ministerio de Desarrollo Económico tuvo en ese mismo año una resolución con la cual creaba un Comité de Integración Económica, así como emitió algunas disposiciones relacionadas con la inversión extranjera. También en este año, algunas resoluciones del INCOMEX relacionadas con mecanismos de la integración sub-regional no aparecieron en el Diario (Banco de la República 1973). Las resoluciones de las Juntas Monetarias, tanto en Colombia como en el Ecuador, no aparecen en los respectivos diarios oficiales. Para todos los tipos de disposiciones que se enumeran en la Tabla N° 9, los Núcleos Nacionales tendrían que contemplar acciones específicas ante los organismos correspondientes para asegurar el arribo regular de los textos legales al Componente.

Ya se ha dicho que la función principal del Componente ha de ser de alerta informativa (current awareness) sobre disposiciones promulgadas en el Grupo Andino. Pero esta función no podría realizarse sin una base de datos original, en relación con la cual las nuevas disposiciones adquieran pleno significado. De igual manera, una base de datos original permitiría realizar búsquedas de alguna profundidad en el Componente. Al pensar en antecedentes de las disposiciones, podemos dividir a éstos en dos periodos.

- a) Disposiciones hasta mayo de 1969, cuando se suscribe el Acuerdo de Cartagena.
- b) Disposiciones desde junio de 1969 hasta 1975 *.

Los antecedentes de estos dos periodos no tienen la misma importancia para el Componente. En primer lugar, debido a que el enfoque del Componente es hacia disposiciones relacionadas con las economías de los Países Miembros desde el momento en que éstas se vinculan por el Acuerdo y no antes. En segundo término, porque las principales leyes reguladoras de la actividad económica se encuentran probablemente en mayor número en el segundo periodo, no obstante su menor extensión. Esto no debe atribuirse a una vida excepcionalmente corta de las disposiciones económicas de los países andinos, sino a una mayor participación del sector público que a partir de 1969, coincidiendo e interrelacionándose con el movimiento de integración, se hizo manifiesta, determinando nuevas leyes de promoción industrial, fomento a las exportaciones, sistemas financiero, aduanas y seguridad social entre otros campos.

* Suponiendo que el Componente pueda comenzar a funcionar a principios de 1976.

TABLA N° 9

Tipos de disposiciones para las cuales los Diarios Oficiales no tienen una cobertura completa

Resoluciones del INCOMEX	I
Resoluciones del Ministerio de Agricultura	N
Resoluciones del Ministerio de Desarrollo Económico	N
Resoluciones del Consejo Directivo de Comercio Exterior	I
Resoluciones de la Junta Monetaria	I
Resoluciones del Consejo Nac. de Política Económica y Social	N
Resoluciones de la Superintendencia de Bancos	I
Resoluciones de la Superintendencia de Industria y Comercio	N
Resoluciones de la Superintendencia de Compañías	I
Resoluciones de la Dirección General Marítima y de Puertos	N
Acuerdos de la Corporación de Turismo	N
Circulares de Ministerios	N
Circulares de INCOMEX	N
Circulares de la Oficina de Cambio	N
ECUADOR	
Resoluciones de la Junta Monetaria	N
Resoluciones de la Superintendencia de Bancos	I
Resoluciones de la Superintendencia de Compañías	N
PERU	
Circulares del Ministerio de Economía y Finanzas	N
Circulares de la Dirección de Aduanas	N
Circulares del Banco Central	N

Clave:

- I : Cobertura incompleta
N : Sin cobertura

Para el período más importante, exactamente desde abril de 1970 en adelante, en la Junta se ha resumido (1970-1973) o indizado (1974-1975) las disposiciones más saltantes. Para las disposiciones anteriores a 1970, así como para completar, si fuera necesario, las disposiciones de 1970 a 1975, se tendría que usar compilaciones así como servicios de hojas sueltas. Puede anticiparse, sin embargo, que la recopilación de la base de datos original para las disposiciones de comercio exterior será bastante laboriosa, pues la Junta no ha cubierto el campo completamente y sólo existen servicios de hojas sueltas en tres Países Miembros (Colombia, Chile y Venezuela),

mientras que las compilaciones están por lo menos atrasadas en tres años en los otros países.

El procesamiento retrospectivo es una tarea ineludible para el componente, pero existen dos maneras de hacer economías al momento de ingresar la base de datos original. En primer lugar, se puede ingresar individualmente tan sólo una parte de los antecedentes (por ejemplo, todos los del segundo período), y cubrir el resto valiéndose de compilaciones que no requerirían ser exhaustivamente indizadas. En segundo lugar tenemos la clase de procesamiento que se puede dar a las disposiciones individuales. Como veremos con mayor detenimiento más adelante, existen tres caminos para esto: texto completo, resúmenes e indización. El camino más barato es indizar, especialmente cuando, como ocurre en el presente caso, un buen número de disposiciones ya ha sido indizado y requeriría solamente una revisión y algunos ajustes en la terminología empleada. Esta opción podría considerarse para la mayoría de los antecedentes y permitiría graduar la exhaustividad en la indización, de acuerdo con la importancia relativa de cada disposición para el Componente y aun procesar algunas compilaciones como unidades, enfatizando únicamente su tema general así como los tópicos principales de su contenido.

Una estimación de la dimensión de la base de datos original que se requeriría resulta muy difícil de realizar sin haber hecho un cuidadoso análisis de la legislación económica de cada País Miembro. Sin embargo, intentaremos llegar a una cifra basándonos en algunos parámetros conocidos, para facilitar el diseño del Componente. Durante el período abril 1970-julio 1975, la Junta ha indizado alrededor de 1,500 disposiciones. Si suponemos que un aumento del 33 por ciento sobre esta cifra permitiría incorporar las disposiciones del lapso junio 1969 - marzo 1970, así como la legislación de importancia que hubiera omitido la Junta, tenemos una cifra de 2,000 disposiciones para todo el segundo período (junio 1969-1975). En cuanto a las disposiciones de comercio exterior, el caso de Venezuela puede proporcionarnos una valiosa referencia para calcular la dimensión de la base de datos requerida. Una compilación que contiene disposiciones hasta 1970 (Banco Central de Venezuela 1971), muestra 64 disposiciones que sería necesario ingresar al Componente. Las leyes que este tipo de compilaciones básicas reúnen, suelen mantener su número a través de algunos años, debido a que la tendencia principal en ellas no es de expansión ni de disminución sino de reemplazo. En consecuencia, podemos suponer que 80 sería el promedio de disposiciones que tendrían las compilaciones básicas, de índole similar a ésta para todos los Países Miembros Colombia y Chile probablemente estarían por encima del promedio). Esto nos da la cantidad de 500 disposiciones para la base de datos original de comercio exterior. Finalmente, parece razonable pensar que las disposiciones económicas fundamentales promulgadas con anterioridad a junio de 1969 (esto es, del primer período), podrían ser representadas adecuadamente si se asigna una cifra de referencia igualmente de 80 disposiciones para cada País Miembro. En esta forma, en total tenemos un estimado de 3,000 disposiciones para la base de datos original del Componente.

5. Viabilidad económica del servicio

Después de haber visto los aspectos básicos del marco para la recuperación de la información legal en el Grupo Andino, podemos volver a las consideraciones iniciales de este capítulo, con el objeto de evaluar la pertinencia de los argumentos expuestos, tanto en favor como en contra del servicio computerizado que tendría el Componente.

En cuanto a las cuatro razones que recomiendan un servicio computerizado, tenemos la siguiente situación en el Grupo Andino:

- a) Las publicaciones legales exhiben serias deficiencias para las necesidades del usuario en el Grupo Andino y específicamente no hay ningún servicio ni combinación de servicios que tenga una cobertura similar a la que se prevé para el Componente.
- b) Una base de datos original de 3,000 disposiciones con un incremento anual de 650 no puede considerarse de gran tamaño ni lo tendría a corto plazo, pero ciertamente presenta problemas para el manejo manual. Un servicio computerizado, además de evitar estos problemas, se hallaría listo para satisfacer los requerimientos a mediano y largo plazo, y en el corto plazo no demandaría tareas laboriosas para el personal en las pequeñas organizaciones participantes en el Componente; además, desde el comienzo brindaría servicios conexos con la base de datos.
- c) La legislación económica fundamental y las normas de integración emitidas durante 1970-1975, que constituirían la mayor parte de la base de datos original, representan el comienzo de tendencias en gran medida comunes en los países andinos y por lo tanto, puede esperarse que tengan una larga vida, lo cual significaría estabilidad para una gran cantidad de almacenaje en la computadora.
- d) Luego de la etapa inicial del Componente (que se estima sería 1976), la cual requeriría ingresar la base de datos original además de las disposiciones seleccionadas ese año (3,650 disposiciones en total, aproximadamente), solamente se necesitaría ingresar anualmente, a corto plazo, alrededor de 650 disposiciones para mantener la base de datos al día. Posteriormente, aunque este número se incrementase, considerado como un porcentaje de la base de datos ya almacenada, el ingreso anual decrecería o por lo menos se mantendría constante.

En relación con los factores económicos que atentan contra la viabilidad económica de un servicio computerizado, la situación en el Grupo Andino es la siguiente :

- a) Si se concibe el Componente ofreciendo servicios solamente a los participantes o a dependencias gubernamentales, entonces es razonable esperar poco uso de la base de datos. Pero la información que

manejaría el Componente es requerida por una amplia gama de organizaciones y personas y no existe razón para impedir que éstas tengan también acceso a ella. Más aún, el propósito de recolectar disposiciones de comercio exterior es el de difundir su contenido a los comerciantes de los seis países. En este caso, si el Componente operara para su audiencia potencial (la cual no tiene un servicio de información alternativo) la frecuencia de uso de la base de datos justificaría ampliamente la elección de un servicio computerizado.

- b) Desde el punto de vista de los usuarios, quizás no en todos los casos el tamaño de la base de datos parezca justificar los precios de un servicio computerizado (por ejemplo, para usuarios con intereses limitados a un aspecto específico de la integración). Sin embargo, el hecho es que debido a la reducida demanda, tal como ocurre con el sistema CREDOC en Bélgica, un servicio computerizado tiene muchas posibilidades de resultar más económico que los servicios convencionales de publicaciones, los cuales ciertamente en el Grupo Andino son bastante caros en su limitada cobertura.

En relación con este último factor, hay otros dos aspectos que anticipan la viabilidad del servicio. El primero, es que las disposiciones relacionadas con la integración económica y el comercio exterior en los Países Andinos, aunque no totalizan un gran número, se hallan dispersas y no son fácilmente accesibles para nadie. El Componente, debido a su status oficial y a sustentarse en la cooperación, sería capaz de obtener y manejar todas ellas. El segundo aspecto es un servicio conexo que la computadora hace factible y que precisamente beneficiaría a los usuarios con intereses en aspectos específicos del proceso de integración: la Difusión Selectiva de la Información. En virtud de este servicio, los usuarios reciben solamente la información pertinente a sus necesidades, sin tener que revisar listas, resúmenes, ni ningún otro tipo de instrumento de referencia. Este servicio se basa en el apareamiento hecho por la computadora de los perfiles de interés de los usuarios con los términos asignados a las disposiciones al ser indizada. Este servicio también determinaría que las disposiciones ingresadas al Componente no tuvieran que ser editadas en su totalidad, sino solamente distribuidas selectivamente a los suscriptores (Connor 1967).

De estas consideraciones podemos concluir que las razones que hacen aconsejable un servicio computerizado para la información legal tienen vigencia en el contexto andino y que los factores económicos involucrados no atenderían contra la viabilidad del servicio, siempre que éste fuera ofrecido a una amplia audiencia e incluyera la difusión selectiva de la información.

Dentro del conjunto de mecanismos de la Estrategia de la Información, el Componente para la Información Legal tiene la peculiaridad de que sus servicios estarían relacionados con todas las provisiones sobre información y con todo el proceso de integración, pues persigue reflejar en detalle los pasos dados por los organismos del Acuerdo y los Países Miembro en su camino al mercado común. Este es el tipo de información fundamental que el Grupo Andino debe asegurar en primera instancia en su Estrategia de la Informa-

ción. De acuerdo con esto, el Componente para la Información Legal tendría que ser el primer mecanismo en diseñarse e implementarse. Hay dos circunstancias que de diferente manera recomiendan el cumplimiento de esta prioridad: la experiencia y logros de la Junta en información legal (a diferencia de cualquier otro tipo de información) y las perspectivas de rentabilidad de los servicios correspondientes (a diferencia, nuevamente, de cualquier otro tipo de información a ser manejada en los mecanismos del Grupo, excepto quizás la información sobre mercados).

La rentabilidad de la difusión selectiva de la información que haga el Componente para una amplia audiencia, aunque no puede esperarse que rinda ganancias, es razonable pensar que alentará, en varios niveles, el desarrollo de los demás mecanismos de información y que además pueda atenuar el impacto inicial de los gastos relacionados con el funcionamiento de la computadora y los canales de comunicaciones de la Red Central.

6. *Alternativas para el procesamiento*

Respecto a la implementación del Componente queda sólo un aspecto importante por examinar: aquel del procesamiento que se daría a las disposiciones. Existen tres posibilidades: texto completo, resúmenes e indización. En la primera posibilidad los textos legales pasan íntegros a la computadora la cual es programada para producir una concordancia, listando cada palabra que aparece en el texto y asignándole un número. Esta concordancia proporciona la base para aparear el texto legal con los perfiles de usuarios o con las formulaciones de búsqueda (Maccanico 1973). Esta posibilidad es especialmente adecuada para jurisprudencia, en la cual el énfasis en las búsquedas se pone en la recuperación (tratando de no omitir ninguna publicación que pueda ser pertinente). El ingreso de los textos completos es la más costosa de las opciones. Si el texto es abreviado o resumido para ser ingresado a la computadora, entonces tenemos la modalidad de resumen o una variante menos utilizada, el análisis formal, en el cual se registra partes del texto bajo encabezamientos determinados. Finalmente, la indización hace que ingresen al sistema, para cada texto, únicamente los términos que representan aspectos básicos de su contenido. Las dos últimas posibilidades pueden emplear el lenguaje natural o un vocabulario controlado. En la primera posibilidad, en el resumen o como términos de indización se emplea palabras extraídas del texto (lenguaje natural). En el caso de un vocabulario controlado, los conceptos más importantes del texto son representados por palabras tomadas de un Tesoro o terminología controlada, especialmente construido para el sistema. Para el Componente, la modalidad de texto completo sólo sería conveniente en el caso de las Resoluciones de la Junta y eventualmente en el de los fallos del Tribunal Andino de Justicia, los cuales corresponden, en diversa forma, al campo de la jurisprudencia. Para el grueso de disposiciones, los resúmenes parecen ser la opción más adecuada, empleando un vocabulario controlado. Esta última resulta ser una elección forzosa, en vista de la existencia de seis distintos vocabularios en las disposiciones de los Países Andinos (siete aún, si consideramos también el vocabulario de las normas de integración), los cuales, aunque individualmente son más o menos limitados, al ser comparados muestran significativas dife-

rencias. Los resúmenes serían particularmente apropiados para la función de alerta informativa del Componente. Sin embargo, no puede esperarse que en la primera etapa del Componente los resúmenes sean confeccionados utilizando un vocabulario controlado. Esto requiere capacitación de los recursos humanos, así como una elaborada terminología y en consecuencia sólo se logrará, probablemente, después de un período en que los resúmenes y la indización se realicen paralela pero independientemente, y en el cual los resúmenes no sirvan todavía para fines de recuperación de la información. Se podrá llegar a este segundo período si desde el inicio el procesamiento se realiza en parte en los diversos núcleos del Componente y no se centraliza en su totalidad. De esta manera, los resúmenes preparados en los núcleos podrían guiar la indización que se haría centralmente, a base de un tesoro que iría creciendo paulatinamente.

La indización es un proceso básico para cualquier sistema de recuperación de la información y por lo tanto demandará mucho cuidado en la etapa de diseño del Componente. La encuesta que se mencionó al referirnos a la elección de participantes para los mecanismos subregionales de información (Capítulo 3.3.2.), destinada a descubrir la situación de los servicios de información oficial en el Grupo Andino así como las características de las necesidades de los usuarios, debería incluir las necesidades de información legal, para que sus resultados brinden elementos para establecer una política de indización para el Componente. Según Lancaster (Lancaster 1968), los resultados de una encuesta pueden proporcionar varios tipos de información útil para la indización:

- a) Un núcleo de vocabulario que refleje los intereses de trabajo y las necesidades de información de los usuarios.
- b) Conocimiento de la clase de servicio que los usuarios necesitan en términos de los grados de recuperación (recall), precisión o de una determinada combinación de ambas.
- c) Una idea acerca del nivel de exhaustividad que se requiere en el proceso de indización.
- d) Una idea acerca del grado de especificidad que requieren los términos a emplearse para la indización.

Esta información acerca de los usuarios podrían combinarse con los conocimientos sobre las características y terminología de los textos legales que existen en la Junta, en virtud de la experiencia ganada con su publicación de resúmenes de leyes y servirá para adoptar una política de indización y diseñar los instrumentos adecuados para la primera etapa de funcionamiento del Componente, todo esto posible de ser subsecuentemente modificado de acuerdo con la retroalimentación (feedback) que proporcionen los usuarios dentro de la mecánica de la difusión selectiva de la información.

II. LA PROBLEMATICA DE LA DIFUSION SELECTIVA DE LA INFORMACION Y LA CONSTRUCCION DE PERFILES DE USUARIOS EN LOS CAMPOS DE LAS CIENCIAS SOCIALES

1. *Difusión selectiva de la información y perfiles de usuarios*

Difusión Selectiva de la Información (DSI) es el nombre que recibe cualquier procedimiento manual o automatizado que intenta proporcionar un servicio personal de alerta informativa (current awareness), seleccionando la literatura de actualidad de posible pertinencia para los intereses de los usuarios. La característica distintiva de la DSI con relación a otros tipos de servicio de alerta informativa, es que en ella a cada usuario se le notifica únicamente de aquellos documentos de interés probable para él, mientras que en otros servicios, al usuario se le informa acerca de todos los documentos que caen dentro de una determinada categoría temática que él ha elegido. En la base del servicio de DSI está el principio de aparear descriptores asignados a los documentos con perfiles de interés de los usuarios; en la literatura usualmente se reserva la denominación para los casos en que el apareamiento se produce mediante el uso de la computadora. En este caso, existe un "file" tanto para los perfiles de documentos como para los perfiles de usuarios, los cuales en conjunto constituyen el área de *input* del sistema. Existe consenso en cuanto a lo que Luhn (Luhn 1961), creador del procedimiento DSI señaló, que la construcción y mantenimiento de los perfiles de usuarios es la "más importante y al mismo tiempo la más crítica" tarea dentro del sistema, debido a que la comunicación entre el hombre y la máquina ocurre por medio de perfiles, en los cuales los campos de interés de los usuarios aparecen descritos en términos que pueden ser entendidos y procesados por el sistema.

En el funcionamiento de un sistema DSI se dan tres fases relacionadas con los perfiles de usuarios: su construcción, la estrategia de apareamiento y su mantenimiento y actualización. En las líneas siguientes trataremos de describir la primera fase, pero es casi imposible realizar esto sin una referencia constante a las otras dos fases, especialmente a la estrategia de apareamiento.

2. *Recolección de información para el perfil*

Básicamente el perfil de usuario es una formulación de los requerimientos de información de un individuo. Puede ser construido por el mismo usuario, con ayuda del personal del servicio, por el personal del servicio o automática o semiautomáticamente mediante la computadora (para conocer los avances en cuanto a construcción automática, que no vamos a examinar ahora, por su problemática aplicación al campo legal, véase Kerr-Waller (1968) y Barker (1972).

El procedimiento más común para recolectar información es el de solicitar al usuario que llene un formulario en el cual expresa sus requerimientos de información e incluye muestras de documentos pertinentes. Acompañando a este formulario, se recomienda (especialmente Garfield, 1970) enviar una

clara explicación de las características del sistema en cuestión. Luego del análisis del formulario llenado por el usuario, en algunos casos un borrador de perfil, es sometido a éste para sus comentarios y eventuales modificaciones o se le solicita información adicional acerca de ciertos puntos o aspectos que él ha mencionado. En uno pocos casos, los usuarios son entrevistados para obtener una información lo más completa posible, debido a que esta medida ha probado ser demasiado costosa (en cuanto a tiempo del personal del Servicio y gastos de transporte). El costo en términos de tiempo y esfuerzo que implica la construcción de perfiles operativos es una de las más grandes preocupaciones de un servicio DSI.

La información más valiosa para el perfil parece ser recogida en los casos en que las formulaciones de los usuarios son en forma narrativa y no en listas ni en clasificaciones (Aitchison 1966), pues de esta manera la relación que guardan los conceptos aparece con claridad. Igualmente, la inclusión de ejemplos de documentos pertinentes proporciona una estimable información. En relación con esto, es menester decir que aunque la parte crucial del contenido de un perfil son los temas, también pueden ser incluidos nombres de autores, instituciones y títulos de publicaciones periódicas como requerimientos de información.

Existe una posición que considera la solicitud de información a los usuarios como una tarea demasiado morosa y además violatoria del concepto de servicio de la DSI, puesto que es muy exigente y plantea una difícil labor al usuario, que él puede no estar capacitado para realizar adecuadamente. Esta posición se manifestó tempranamente en el trabajo de Luhn y que enfatiza la necesidad de métodos automatizados o semiautomatizados para la construcción de perfiles.

Todos los métodos para la construcción de perfiles hacen hincapié en el hecho de que una vez traducidos, de una manera u otra, los requerimientos de información a perfiles, éstos deben ser considerados meramente tentativos y modificados en la medida que sea necesario de acuerdo con su rendimiento y con la retroalimentación proporcionada por los usuarios.

3. *Los sistemas de lenguaje natural*

En el caso de sistemas que no usan una terminología controlada para hacer los perfiles de los documentos, sino palabras clave o fragmentos extraídos de los documentos, el proceso de construir un perfil de usuario es particularmente laborioso. El usuario formula sus requerimientos de información teniendo a la mano una lista de palabras que contiene todas las palabras que han aparecido a través de un período largo en el área de *input* del sistema y su frecuencia. En esta formulación original se identifica los conceptos principales y luego se les expande agregándoles términos más amplios, más específicos y afines, con el auxilio de "guías de búsqueda", especialmente diseñadas, que muestran la relación existente entre los términos. La expansión de la formulación es particularmente importante pues todas las posibles formas de presentación de los conceptos deben estar previstas, incluyendo sinónimos, formas en singular y plural y variaciones en la ortografía. La última etapa de

la construcción es la de refinamiento, en la cual los resultados de la expansión son reducidos y simplificados según el grado de utilidad que parezca ofrecer cada uno y el tamaño que se juzgue conveniente para el perfil. De la más alta importancia en el refinamiento es la introducción de términos truncados que sirven para detectar, mediante una raíz común, una gama de conceptos derivados y relacionados. Términos truncados como "CEL" pueden ser truncados en uno o en sus dos extremos (por ejemplo, para detectar términos como *celular*, *Brucella*), indicando esto con asteriscos (CEL*, *CEL*). Algunos sistemas tienen también como auxilio índices de letras clave que ilustran respecto a los términos que podrían ser detectados con determinadas raíces (Veal 1968).

Luego del refinamiento, viene la codificación del perfil. Usualmente, en los sistemas de lenguaje natural cada concepto está representado por varios términos y estos aparecen reunidos formando "parámetros". Los parámetros a su vez están vinculados por medio de formulaciones de lógica Booleana o por "pesos" para efectos del apareo con los documentos.

Varios sistemas emplean solamente palabras clave tomadas de los títulos de los documentos. En estos casos al usuario se le pide que describa sus intereses en un conjunto de palabras que en su opinión podrían aparecer en los títulos de los documentos pertinentes. Estas palabras son posteriormente confrontadas por el personal de servicio con índices de títulos (Rice 1965). El problema de emplear títulos es que el sistema no funciona en ciertos aspectos como un servicio de difusión selectiva de información, pues los pedidos de información no pueden ser hechos tan específicos como sería deseable en un servicio individualizado. Sin embargo, es necesario decir que esta situación varía según el campo temático de que se trate. En algunos campos parecería ser cierto que los títulos son "los medios más simples con los cuales se puede representar los documentos y aún satisfacer las necesidades de los usuarios", formulación clave que hace Vickery (Vickery 1970) con relación al problema de la adecuada representación de los documentos en los sistemas de información. En la práctica, se ha recogido evidencias para demostrar que los títulos son tan efectivos como los resúmenes para la recuperación, como también para demostrar lo contrario (Connor 1967).

Contra el empleo de sistemas de lenguaje natural puede argumentarse que en su funcionamiento la causa principal para apareamientos errados (entre las palabras clave de los documentos y los perfiles) radica en la total falta de control sobre el vocabulario que entra al sistema. Debido a esto, se ha diseñado varios mecanismos para disminuir la frecuencia de apareamientos errados y mejorar la exactitud en la recuperación de la información, como veremos más adelante.

4. *Sistemas de vocabulario controlado*

La mayoría de los sistemas de DSI emplean un vocabulario controlado para los perfiles de documentos y de usuarios, Cooper (Cooper 1968) encontró en EE. UU. que de doce servicios DSI, siete tenían un Tesaurus o algo similar, mientras que sólo tres empleaban el lenguaje natural.

En estos sistemas, la formulación del usuario se utiliza para hacer una lista de los conceptos pertinentes, los cuales son expresados como descriptores, tomados de una lista normalizada de descriptores o de un Tesauro DSI, el cual se emplea también para indizar los documentos que entran al sistema. En algunos casos, al usuario se le devuelve la introducción de su formulación original a descriptores y, en tanto que la traducción hecha por los especialistas del servicio no puede ser modificada, al usuario se le pide en cambio que elimine, corrija o agregue términos en una sección especial que contiene otros descriptores, los cuales representan conceptos más específicos o relacionados con los conceptos originales, de probable pertinencia (Aitchison 1966). Los descriptores finalmente seleccionados por el personal conforman un borrador de perfil y aparecen unidos, como en el caso de lenguaje natural, por formulaciones lógicas o por "pesos". El producto final tiene la misma estructura que en los sistemas de lenguaje natural, esto es un perfil con dos partes, a saber: una lista de descriptores, los que se hallan en algunos casos agrupados en "parámetros" y la especificación de "pesos" o la formulación de lógica Booleana que en cada caso expresa los tipos de relaciones entre los descriptores o parámetros que serían aceptables para el usuario.

El argumento a favor de un vocabulario controlado es que mientras para la formulación de las necesidades de los usuarios indudablemente resulta mucho más apropiado el lenguaje natural, la indización es mucho más efectiva con una terminología controlada, de manera que para hacer compatibles ambos vocabularios, la transformación del lenguaje natural a descriptores debe hacerse en algún momento.

Una tercera alternativa la constituyen los esquemas de clasificación, los cuales proporcionan ciertas ventajas pero traen también un número de dificultades. En la mencionada encuesta hecha por Cooper, solamente dos de doce servicios DSI habían optado por esta alternativa.

5. *Mecanismos de apareamiento: formulaciones lógicas y términos ponderados*

La parte central del funcionamiento de un servicio DSI la constituye el apareamiento de los perfiles de usuarios y documentos, del cual se deriva directamente la difusión de la información. En cualquier sistema, para considerar que un documento es de interés para un usuario, no es necesario que todos los términos del perfil del usuario aparezcan en el documento. La estructura interna del perfil del usuario refleja de diversas maneras las combinaciones de términos que son aceptables para el usuario. Estos criterios de apareamiento pueden ser expresados a través de varios mecanismos que van desde la simple coordinación de un número especificado de términos y son principalmente los siguientes:

- a) Términos ponderados o con "peso" asignado: éste es el criterio más popular y puede ser implementado de diversas maneras, siendo típica la asignación de "pesos" a cada término del perfil según su importancia relativa, y la fijación de un "peso mínimo" total que al ser igualado o superado determina que el documento sea seleccionado.

En este procedimiento, para cada documento, la computadora identifica los términos coincidentes y suma sus "pesos", comparando seguidamente el resultado con el "peso mínimo" establecido.

- b) Formulaciones de lógica Booleana: este mecanismo sigue en popularidad al anterior y consiste en términos o grupos de términos conectados por la conjunción ("y") por la disyunción ("o") o calificados por la negación ("no"). Los documentos en los que los términos llenan las condiciones de las formulaciones Booleanas del perfil, son seleccionados.
- c) Estrategia de términos indispensables, probables y proscritos: los términos del perfil van acompañados de las calificaciones "no" (documentos con estos términos no deben seleccionarse), "debe" (documentos con estos términos deben seleccionarse siempre que otro de sus términos no tenga un "no"), "quizás" (un documento es seleccionado si reúne un cierto número o porcentaje de estos términos de probable pertinencia). Similar a esta estrategia es aquella que califica los términos como "imperativos" (cualquier documento con uno de estos términos es pertinente) y "esenciales" (solamente los documentos que incluyen estos términos son pertinentes).
- d) Elementos de la "lógica de búsqueda", según la denominación dada por Vickery (Vickery 1970): los documentos son seleccionados solamente si los términos del perfil aparecen en un orden determinado, que puede ser: consecutivo (A/B) o dentro de límites especificados (esto es dentro de determinadas sub-unidades del documento, como párrafos u oraciones).

De todos estos mecanismos, que no son mutuamente excluyentes, debe señalarse que el citado en tercer orden es significativamente menos efectivo, debido a la exagerada importancia que da a términos aislados, y que los dos mecanismos mencionados en cuarto término son evidentemente de carácter complementario (especialmente adecuados para sistemas que emplean un lenguaje natural) y no pueden usarse por sí solos. Por lo tanto examinaremos con mayor detalle solamente los dos primeros mecanismos.

Bloch (Bloch 1968), menciona dos razones básicas para preferir las expresiones Booleanas a los términos ponderados en un sistema DSI:

- a) Se obtiene un procesamiento más rápido, pues en cuanto un elemento que aparece unido por la conjunción ("y") no es apareado en el documento, la búsqueda para el correspondiente perfil se detiene instantáneamente.
- b) Al ser consultado, los usuarios prefieren las expresiones Booleanas, pues éstas les proporcionan un cierto sentido de control sobre sus perfiles.

Vickery, por el contrario, sostiene que en general la ventaja de los términos ponderados sobre las ecuaciones lógicas, es que con los primeros se acelera considerablemente el procesamiento en la computadora (evidentemente Bloch y Vickery tienen en mente diferentes momentos del funcionamiento de la computadora en relación con el perfil; como ninguno de los dos menciona parámetros concretos para sustentar sus opiniones, éstas no pueden ser debidamente contrapuestas). Pero el mismo Vickery deja en claro que los términos ponderados no pueden realizar con igual perfección las funciones que cumplen las formulaciones lógicas.

El hecho de que tanto las ecuaciones lógicas como los términos ponderados poseen ventajas y debilidades, que al ser evaluadas no arrojan un balance que se incline claramente a favor de uno de ellos, puede ser confirmado con los hallazgos de Cooper, quien encontró que de doce servicios DSI siete usaban términos ponderados (en dos casos en combinación con otros mecanismos) y cinco empleaban ecuaciones lógicas (en tres casos en combinación con otros mecanismos). Vickery señala que la mejor elección parecería ser una combinación de términos ponderados y formulaciones lógicas; sin embargo, debido a la complejidad que tiene la combinación de estos mecanismos, se encuentra que no es muy popular en la práctica: Cooper encontró que sólo un servicio había implementado esta combinación (en la encuesta de Cooper ocho de doce servicios usaban un solo mecanismo).

Cuando recién se estaban comenzando a desarrollar los sistemas DSI, Tritschler (Tritschler 1964) publicó un vigoroso argumento en favor del uso de términos ponderados. En primer lugar, decía que los términos ponderados son particularmente útiles porque permiten tomar en cuenta la indización imperfecta que es un componente normal de todo sistema de información (ningún indizador le asigna a un documento todos los términos que requiere y, por otro lado, no hay dos indizadores que asignen el mismo conjunto de términos a un determinado documento). Esto motiva que el conjunto de descriptores que representa a cada documento sea necesariamente incompleto. Ahora bien, mientras el mecanismo del "peso mínimo" total permite seleccionar aproximaciones a un tópico deseado (que en la realidad podrían ser el tópico mismo, inadecuadamente representado por una indización imperfecta), las formulaciones lógicas de carácter simple, que usan solamente la conjunción o la disyunción, no actúan satisfactoriamente en estos casos (las formulaciones con conjunción no seleccionan documentos a los que les falta un solo elemento y las formulaciones con disyunción seleccionan una gran cantidad de material no pertinente). Sin embargo, el mismo Tritschler reconoce que el problema puede ser resuelto con formulaciones lógicas que tengan tanto conjunciones como disyunciones. Pero considera inconveniente en la práctica realizar la difícil tarea de reducir las necesidades de información a este tipo de formulación precisa, más aún cuando un resultado equivalente puede obtenerse mediante la simple técnica de apareo y cuenta, en la cual se basa el mecanismo de términos ponderados. Podemos darnos cuenta de que Tritschler no está completamente en lo cierto, pues los términos ponderados no resultan un sustituto perfecto de las formulaciones lógicas (cuando éstas incorporan conjunciones y disyunciones). Los términos ponderados brindan tan sólo una útil forma de lógica condicional; también pueden ser usados satisfacto-

riamente (los "pesos negativos" en concreto) en lugar de la calificación negativa de las formulaciones lógicas; por último, son particularmente recomendables para satisfacer la típica solicitud de información relacionada con "cualquier material que tenga dos de tres (determinados) temas", como señala Veal (Veal 1968).

Un segundo argumento de Tristschler es que los términos ponderados facilitan la confección de listas graduadas de los documentos seleccionados, las cuales resultan un valioso instrumento para el usuario, pues lo ayudan a decidir acerca de la cantidad y grado de pertinencia de los documentos de los que quiere ser periódicamente informado. Este es un detalle fundamental, como veremos posteriormente.

6. *Grados de recuperación y precisión con respecto a los mecanismos de apareamiento*

Los servicios DSI no pueden obtener una total efectividad, debido a factores relacionados con las imperfecciones del lenguaje y de la indización. Sin embargo, es posible analizar el rendimiento de un sistema en términos de sus grados de precisión y recuperación (o "recobro" como lo denominan algunos autores), comparando en este proceso el comportamiento de los términos ponderados y las ecuaciones lógicas como mecanismos de apareamiento. Connor dice que en general las ecuaciones lógicas pueden obtener un alto grado de precisión con algún control sobre la recuperación al mismo tiempo, mientras que los términos ponderados permiten una alta recuperación con algún sacrificio de la precisión, brindando en conjunto un mejor rendimiento que las ecuaciones. Es oportuno recordar que el grado de recuperación dentro de un sistema de información se puede medir dividiendo el número total de documentos existentes en el sistema, que son pertinentes a una determinada consulta, entre el número de documentos pertinentes a la consulta que en una búsqueda se puede recuperar. El grado de precisión se obtiene dividiendo el número total de documentos recuperados para una determinada consulta entre el número de documentos que resultan efectivamente pertinentes a la consulta (una fórmula matemática para el grado de recuperación sería $R = TDR/TDP$, donde R es el grado de recuperación, TDP es el total de documentos pertinentes recuperables y TDR el total de documentos efectivamente recuperados; para el grado de precisión la fórmula sería $P = DP/TDR$, donde P es el grado de precisión, TDR el total de documentos recuperados y DP el número de documentos efectivamente pertinentes a la consulta que se ha recuperado).

Connor contrasta estos dos mecanismos con la estrategia de términos probables, indispensables y proscritos, con la cual es posible obtener un grado más o menos alto de precisión a expensas de significativas pérdidas en la recuperación. Aitchison, por otro lado, evaluó los grados de precisión y recuperación con un acopio de 700 documentos y encontró que las ecuaciones lógicas dieron una combinación óptima de recuperación y precisión (ambas en el más alto grado posible) superando en rendimiento a un mecanismo de apareamiento de hasta 4 descriptores, el cual es básicamente el mecanismo de los términos ponderados.

La función principal de un servicio DSI es la de proporcionar información pertinente a los usuarios (recuperación) y su segunda finalidad es la de preservar el tiempo del usuario, eliminando para él toda la información no pertinente (precisión). Además hay consenso respecto a que proporcionarle al usuario demasiado material resulta tan inconveniente como proporcionarle demasiado poco. Una marcada prioridad de la recuperación motivaría la situación de exceso de material y un exagerado énfasis sobre la precisión traería por consecuencia demasiado poco material, de manera que ambos extremos deben evitarse. Ahora bien, se sabe que una mejora en el grado de recuperación se hace posible a expensas de la precisión y viceversa; por lo tanto, debe hallarse un punto de equilibrio para cada servicio y aun para cada usuario. La situación exacta del punto de equilibrio depende evidentemente de los usuarios (los cuales deben escoger un punto en la "parábola del rendimiento en términos de recuperación-precisión" según la denominación de Bloch), quienes pueden preferir recuperación, precisión o una cierta combinación de ambas. Suplementando esta fundamental opción que tienen los usuarios, los documentos pertinentes pueden serles notificados en la forma de listas graduadas o se les puede solicitar que especifiquen con antelación la cantidad de documentos que les gustaría conocer periódicamente. En concordancia con estas especificaciones dadas por los usuarios, es posible elegir los mecanismos individuales o combinados a ser usados en los perfiles para el proceso de apareamiento.

7. *La construcción de perfiles en los campos de las ciencias sociales*

En los campos de las ciencias sociales no existen propiamente hasta la fecha servicios DSI totalmente operativos. Existen solamente servicios experimentales y en general de alerta informativa (Current Awareness). Hay varias razones para explicar esta situación, como lo señala un informe de la Universidad de Bath (Bath University 1970). En Inglaterra existe el servicio de alerta informativa de la Universidad de Bath, el cual incluye un pequeño y más bien informal servicio DSI y también el proyecto de la Biblioteca de la Cámara de los Comunes y el Laboratorio Culham (Hall 1972), el cual es en realidad un servicio de alerta informativa que difunde periódicamente listas temáticas. En los Estados Unidos estuvo en funcionamiento un servicio DSI experimental en el campo del estudio internacional de valores en política, que ha sido descrito por Janda (Janda 1968). En realidad en este caso la DSI era una extensión de un servicio de recuperación de información que usaba el Método TRIAL (Technique to Retrieve Information from Abstracts of Literature). Los usuarios, profesores de la North Western University, preparaban ellos mismos listas de palabras claves que describían sus intereses sobre cuestiones de actualidad. Estas listas, que en ningún caso excedían de diez palabras, eran generalmente modificadas en muy pequeña medida por el personal del Servicio, que introducía cuando era conveniente, términos truncados. Las listas eran usadas como perfiles para ser confrontadas periódicamente con un *input* de resúmenes especiales de documentos y publicaciones. Las palabras clave empleadas no formaban parte de un vocabulario controlado y podían ser términos simples o compuestos y siglas. Por otro lado, los resúmenes TRIAL (el texto completo no se usaba, aunque se consideró inicialmente la posibi-

lidad, por razones de costo), de tipo informativo (no meramente descriptivo), consistían en varias tarjetas cada uno. Una de éstas, el "Sumario" describía el documento en función de encabezamientos fijos, tales como "problema", "diseño de la investigación", "conclusiones", "sugerencias para la investigación". Otra tarjeta, la "formulación", expresaba toda premisa o hallazgo mencionados en el resumen en sus palabras clave fundamentales. Finalmente, la "elaboración" reproducía las palabras originales para proporcionar los elementos contextuales y permitir evaluar adecuadamente cada proposición. El apareamiento, por medio de una combinación de ecuaciones lógicas y términos ponderados, ocurría entre la tarjeta de "formulación" y el perfil del usuario. Como rasgo particular, cabe mencionar que la computadora podía ser programada para la búsqueda de frases y términos compuestos en los resúmenes (en este caso las frases y términos compuestos aparecían unidos por asteriscos en los perfiles).

Por lo que apreciamos de este servicio experimental, podemos fácilmente inferir que en cuanto a las técnicas de construcción de perfiles de usuarios, los campos de las ciencias sociales requieren un enfoque diferente a los que se dan en las ciencias puras y aplicadas. Sin embargo, hay que señalar que en el servicio de la North Western University este hecho parece no haber sido totalmente reconocido. En efecto, mientras los documentos eran representados de manera bastante compleja (los resúmenes TRIAL), esta complejidad no hallaba una contraparte en la construcción de perfiles de usuarios. Por el contrario, los perfiles eran construidos aun de manera menos sofisticada que en los servicios de información científica y tecnológica. Evidentemente, existía una marcada falta de proporción entre las escasas palabras claves tomadas del lenguaje natural que empleaban los usuarios y los elaborados resúmenes. Por ejemplo, en los resúmenes se reconocía la importancia del contexto original para aclarar el significado de las palabras clave, pero en los perfiles de usuarios no había ningún elemento que pudiera ser confrontado con esta porción de los resúmenes (la especificación, dentro de los perfiles de usuarios, de funciones o roles aceptables para las palabras clave hubiera sido un elemento adecuado para este propósito). De todas maneras, el procesamiento que se daba a los documentos en este servicio indica que los elementos del perfil de usuario y su construcción tienen que ser más sofisticados e incorporar mecanismos adicionales en un servicio DSI para cualquier rama de las ciencias sociales, a diferencia de los servicios que operan para ciencia y tecnología.

Lo que sí se implementó en este servicio experimental y resulta muy significativo es el uso combinado de términos ponderados y ecuaciones lógicas, el cual, como hemos visto, no se da frecuentemente en los servicios para ciencia y tecnología. En la misma forma, la mayor complejidad del lenguaje empleado en las ciencias sociales, motivó que se debiera emplear gran cantidad de frases y términos compuestas como palabras clave en los perfiles. Esta parece ser la razón por la cual el servicio que estamos analizando eligió el lenguaje natural, no obstante resultar muy claro que hubiera sido ideal, aunque imposible en la práctica, un control de la terminología.

En la descripción que hace Janda del servicio, no menciona nada acerca del equilibrio que se logró entre los grados de recuperación y precisión, pero parece lógico pensar que en cualquier campo de las ciencias sociales sería muy peligroso para un servicio de información propender a un alto grado de precisión. Aun a nivel de usuarios individuales, se puede conjeturar que la posición extrema fue probablemente la de elegir una combinación de precisión y recuperación, aunque siempre con preferencia para esta última. Consecuentemente, podemos suponer que el uso de términos ponderados (de mejor rendimiento en términos de precisión) resultaba determinante para el apareamiento de perfiles y documentos.

El informe de la Universidad de Bath dedica bastante atención a los problemas especiales de las ciencias sociales que dificultan el funcionamiento de servicios DSI en sus ramas. Hay dos de éstos que se relacionan significativamente con la construcción de perfiles de usuarios:

- 1) La falta de rigor del lenguaje de las ciencias sociales y las variaciones en la acepción de los términos complica enormemente la tarea de traducir las necesidades de información a perfiles operativos.
- 2) El continuo cambio en los intereses y necesidades de los usuarios es un factor adverso. Por su naturaleza, los temas sociales están estrechamente vinculados a los acontecimientos económicos y políticos y, por lo tanto, se ponen en boga por un tiempo y luego decaen. Estos cambios coyunturales no pueden predecirse en ninguna medida al momento de recoger la información para el perfil.

Un tercer problema podría agregarse a estos dos, derivado lógicamente de ellos, e insoslayable cuando se conoce la manera como el servicio DSI funciona en la Universidad de Bath. En este servicio la actualización de los perfiles, en constante cambio, sólo es posible merced a un contacto estrecho y permanente del personal del Servicio con los usuarios, que constituyen un grupo reducido. Esta situación no podría darse en un sistema convencional, de mayores dimensiones, más aún si tomamos en cuenta los testimonios que aparecen en la literatura (Bloch y Garfield, especialmente) en el sentido de que los usuarios no se muestran de ordinario muy inclinados a proporcionar información para un servicio DSI.

BIBLIOGRAFIA

AITCHISON T. M. and Clague. P. (1966) *Selective dissemination of information project*. London, National Electronics Research Council, 1966.

BANCO CENTRAL DE VENEZUELA (1977) *Disposiciones legales que regulan las importaciones y las exportaciones de Venezuela*. Caracas, Imprenta del Banco Central, 1971.

BANCO DE LA REPUBLICA DE COLOMBIA (1973) *Disposiciones Económicas de 1972*. Bogotá, Talleres Gráficos del Banco de la República, 1973.

BARKER ET AL. (1972) Towards automatic profile construction. *Journal of Documentation* (1) (1972).

BATH UNIVERSITY OF TECHNOLOGY LIBRARY (1970) *Experimental information officer in the social sciences report to OSTI*. Bath, february, 1970.

BLOCH, U. and OFER, K. D. (1968) Experiments with an SDI system. In *FID Mechanized information storage, retrieval and dissemination*. Amsterdam, 1968.

CARTA ECONOMICA, *publicación semanal de CEDE especializada en asuntos económicos y financieros*, 2, 3, 4 (1970, 1971, 1972).

CONNOR, J. H. (1967) SDI: A review of the literature and the issues. *Library Quarterly*. 37 (4) 221 — 245 (1967).

COOPER, M. (1968) Current information dissemination: ideas and practices. *Journal of Chemical Documentation*, 8 (5) (1968).

DIARIO OFICIAL DE LA REPUBLICA DE CHILE (1972, 1973).

DIARIO OFICIAL (de Colombia) (1972, 1973).

EL PERUANO, Diario Oficial (del Perú) (1972, 1973).

GACETA OFICIAL de Bolivia (1972, 1973).

GACETA OFICIAL de la República de Venezuela (1973).

GARFIEL, E. (1970) The role of man and machine in an international SDI. In, *FID Users of documentation*. Buenos Aires, 1970.

HALL ET AL. (1972) An experimental current awareness service in the social sciences. *Journal of Documentation* (1) (1972).

JANDA, K. (1968) *Information retrieval; applications, to political science*. New York 1968.

KEER-WALLER R. D. (1968) Automated information dissemination system (AIDS). In,

- FID *Mechanized information storage, retrieval and dissemination*. Amsterdam, 1968.
- LANCASTER, F. W. (1968) *Information retrieval systems: characteristics, testing and evaluation*. New York, Wiley, 1968.
- LEGISLACION ECONOMICA; *Compilación progresiva de la totalidad de las disposiciones económicas*, 40, 41 (1972).
- LEGISLACION ECONOMICA Y SOCIAL DE LOS PAISES MIEMBROS (1971, 1972, 1973).
- LUHN, H. P. (1961) Selective dissemination of new scientific information with the aid of electronic processing equipment. *American Documentation* (2) (1961).
- MACCANICO, A. (1973) Legislative-legal information retrieval for the Italian Parliament. In, ASIS A. *world of information*. 1973.
- MYERS, J. M. (1973) Computers and the searching of law texts in England and North America: A review of the state of the art. *Journal of Documentation*, 29- (2) 212-226 (1973).
- REGISTRO OFICIAL: *Organo del Gobierno del Ecuador*, 1, 2, (1972, 1973).
- RICE, C. N. (1965) A computer based alerting system for "chemical titles". *Journal of Chemical Documentation*, 5 (3) (1965).
- TAPPER, C. (1973) *Computers and the law*. London, Weidenfeld and Nicolson, 1973.
- TRISCHLER, R. J. (1964) Effective information-searching strategies without "perfect-indexing". *American Documentation* (2) (1964).
- VEAL, D. C. (1968) An experiment in the selective dissemination of chemical information. In, Houghton, B. ed. *Computer based information retrieval systems*. London, Bingley, 1968.
- VICKERY, B. C. (1970) *Technique of information retrieval*. London, Butterworths, 1970.