

LAT 708

3330

CINTERFOR

Centro Interamericano de Investigación y Documentación sobre Formación Profesional

BIBLIOTECA



CENTRO UNIVERSITARIO
DE INVESTIGACIONES
BIBLIOTECOLÓGICAS

G. K. Thompson

CUATRO CONFERENCIAS SOBRE DOCUMENTACIÓN

Montevideo
1972

INFOBILA

No. Lat.	<u>0001/08</u>
No. Adq.	_____
No. Sist.	_____
Tipo de Adq.	<u>Donación</u>
Fecha	<u>02/09/2019</u>

I. ENFOQUE GENERAL

Las personas se comunican de maneras muy distintas y a niveles muy dispares, pero el objetivo común de toda clase de comunicación es transmitir informaciones. Éstas se pueden transmitir por medio de palabras orales o escritas, de símbolos, de mapas, de ilustraciones, de cintas magnéticas, películas, etc., pero lo que sea, el medio por el cual se transmite esta información, el simple hecho de transmitirla, no hace que esta información sea útil. Para serlo, primeramente la información tiene que ser recibida por una persona que sepa aprovecharla, que sepa evaluarla, comprenderla, utilizarla y determinar si es una información significativa o trivial y, además, si es una información que repite algo ya sabido.

Desde hace varios años, el problema de encontrar, evaluar y canalizar informaciones científicas ha adquirido suma importancia. Se calcula que anualmente los países industrializados gastan alrededor de veinte mil millones de dólares en investigaciones científicas y tecnológicas. Los resultados de estas investigaciones se publican en unas 30.000 revistas científicas y también en innumerables libros, informes técnicos, etc. En total, se calcula que un millón de documentos nuevos, que se pueden calificar de importantes, se publican anualmente. Nos referimos a documentos en el sentido palpable. Hay todavía más, conservados en forma de cintas magnéticas, datos estadísticos no publicados y series estadísticas exportables, si se conoce su existencia y si existe la manera de encontrarlas y utilizarlas. Estas informaciones son los resultados de investigaciones, pero igualmente sirven de materia prima para iniciar otras investigaciones. Y en teoría, por lo menos, todas estas informaciones publicadas, son accesibles a todos. Pero sabemos bastante bien cómo pasan las cosas en la realidad. Volvamos un instante a los países industrializados. En algunos países podemos utilizar buenas bibliotecas, existen servicios de documentación, servicios bibliográficos, de reseñas bibliográficas, de resúmenes analíticos, etc., que facilitan el acceso a una parte de este millón de documentos. Pero todos estos servicios no lo hacen sino parcialmente, de manera bastante superficial, y según ciertas personas, demasiado lentamente. Se calcula que pueden pasar de 30 a 45 meses para que un artículo científico publicado por una revista científica, desde que el autor escribe la versión preliminar de los resultados de su investigación hasta que se publica el resumen analítico en una revista de resúmenes. Y en vez de mejorar, esta situación se agrava cada día más. Porque no significa simplemente el aumento de la cantidad del número de los documentos producidos. Es mucho más. Por un lado existe una mayor

demanda de las necesidades de los investigadores, y por otro es más difícil todavía: hay una falta enorme de personal calificado para trabajar en estos organismos bibliográficos.

Tenemos necesidades nuevas y al mismo tiempo sentimos que nos basta indentificar un estudio sólo por un número de una clasificación decimal universal u otra, o con dos o tres encabezamientos de materia como lo hacen los bibliotecarios desde hace varias generaciones. Los investigadores mismos exigen más. Exigen un análisis un poco más profundo de lo que solemos hacer en situaciones clásicas y, además un todas las disciplinas, sean ciencias naturales, aplicadas o en ciencias sociales, se siente cada día más la necesidad de atravesar diversas disciplinas. Este fenómeno se nota sobre todo en ciencias sociales y es un problema que nos toca a todos los que trabajamos de cerca o de lejos; es un problema de formación profesional.

Se puede decir que la profesión de documentalista nació de la necesidad de tener un vínculo entre los productores de informaciones y los que tienen que utilizarlas. Hemos hablado algo de los problemas de los países industrializados, pero existe un problema aún más grave: es el que sufren lo que llamamos aquí en Europa, los países del Tercer Mundo, países que ni siquiera pueden soñar un día alcanzar el nivel que tienen algunos países industrializados hoy. No disponen de buenas bibliotecas, no tienen infraestructuras, y sobre todo carecen de personal calificado. Hablamos mucho, se lee mucho sobre el problema de lo que llamamos en inglés "Technology gap", "écart technologique" en francés, (no sé la expresión equivalente en español; supongo que es "brecha tecnológica" o algo así). Pero me parece que es peor todavía el problema de la brecha documental, por la razón siguiente: en los países que podemos llamar países bien organizados bibliográficamente se produce mucho; pero el hecho de tener servicios bibliográficos permite a estos mismos países producir todavía más, porque es más fácil el acceso a las informaciones ya publicadas, lo que determina un progreso todavía más rápido.

Si vamos a estudiar el problema de cómo salir de esta situación en la cual nos encontramos, me parece que debemos tratar de encontrar soluciones nuevas. Pero antes de encontrarlas hay que saber plantear exactamente el problema. Desgraciadamente, no sabemos muy bien cómo plantear el problema del costo de las informaciones científicas. En el mundo actual tratamos siempre de dar un precio para tal o cual tipo de trabajo, pero todavía no sabemos cómo establecer un precio al valor de la información científica.

En este enfoque general vamos a terminar mirando un momento otra vez la situación en los países industrializados y ver cómo se ha tratado recientemente de encontrar soluciones nuevas en el plano nacional a estos problemas. Se puede decir que este enfoque va a ser algo pesimista. El problema es un poco diferente; aún en un país, en un continente como Europa, con cierta cantidad de servicios bibliográficos bien organizados, se hace muy poco en el plano inter

nacional y la comunicación entre los diferentes países crea un problema gigante. Existe el problema de comunicación oral y escrita, porque casi todos los países tienen su idioma o idiomas nacionales. En América Latina, por lo menos, es claro el obstáculo a la comunicación científica, pero por lo menos existe el idioma español como medio de comunicación accesible a la mayor parte de los países del continente.

Miremos otra vez lo que se hace o lo que se ha tratado de hacer en algunos países en años recientes. Primeramente asistiremos a una cosa un poco curiosa: al nacimiento de una cantidad de organismos nuevos que se preocupan de problemas de documentación, de información científica. No quiero hacer aquí un juego semántico, pero pasamos por épocas. Se habló mucho, hace varios años, de biblioteconomía, luego empezamos a hablar de documentación y ahora hablamos de información científica. No vamos a tratar de explicar los diversos matices, en este sentido, de un concepto u otro, pero si surge este concepto de información científica es porque desde hace algún tiempo, no son solamente los bibliotecarios y los documentalistas, los que se preocupan por el problema de la transmisión del análisis de informaciones científicas. En algunos casos son los científicos mismos que tratan de encontrar soluciones nuevas. Hace un par de años se hablaba mucho de la utilización de máquinas electrónicas, de computadoras electrónicas, para sacarnos de este apuro. Y los propios fabricantes de estas máquinas trataban de convencernos de que el problema era muy fácil, que se podía, por ejemplo, utilizar las propias palabras contenidas en los títulos de los artículos, de los informes técnicos, para indizar, para analizar los documentos. Y que se podrían utilizar las máquinas electrónicas para preparar índices de todas clases. En la mayor parte de los casos hemos visto que esos índices no resolvían nada, que era una forma de no encontrar, de no mirar de frente el verdadero problema tal como se debía plantear. Pero es indudable que si sabemos utilizarlas, estas máquinas electrónicas pueden ser de una gran utilidad para el desarrollo, no solamente de servicios de documentación en el plan nacional, sino sobre todo para la diseminación rápida de informaciones científicas a través del mundo.

Voy a hablar, pues, de técnicas nuevas. Pero hablar de técnicas me parece un poco secundario. Es más importante tratar de comprender cuáles son los objetivos de este trabajo. Me he referido a la biblioteconomía, de documentación de información científica. Creo que primeramente tenemos que decir cuáles son los objetivos más importantes que el documentalista debe tratar de alcanzar. Primeramente, llevar rápidamente a conocimiento del investigador todas las informaciones nuevas pertinentes que llegan a nuestras manos. Y esto de la manera que insuma el menor tiempo posible al investigador. Segundo: tendremos que organizar nuestras actividades de manera tal que, cuando tengamos que hacer una investigación retrospectiva, sabremos encontrar todos los documentos importantes y pertinentes al trabajo de nuestros investigadores. Para alcanzar estos dos primeros puntos, necesitamos una mayor sistematización del trabajo, una mayor racionalización de nuestra manera de trabajar, y posiblemente, una

...ación, una automatización de ciertos procedimientos. Ante todo, nos
...nular nuestros propios métodos de trabajo y plantear varios interroga
...es. Supongo que habrá muy pocas soluciones, pero me parece que es muy impor
...ante formular ciertas preguntas acerca de lo que hacemos, y cómo y por qué lo
...hacemos de esta manera.

Si podemos encontrar una respuesta a algunas de estas preguntas podremos enton
...es tratar de soñar con hallar una manera eficaz de realizar estos objetivos.
Si pensamos un momento en cada cosa que compone nuestro trabajo, cada elemento
de nuestro trabajo cotidiano, veremos una cierta cantidad de cosas bastante
anacrónicas. Quiero hablar para ilustrar este punto de una sola cosa: el catá
logo, la preparación de nuestro catálogo, nuestros ficheros de trabajo (puesto
que supongo que la mayor parte de nosotros tiene un fichero, un catálogo) y
quiero hablar sobre todo de hacer el catálogo descriptivo, es decir la descrip
ción bibliográfica, la descripción física de los documentos que tenemos en
nuestras colecciones.

En los países de habla inglesa, el 50% de la producción literaria, la produc
ción de libros, se vende a bibliotecas. Es muy frecuente que el mismo libro
se encuentre en 100 o en 1000 bibliotecas distintas y que esté catalogado de
100 ó 1000 maneras diferentes y, peor todavía, por personal calificado. Todo
esto en una época donde falta personal calificado por todas partes. ¿Cómo po
demos justificar este estado? ¿Cómo podemos justificar en el año 1966 que un
mismo libro deba catalogarse de 100 o de 1000 maneras distintas en 100 ó 1000
bibliotecas o centros de documentación distintos? Los bibliotecarios siempre
encuentran una razón diciendo: "Bueno ... en los países de habla inglesa, por
ejemplo, se imprimen fichas de catálogo en la Biblioteca del Congreso en
Washington, la Bibliografía Nacional Británica imprime fichas". Pero ellos
nos dicen: "Esto es demasiado lento, hay que esperar las fichas". O sea que
las reglas de catalogación no corresponden de una biblioteca a otra. Siempre
encuentran una razón para explicar (o tratar de explicar) este estado de co
sas. Pero estoy muy poco convencido con todas estas razones. ¿Y cómo pod
mos explicar, justificar este estado de cosas a miembros de la comunidad cien
tífica?

Considerada desde el exterior esta situación es absolutamente absurda. Por
todos lados en el mundo encontramos o estamos empezando a encontrar normas
internacionales. Un ejemplo que conocemos todos es el sistema métrico, pero
hay muchos ejemplos en el mundo entero. Hasta la propia Inglaterra va a aban
donar sus libras, peniques y chelines para tener una moneda decimal. ¿Y qué
hemos fabricado nosotros, bibliotecarios, en varias generaciones? Reglas pa
ra establecer fichas de catálogo que no son reglas, que no son nada más que
una compilación de diversas prácticas en diversas bibliotecas grandes. ¿Cómo
podemos soñar con una colaboración internacional cuando aún en el plan nacio
nal tenemos una situación tan absurda?

Para mí, estas reglas de catalogación representan casi el fracaso intelectual de una profesión. La normalización internacional y la simplificación de estas reglas de catalogación son la única manera de poder utilizar esos nuevos instrumentos de trabajo que son las máquinas electrónicas. Tratar de dar, de hacer tragar por una máquina electrónica las reglas de catalogación utilizadas en la mayor parte de las bibliotecas sería dar, no solamente un dolor de cabeza enorme a la máquina, sino un dolor de cabeza insoportable a nosotros mismos. Y es solamente cuando miramos en el espejo que nos decimos: "¿Por qué hacemos esto?" Una vez que se hace esta pregunta, entonces solamente podremos tratar de encontrar una solución.

Han calculado, por ejemplo, en el solo estado de Nueva York y únicamente para los bibliotecas municipales, las bibliotecas populares, que si se podía catalogar los libros nuevos de manera central, para todo el estado de Nueva York, podrían ahorrar en salarios de bibliotecarios profesionales más de un millón de dólares anuales. Y esto solamente para las bibliotecas populares del estado de Nueva York.

La tecnología moderna nos permite hacer mucho, pero no nos lo permite con reglas imposibles y anacrónicas como las que tenemos que utilizar en la mayor parte de las situaciones.

He hablado de estos libros ingleses y he dicho que se calcula que el 50% de los libros nuevos producidos en Inglaterra y en EE.UU. se vende a la biblioteca. Y entonces, ¿por qué no podemos soñar con una catalogación de estos libros por los propios editores? Al fin y al cabo, cuando compramos una lata de leche condensada existen impresas en la etiqueta instrucciones para el uso. ¿Y por qué nosotros, los bibliotecarios no podemos hacer igual? ¿Imprimir fichas de catálogo con cada documento científico nuevo, con cada libro nuevo que sale? ¿Y por qué no pueden hacerlo los propios editores para evitar un desperdicio gigantesco de nuestros recursos humanos que podríamos utilizar para dar mejores servicios a nuestros lectores?

Quiero terminar esta primera parte con una idea general que va a ser más o menos la idea de base de todo lo que va a seguir. Si vamos a encontrar soluciones nuevas, si vamos a saber cómo plantear el verdadero problema, cómo alcanzar estos distintos objetivos, tenemos que hacerlo todo tomando en cuenta las necesidades de nuestros lectores. Y lo que tenemos que hacer no es solamente un trabajo introspectivo. Tenemos que interiorizarnos, ir, discutir, dialogar con los lectores, con los investigadores, con los científicos que utilicen o podrían utilizar algún día nuestros servicios, para saber cuáles son sus necesidades y cómo transformarlas en realidad práctica.

II. PROBLEMAS DE IDENTIFICACIÓN, ADQUISICIÓN, CATALOGACIÓN, CLASIFICACIÓN, DISEMINACIÓN, "INFORMATION RETRIEVAL"

Muchos investigadores tienen su propia biblioteca personal que contiene revistas, libros e informes. Esta pequeña biblioteca particular es un modelo en miniatura de bibliotecas más grandes, de centros de documentación. Idealmente cada investigador debería seleccionar y leer cada nuevo documento pertinente a su campo de trabajo. La imposibilidad absoluta de hacer esto es la mayor justificación para la existencia de un centro de documentación, de un servicio de información científica. Y explica, al mismo tiempo, una de las dos funciones principales de dichos centros de documentación: escoger entre todo el material recibido lo que es pertinente al trabajo de cada investigador.

Este proceso comprende la selección de informaciones pertinentes y la condensación de estas informaciones en un mensaje que el investigador individual pueda comprender. En muchos casos, los servicios gubernamentales con institutos de investigaciones, pueden tener un pequeño centro de documentación, alguna biblioteca en manos de una persona calificada. Y entonces esta persona prepara una biblioteca de trabajo para el personal de su centro, de su servicio, de su instituto. Si el número total de libros y revistas que viene a este centro de documentación puede ser tratado por una sola persona, entonces se puede conseguir muy bien un buen servicio personalizado para cada investigador. Si el bibliotecario o documentalista conoce íntimamente su colección de libros y revistas así como las necesidades de cada investigador, entonces, si vale algo, será capaz de presentar nuevos documentos pertinentes a cada investigador individualmente.

Pero desde el momento en que el volumen de material bibliográfico recibido en el Centro sea demasiado grande para una sola persona, entonces se multiplicarán los problemas, porque hay que tener una división de trabajo y eventualmente una división por sectores.

Básicamente, entonces, hay seis problemas que hay que tratar de resolver. El primero es un problema de identificación. Claro, antes de poder adquirir documentos nuevos es necesario saber lo que existe. Los científicos se enteran en la mayor parte de los casos, a través de comunicaciones personales, o revistas científicas, de la existencia de nuevos estudios. Los bibliotecarios y documentalistas además, pueden fiarse de ciertos instrumentos bibliográficos tales como catálogos de editores, bibliografías nacionales, y una gran

cantidad de otros instrumentos que conocemos todos. Es obvio que sigue siendo muy difícil la identificación de ciertos tipos de documentos, tales como publicaciones oficiales, documentos de organizaciones internacionales (aquí puedo abrir un paréntesis para decir que en la OIT sabemos algo del problema de tratar de identificar los propios documentos producidos por la misma OIT) e informes técnicos.

Es particularmente difícil identificar estos tipos de documentos, aún cuando dispongamos de índices adecuados a ese tipo de material. Surge pues el primer obstáculo, el primer problema: el de la identificación. Una vez identificado el documento, es necesario decidir primeramente si vale la pena adquirirlo. Y en segundo lugar, cómo adquirirlo. Se puede decir que en muchas disciplinas, especialmente en formación profesional, se escribe muchísimo. Pero de todo lo que se escribe y se imprime, ¿cuántos documentos vale la pena adquirir, analizar o conservar?

En nuestros días ninguna biblioteca puede ser completa. No hay ninguna biblioteca imaginable que pueda contener todos los documentos producidos sobre tal o cual materia. Entonces, nos hacen falta criterios de selección. Una vez que decidimos, pues, cuáles son los documentos que queremos adquirir, queda la otra parte del problema, es decir la adquisición física del documento. El tercer problema, pues, es la catalogación. Es necesario preparar una referencia bibliográfica según ciertas normas universales reconocidas para permitir después una identificación inequívoca del documento por autor o título, de todos los documentos contenidos en la colección. Se prepara, pues, una especie de tarjeta de identidad del documento que contiene estas informaciones y se añade, en general, el nombre del editor, la fecha, el lugar de publicación y otros datos. Esto para hacer lo que llamamos catalogación descriptiva de los documentos.

Un cuarto problema (problema que nos interesa sobre todo a nosotros) es el de la clasificación. Si vamos a preparar un mensaje para los investigadores para utilización futura de las colecciones, mensajes que pueden transmitir ideas acerca del contenido intelectual de un documento, es necesario formular estos mensajes en un lenguaje que el lector pueda comprender. Si utilizamos el lenguaje natural del lector para la descripción de documentos, se crearán dificultades serias, nacidas del hecho de que el lenguaje natural es demasiado ambiguo para permitir una selección científica de documentos más adelante. A través de las generaciones, entonces, se ha preparado un cierto número de sistemas de clasificación cuya mayoría es errática. El más difundido ha sido el sistema de clasificación decimal universal llamada clasificación de Bruselas, que deriva de una clasificación norteamericana, preparada hace un siglo por Melville Dewey. Estas clasificaciones se dicen universales en carácter, y se pueden utilizar con cierto éxito en diversas situaciones, pero son especialmente difíciles de utilizar en distintos sectores de las ciencias sociales, donde es preciso subrayar ciertos enlaces conceptuales. Hoy en día, pues, la tendencia es hacia la utilización de descriptores, escogidos de un vocabulario bien definido o thesaurus, como suele decir.

Un quinto problema es el de diseminación de estas informaciones. Una vez identificadas, adquiridas, catalogadas y clasificadas, hay que llamar la atención del investigador hacia el documento, lo que se puede hacer de varias maneras. Se puede circular el documento mismo, se puede enviar una notificación personal, se puede preparar un boletín, una lista de adquisiciones nuevas, que se imprimen y circulan. Hay varias maneras de realizar esta diseminación. Y, "last but not least", como decimos, una fase muy importante en el problema de "retrieval", es decir, de localización, de selección de informaciones: una vez terminado el ciclo precedente, el documentalista debe estar equipado para encontrar documentos que han sido registrados.

He evocado los seis problemas mayores que componen nuestro trabajo de documentalista. Antes de seguir quiero hablar otra vez del ciclo total de información y del ciclo de documentación. Se puede estimar que una investigación seria puede tomar, digamos, nueve meses de estudios, después de lo cual el investigador escribe un informe, un artículo. Este trabajo puede durar, digamos, tres meses más. Entonces este informe, este artículo se presenta a una conferencia de especialistas: otros tres meses. Deciden, después de la conferencia, editar las actas de la conferencia, y esto puede tomar, supongamos, dieciocho meses más. Entonces, un servicio de reseñas bibliográficas, un servicio de "abstracts", se entera de la existencia del documento, lo analiza y lo publica. Esto lleva, término medio, ocho meses más.

Entonces uno de nosotros, bibliotecario o documentalista, ve el resumen y pide el documento. Un mes más. Luego el documento llega, tenemos que catalogarlo, clasificarlo, etc. Tres meses más. Más tarde señalamos la existencia del documento en una lista de adquisiciones: otro mes. Si ustedes han hecho el cálculo aritmético conmigo, verán que el total de todo este ciclo de información suma cuarenta y seis meses, lo que no es una exageración. En la mayoría de los casos, si las cosas pasan así, está bien, porque hay muchos campos en que el ciclo será aún mayor que cuarenta y seis meses. Pero a partir del momento en que el investigador ha escrito su primer informe, pasaron, hasta llegar a las manos de otro investigador, 34 meses. Me parece que nuestro problema es cómo reducir este lapso de tiempo de 34 meses al mínimo posible.

Primeramente el problema de identificación. Ya existe cierto número de instrumentos de trabajo para los libros. Ciertos países publican catálogos de libros que van a ser editados. De esta manera se puede saber de antemano, con tres o cuatro meses de anticipación, cuáles son los libros nuevos que van a salir. Y existen índices por materias en estos catálogos de nuevos libros.

El libro, sobre todo en formación profesional, es un instrumento bastante lento y nos interesa, me parece mucho más, la revista y el informe técnico. La necesidad de trabajar rápidamente se ha hecho sentir, sobre todo, en las ciencias físicas, donde la materia cambia a una velocidad extraordinaria. Aún la revista científica, para la mayor parte de los investigadores en ciencias fí-

sicas, es demasiado lenta y han tratado de encontrar nuevas maneras de diseminación muy rápida de informaciones científicas nuevas. En física, por ejemplo, existen dos servicios semanales, uno que se llama Physics Letters y otro que se llama Physical Review Letters que son, como dice su título, cartas escritas por científicos sobre tal o cual tema. Estas cartas no pueden tener más de 800 palabras o dos páginas de texto y el científico, el autor mismo del trabajo, tiene que pagar U\$S 60 por página para editar su carta en la revista. Además, el autor tiene que pagar U\$S 10 para la inclusión de su artículo, de su carta en la revista de resúmenes analíticos, Physics Abstracts. Entonces paga U\$S 60 por página y esto le da derecho a 100 ejemplares de tiradas aparte del artículo, que puede distribuir entre sus amigos, entre sus conocidos, en el campo científico que le interesa.

Esta idea de una nueva forma de revista bajo la forma de cartas, Physics Letters, es una manera de tratar de eliminar un poco este lapso necesario para la impresión y la difusión de nuevos conocimientos científicos. Existen también otras maneras de hacerlo. Por ejemplo, en ciertos campos científicos, se imprimen de antemano los sumarios de nuevos números de revistas y se puede difundir este sumario varias semanas antes de la publicación de la revista misma.

De todas maneras, en las revistas científicas, es normal que todos los artículos en un número dado de la revista, no puedan interesar a todo el mundo. Cada investigador lee solamente los artículos que le interesan a él. Es así que existe cierta cantidad de revistas generales, para hablar otra vez del campo científico, revistas como Scientific American, Nature o Science que son revistas generales para el mundo científico en general, para tener a los científicos informados de los nuevos conocimientos en distintos campos del trabajo. Y ahora hay otra tendencia también que se nota sobre todo en Francia. Es decir, los editores de revistas científicas se dan cuenta de que no todo el mundo se interesa por todo. Por esta razón, empiezan a editar la revista científica, no como un cuaderno entero, sino por fascículos separados, que se pueden dar a varias personas distintas. Esta es tal vez una tendencia nueva que, claro, va a complicarnos a nosotros la existencia, puesto que nosotros los bibliotecarios tenemos siempre la idea de archivos, de conservación de una revista científica encuadrada, y la idea de editar una revista científica por fascículos separados nos va a crear cierto número de problemas para resolver en cuanto a su conservación.

Otra cosa bastante reciente también es la profusión de fotocopias. Ahora no es costoso sacar fotocopias de artículos científicos. Entonces el problema de circular, de conservar números enteros, años enteros o volúmenes enteros de revistas se plantea un poco menos, dada la posibilidad de hacer con poco gasto copias en gran escala, fotocopias de artículos. Otra cosa que hace cierto número de centros de documentación es fotocopiar las tablas de materias de nuevas revistas recibidas que luego circulan entre los investigadores, en vez de circular las revistas mismas.

Ha hablado nada más que de la revista científica y de ciertas tendencias nuevas, ciertas formas nuevas de presentar informaciones científicas para tratar de luchar contra este problema de esperar 34 meses, a partir de la fecha de publicación de un artículo hasta el momento de su utilización efectiva en el laboratorio. Sin embargo, hay un peligro muy grande cuando se habla de editar fascículos separados y cuando se habla de mandar tiradas aparte, (o en física nuclear, lo que pasa es que ni siquiera hacen tiradas aparte, no publican el artículo del todo, es decir hacen lo que se llama "reprints" que se distribuyen entre otras personas interesadas de la comunidad científica). Y estamos asistiendo pues, a la creación de aristocracias nuevas compuestas por científicos que trabajan en el mismo campo y el forastero puede difícilmente penetrar dentro de este círculo aristocrático. Entonces no hemos visto el final de los problemas creados por nuevos tipos de publicación o de difusión de artículos científicos o informes técnicos.

Hemos hablado un poco de la revista, del informe técnico. No quiero terminar esta segunda parte sin evocar una categoría importantísima para nosotros, para todos los servicios que se ocupan de formación profesional y aprendizaje, es decir, de las publicaciones oficiales, de las publicaciones gubernamentales. Es evidente que los distintos servicios nacionales producen una cantidad impresionante de documentos importantes y útiles para todos, pero además es necesario, para el propio servicio nacional, disponer en su propia biblioteca o centro de documentación, de documentación oficial sobre otros temas, para poder apreciar el cuadro general en que se inserta todo el trabajo de formación profesional y aprendizaje. Por ejemplo, me parece normal que un servicio así disponga, por lo menos, del plan nacional de desarrollo económico y social, del presupuesto de la nación, del informe del Ministerio; sobre todo informes del Ministerio de Educación, del Ministerio de Trabajo, de encuestas sobre la planificación de la mano de obra, anuarios estadísticos, etc., etc. Se podría hacer una lista bastante larga de lo que podríamos llamar documentos de base. Esto sí es necesario, es absolutamente urgente tenerlo. El plan nacional presenta para nosotros trabajando en el dominio internacional otro tipo de problema, es decir, fuera del país mismo es a menudo muy difícil saber la existencia de estas diferentes clases de documentos. Luego de conocer la existencia de esta clase de documentos se plantea otro problema que he mencionado anteriormente: adquirirlos. Pero estos son ya problemas que ustedes conocen mejor que nadie y si los menciono es que me parece primordial para el desarrollo de todo nuestro trabajo de formación profesional el poder fertilizarnos mutuamente (si fertilizar es la palabra digna de este pequeño curso) con informaciones nacionales de cada uno de esos países. Porque en definitiva, cada uno tiene algo o mucho para contribuir al progreso y al desarrollo general de toda la comunidad científica. Y espero que a través de esta colaboración nueva, nacida en parte gracias a la iniciativa de CINTERFOR, es que todos podremos disponer en el futuro de mejores servicios bibliográficos y documentales en el dominio de la formación profesional.

III. INTRODUCCIÓN AL EMPLEO DE DESCRIPTORES

Volvamos un momento diez años atrás. Un investigador de la casa IBM, el Ing. Hans Peter Luhn, lanzó la idea del índice KWIC. ¿Qué quiere decir el índice KWIC y cómo pensaban los científicos hace diez años? Esto podría ser útil para los documentalistas y el mundo científico en general. Tenían la noción de que podían utilizar, de manera general, el título de artículos científicos para indizar el contenido intelectual del artículo, y pensaban hacerlo de la manera siguiente: utilizando computadoras electrónicas podían sacar del título, cada palabra y considerar cada palabra como una palabra clave y permutar el título del artículo alrededor de aquella palabra clave. No tomaron en cuenta, por supuesto, ciertas palabras que no tenían sentido, como artículos, preposiciones, etc. Pero en fin, en el idioma inglés pensaban que no había más que una docena de palabras que verdaderamente no tenían sentido. De ahí los índices KWIC (que quería decir: "key word in context", o sea palabra clave en el contexto). Se preguntaron sobre una manera de índice permutado por orden alfabético de palabra clave. Era evidente que no tenían más valor que el valor del título mismo y sabemos muy bien que es muy raro en ciencias sociales que el título tenga gran valor para revelar efectivamente el contenido intelectual de un documento.

Pero era claro que esa técnica de permutación del título implicaba poco gasto en su aplicación y costaba menos dinero que utilizar los servicios de un documentalista profesional para indizar los mismos documentos. Estos estudios sobre la utilización de palabras claves también se vincularon con otros estudios sobre la posibilidad de indización automática a partir del lenguaje natural y también de traducción automática. Se puede decir que durante varios años, la IBM y otras casas y organismos gubernamentales gastaron varios millones de dólares en investigaciones sobre la posibilidad de realizar traducciones automáticas, indización automática por máquinas electrónicas. Estas investigaciones revelaron que por el momento no parece ser una cosa que podemos fácilmente realizar, pero aprendimos mucho acerca de problemas de construcción de lenguaje natural, que tal vez algún día nos permita, dentro de una generación o dos, ir mucho más allá de lo que somos capaces de hacer. No por la tecnología que tenemos y es buena, sino por la teoría lingüística que tenemos hoy en día, que no es adecuada. Pero de todo esto salió una cosa evidente: la noción de que podíamos utilizar sólo palabras claves sacadas de un texto en lenguaje natural no era suficientemente concreta como para permitir una indización buena del contenido intelectual de su documento. Por ejemplo, muchos conceptos en que

nuestra mente se revelan directamente como tales, como por ejemplo el caso de formación profesional, que es una noción de dos palabras.

¿Cómo una máquina electrónica puede, leyendo un texto en lenguaje natural, descifrar que se trata de un solo concepto y no de dos? Porque formación es una buena palabra, profesional también. En el sistema de palabras claves en los índices KWIC, etc., si había la noción formación profesional, lo considerarían como dos palabras clave para el índice. Pero debemos convenir en que formación profesional es una sola noción, un solo concepto. Esto no revela nada nuevo para los bibliotecarios, porque tenemos desde hace ya muchos años, encabezamientos de materia, que utilizamos en nuestras fichas de catálogo, así como los índices de la propia clasificación decimal universal que son nociones, son conceptos enteros. El problema está, pues, en las máquinas electrónicas: no podemos utilizar las palabras solas como medio de indizar un documento. Había que tratar de utilizar conceptos unitarios que no eran tomados de un vocabulario abierto. Veíamos muy bien la necesidad de construir un vocabulario para cada disciplina que queremos estudiar. De modo que con el tiempo, se han ido revelando los diferentes componentes que hoy nos permiten disponer de un vocabulario de base sobre los problemas que tratan la formación profesional.

Y para ponerse de acuerdo con la terminología que queremos utilizar, diría que a estos conceptos les damos el nombre general de descriptores. Haremos entonces una pequeña investigación sobre la manera de utilizarlos. Primeramente existe el problema de construcción de un vocabulario, de un "thesaurus" que podamos utilizar para el proceso de indización. Por regla general, para empezar un nuevo sistema con descriptores, partimos de cero, o sea que decidimos cuál es el cuadro dentro del cual queremos tratar de trabajar. Buscamos así una colección de documentos que corresponda a este cuadro. Tratamos de decir, pues, cuáles son los temas que queremos tratar en un servicio de documentación sobre formación profesional.

Empezamos por tratar de definir lo que entendemos nosotros por formación profesional, que puede variar de un país a otro, así como la noción de aprendizaje puede variar, al igual que casi todas las nociones. Por ejemplo, ¿qué quiere decir la noción de joven, de niño o de adulto? Puede haber variaciones bastante grandes y sorprendentes de un país a otro, como por ejemplo la definición (que nos parece fácil, pero que no lo es tanto) de palabras como libro o documento o revista. ¿Qué constituye un libro? ¿A partir de cuántas páginas es una obra un libro? Existen normas, claro, pero que no se aplican de la misma manera en todas partes.

Entonces es evidente que la primera cosa que necesitamos para construir nuestro vocabulario son pequeñas definiciones. No llegaría a decir que nos hacen falta definiciones de diccionario, pero por lo menos en el trabajo de indización y sobre todo en la selección de documentos nos hacen falta instrucciones

para el uso de estos descriptores, para evitar toda clase de confusiones. Si por ejemplo, para una noción como la de aprendizaje pueden existir varias definiciones diferentes, vamos a tomar la definición que a nosotros nos parece corresponder a nuestras necesidades como la mejor, estableciendo por escrito cuál es, a nuestro juicio, la utilización del concepto que nosotros queremos en nuestro servicio de documentación y sabremos (y los documentalistas que van a utilizar el servicio sabrán) que para nosotros aprendizaje quiere decir tal y tal cosa y no otra. De esta manera, y solamente de esta manera podremos eliminar del sistema de documentación toda ambigüedad. La ambigüedad es uno de los peligros mayores que debamos afrontar en la construcción de un buen sistema de documentación, al igual que la sinonimia. Tenemos que descartar definitivamente los sinónimos utilizando referencias, u otra cosa. De otra manera vamos a tener un vocabulario enorme que sería muy difícil de aplicar.

La idea de base de tener estos descriptores es la siguiente: cuando lleguen a nuestras manos documentos nuevos, sobre tal o cual materia, el documentalista va a mirar el documento, lo hojeará, si es corto, hasta puede leerlo y tratar de sacar del documento suficientes nociones básicas que podría traducir en la terminología de nuestro sistema de documentación, es decir, en los descriptores. Y para cada documento va a insertar la tarjeta, la hoja de papel con los seis, ocho, diez descriptores más importantes que para él corresponden a las materias sobresalientes de este documento. Estos descriptores se van a registrar según varios sistemas (aquí no voy a entrar en las distintas técnicas para registrar estos descriptores). Pueden ser técnicas completamente manuales que funcionan muy bien, o pueden ser como es nuestro caso en la OIT en Ginebra, un sistema base de computador electrónico. Porque poco importa, como he dicho anteriormente, la técnica, la tecnología, sino que es un problema de dimensiones y de objetivos finales más que otra cosa. Se puede hacer todo esto sin una máquina electrónica. Lo que nos interesa es la teoría de cómo se construye el vocabulario de descriptores y cómo se utiliza en la práctica.

Existen en el comercio tarjetas que uno puede comprar para utilizar en un servicio manual, o sea un servicio que trabaja sin máquinas electrónicas. Hay dos sistemas más conocidos que otros: uno que se llama tarjeta uniterm, que es nada más que una tarjeta en la cual se escriben el nombre de un descriptor y los números de todos los documentos analizados que contienen este descriptor en su descripción analítica. Entonces cuando el documentalista hace una búsqueda de informaciones comparando dos tarjetas, busca dos descriptores distintos, por ejemplo: para saber lo que tiene el Centro sobre aprendizaje en Chile, mirando la tarjeta de aprendizaje y la tarjeta de Chile, trata de ver comparando las dos tarjetas si el mismo número de documento se encuentra a la vez en las dos tarjetas distintas.

Se puede hacer la misma cosa perforando un agujero en el lugar indicado en la tarjeta para el número de documento. Esto es lo que llamamos en inglés el sistema "punch-a-book" y entonces se puede hacer la búsqueda de informaciones colocando una tarjeta encima de la otra y mirando frente a una luz eléctrica cuando son los agujeros perforados a la vez en las dos o tres tarjetas (lo que se puede ver mirando la luz que atraviesa los agujeros). Son dos sistemas manuales que no cuestan casi nada y que pueden ser adecuados para el documentalista en un pequeño servicio de documentación.

Podemos, sin embargo, volver un momento al sistema electrónico que nos permita hacer un poco más. Otro sistema desarrollado por la misma casa IBM es lo que ellos llaman selective dissemination of information, es decir, deseminación selectiva de informaciones. La idea es muy sencilla: si vamos a indizar documentos utilizando descriptores en un vocabulario fijo, podemos también indizar personas con los mismos descriptores. Es decir, vamos a preparar un perfil personal para cada investigador, perfil que va a corresponder a sus necesidades documentales. Si hay una persona que se ocupa de la formación de técnicos en la industria química de esto podemos sacar varias nociones distintas: la noción de formación, la de técnicos y la de industria química. Son tres nociones distintas. Esta misma persona también puede preocuparse de otros asuntos. Todo esto, todos sus proyectos de investigación y sus necesidades personales se pueden indizar como se podría indizar un documento.

La idea, el sistema SDI, es nada más que comparar electrónicamente el perfil personal de los investigadores con los perfiles en descriptores de los documentos que llegan al servicio de documentación. De esta manera, cada vez que la máquina encuentre una correspondencia entre el perfil personal y el perfil del documento, imprimirá una tarjeta que se manda al investigador y cuya llegada se anuncia al servicio de documentación, dando su título, autor, etc. La idea de base, pues, es muy sencilla. Es algo que se puede hacer utilizando máquinas electrónicas. Hasta hoy no se ha utilizado mucho porque es bastante largo; aún para la máquina electrónica, las diferentes operaciones son bastante largas. Pero en fin, es una manera, un medio de atacar, sobre todo en una escala bastante grande, la personalización de un servicio de documentación para entrar en un contacto directo con los clientes, lectores.

Se habló hace varios años y seguimos estudiando el problema de lo que los norteamericanos llaman information retrieval. No sé exactamente cómo traducir esta noción al castellano, supongo que es algo como localización selectiva de informaciones. En realidad, lo que llamamos "information retrieval" significa más bien, selección de documentos sobre las referencias bibliográficas, porque en la mayor parte de los sistemas existentes es exactamente esto lo que se hace. Los sistemas no permiten revelar los datos contenidos en los documentos, sino que facilitan solamente el acceso a referencias bibliográficas o a los documentos. Pero organizar las estrategias de selección de estos documentos constituye un problema intelectual bastante nuevo, porque por un la-

do al analizar el documento, la persona que prepara los descriptores -que tratan de revelar el contenido intelectual de un documento- no puede saber en el momento de hacer el trabajo de indización cómo esta información va a ser utilizada dentro de dos o tres meses, dos o tres años o dentro de veinte o treinta años. Entonces está forzado a hacer este análisis de manera bastante general. También hay otro problema: no hay dos documentalistas que mirando el mismo documento lo interpreten exactamente de la misma manera, es decir, aún con un vocabulario bien definido. Y tratándose de documentos sobre un tema general común, tal como la formación profesional, es muy raro que dos personas vean exactamente el mismo documento de la misma manera y no es sorprendente que los descriptores que asignarán al documento no sean exactamente idénticos.

Entonces, al principio, hay un problema de precisión o de imprecisión y hay que tomar esto en cuenta al formular las preguntas para utilizar el sistema después. Luego hay que formular las preguntas de manera tal que, por una parte, no se echen a perder referencias válidas y que por otra parte, no tengamos como respuesta docenas de referencias inútiles que no corresponden exactamente a lo que buscamos.

Otras veces, el problema es el de organización intelectual del trabajo y desgraciadamente todavía sabemos muy poco. Nos falta la parte teórica para orientarnos en la organización científica de este trabajo. Sin embargo, la máquina electrónica podrá ser muy útil para guiarnos en la formulación de esta teoría de la información. En la cuarta parte veremos una realización concreta del Servicio Central de Biblioteca y de Documentación de la OIT y cómo es posible utilizar -para tratar de encontrar maneras nuevas de solucionar el problema- las computadoras electrónicas.

IV. EL SISTEMA ISIS DE LA OIT

Para esta parte hará falta mirar dos documentos que he mandado hacer varios días por avión a Montevideo: uno que se llama Utilización de computadores para la selección de informaciones en el sector de las relaciones de trabajo y el segundo, un documento que lleva el No. L D No. 33 y que se llama Table-dex y 100 index. Nota explicativa (1).

Al iniciar nuestros trabajos de estudios sobre una mecanización eventual del Servicio de Documentación de la OIT miramos a nuestro alrededor para tratar de cerciorarnos de lo que hacían los servicios ya existentes y que disponían de aparatos electrónicos. Nos dimos cuenta de una cosa un poco curiosa, es decir, parecía surgir por todas partes la noción general de descriptores. Sin embargo, y a pesar del hecho de utilizar sistemas de descriptores, unos y otros seguían escribiendo, preparando resúmenes analíticos o "abstracts" para los mismos documentos y hay casos, existen todavía, en que, para un solo documento, el Servicio de Documentación da descriptores, prepara una breve anotación y además prepara un resumen analítico. Tres maneras diferentes que un solo servicio de documentación utiliza para registrar informaciones sobre un solo documento. Esto nos pareció un poco exagerado y además, sabíamos que había un doble problema con la preparación de "abstracts". El primer problema era el costo de prepararlos, el segundo el factor tiempo: llevaba bastante encontrar un experto para escribir el "abstract" sobre tal o cual tema.

Entonces tratamos de ver si no era posible, utilizando aparatos electrónicos, encontrar un medio camino entre una simple lista de descriptores y un resumen clásico, entero. Encontramos que una fábrica de productos farmacéuticos en Basilea había introducido en su Servicio de Documentación un sistema así, que ellos llamaron Code scanning y que consistía en escribir frases en lenguaje casi abierto, en inglés, dentro de las cuales se encontraban los descriptores. Después de haber estudiado el sistema y los demás, decidimos que esto valía la pena investigarse un poco más. Y ahora en la OIT estamos en el cuarto año de un sistema que funciona a base de un tal sistema de análisis documentario, y si ustedes quieren mirar el pequeño folleto impreso, en la página 163, verán una anotación preparada por nuestro Servicio de Documentación y que utiliza este sistema.

(1) Ver páginas finales.

El primer problema era el de vocabulario. No queríamos inventar la rueda, como dije anteriormente. Examinamos, antes de empezar, veinte terminologías existentes, inclusive una terminología existente en el CIRF y decidimos tratar de incorporar a nuestro nuevo sistema de descriptores lo mejor que podíamos encontrar en estos veinte vocabularios existentes. De modo que no partimos, como ya dijera, de cero, sino a partir de algo, a partir de reflexiones de trabajo que otros servicios de documentación habían realizado antes. Pero teniendo en cuenta una cosa muy importante: la naturaleza misma de un sistema de descriptores en ciencias sociales, donde hay muchos cambios, muchas novedades, estaba en que el sistema de descriptores nuestro tenía que ser flexible y permitir una expansión, tenía que permitir cambio o intercambios de descriptores. O sea que, en otros términos, no queríamos empezar con una clasificación rígida, sino tener un sistema flexible. Y como no había estructura definida entre los descriptores, el documentalista individual tiene que decidir a qué nivel quiere indizar el documento, si quiere utilizar descriptores generales o más específicos. Además, le toca al investigador individual cuando presenta preguntas al Servicio de Documentación el decir si se interesa en puntos específicos o en problemas generales más amplios.

Mirando el texto de esta anotación ustedes notarán varias cosas: primeramente, los descriptores se señalan en el texto entre líneas diagonales. Esto no es nada más que un medio de señalar a la computadora la existencia de un descriptor. Por ejemplo, en la primera línea, hay un descriptor: "Manual workers", obrero manual. El descriptor no puede saber que "obrero manual" es un solo concepto, un solo descriptor. Por esta razón hay que indicarlo entre diagonales para que la máquina sepa reconocer dónde empieza el descriptor y dónde se acaba. Dentro del texto de esta anotación hay como siete u ocho descriptores utilizados. Los primeros tres están subrayados. Esto no es nada más que un pequeño truco interno nuestro que utilizamos para el fichero público. Los descriptores subrayados en el texto son los que archivamos en nuestros catálogos públicos, de los cuales tenemos dos en Ginebra, en dos lugares distintos. Pero todo está registrado en cinta magnética y en discos magnéticos si pueden utilizarse para investigaciones, pero únicamente los que están subrayados pueden servir para investigaciones manuales en los ficheros de la biblioteca.

En estas pequeñas anotaciones tratamos de hacer dos cosas: es evidente que tratamos de indicar en pocas palabras el contenido intelectual del documento y además, cuando es posible, tratamos de indicar las conclusiones del autor. Con la mayoría de los documentos esto es imposible pero ocurre, sobre todo, con artículos de revistas que es posible señalar, en líneas generales, las conclusiones del autor. De todas maneras, la filosofía básica que sustenta estas anotaciones es la siguiente: un resumen analítico de tipo clásico de todas maneras no puede reemplazar el artículo del documento original. Si esto fuera posible ¿por qué guardaríamos los originales? El resumen analítico clásico no puede hacer más que guiar al lector. Si podemos guiar al lector, un poco más de lo que se puede hacer con los métodos tradicionales de

biblioteca, si podemos realizar esto sin los gastos de tiempo y de dinero que insumen los resúmenes completos, nos parece que valdrá la pena intentar realizarlo.

De todas maneras existe, y va a existir todavía durante muchos años, el trabajo inicial de análisis del contenido intelectual del documento, sea para escribir resúmenes analíticos clásicos, sea para escribir anotaciones del tipo OIT. Este trabajo inicial hay que hacerlo. La idea de registrar todo esto en cinta magnética es para poder utilizar en distintos rumbos y para diferentes aplicaciones, esta información básica que registramos una sola vez. El pequeño folleto contiene una explicación mucho más detallada de cómo funciona el sistema y es por esto que no es necesario que ahora yo explique esto nuevamente, pero lo que sí trataré de explicar con más detalles es cómo funcionan los índices de tipo nuevo que hemos estudiado en la OIT y para esto les pido mirar el folleto que lleva el No. L D No. 33, empezando con la última página que lleva el descriptor automatización encima de la página.

Este índice de aspecto un poco primitivo no es, como podrán imaginarse, un índice KWIC clásico sobre los títulos de un documento. Se fabrica de la manera siguiente: la computadora electrónica lee cada anotación y saca de la anotación en el contexto de la anotación cada descriptor y ordena los descriptores por orden alfabético. Para cada descriptor en cada documento imprime una línea de índices de la manera siguiente: tomando el ejemplo automatización, ustedes verán a la izquierda dos cifras: 65, 64, 63. Estas cifras representan el año de publicación de un documento en orden invertido, es decir, primeramente los años más recientes. Al lado del año de publicación figura el número de documento en el servicio de documentación, el número del resumen de la anotación y luego el idioma en que está escrito el documento. Hay algunos casos aquí de documentos escritos en tres idiomas o más. En este caso el computador tiene la instrucción de imprimir la misma línea dos veces para poder incluir los otros idiomas. Y entonces sale el índice propio que se compone de la manera siguiente: (es algo similar al índice KWIC) tomando la palabra "automatización"; automatización en cada anotación se saca del contexto y se imprime una sola vez encima de la página, pero dentro de cada anotación el descriptor "automatización" se inserta en el espacio central blanco. En el medio de la página hay una columna blanca, que corresponde al sitio del descriptor "automatización" en una anotación. A la izquierda de esta columna blanca se encuentran todos los descriptores que se encuentran en la anotación original a la izquierda del descriptor "automatización", y a la derecha se encuentran los descriptores que en la anotación original aparecían después del descriptor "automatización". A la izquierda como a la derecha hay sitio para cincuenta caracteres de información; es por eso que lo llamamos índice. Y todo eso se conserva en el contexto original y tratándose aquí sólo de descriptores y no del lenguaje natural es muy fácil cuando dispongamos de glosarios en otros idiomas para el vocabulario de descriptores, instruir a la máquina para que imprima esos índices en francés o en castellano o en cualquier otro idioma que se podrán consultar exactamente de la misma manera.

En forma inversa, otra cosa es posible. En el sistema actual de la documentación de la OIT en Ginebra utilizamos solamente inglés como idioma de entrada en el sistema, pero es evidente que existiendo un vocabulario de descriptores en otros idiomas se podría hacer el trabajo de análisis y anotaciones en otros idiomas y con la misma máquina electrónica producir todos los índices también en inglés, o sea hacer el proceso contrario.

Volviendo un momento a este índice se verá también otra cosa: que encima de la página se encuentran todas las referencias de los documentos que tratan del problema de la automatización en general y que después tenemos el descriptor "automatización" subdividido por países, Bélgica, Canadá, países en vía de desarrollo, Europa, Francia, Alemania, etc. Eso nos permite muy a menudo con un vistazo consultar este índice para hacer investigaciones bastante complicadas y eso suponemos que nos va a permitir evitar de correr a la máquina cada vez que tenemos una pregunta un poco complicada para formular: deberíamos poder utilizar estos índices para contestar la mayoría de las preguntas que hemos recibido hasta hoy. Evidentemente la idea clave de todo esto es utilizar estos índices en sitios donde es imposible tener acceso a un aparato electrónico.

En la página anterior hay otro índice de aspecto todavía más primitivo que el índice en letras: es lo que llamamos el Tabledex. Aquí el problema es un poco diferente y una vez que se sabe manejar, este índice tiene que resolver otro tipo de preguntas. En el índice de 100 caracteres que hemos visto es posible perder ciertas informaciones a la izquierda o a la derecha, porque claro, en cada línea de texto hay sitio solamente para 100 caracteres de información, mientras que aquí se encuentra para cada documento todos los descriptores. Pero hay un pequeño problema: el descriptor se reemplaza en la máquina por un código de cuatro cifras. Entonces, para poder utilizar el índice hay que tener las tablas con el código para encontrar el número correspondiente a cada descriptor. El índice se consulta de la manera siguiente: supongamos el caso del investigador que se interesa en el problema de la formación de técnicos en la industria química. Consultamos las tablas para ver cuáles son los números código correspondientes a formación profesional, industria química, técnicos y, si queremos añadir un cuarto, puede ser Venezuela. Tenemos los números códigos de estos cuatro conceptos y miramos la tabla que tenga el número más bajo y estos números se dan por frecuencia de utilización. Es solamente la máquina electrónica que puede detectar esto, es decir, los conceptos utilizados con mayor frecuencia tienen los números más grandes, pero mirando bajo el número del código menor, en este ejemplo OO 15 A F A, se puede encontrar probablemente mirando una sola página de un índice. Aún para una colección de documentos de varios miles o docenas de miles de documentos, uno debería encontrar las respuestas a una pregunta mirando una sola página o dos. Es pues la manera de clasificar los documentos por orden de frecuencia de utilización. Se ve, por ejemplo, aquí en este caso he hecho un círculo alrededor de dos cifras, se supone que buscamos tres conceptos, uno que es el OO 15, Africa, otro que era 168 y otro el 252. Se ve mirando la tabla que

Hay sólo tres documentos que corresponden a estos tres criterios, hay varios más que contienen dos de los tres, pero hay solamente tres que tienen estos tres números corridos. Los problemas con este índice son: primeramente supongo que sólo un documentalista tendrá la paciencia de utilizarlo. Otro problema: es evidente que hay que traducir el descriptor en un número código, pero además los descriptores aquí están ordenados en orden numérico ascendente, es decir que no están en el contexto original de la anotación. Y supongamos también que vamos a encontrar de vez en cuando documentos que son documentos falsos, que tengan los tres números de código que buscamos pero que son documentos que tratan de otra cosa, simplemente porque es la combinación de esos tres descriptores que por casualidad se encontraron en la misma anotación, lo que probablemente se va a producir menos con el otro tipo de índice. Estos son dos ejemplos de tipos completamente nuevos de índice que podemos realizar a partir de un solo trabajo intelectual original, es decir el trabajo de análisis del documento bajo la forma de estas anotaciones, utilizando un vocabulario bien definido de descriptores. Es evidente que con un poco de imaginación se podrían concebir otros tipos de índice o de servicio.

Quiero finalizar evocando una sola posibilidad para el futuro, que para algunos puede parecer bastante lejano pero que técnicamente podríamos hacer mañana por la mañana. Es lo siguiente: todos estos sistemas de documentación son sistemas indirectos; es decir, el investigador puede tener acceso a estos documentos sólo a través de un servicio de documentación o a través de una biblioteca, pero sería formidable para el investigador que él pudiera dialogar directamente con la máquina electrónica para tratar de hallar las respuestas a sus preguntas, lo que no es un sueño imposible. Veremos en los meses siguientes cómo organizar una experiencia con el sistema existente en la OIT; justamente del diálogo hombre-máquina de este tipo. Creo que es falso hablar de diálogo hombre-máquina, es mucho más. Es un diálogo entre el investigador y todos los documentalistas que van haciendo este trabajo intelectual de análisis de los documentos. Es increíble que vamos a poder disponer dentro de poco tiempo de un instrumento que permita un diálogo efectivo entre todos nosotros y ya sabemos cuántos investigadores en el mundo entero. Es una nueva era que aún para mí -tengo que confesar- de vez en cuando se parece un poco a la ciencia ficción, pero no lo es. Disponemos hoy en día de los aparatos; no cuesta tanto como lo que se pudieran imaginar, nos hacen falta años y años y años de intenso trabajo intelectual para saber cómo hacerlo y para quién y con qué objetivo. Hay mayor cantidad de preguntas que de respuestas, pero finalmente esto importa poco. Nuestra generación vivirá una experiencia absolutamente nueva. Entramos en una verdadera era de participación activa de nuestra profesión, como profesión en la aventura humana. Nunca hemos tenido mayor oportunidad de participación. Esperamos aquí en Ginebra contar mucho con su colaboración en esta aventura y les ofrecemos la nuestra. Muchísimas gracias y hasta la vista.

NOTA EXPLICATIVA

a) TABLEDEX

El "Tabledex", compuesto con la ayuda de la computadora, permite suministrar referencias en las que se encuentran varios descriptores. Se establece un número N de tablas (N representa la cantidad total de descriptores utilizados en una lista dada), de forma tal que cada línea de cada tabla muestre la siguiente información, por cada documento inscrito en el sistema:

- Fecha de publicación del documento
- Número de referencia en el boletín
- Idioma(s) de la publicación.

y, al interior de la tabla, todos los números-código de los descriptores de cada anotación numéricamente superiores al número de código de la tabla.

En el ejemplo que figura en el anexo, los descriptores están clasificados por orden numérico ascendente de código en cada línea y ordenados de tal forma que los números de la primera columna de la tabla sean igualmente ascendentes. Toda anotación que comprenda el número-código del descriptor-clave, es decir, en el ejemplo que nos ocupa, el número 0015, está representado en la tabla.

He aquí la forma de consultar dicha tabla. Supongamos que nos interese encontrar documentación sobre el empleo en los países africanos, en el campo de los planes nacionales de desarrollo:

La primera operación será la de tomar los números-código de cada descriptor (plan nacional = 0168, empleo = 0252, Africa = 0015). Lo que nos proporciona en orden numérico ascendente: 0015, 0168 y 0252. Consultemos a continuación el Tabledex. Sólo la tabla que porta el número más bajo nos interesa con relación a esta demanda. En la tabla 0015 examinemos pues solamente las líneas cuyo primer número-código es inferior o igual al segundo número, o sea 0168. Para localizar el tercer número-código no nos queda más que examinar las líneas que contienen el número 0168 para ver si el número 0252 figura también en ella.

En el modelo presentado podemos comprobar que varias anotaciones comportan 2 de los 3 números-código requeridos (0015 + 0168 ó 0015 + 0252) pero veremos que sólo tres anotaciones comportan el conjunto de ellos. Por tanto será inútil de examinar la tabla más allá de la línea 18, ya que el número-código que figura al comienzo de la línea 19 es superior al número 0168, necesario a nuestra operación.

b) 100-INDEX

El "100-Index" es un índice permutado del tipo KWIC, consistente en una línea de texto por cada aparición de un descriptor dado en una anotación. Se retira el descriptor-clave de la anotación para colocarlo en el encabezamiento de la página, por ejemplo, en el caso que figura en el anexo II, el descriptor "Automation". A la izquierda del espacio central libre, se podrán encontrar los descriptores que, en la anotación, preceden inmediatamente al descriptor-clave "Automation" hasta el límite de 50 pulsaciones. A la derecha de dicho espacio central figurarán todos los descriptores que siguen al descriptor "Automation" en la anotación, igualmente dentro de los límites de 50 pulsaciones.

En este tipo de índice, la computadora no imprime las palabras que unen los descriptores entre sí; la permutación se realiza tan sólo sobre los descriptores citados. Para evitar márgenes de desigualdad puede ocurrir que sea necesario seccionar un descriptor, es decir que las primeras letras de dicha palabra puedan completar una línea y que el resto de la palabra se encuentre al comienzo de la misma línea. Vea la segunda línea del ejemplo que figura en el anexo. Es fácil deducir que añadiendo "ODUCTION" a las letras "PR" con que termina la línea leamos la palabra "Production", descriptor que figura en la anotación. Toda anotación comienza con el signo menos (-) y termina con el signo igual (=). En primer lugar figuran las obras generales, en orden cronológico inverso, comenzando por las más recientes. Las demás publicaciones están colocadas por orden alfabético de país (Bélgica, Canadá, etc.). Este tipo de índice permutado comporta además en cada línea la fecha de publicación de la obra citada, el número de referencia del boletín y el idioma (o idiomas) en que ha sido publicada la obra. Si la edición ha sido hecha en varios idiomas, la entrada vuelve a repetirse en otra segunda o tercera línea, a fin de que todas ellas estén incluidas. Véase por ejemplo la entrada No. 00183, repetida por tres veces en el modelo (anexo II).

Supongamos que alguien trate de hacer una investigación sobre las implicaciones económicas de la automatización en los países en vía de desarrollo y busque documentación sobre dicho tema. Consultando la tabla "Automation", y la subdivisión "Developing country" búsquense las entradas donde figure el descriptor "Economic implication". Podrán encontrarse seguidamente los números de referencia de dos documentos, o sea el número 00850 y el número 00059. El primero representa una publicación de las Naciones Unidas editada en inglés en 1964, y el segundo, otra del Consejo de Europa, igualmente editada en 1964, y publicada en inglés y en francés.

El "100-Index" reproduce los descriptores en el mismo orden que figuran en la anotación original. Ello permite consultar manualmente las tablas del índice sin el concurso de la computadora procediendo así a una clasificación sumaria que permita eliminar a continuación las entradas no pertinentes. Es de esperar que en el futuro gracias a los índices "Tabledex" y "100-Index" será posible satisfacer un gran número de demandas en el dominio de la documentación sin necesidad de recurrir a la ayuda de la computadora.

120.7.71

lvs

TABLE 0015 AFRICA

Anexo I

65	00355	ENGL	0016	0024	0690	0788	0796	0992	1051	1063	1079								
64	00287	ENGL	0024	0048	0252	0335	0475	0673	1060										
64	00215	ENGL	0024	0054	0195	0212	0223	0252	0270	0562	0673	0771	0841	0842	0960	1101			
64	00353	ENGL FREN	0024	0056	0404	0475	0824	1079	1123										
64	00357	ENGL FREN	0024	0102	0334	0475	0664	0760	0796	1051	1079								
64	00014	ENGL FREN	0058	0270	0624	1033	1087												
64	00259	ENGL	0078	0168	0443	0447	0486	0966	1026	1113									
64	00032	ENGL	0082																
63	00147	FREN	0082	0577	0771	0808													
63	00538	ENGL	0082	0803	0891														
64	00369	ENGL	0082	0803	0891														
65	00309	ENGL FREN SPAN	0120	0169	0169	0174	0242	0252	0303	0447	0458	0547	0683	0798	0814	0899	1016	1084	1116
65	00670	ENGL FREN SPAN	0149	0229	0389	0444	0447	0545	0549	0834	1053								
64	00313	ENGL	0168	0170	0212	0993	1079												
64	00354	ENGL FREN	0168	0170	0219	0252	0259	0667	1079	1113									
64	00313	ENGL	0168	0170	0303	0355	0466	0474	0495	0590	0897	0951	1076	1079					
64	00155	ENGL	0168	0170	1079														
64	00082	ENGL	0168	0253	1051	1079													
62	00293	ENGL FREN	0170	0979	1021	1087	1092												
64	00601	ENGL	0216	0417	0474	0671	0786	0992	1027										
64	00364	GERM	0218	0669	0834	1053													
64	00013	ENGL FREN	0219	1067															
64	00174	FREN	0244	0313	0334	0495	0602	1051											
64	00159	ENGL	0252	0960	1076	1079													
64	00148	ENGL	0255	0273	0405	0956	1019	1027	1113										
64	00303	ENGL	0270	0891	0899														
64	00602	ENGL	0474	0899	1076														
64	00485	ENGL	0519	0771															
64	00473	FREN	0775	0853	1053														
64	00793	GERM	0908	1169															
65	00350	ENGL	0992	1026	1077	1079													
64	00156	ENGL	1079																
64	00157	ENGL	1079																
64	00158	ENGL	1079																
64	00322	ENGL																	

