

# Sistemas de recuperación de la información en los archivos: un análisis de situación y perspectivas

Ramón Martín Suquía

De entre las diferentes funciones de trabajo del archivero la de tratamiento de la información tiene una importancia central debido, sobre todo, a que ésta es la función “puente” entre la de conservación y la de difusión, sin la cual no tiene mucho sentido la primera y no puede realizarse la segunda<sup>1</sup>. El proceso general de tratamiento de la información se subdivide en los procesos de análisis y de recuperación de la información. Mientras que en el primero de estos se describe el documento, en el segundo se pretende su localización. Ambos procesos están directamente relacionados de manera que el análisis de la recuperación de la información es, en gran medida, el resultado del éxito o el fracaso de las decisiones tomadas en el proceso de descripción.

El análisis de las diferentes formas y medios de recuperación de la información disponible en un servicio de archivo no ha merecido en los manuales “clásicos” de archivística un estudio individualizado. Esta falta de literatura archivística dedicada a la recuperación de la información<sup>2</sup> exige, solo con el ánimo de poder delimitar este artículo, proponer una definición que, como creo que no puede ser de otro modo, tomo de nuestros colegas los documentalistas: “La *recuperación de información* es el proceso de selección, en un fondo documental, de la información más adecuada a las demandas de los usuarios”<sup>3</sup>. Sin embargo no se puede decir que no haya existido una preocupación permanente por este proceso, ha quedado más bien imbuido en dos áreas de análisis mucho más estudiadas, la descripción de los documentos de archivo, y las cuestiones relativas al acceso a la documentación.

Este artículo pretende estudiar el estado de esta cuestión sobre todo en el ámbito de los países latinos, aunque se tendrá en cuenta el ámbito anglosajón como exponente de las diferentes posibilidades que se le presentan al responsable de archivo.

## 1. El archivero: gestor de información

En el día a día de los archivos, se puede afirmar sin temor a exageración, la recuperación de la información ha sido y sigue siendo una de

las principales áreas del quehacer diario del archivero. ¿Qué ha hecho y sigue haciendo sino el archivero todos los días en su archivo? Poner la información a disposición de los que la necesitan<sup>4</sup>. Todas las funciones del archivo (la conservación, la descripción,...) tienen como objetivo final la consulta de información.

El archivero, como gestor de información, no puede tener en cuenta la posibilidad de la descripción de documentación de archivo y la posterior recuperación de esta información si no es por medio de la utilización de recursos informáticos. La capacidad de tratamiento de masas de información que de otra manera no son aprehensibles de forma inmediata, la rapidez en la localización de dicha información, las posibilidades de un tratamiento y una mejora continuada de la información que se hace llegar a los usuarios, se convierten en razones determinantes para tener que utilizar estos medios. La informática no es la razón de ser del trabajo del archivo, pero es la herramienta sine qua non.

Por el contrario –creo que no debemos de olvidarlo nunca– de la experiencia acumulada hasta la actualidad parece que se puede colegir el que, más que un cambio de la función de tratamiento de la información en sí, lo que está cambiando son los medios, la forma de hacer el trabajo diario. El archivero al describir fondos de archivo interpreta<sup>5</sup>, y en la medida en que su trabajo es de interpretación, este no es sustituible por medios totalmente automáticos. Es cuando finalizada este proceso de interpretación cuando los medios informáticos le ayudan a él, y a sus usuarios, a recuperar la información que ha generado. La utilización de medios informáticos en el tratamiento de la información de archivo no da mayor valor añadido al trabajo de descripción en sí, que será bueno o malo en la medida en la que el archivero sepa reflejar el contenido y el contexto del documento, pero mejora de manera evidente las posibilidades de explotación de ese valor añadido. Acudiendo a un análisis económico clásico, “producir” cualquier cosa exige saber y hacer bien el trabajo, y luego saber servirlo, el análisis de la recuperación de la información nos enseña a servir; saber y hacer bien el trabajo es otra cosa.

La principal idea que se pretende transmitir en este artículo es, pues, que el proceso de tratamiento de la información de archivo se puede resumir como el proceso de suma de un activo intelectual a la información “bruta” de los documentos, y que esta suma está directamente condicionada por el método por el cual se transmite, y por la capacidad del usuario de extraerle el mayor provecho posible. El análisis detallado del modo de transmisión de la información, y de la cantidad y calidad de la información que llega al usuario es la razón de ser del análisis de la recuperación de la información en los archivos. Ni que decir tiene que, desde mi punto de vista, el análisis de los métodos y los resultados relacionados con la recuperación de la información en los archivos deben ser básicos en el trabajo de los archiveros. Cómo hacer que este activo intelectual que proporciona el archivero se transmita al completo, sin

perdida de calidad, con la mayor facilidad, y siempre teniendo en cuenta las posibles necesidades de nuestros usuarios es uno de nuestros quehaceres fundamentales.

## 2. La recuperación de la información en la política de gestión de archivo

En este comienzo de siglo se han combinado diferentes factores, externos e internos a los propios archivos, que exigen replantearse una y otra vez las formas de recuperar la información archivística. Estos factores hacen que esta función se haya convertido en una de las claves más importantes de los servicios de archivo entendidos no solo como custodios de una documentación sino, y sobre todo, como servicios atentos a las necesidades de sus usuarios. No parece, por el contrario, que la mayoría de los archivos hayan modificado sus estructuras para adecuarse a estos cambios, lo que constituye, como poco, un indicador de la falta de análisis permanente de la política de gestión del archivo por parte de los servicios.

Factores externos que han madurado ya como la democratización en el acceso a los archivos<sup>6</sup>, y la mejora en la capacidad de reproducir los documentos –al principio en papel o soportes fotográficos, luego, y coexistiendo con los anteriores, en soportes digitales–; y sumado a ellos, un factor más reciente pero extendido con una asombrosa rapidez, la expansión de los recursos informáticos para el tratamiento de la información de cualquier tipo<sup>7</sup>, han modificado el modo de hacer el trabajo diario de los servicios de archivo, tanto por la capacidad de tratar masivas cantidades de información, como porque esta sea consultada de manera ágil, y más recientemente, incluso a distancia por la capacidad de permitir ese tratamiento y consulta vía Internet.

La normalización de la descripción archivística constituía un factor interno pendiente en el ámbito de los archivos, y ha sido el punto de atención de los archiveros hasta hace muy poco tiempo. Una vez resuelta esta cuestión la recuperación de la información queda accesible al análisis, cosa que anteriormente, por la dispersión de esfuerzos y soluciones a la hora de definir la descripción se hacía imposible. Por medio de normas como ISAD (G), ISAAR (CPF) o EAD, disponemos de un formato normalizado sin el que nunca se hubiese podido plantear una política de análisis y recuperación de información bien delimitada. La existencia de la norma nos exige replantear una recuperación de información que debe, por ejemplo, ser capaz de expresar el concepto de descripción multinivel de manera inteligible para los usuarios<sup>8</sup>. O enviar y reenviar a los usuarios a través de información que describe por una parte documentos o grupos de documentos, por otras instituciones, por otro nombre de personas o lugares... y todo ello sin perderse; ¡Formidable trabajo;.

Esta combinación de factores supone un "cierre del círculo". En estos momentos, y por pura acumulación de razonamientos, cualquier ciudadano se siente sujeto de derecho para conocer de la existencia de un documento por medio de un sistema de información ágil y a distancia; para consultarlo in situ o en su entorno personal a través de una copia; para, incluso, solicitarlo sin llegar a pisar nunca el servicio de archivo donde se le atiende; y –no hacen falta grandes dotes adivinatorias– para que en un futuro pueda pretender consultar la imagen de ese documento en tiempo real desde el lugar en el que se encuentre. Si, además de ello, este ciudadano forma parte de la misma organización a la que atiende el servicio de archivo aún serán mayores las exigencias. Podríamos resumir esta situación como la expresión de una contradicción, la intermediación entre el documento y el usuario se va haciendo cada vez "menos personal" y por el contrario queremos que sea más personalizada. Solamente una buena definición de los sistemas de recuperación de información puede resolver esta contradicción. ¿Pero, han sabido responder los archivos a este reto? ¿Han sabido adecuar sus sistemas de información a estas nuevas circunstancias?

Un análisis de la situación de los sistemas de información de archivos deja bastante claro que en ciertas áreas geográficas, sobre todo en la Europa latina, los archivos han sido poco ágiles en el aprovechamiento de las nuevas tecnologías<sup>9</sup>. Cuando las han utilizado lo han hecho casi exclusivamente aplicándolo a los procesos de descripción, de gestión de los ingresos en archivos históricos y administrativos, y a los procesos de préstamo y control de usuarios, siendo pocos los casos de experiencias de informatización integral de los servicios de archivo<sup>10</sup>. Además en muchos casos las posibilidades de utilización de este recurso tienen un carácter de uso interno, y en muy pocos casos los usuarios los pueden utilizar como una forma de consulta directa. En lo relativo a la recuperación de información esta realidad es muy patente y requiere un decisivo cambio de paso por parte de muchos responsables de los servicios de archivo, un salto de una gestión dirigida a satisfacer necesidades internas a una gestión en función de las necesidades externas. No es de recibo que, por ejemplo, grandes archivos implanten costosos sistemas de control del préstamo sin que, en paralelo, no se haya hecho nada por facilitar la localización y el acceso a la información que custodian.

Se podrá argumentar que los servicios de archivo no disponen de los medios económicos y humanos para llevar adelante estas tareas. Pero, parece claro, ha existido en paralelo a esta falta de recursos una falta de capacidad de reestructuración de los medios económicos y humanos de que disponen muchos servicios de archivo.

Al contrario de lo que ocurre con el formato de descripción, en la recuperación no hay "recetas" únicas, cada servicio de archivo requiere un análisis diferenciado y la adopción de una política determinada; incluso esta política puede, y en ocasiones debe, ser diferente en los

diferentes fondos conservados, según sus características. Es necesario que, desde la dirección, se tengan en cuenta las circunstancias del servicio, de manera que se valore adecuadamente la asunción de una técnica determinada no por la posible sensación de adelanto con respecto a otros servicios sino como respuesta a las necesidades concretas que tiene este. La toma de decisión en el ámbito del análisis de los sistemas de búsqueda y recuperación de la información –directamente relacionada con la política de descripción<sup>11</sup>, anterior, y la política de difusión, posterior–, requiere la evaluación de aquellos factores que la determinan.

Se pueden señalar como factores que determinarán la elección de una política de recuperación de información en el archivo los siguientes:

1. El tipo de información conservado. Documentación que contiene información administrativa, muy estructurada y previsible o, por el contrario, información cognitiva, muy poco estructurada.
2. El soporte o legibilidad automática del documento. Permite la indización automática o no de la información, la reproducción en ficheros gráficos fácilmente transportables, etc.
3. Recursos económicos del servicio.
4. Recursos humanos del servicio.
5. El destino, en cuanto a conservación o no, de los documentos descritos.
6. El universo posible de usuarios.
7. Los tipos de usuarios. Son necesarias diferentes políticas para usuarios de archivos administrativos o de archivos históricos, usuarios expertos o inexpertos, etc.
8. Antigüedad de la documentación, que incide en la mayor o menor intensidad de problemas lingüísticos, de necesidad de explicación del contexto, etc.
9. El tamaño del fondo, que obliga a seleccionar niveles de descripción.

Se puede utilizar el análisis de las fases en las que se plasma el proceso de recuperación de la información como medio para explicar las diferentes políticas posibles en el ámbito de la recuperación de la información.

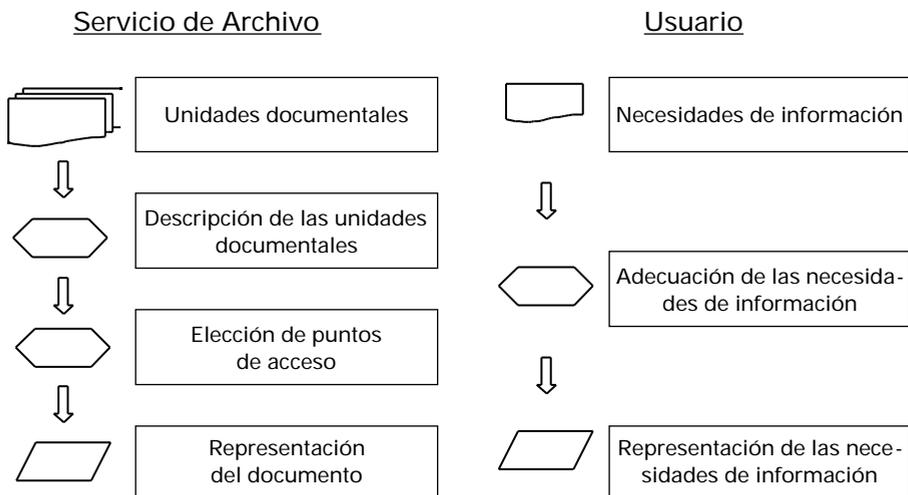
### 3. El proceso y las técnicas de recuperación de la información de archivo

Analizar el proceso que va desde la necesidad concreta que tiene un usuario de archivo hasta el o los documentos que realmente consulta es analizar una carrera de obstáculos, una carrera que solo en ocasiones alcanza llegar a su meta. En muchas ocasiones las responsabilidades sobre el éxito o fracaso de tal carrera se atribuyen exclusivamente a la

habilidad –o torpeza– del usuario, o a la habilidad –o torpeza– del archivero, obviando de esta manera un análisis mucho más estricto, que es el que realmente nos puede permitir encontrar las claves del éxito o del fracaso. La necesidad de este análisis se acentúa en la medida en la cual el contacto directo entre el usuario y archivero se hace más lejano. Y esta es, cada vez con más frecuencia, la situación que se produce en los archivos, bien porque los servicios permiten la consulta de los instrumentos descriptivos a distancia, bien por las limitaciones de personal de muchos de nuestros archivos, o por el natural reacio de muchos investigadores a transmitir sus necesidades de información al archivero.

La eficacia en la recuperación de la información en los archivos es el resultado de un contraste. El contraste entre la suma final de las decisiones tomadas en el proceso de descripción por parte del gestor de información, el archivero, y la suma final de las decisiones tomadas por el usuario en el momento de hacer la consulta. Se puede seguir un sencillo esquema de análisis basado en la descomposición e identificación de los diferentes elementos que intervienen en esta recuperación de la información<sup>12</sup>, analizando los elementos responsabilidad del archivero, y los que son responsabilidad del usuario. Este esquema nos servirá para un análisis del estado de situación de la recuperación de información en los archivos en la actualidad, y para plantear algunas proyecciones de futuro.

El modelo conceptual de los sistemas de recuperación de información está formado por los siguientes elementos<sup>13</sup>: las unidades documentales, las necesidades de información, la representación de las unidades documentales, la representación de las necesidades de información y la comparación de estas representaciones. De manera gráfica el proceso es el siguiente:



### 3.1 Las unidades documentales

Es en el comienzo del análisis conceptual de la recuperación de la información donde surge la primera toma de decisión del servicio. A diferencia de otros centros de información (bibliotecas o centros de documentación) en el archivo todas las unidades documentales no son iguales, se produce información jerarquizada, no toda "vale" lo mismo, por lo que el archivero tiene que discriminar. Debe de escoger un nivel de descripción, una intensidad de descripción... Hay que tomar la decisión de qué se describe<sup>14</sup>, pensando en qué se recupera y como, algo, quizás, sencillo de entender para los archiveros, pero muy difícil para los usuarios. El objeto de análisis del archivero es la unidad documental, que engloba tanto al documento simple como al documento compuesto (expediente), como a las agrupaciones de estos debidas a la procedencia (los fondos, las secciones, las series,...). Para complicar esta situación estas unidades documentales están relacionadas entre sí, y estas relaciones son "explicativas", son necesarias para entender la naturaleza de las propias unidades documentales analizadas. No es suficiente describir el documento, la unidad documental, debemos describir sus relaciones. La consideración de esta característica condiciona cualquier sistema de recuperación de información archivística que pretenda ser verdaderamente eficaz<sup>15</sup>. Y, a su vez, esta decisión va a condicionar claramente la política de difusión.

Esta concepción de información jerarquizada, que toma carta de naturaleza normativa con la norma ISAD, estaba imbuida en la manera de hacer de cualquier archivero que considerara el principio de procedencia como la base de cualquier trabajo de archivo, pero se ha reflejado poco en los instrumentos descriptivos que este ha generado. En determinadas áreas, como en España o Francia, la necesidad de descripción de las unidades documentales superiores (fondos, secciones, series...), aquellas que no se refieren directamente a los documentos o a los expedientes, y la descripción de ciertas autoridades (organismos productores,...) va a requerir un importante cambio de actitud en lo relativo a qué se describe<sup>16</sup>. Y es probable que durante un período importante de tiempo muchos instrumentos descriptivos automatizados carezcan de una información de calidad en estos niveles.

### 3.2 Proceso de representación de las unidades documentales

Una vez de haber seleccionado que se describe, el archivero opta por como describir. Es evidente que la expresión del contenido de los documentos es uno de los elementos claves que determinaran la recuperación de éste. La existencia de normas de descripción facilita la toma de la decisión en cuanto al formato de descripción, pero el problema se encuentra en la aplicación de este formato<sup>17</sup> y en las áreas

que todavía no han sido normalizadas (niveles de descripción, autoridades, etc.)

Esta descripción puede ser automática, o intelectual –inteligente–. La descripción automática consiste en la detección automática de los elementos que permitan identificar el documento, o en el aprovechamiento de la descripción realizada por los propios creadores o usuarios de los documentos. En el primero de los casos, que se utiliza exclusivamente en los archivos de gestión y administrativos, los problemas de localización de la información derivada de estas descripciones son evidentes, pero puede resultar un método eficaz. No existe una consistencia en la descripción debido a que es el propio creador del documento quien lo describe, normalmente de una manera muy vaga, y como instrumento de una recuperación personal, no pensada para usuarios “externos”, sin pretender extraer la información que el creador del documento da por conocida. En estos tipos de archivos los sistemas de codificación, basados en la clasificación, en contraposición a la utilización de técnicas de interrogación tradicionales, se consideran como elementos más fiables para la recuperación. Puede ser un sistema perfectamente válido para muchos servicios de archivo, ya que no tiene sentido, por ejemplo, describir inteligentemente documentación que en los cuadros de conservación aparezca como expurgable, o que tenga una información muy estructurada y repetitiva. Es el archivero quien tiene que tomar la decisión de si este sistema de descripción es suficiente para atender a las necesidades futuras de recuperación de información que se presentan a los usuarios.

La necesidad de una descripción “inteligente”, la realización de un resumen del documento, el hecho de que haya que describir el contenido de los documentos de manera que se extraiga tanto la información evidente como la que no lo es<sup>18</sup>, parece que va a seguir siendo necesaria por mucho tiempo en los archivos<sup>19</sup>. Este aspecto es evidente en todos los niveles superiores a documentos o expedientes, ya que estas descripciones sencillamente no existen. La cantidad de información generada hasta la actualidad, la imposibilidad de utilización de sistemas de reconocimiento de caracteres para manuscritos no contemporáneos y el inmenso volumen que esta adquiriendo esta, obligan a optar por este tipo de descripción. Como decíamos anteriormente existen factores que determinan claramente la necesidad de esta descripción si se quieren recuperar adecuadamente la información: documentos que contienen información no estructurada, no estrictamente actuales, o documentos de acceso para usuarios heterogéneos (internos y externos) exigen sin duda una descripción inteligente.

Solamente en manuales clásicos de archivística se analizan las cuestiones relacionadas con la forma concreta del resumen: que elementos se recogen, con que formas, etc. Establecer un manual de uso sobre la forma de expresión del contenido de los documentos debería ser una praxis mucho más habitual de lo que lo es en la actualidad. La forma de

descripción establecida deberá en todo caso basar su lógica en la propia necesidad de recuperación de la información.

### 3.3 Proceso de control de puntos de acceso a la información

Aún pudiendo considerar la elección de puntos de acceso como una parte más de la descripción es necesario estudiarla aparte, por su complejidad y por ser el elemento clave en el proceso de recuperación.

La elección de los puntos de acceso, los puntos sobre los que se basa el sistema de recuperación, ha sido muy poco estudiada en nuestra disciplina<sup>20</sup>, al contrario de lo que ha sucedido en biblioteconomía o documentación. La norma ISAD (G) ha dado carta de naturaleza a los puntos de acceso definiéndolos como “los nombres, palabras clave, términos de indización, etc. por los cuales una información puede ser buscada identificada y recuperada”<sup>21</sup>. En todo caso la elección de los puntos de acceso ha sido un trabajo habitual en los archivos, donde la realización de índices o la utilización de descriptores han sido norma general.

El carácter interrelacionado de la descripción de la información de archivo impuesto por la misma norma ISAD nos obliga a prever dos tipos de puntos de acceso, que deben verse reflejados en la recuperación de la información. Unos puntos de acceso diferencian una unidad documental concreta de las demás, son los elementos que nos permiten localizar individualizadamente una información (son las autoridades o los descriptores). Otros puntos de acceso son, en cambio, necesarios para poder localizar la unidad documental dentro del fondo en el cual se encuentra. Son los puntos de acceso que nos deben permitir “navegar”, movernos a través de la estructura que determina la procedencia, y situar el documento en su contexto. Estos conceptos han sido ya descritos en el área anglosajona, donde se les denomina punto de acceso principal y secundario<sup>22</sup>. La definición de la estructura de información, los campos de la base de datos, deberá tener en cuenta estos dos tipos de puntos de acceso.

Podríamos denominar a estos sistemas como de selección de puntos de acceso para el texto, y puntos de acceso para el contexto.

#### 3.3.1 Puntos de acceso para el texto

Se puede relacionar directamente la elección de puntos de acceso para el texto con la indización<sup>23</sup>. Hasta la actualidad se han utilizado en los archivos básicamente dos técnicas diferentes a la hora de establecer los puntos de acceso en la información no contextual: la indización automática, y la indización manual. Desde hace muy pocos años existe una tercera técnica de indización, la indización contextual, que supone un

importante avance en alguna de las cuestiones relacionadas con la recuperación.

a) La indización automática<sup>24</sup> está en estos momentos muy extendida en muchos servicios de archivo de gestión o administrativos, ya que rentabiliza de manera muy rápida el trabajo de descripción realizado, de manera que se ha convertido en un sistema perfectamente válido.

Por el contrario, la indización automática de instrumentos descriptivos de archivo no estrictamente contemporáneos provoca el problema de que el lenguaje y la forma de descripción utilizada en otras épocas provoque problemas de recuperación de esta información. La única posibilidad de recuperación de esta información es la utilización de técnicas de lenguaje controlado de manera que existan relaciones de no-descriptores llamémosles "de época" que envíen a descriptores con terminología actual.

b) La indización manual o controlada<sup>25</sup> se ha mostrado como necesaria en cualquier archivo que rebase la fase administrativa, debido a que los problemas de ruido documental derivados de una recuperación no controlada se acentúan de manera clara en la medida en la cual el usuario no conoce el contexto, bien organizativo, lingüístico, o institucional.

La forma de indización más común son los descriptores<sup>26</sup>, estos se convierten en elementos básicos del proceso de recuperación de la información ya que permiten reducir las posibilidades de expresión de las ecuaciones de búsqueda por parte de los usuarios, de manera que se pone en relación el proceso de indización con el de recuperación. Han sido, incomprensiblemente, más usados para describir que para recuperar información, ya que muchos de los sistemas de recuperación de la información no incluyen la visualización de estas listas de descriptores –sea cual sea la forma que tomen–.

La gran mayoría de los esfuerzos relacionados con la indización han ido orientados al establecimiento de identificadores de topónimos, nombres de personas o instituciones –y la norma ISAAR(CPF)<sup>27</sup> es un claro reflejo de ello–. La elección de descriptores de materias es, en cambio, la cenicienta de los esfuerzos de indización a pesar de que es este recurso el que mayor valor añadido produce en el trabajo de descripción. La información de los datos específicos de nombres de personas, topónimos, etc. es más fácil de localizar por medio de las búsquedas por texto libre pues, aún existiendo posibles variaciones gramaticales, el grado de variación es infinitamente menor al de la capacidad de los idiomas para expresar con palabras diferentes objetos o ideas parecidos. En el campo de la normalización de descriptores de materias y en el de su control queda en los archivos aún muchísimo trabajo por hacer.

c) La indización contextual. Este tipo de indización permite no solo la interconexión de términos de manera automática, sino la localización a

través de ellos de nuevos elementos de interés para el usuario<sup>28</sup>. La capacidad de algunos nuevos sistemas de información de analizar el contexto que rodea a un término determinado y, en función de este, lanzar una búsqueda que no requiere de una nueva identificación de necesidades por parte del usuario puede suponer un gran avance en los resultados de recuperación de la información, y deberá ser tenido en cuenta por los responsables de los archivos<sup>29</sup>. La utilización de estas técnicas consiguen suplir de alguna manera las capacidades de recuperación de lenguajes controlados como el tesauro. Es previsible una mejora en estos sistemas de manera que deberán ser estudiados detalladamente como alternativa<sup>30</sup>.

La necesidad de consistencia en la indización, o más bien la valoración del beneficio obtenido con relación a la cantidad de trabajo que ella requiere, es discutida por parte de muchos archiveros. En algunos casos se ha considerado que la recuperación de la información se mejoraría poniendo más énfasis en aumentar el número de puntos de acceso y menos de obtener consistencia en la indización<sup>31</sup>. No parece que pueda extraerse una norma general sobre su conveniencia o no; es el tipo de documentación, y el tipo de información que se encuentra en ella, el que determina la necesidad o conveniencia de la utilización de un lenguaje controlado. Cuanto mayor es la desestructuración de la información que se trata, más necesaria será la utilización de un lenguaje controlado que permita limitar a los usuarios los posibles términos de búsqueda en un número que le resulte aprehensible. Determina esta necesidad asimismo la distancia cronológica con respecto a los documentos, a mayor distancia cronológica mayor necesidad de control de vocabulario<sup>32</sup>. Y, por último el tipo de usuario, su conocimiento de la institución, de la documentación, puede exigir la utilización o no de instrumentos de control de lenguaje. Será función de los responsables delimitar la política descriptiva del archivo, la toma de decisión sobre la necesidad o no de utilizar estos instrumentos de control, o reducir las posibilidades a las que ofrece la consulta del lenguaje natural<sup>33</sup>.

Es posible una aplicación de la teoría de las tres edades al análisis de la recuperación de la información. En los documentos con valor administrativo vigente los puntos de acceso necesarios están sobre todo relacionados con la procedencia, de tal manera que los sistemas de clasificación suponen un elemento básico en la mayor parte del proceso de recuperación de la información. En la medida en la que los documentos adquieren valor histórico y van perdiendo el administrativo, se diversifican los puntos de acceso que pueden resultar interesantes para los usuarios. La progresiva aparición de valores históricos del documento abre el abanico de usuarios de este. Es posible que el usuario no tenga en cuenta el criterio de procedencia al efectuar el proceso de búsqueda –lo que no significa que este usuario no deba posteriormente “entender” el documento como testigo de un proceso administrativo determinado–, en estos casos el investigador busca en documentos creados con

un fin determinado valores informativos que el creador del documento no pretendía reproducir. Estas circunstancias obligan a no utilizar criterios de búsqueda por navegación a través de la procedencia sino ecuaciones de búsqueda directa de términos que interesen al usuario. En todo caso esto se produce sobre todo en la descripción de documentos de tipo histórico. En este aspecto la descripción de documentos de archivo se aleja claramente de la descripción de documentalistas y bibliotecólogos, encargados de transmitir solamente el contenido expreso de un documento, el deseado por el creador del documento; el archivero puede, y debe, sobrepasar este umbral proponiendo puntos de acceso que reflejen valores contextuales no presentes en el momento de la creación.

Un caso muy claro de esta evolución es el de la documentación judicial, creada para resolver un litigio no pretende presentar un aspecto concreto de la vida, sino resolver un problema legal determinado. Cuando estos documentos dejan de tener un sentido administrativo, y pasan a tenerlo histórico pueden interesar a algunos usuarios por el conocimiento de cuestiones relacionadas con la justicia, pero estos son los menos, la mayoría de los usuarios relacionados con la investigación en documentación judicial busca información que en principio no tiene que ver con el propio proceso sino que reflejan un aspecto de la realidad que ha quedado recogido de manera indirecta en esta documentación y que son cuestiones que en el momento del proceso administrativo concreto son absolutamente irrelevantes.

La indización controlada trata de evitar los problemas que se producen en la recuperación de la información indizada automáticamente, y básicamente se ha utilizado en los archivos en forma de tres técnicas diferenciadas: las listas de descriptores, los tesauros, y las autoridades.

#### a) las listas de descriptores

Es la forma más simple de control de lenguaje y consiste en el establecimiento de listas de términos admitidos como descriptores. Se suelen presentar en forma alfabética. Son deudores del inmenso trabajo realizado por muchos archiveros a lo largo de la historia en la realización de índices de los documentos descritos. El hecho de que en los archivos no se hayan usado formas de indización normalizadas no puede suponer que la experiencia y el trabajo acumulados en la atención a los usuarios (que se refleja en los instrumentos descriptivos) no se puedan utilizar a la hora de recuperar los documentos. Tratar de reconvertir y aprovechar los índices existentes, o crear listas de descriptores, pueden ser trabajos de una alta rentabilidad, de manera que debe ser valorado como una posibilidad. En todo caso deberían ser accesibles al usuario como elemento clave en la toma de las decisiones relacionadas con la consulta.

## b) Los lenguajes controlados: tesauros

El tesoro es, al contrario de la lista de descriptores, el resultado de una estrategia de descripción muy elaborada. Ha sido muy poco utilizado en los archivos aunque existen diferentes proyectos comunitarios, aplicables a diferentes archivos, que presentan una gran rentabilidad<sup>34</sup>. Existen algunos proyectos como los de los gobiernos australiano o canadiense, interesados en la realización de un tesoro común de términos administrativos, que suponen la adopción del control del lenguaje como política de estado<sup>35</sup>. No parece que este tipo de proyectos sea fácilmente realizables en ámbitos administrativos tan poco estructurados como el del caso español. Se ha planteado, incluso, la posibilidad de su aplicación a la documentación administrativa de carácter muy específico<sup>36</sup> (departamentos especializados, etc.), pero las circunstancias de muchos archivos lo hacen semi-imposible.

Es muy importante valorar las posibilidades reales de un servicio de archivo a la hora de crear y mantener un tesoro, que es un trabajo intenso y costoso. En todo caso los beneficios que reporta un lenguaje controlado y estructurado como el tesoro, tanto en el momento de la descripción como en el de la recuperación, son innegables. Por lo que se debe tener en cuenta la posibilidad de aprovechar tesauros implantados en otros servicios como fuente para el control del lenguaje. En este contexto existen atractivos planteamientos<sup>37</sup> que propugnan la utilización de los tesauros de manera interactiva de forma que estos vayan respondiendo tanto a las necesidades de los indizadores como de los usuarios.

## c) el control de autoridades

La norma ISAAR (CPF) define el control de autoridades como "el control de términos normalizados que comprende los nombres utilizados como puntos de acceso", y lo aplica sobre los nombres de personas, los de entidades y los de familias. Ha sido estudiado a fondo recientemente<sup>38</sup>, y no entraremos a analizarlo como técnica. Cabe decir que tanto la aprobación de la propia norma, como parte de la literatura científica han hecho que el control de autoridades sea un objetivo prioritario para un buen número de proyectos de descripción de archivo.

Tras un análisis de la situación real de los archivos, por lo menos de aquellos de la llamada Europa latina, se puede deducir que la aplicación completa de las normas ISAAR (CPF) es muy complicada, debido a que la normalización de estos accesos es compleja y costosa. Se puede aplicar a las autoridades relacionadas con los niveles superiores de descripción (fondos o secciones), pero una información tan detallada es, desde un punto de vista práctico, irrealizable para aquellos que han centrado más sus esfuerzos en la descripción de las unidades documentales menores, dejando un poco de lado la descripción y caracterización

de los organismos productores, y de los tipos documentales creados por estos<sup>39</sup>. Parece evidente la necesidad de que estos proyectos sean desarrollados de manera conjunta entre diferentes archivos, de manera que la rentabilidad sea la máxima<sup>40</sup>. Es necesaria una gran flexibilidad por cuanto habrán de convivir puntos de acceso muy controlados (por medio de control de autoridades) con otros muchos menos controlados (por medio de descriptores) pero muy útiles en el proceso de recuperación de la información.

### *3.3.2 Puntos de acceso para el contexto*

Una asunción real del archivero de la norma ISAD requiere del sistema de información de archivo la recuperación de la información archivística en contexto<sup>41</sup>, es decir, no es suficiente con presentar al usuario el resultado directo de su consulta, es necesario que pueda en cualquier momento conocer la razón de ser del documento que consulta como parte de un conjunto de documentos, debe de dar a conocer su procedencia. Como ya he comentado este concepto ha sido ya descrito en el área anglosajona, pero se refiere solamente al concepto de productor. Esta procedencia puede ser completada entendiéndola no solo como el punto de acceso a la descripción del organismo productor, sino también describiendo la situación del documento dentro del contexto estructural del fondo, en que parte de este fondo se encuentra. En alguna de las experiencias de descripción automatizada que se está dando en el estado español se ha trabajado ya con este concepto de puntos de acceso de contexto, por medio de elementos de descripción que permiten la identificación del lugar concreto que ocupa la unidad documental descrita dentro del cuadro de clasificación, y de puntos de acceso a los niveles de descripción superior o inferior a la unidad documental que se está describiendo, de manera que se pueda efectuar una navegación<sup>42</sup>.

Una de las circunstancias que dificulta más el proceso de recuperación de la información respetando las características multinivel es la denominada herencia<sup>43</sup>. Si se pretende que esta sea explícita en cada uno de las descripciones de las unidades documentales<sup>44</sup> se puede dar la paradoja de que los usuarios, sobre todo aquellos que utilizan técnicas de localización difusa – donde tienen cabida todos aquellos que no han estado implicados en la organización concreta que ha dado lugar al fondo – se vean disuadidos de la consulta de los campos de descripción del contexto por su repetición continua. Parece mucho más apropiada la utilización de técnicas de navegación hipertexto para la solución de estos problemas, de tal manera que se permita subir y bajar entre los diferentes niveles de información a través de enlaces hipertexto, lo mismo que acudir a los encabezamientos de materia autorizados a través de estos enlaces, sin necesidad de visualizar permanentemente estos datos.

### **3.4 Proceso de identificación de las necesidades del usuario con lenguaje libre**

Hemos analizado detenidamente la parte del proceso de recuperación de información que es responsabilidad directa del servicio de archivo analicemos ahora el proceso desde el punto de vista del usuario. Esta cuestión ha sido muy poco estudiada en el campo de los archivos pero es un elemento clave a tener en cuenta.

Como se ha señalado con anterioridad el tipo de usuario debe determinar el sistema de recuperación a implantar. En este aspecto los dos tipos básicos de usuario serán aquel, cualquiera sea su relación con la documentación, parte del organismo de procedencia, o concededor de este; y el usuario externo que no forma parte ni tiene relación con el productor. En el primero de los casos la identificación de las necesidades del usuario suele ser exacta, aunque esta exactitud se pierde en la medida que haya pasado el tiempo desde la creación del documento. En estos casos el sistema de recuperación no requiere lenguajes controlados, de manera que la indización automática, reforzada por sistemas de recuperación por rangos o booleanos es suficientemente eficaz.

Por el contrario el usuario externo presenta un acercamiento a la demanda de información muy diferente. La experiencia de muchos archiveros en la atención a sus usuarios no ha quedado recogida, por desgracia, en ningún análisis de este aspecto. De esa experiencia creo que se puede colegir que el usuario externo que no forma parte de ninguna manera del negocio jurídico o informativo (sea cual sea su tipo) que refleja un documento, utiliza normalmente una lógica difusa que hace que utilice en primera instancia términos genéricos para expresar sus deseos, para, y posteriormente, pasar a concretarlos en la medida en que la primera respuesta sea o no satisfactoria. Esta lógica difusa exige que los instrumentos de recuperación que se propongan a estos usuarios dispongan de sistemas que faciliten la búsqueda. Es función de los responsables del archivo escoger de entre las diferentes técnicas existentes, siempre que sean aplicables, las más apropiadas.

### **3.5 Proceso de adecuación de las necesidades del lenguaje libre al lenguaje de interrogación**

El primer examen de un sistema de recuperación se produce en este momento, ¿es capaz el sistema de orientar al usuario a la hora de acomodar la estrategia de búsqueda?

En este punto nos encontramos con un grave problema consecuencia de la complejidad de la descripción archivística. Como se ha dicho es necesaria una combinación de recursos de texto –búsqueda– y navegación –contexto–, y ambos se basan en dos sistemas lógicos divergen-

tes<sup>45</sup>. El primero es un sistema visual e intuitivo, basado en inferencias rápidas de ensayo-error, que funciona por medio de desplazamientos virtuales, es, por lo tanto, fácil de usar sin entrenamiento previo. La búsqueda se basa, al contrario, en la planificación y el diseño de una estrategia con fórmulas precisas (aunque también exista el ensayo-error), y se requiere un cierto conocimiento, no es fácil de usar. Un modelo apropiado deberá compaginar ambos sistemas, tanto en el momento de realización de la consulta como en la presentación de resultados, lo cual es más fácil de exponer que de plasmar en, por ejemplo, el pequeño espacio de una pantalla de ordenador.

En el ámbito de la búsqueda los sistemas de adecuación del lenguaje libre son sobre todo pertinentes, como es lógico, en el caso de los usuarios que hemos denominado externos. En todo caso la política de descripción establecida condiciona la existencia o no de posibilidades de adecuación de las necesidades expresadas en lenguaje libre a formas más estrictas. El usuario podrá adaptar sus expresiones en lenguaje libre siempre que la estructura de datos establecida le permita utilizar diferentes sistemas de búsqueda. Siguiendo la lógica anteriormente establecida esta adecuación del lenguaje puede estar relacionada con la búsqueda por texto o por contexto. En lo relativo a la información textual, en caso de haberse utilizado cualquier sistema de control del lenguaje, el usuario puede acomodar las necesidades expresadas en lenguaje libre a los términos admitidos. En el ámbito de la recuperación esta posibilidad se refleja en la presentación de tesauros o listas de descriptores en el momento de la consulta. Estos sistemas están extendidos en las aplicaciones in situ de muchos archivos aunque apenas se ha ofertado en la recuperación vía Internet. La presentación de listados de términos, bien en forma alfabética o jerarquizada ha sido una de las principales áreas de trabajo para los profesionales de la difusión de la información de manera que en la actualidad existen diferentes sistemas de presentación de términos de recuperación, los llamados mapas conceptuales o lexicográficos<sup>46</sup>, que facilitan en gran manera la consulta de los términos utilizados en la indización. No conozco ningún caso de utilización de estas técnicas en la recuperación de información de archivo.

La recuperación de información de archivo de tipo textual requiere unos mínimos que deberían ser respetados, como la búsqueda entre intervalos de fechas, la posibilidad de búsqueda sobre campos específicos, la posibilidad de utilización de operadores booleanos, la posibilidad de visualización de los términos controlados (en su forma de tesauro, lista de descriptores,...) o la posibilidad de búsqueda en texto libre.

A pesar de ello los sistemas más modernos no presentan, inexplicablemente, alguna de estas posibilidades. Con la aplicación de EAD se ha comenzado a plasmar las normas generales de la descripción multinivel. Sin embargo, en algunos de los ejemplos de aplicación de aplicación de EAD los sistemas de búsqueda pierden las posibilidades de recuperación de la información que podemos llamar complejas, aquellas relacio-

nadas con las bases de datos. Esto se debe a que al presentarse la información como un texto plano la información debe ser localizada de una manera secuencial, no pudiéndose utilizar técnicas de búsquedas booleanas. Puede darse incluso el caso, fácilmente comprobable en este momento en la web, de que existan recursos archivísticos que no tengan posibilidades de búsqueda específica de rangos cronológicos, elemento básico, a mi entender, de cualquier sistema de recuperación de información de archivo.

Por último, como ya hemos comentado, en los casos más avanzados el sistema puede disponer de un motor de búsqueda capaz de realizar búsquedas difusas, que adecuan la expresión concreta utilizada por el usuario, contextualizando de manera automática tal término, y realizando la búsqueda no solamente por el término propuesto, sino también por aquellos que se encuentran cercanos a su área semántica o contextual.

En lo relativo a la información contextual, la organización de los fondos por procedencia constituye un eje básico, de manera que, al contrario que en los sistemas de recuperación clásicos en documentación, en la praxis de recuperación de información en archivos la selección de información por medio del propio cuadro de clasificación debe de convertirse en un recurso de acceso sine qua non. La utilización de técnicas de navegación a través del cuadro de clasificación de los fondos debería constituir uno de los elementos de cualquier sistema de consulta de manera que se adapten las necesidades del usuario a la forma concreta de organización adoptada.

El establecimiento de políticas de formación de usuarios o el desarrollo de potentes instrumentos automatizados de ayuda en los sistemas de recuperación pueden constituir importantes elementos de refuerzo para asegurar una adecuada traslación de las necesidades del usuario a la estrategia de la búsqueda.

### **3.6 Proceso de comparación entre las necesidades del usuario y las informaciones representadas**

El proceso de recuperación de información desemboca en el resultado de la búsqueda, verdadero test de calidad del sistema. No se conocen experiencias de valoración de resultados de búsquedas aplicados a la localización de información en archivos. La evaluación debería tener en cuenta la existencia de dos tipos de resultados: el textual y el contextual. Los resultados textuales, las referencias localizadas, deben ser evaluados en cuanto a su exhaustividad y su pertinencia. En este ámbito también se demuestran las diferencias entre las necesidades de archivos de gestión y administrativos que deben de ofrecer niveles de pertinencia y exhaustividad muy elevados, mientras que en los archivos his-

tóricos la pertinencia es mucho más difícil de valorar debido a la lógica difusa de interrogación de gran parte de las demandas.

La evaluación de la presentación de los resultados contextuales es necesaria en los archivos. Como ya hemos comentado la necesidad de compaginar los sistemas de navegación y búsqueda exige una presentación de resultados que permita la visualización de las referencias seleccionadas y de los enlaces que permitan contextualizar estas referencias. Instrumentos de descripción en formato XML, a través de EAD, solucionan ya esta cuestión, proporcionando el efecto de navegación que permite moverse a través de los cuadros de clasificación. Evaluar el aprovechamiento que los usuarios hacen de estos sistemas de enlaces, detectar sus defectos, y mejorar los sistemas de recuperación se convierte en una obligación para aquellos archivos que pretendan ofrecer un buen servicio

## 4. Conclusiones

De todo lo expuesto hasta el momento se pueden concluir por una parte una serie de consideraciones generales, por otra una serie de características básicas de un sistema de recuperación de información de archivo y, por último, una serie de repercusiones a tener en cuenta en el caso del establecimiento de un sistema avanzado de recuperación de información.

### a) Consideraciones generales.

Las políticas de descripción y recuperación de la información en los archivos deben de ser políticas desiguales, utilizando diferentes recursos y proponiendo diferentes resultados para cada uno de los casos que tenga que gestionar el archivo. Estos casos están directamente relacionados por las características específicas de cada fondo. Los condicionantes de tamaño del fondo, tipo de usuarios, tipo de información (estructurada, cognitiva,...) son los que deben de orientar a los directivos en esta toma de decisión.

Es tarea del archivero analizar la capacidad o incapacidad de las series documentales como fuentes de información diferente a la que intrínsecamente se deduce de su procedencia, y aplicar los criterios de indización de maneras diferentes (con mayor o menor intensidad) en cada una de ellas.

El proceso de recuperación de la información en archivos exige la existencia en paralelo de sistemas de búsqueda a través de puntos de acceso relacionados con el texto (índices, descriptores,...), y con el contexto (procedencia y clasificación). Parece evidente la necesidad de combinar técnicas de bases de datos documentales y relacionales en la des-

cripción archivística, de manera que se puedan resolver de manera paralela las necesidades de recuperación de la información con técnicas de búsqueda documental y, a su vez, se pueda utilizar las posibilidades relacionales como enlaces para la posterior utilización de técnicas de navegación hipertexto a través de cuadros de clasificación, descripciones, autoridades, etc.

b) Requisitos para un sistema de recuperación de información archivística:

- Debe, ante todo, respetar las normas de descripción multinivel.
- Debe permitir el acceso a registros de unidades documentales situados en diferentes niveles de información.
- Debe de permitir la navegación entre estos vínculos de manera que se pueda acudir a recoger información a los niveles superiores o inferiores de manera automática (sin necesidad de realización de nuevas búsquedas)
- Debe de permitir la visualización gráfica de los cuadros de clasificación, de manera que el usuario pueda seleccionar información no solo a través de funciones de búsqueda sino meramente de navegación.
- Debe de soportar tanto índices automáticos como controlados –en el caso de que existan– teniendo estos mayor o menor peso específico de acuerdo a los perfiles de usuario del servicio en concreto.
- En los sistemas de recuperación de información histórica se deben de permitir tanto las búsquedas libres como por medio de lenguaje controlado.
- Debe de permitir la utilización de sistemas de búsquedas por rangos de fechas, booleanas, truncamientos, etc.
- Debe de permitir la consulta a distancia por Internet.
- Debe de ser capaz de aprovechar el trabajo de descripción realizado con anterioridad, de manera que no sea necesario una nueva descripción de toda la documentación.
- Es recomendable una navegación entre descriptores en el caso de que existan, a través de la estructura del tesoro, o a través de listas.
- Debe de prever la posibilidad de enlace con imágenes de los documentos.
- Debe de permitir la selección por parte del usuario de universos de búsqueda relacionados con el cuadro de clasificación, de manera que las búsquedas se realicen contra el grupo de unidades documentales relacionadas con un determinado nivel de descripción.
- Debe de presentar interfaces de consulta amigables.
- Debe de permitir la generación automática de instrumentos descriptivos.
- Debe de prever la posibilidad de que los usuarios guarden sus perfiles de búsqueda<sup>47</sup>

### c) Repercusiones en la gestión.

La implantación de una política expansiva en la difusión, consecuencia final de una política de recuperación determinada, provocará de manera previsible una serie de cambios en la gestión del servicio. El principal de los cambios es la posibilidad de ampliación del universo de usuarios. La característica del documento de archivo como único, combinada con los nuevos recursos de localización de la información, y la posibilidad de reproducir el documento a precios asequibles hacen que el universo de usuarios pueda ser mucho mayor del que hasta la actualidad se ha considerado. La consulta y la atención a distancia no son ya utopías para los archiveros.

La potenciación de los servicios de referencia archivística y, sobre todo, de los servicios de reproducción de documentos se constituye en premisa básica a la hora de establecer un sistema de gestión de archivo adecuado. Si pretendemos que nuestros usuarios sean capaces de acceder a la información incluso a distancia tenemos que prever la consecuencia de ello, la atención a distancia tanto de los servicios de ayuda en la referencia como de reproducción de los documentos, y los problemas de todo tipo que se derivan de ellos (personal, servicios, derechos de reproducción,...).

Será necesario dedicar una mayor cantidad de recursos a los procesos de presentación de resultados, tanto in situ como a distancia.

Será necesario el establecimiento de una política de evaluación de los sistemas de recuperación, que permita localizar los puntos fuertes y débiles del proceso de localización y comunicación de datos, y adecuarlos permanentemente<sup>48</sup>.

Por último el valor añadido en el ámbito de la descripción deberá seguir siendo orientado a un conocimiento del contexto, un conocimiento histórico, que continua siendo el elemento que permite entender el documento.

## Notas

- 1 COUTURE, C.; ROUSSEAU, J. *Les Archives au Xxe siècle*. Montreal: Université de Montreal, 1982, pág. 257.
- 2 NÚÑEZ FERNÁNDEZ, Eduardo. *Organización y gestión de archivos*. Gijón: Trea, 1999, pág. 579.
- 3 ROVIRA, Cristófol. "Teoría de la recuperación de la información". En: *Curso de Documentación Digital*. Barcelona: Universidad Pompeu Fabra, 1999. CD.
- 4 DUCHEIN, Michel. "El archivero en la era de la comunicación: especificidad y expansión de la archivística". En: *Irargi. Artxibistika aldizkaria - Revista de Archivística*, II, (1989), pág. 32.
- 5 HENSEN, Steven L. *Papers and Personal Manuscripts. A cataloguin manual for Archival Repositories, Historical Societies, and Manuscript Libraries*. Chicago: SAA, 1989.

- 6 TARRAUBELLA, X. "Los archivos y sus usuarios". En: Conferencia Europea de Archivos (5ª. 1997. Barcelona). V Conferencia Europea de Archivos : las bases de la profesión. Barcelona : Associació d'Arxivers de Catalunya, 1997.
- 7 GIBAJA VELÁZQUEZ, J.C.; HUGUET SANTOS, M. "La Historia en casa: Nuevas tecnologías y archivos domésticos". En: *Congreso Internacional sobre Sistemas de Información Histórica*. Vitoria: Juntas Generales de Alava, 1997.
- 8 MAFTEI, Nicolás. "Requisitos de una aplicación informática para descripciones multi-nivel y representaciones en contexto". *Tabula*, Salamanca, 4, (1999), pág. 172,
- 9 CRUZ MUNDET, José Ramón. *Manual de Archivística*. Madrid: Fundación Germán Sánchez Ruipérez, 1994, pág. 300.
- 10 Sirva el caso de Cataluña como ejemplo, donde, aún estando adelantados en la aplicación de recursos informáticos en los archivos con respecto a las demás partes del estado español, el resumen que se puede hacer es este. PERPINYA I MORERA, Remei. "L'aplicació de les tecnologies de la informació als arxius catalans". *BiD*, núm. 5 (deseembre 2000).
- 11 AGUINAGALDE OLAIZOLA, F. Borja de. "Esperienze archivistiche nella descriziones e nell'automazione". En: *Gli Standard per gli archivi europei: esperienze e proposte*. Roma: Ministero per i beni culturali e ambientali, 1996, pág 254.
- 12 ROVIRA, Cristófol. "Teoria de la recuperación de la información". En: *Curso de Documentación Digital*. Barcelona: Universidad Pompeu Fabra. 1999. CD.
- 13 Ibídem.
- 14 AGUINAGALDE OLAIZOLA, F. Borja de. "Esperienze archivistiche nella descriziones e nell'automazione". En: *Gli Standard per gli archivi europei: esperienze e proposte*. Roma: Ministero per i beni culturali e ambientali, 1996, pág 270.
- 15 MACNEIL, Heather. "The context is all: describing a fonds and its parts in accordance with the Rules for the Archival Description". En: *The archival fonds: from theory to practice = Les fonds d'archives de la théorie à la pratique*. Ottawa: Bureau canadien des archivistes. Comité de planification sur les normes de description, 1992.
- 16 Este elemento debería afectar la política de formación de profesionales de archivos, donde en la medida en que se acentúan los perfiles de gestor de información se pierden los perfiles de especialistas en organizaciones (en su historia, en sus estructuras,...), cuando el perfil que impone la aplicación de las reglas multinivel es precisamente el del coocedor de topo de tipo de estructuras de organización (públicas, privadas,...).
- 17 Existen ya algunos trabajos que han analizado esta cuestión, en el estado español destaca: BONAL ZAZO, J.L.; GENERELO LANASPA, J.J.; TRAVESÍ DE DIEGO, C. *Manual de Descripción Multinivel : Propuesta de adaptación de las normas internacionales de descripción archivística*. [Valladolid] : Junta de Castilla y León, Consejería de Educación y Cultura, 2000.
- 18 DEL VALLE PALMA, Mª Teresa. "Técnicas y métodos para mejorar la calidad de indización y su recuperación en bases de datos documentales en Ciencias Sociales y Humanidades" En: *5es Jornades Catalanes de Documentació*. Barcelona: Socadi, 1995.
- 19 PINTO, M.; LANCASTER, F.W. "Abstracts and abstracting in knowledge discovery". *Library Trends*, 48 (1), (1999), pág. 234-238.
- 20 MACNEIL, Heather. "La costruzione di sistemi archivistici di indicizzazione per soggetto". En: *Gli Standard per gli archivi europei: esperienze e proposte*. Roma: Ministero per i beni culturali e ambientali, 1996, pág 270.

- 21 ISAD(G) : Norma Internacional General de Descripción Archivística : adoptada por el Comité de Normas de Descripción, Estocolmo, Suecia, 19-22 septiembre 1999 / [Versión española de Asunción de Navascués Benlloch...]. Madrid : Subdirección de los Archivos Estatales, 2000.
- 22 STIBBE, Hugo. "L'application de la notion de fonds: un article qui traite de la catégorie d'accès principale, de la description à plusieurs niveaux et du contrôle d'autorité". *Archives*, vol. 25-1, (1993).
- 23 PINTO MOLINA, M. Análisis documental. Fundamentos y procedimientos. Madrid: Eudema Universidad, 1991, pág. 208.
- 24 Ibidem. Pag. 142. Citando a Van Slype. Lenguajes documentales.
- 25 PINTO MOLINA, M. "Tratamiento de los contenidos en la sociedad de la información". En: *La sociedad de la información*. Madrid: Centro de Estudios Ramón Areces, 2.000
- 26 ORTEGO DE LORENZO-CÁCERES M<sup>a</sup> Pilar, BONAL ZAZO José Luis. "El control de autoridades y la normalización de los puntos de acceso en los archivos". *Tabula*, Salamanca, 4, (1999), pág. 41-69.
- 27 ISAAR (CPF) : norma internacional sobre los encabezamientos autorizados archivísticos relativos a entidades, personas y familias. Madrid : Ministerio de Educación y Cultura, Dirección General del Libro, Archivos y Bibliotecas, 1997.
- 28 GIMÉNEZ PERELLO, Javier. "Sistemas de indización aplicados en bibliotecas: clasificaciones, tesauros y encabezamientos de materia". En: *Tratado básico de Biblioteconomía*, Jose Antonio Magán (Coord.). Madrid: Editorial Complutense, 1998, pág. 207.
- 29 La existencia de sistemas como los de "mapas conceptuales", que consiguen representar gráficamente estas relaciones, facilitan aún más los procesos de recuperación de información.
- 30 SALTON, G. *Automatic text procesing: the transformation, analysis and retrieval of information by computer*. New York: Addison Wesley, 1989. Citado por María Pinto Molina en *La sociedad de la información*. Madrid: Centro de Estudios Ramón Areces, 2.000.
- 31 BEARMAN, David A. "Authorithy control issues and prospects". En : *The American Archivist*. Chicago, V. 52, (Summer 1989), N. 3, pág. 286-299.
- 32 Aunque existen múltiples trabajos que proponen la utilización de lenguajes controlados: GRUPO DE ARCHIVEROS MUNICIPALES DE MADRID. *Materiales para un Tesaurus de Archivos Municipales*. Madrid: Comunidad de Madrid, Consejería de Cultura, 1999. DE SANTOS CANALEJO, Elisa Carolina. *Tesaurus Archivo*. Madrid: Ministerio de Asuntos Exteriores, Secretaría General Técnica, Archivo General, D.L. 1994; DIRECTION DES ARCHIVES DE FRANCE. *Thesaurus w.: vocabulaires normalises pour la description et l'indexation des archives administratives locales contemporaines*. [Paris]: Direction des Archives de France, 1989. DOMENECH I FERNÁNDEZ, Silvia. *Tesaura BIMA: Arxiu Municipal de Barcelona*. Barcelona, Ajuntament, 1997.
- 33 Que en la información de archivos suele estar ya "reducido" debido a que se utilizan como elemento de recuperación las descripciones realizadas por el propio servicio.
- 34 Ver nota 30.
- 35 <http://www.naa.gov.au/recordkeeping/control/summary.html>
- 36 LÓPEZ CARREÑO, Rosana. "Necesidad del uso de tesauros en el manejo de documentación administrativa en los archivos y centros de documentación". *Scire*, 5:2, (jul-dic, 1999), pág. 115-130.

- 37 ROSA, Antonio de la. "Instrumentos terminológicos en el WWW : XML". *El profesional de la información*, Vol. 8 : Num. 10, (1999), pág. 14-36.
- 38 ORTEGO DE LORENZO-CÁCERES M<sup>a</sup> Pilar, BONAL ZAZO José Luis. "El control de autoridades y la normalización de los puntos de acceso en los archivos". *Tabula*, Salamanca, 4, (1999), pág. 41-69.
- 39 En esta cuestión las diferencias de, por ejemplo, las tradiciones italiana y española de política descriptiva son claras, y la aplicación de la normas en ambos países tendrá que tener, por pura lógica, un diferente tratamiento.
- 40 SANTOS, Elisa Carolina de. "La indización en la recuperación de la información". *Lligall*, 12, (1998).
- 41 MAFTEI, Nicolás. "Requisitos de una aplicación informática para descripciones multi-nivel y representaciones en contexto". *Tabula*, Salamanca, 4 (1999), pág. 172.
- 42 Irargi. Centro de Patrimonio Documental de Euskadi. [www.irargi.org](http://www.irargi.org)
- 43 ISAD(G) : Norma Internacional General de Descripción Archivística : adoptada por el Comité de Normas de Descripción, Estocolmo, Suecia, 19-22 septiembre 1999 / [Versión española de Asunción de Navascués Benlloch...]. Madrid : Subdirección de los Archivos Estatales, 2000.
- 44 MAFTEI, Nicolás. "Requisitos de una aplicación informática para descripciones multi-nivel y representaciones en contexto". *Tabula*, Salamanca, 4, (1999), pág. 174
- 45 CODINA, Lluís. "Investigación en línea: cómo funcionan los servicios de obtención de información en Internet". En: *Curso de Documentación Digital*. Barcelona: Universidad Pompeu Fabra. 1999. CD.
- 46 GÓMEZ CAMARERO, C. "Descripción de recursos electrónicos sobre programas de análisis de contenido y visualización de la información". En: *La sociedad de la información*. Madrid: Centro de Estudios Ramón Areces, 2.000.
- 47 <http://www.essexcc.gov.uk/heritage/>
- 48 COUTURE, C., "La evaluación de los archivos. Estado de la cuestión," *Lligall*, 1999. [http://www.arxivvers.com/html/articles/pdf/cas/06\\_evalua.pdf](http://www.arxivvers.com/html/articles/pdf/cas/06_evalua.pdf).

## Resum

La recuperació d'informació és per als arxivers una tasca quotidiana que, sovint, no ha merescut una anàlisi detallada per part dels mateixos arxivers. Tanmateix, entre les diverses funcions que configuren la feina de l'arxiver la de recuperació de la informació adquireix una importància fonamental, sobretot pel fet que és una funció "pont" entre la descripció i la consulta del document, i que sense ella la primera no té sentit i la segona no es pot portar a terme.

De la confluència de diversos factors que es resumeixen en la necessitat d'oferir atenció als usuaris sense la intervenció directa de l'arxiver com a intermediari, en resulta la conclusió que les qüestions relacionades amb el tractament de la informació, i de manera especial les relacionades amb la recuperació de la informació, tenen una importància cada vegada més gran. La principal evidència d'aquesta importància creixent és l'aplicació massiva de recursos informàtics als recursos de descripció. Però fins ara, aquests recursos informàtics s'havien emprat més aviat com una eina per a la creació d'instruments descriptius que no pas com una eina de recuperació de la informació per part dels usuaris.

L'anàlisi dels sistemes de recuperació de la informació que gestiona el servei d'arxiu, així com la seva adequació a les circumstàncies actuals, s'han d'entendre com un component més del programa descriptiu del servei d'arxiu.

Per bé que són escasses, les experiències realitzades en aquest àmbit permeten identificar les diferents possibilitats que té a l'abast l'arxiver a l'hora de definir els sistemes de recuperació. Les més conegudes són les relacionades amb l'ús de tècniques de control lingüístic, com ara les d'indexació i control d'autoritats: descriptors lliures, llistes de descriptors, llenguatges controlats (thesaurus, etc.). Tot i així, l'aparició d'Internet ha fet palesa la necessitat de complementar aquests recursos amb sistemes de navegació que permetin als usuaris moure's amb eficàcia entre les diverses informacions.

## Résumé

La récupération d'information est pour les archivistes une tâche quotidienne qui, bien souvent, n'a pas mérité une analyse détaillée de leur part. Cependant, parmi les différentes fonctions du travail de l'archiviste, celle de récupération de l'information a une importance essentielle par le fait, principalement, qu'il s'agit d'une fonction "pont" entre celle de description et celle de consultation du document, sans laquelle la première n'a pas beaucoup de sens et la deuxième ne peut pas être réalisée.

Il s'avère que, par la confluence de différents facteurs qui se résument à la nécessité d'accueil d'utilisateurs sans la médiation directe de l'archiviste, les questions rattachées au traitement de l'information et, en particulier, celles liées à la récupération de l'information, sont de plus en plus importantes. La principale évidence de cette importance croissante est l'application massive de ressources informatiques à celles descriptives. Mais ces ressources informatiques ont été utilisées jusqu'à maintenant davantage comme un outil pour la création d'instruments descriptifs que comme outils de récupération de l'information de la part des usagers.

L'analyse des systèmes de récupération de l'information gérés par le service des archives et leur adaptation aux circonstances actuelles doivent être entendues comme une partie de plus du programme descriptif du service des archives.

Bien que les expériences dans ce domaine soient rares, celles-ci permettent d'identifier les différentes possibilités qui seront présentées à l'archiviste à l'heure de définir les systèmes de récupération. Les plus connues sont celles d'utilisation de techniques de contrôle linguistique, comme le sont celles d'indexation et de contrôle d'autorités: descripteurs libres, listes de descripteurs, langages contrôlés (thesaurus...). Néanmoins, la présence d'Internet a mis en évidence la nécessité de compléter ces ressources par des systèmes de "navigation" permettant aux usagers d'évoluer au milieu de l'information d'une manière efficace.

## Summary

For archivists, the recovery of information is an everyday task which often has not merited detailed analysis by them. However, among the different functions of the archivist's work, that of recovering information has central importance due, above all, to the fact that it is the 'bridge' function between the description and the consultation of the document, without which the first makes little sense and the second cannot be carried out.

The confluence of different factors, which can be summarised as the need for attention to users without the direct intervention of the archivist, has the result that issues related with the treatment of information, and particularly those related with the recovery of information, are of increasing importance. The principal evidence of this growing importance is the massive application of IT resources in description resources. But these IT resources have been used to date more as a tool for creating descriptive instruments than as tools of information recovery by the users.

The analysis of the information recovery systems managed by the Archives Service, and its adaptation to the present circumstances, must

be understood as one more part of the Descriptive Programme of the archives service.

Although experiences in this field are still scarce, they make it possible to identify the different possibilities which offer themselves to the archivist when defining the recovery systems. The best known are those which use linguistic control techniques, such as indexing and authority control: free descriptors, lists of descriptors, controlled languages (thesauruses, etc.) However, the appearance of Internet has manifested the need to complement these resources with systems of 'surfing' which enable users to move effectively through the information.