



- INICIO
- ZONA DE NOTAS
- NÚMEROS ANTERIORES
- SOBRE EPI
- DISTRIBUCIÓN/BASES DE DATOS
- SUSCRIPCIONES
- PUBLICIDAD
- INFORMACIÓN AUTORES
- AGENDA
- IWETEL
- THINKEPI
- COPYRIGHT
- EQUIPO DE LA REVISTA

- CONTACTO



 WWW EPI

Noviembre 1999

Evaluación de sistemas de recuperación de información: aproximaciones y nuevas tendencias

Por **M^a Dolores Olvera Lobo**

Resumen: Las principales tendencias de investigación en el ámbito de la evaluación de la recuperación de información (RI) se resumen en dos: la aproximación tradicional o algorítmica -cuyos referentes principales son el proyecto Cranfield y la más reciente conferencia Trec- frente al modelo cognitivo -centrado en las estructuras cognitivas del individuo-. Estas corrientes de investigación, no obstante, se complementan, por lo que una aproximación integradora debería presentar una nueva propuesta metodológica válida que contase con un sustento teórico sincrético, pero no ecléctico.

Palabras clave: Evaluación de la recuperación de información, Aproximación tradicional, Aproximación algorítmica, Aproximación cognitiva, Modelo cognitivo, Aproximación integradora.

Title: Evaluation of information retrieval: approaches and new trends

Abstract: Research on the evaluation of information retrieval (IR) can be classified into two main streams: a) the traditional or algorithmic approach, whose main referents are the Cranfield project and, more recently, the Trec conference, and b) the cognitive model, which focuses on the cognitive structures of individual users. These tendencies, however, are complementary. Thus, both of them should be integrated in a new syncretistic, but not eclectic, theoretical model.

Keywords: Evaluation of information retrieval, Traditional approach, Algorithmic approach, Cognitive approach, Cognitive model, Syncretistic approach.

Olvera Lobo, M^a Dolores. "Evaluación de sistemas de recuperación de información: aproximaciones y nuevas tendencias". En: *El profesional de la información*, 1999, noviembre, v. 8, n. 11, pp. 4-14.

La evaluación de sistemas de recuperación de información (SRI) se encuentra en un momento crucial en el que se suceden los cambios, se realizan constantes aportaciones y se emprenden rigurosas investigaciones. Sus antecedentes, los nuevos proyectos, prototipos y experimentos de evaluación que actualmente se llevan a cabo, se remontan a las décadas de los años 50 y 60, lo que Lesk (1995) denomina la infancia y la edad escolar, respectivamente, de las siete edades de la recuperación de la información. Desde entonces se pueden detectar varias aproximaciones al problema.

Dervin y Nilan (1986) plantearon las tendencias de investigación existentes como la centrada en sistemas frente a la orientada a usuarios provocando un creciente interés por incorporar a éstos en el proceso de evaluación. Por su parte, varios años después, Ingwersen (1992) sintetiza las líneas de estudio en tres corrientes: la clásica o algorítmica, la orientada a usuarios (*user-oriented*) y la cognitiva, que pueden resumirse en dos: la aproximación tradicional frente al modelo cognitivo (Harter y Hert, 1997).

La primera se centra en los algoritmos y estructuras de datos necesarios para optimizar la eficacia de las búsquedas que pueden realizarse en bases de datos textuales. Mientras, la segunda analiza el papel del usuario y de las fuentes de conocimiento implicadas en la recuperación de información (RI) (Ingwersen y Willett, 1995). No obstante, estas corrientes de investigación no son incompatibles entre sí (Ingwersen, 1992) sino que, por el contrario, pueden (y deberían) complementarse.

La aproximación tradicional o algorítmica

La evaluación de SRI resulta de la máxima importancia para garantizar su adecuado funcionamiento, es decir, la recuperación de información pertinente y



una correcta adaptación a las necesidades de los usuarios: facilidad de uso, respuestas rápidas, coste razonable, etc. (Salton y McGill, 1983). Casi todos los problemas detectados en este proceso se derivan de la falta de acuerdo sobre las características de la recuperación misma (Tague, 1992; 1996).

En sentido genérico, se habla de un sistema como una colección de componentes interrelacionados que interactúan en la realización de una función específica con un propósito concreto, por lo que casi cualquier entidad puede considerarse como tal (Pao, 1989).

Un SRI, sin embargo, en su forma más simple, puede verse como una "caja negra" que acepta *inputs* y produce *outputs* (Harter y Hert, 1997). Durante este proceso realiza actividades que incluyen: el reconocimiento de la estrategia de búsqueda planteada, la aplicación de diferentes algoritmos de recuperación y de ordenación de los resultados según su relevancia o utilidad probable para el usuario, la selección de ítems de información (documentos o su representación), etc. Intentan localizarlos y recuperarlos tan veloz y económicamente como sea posible, por lo que su valor depende de su capacidad para identificar rápida y correctamente la información útil, de su facilidad para rechazar los ítems extraños o irrelevantes y de la versatilidad de los métodos que emplea (Salton y McGill, 1983).

«Las críticas a la aproximación algorítmica se centran en la ausencia de los usuarios y en que no se consideran sus estructuras cognitivas»

Algunos precedentes de la aproximación clásica a la evaluación de SRI se sitúan en 1953 cuando se desarrollan simultáneamente el test *Astia-Uniterm*, en la *Armed Services Technical Information Agency* de los Estados Unidos, y el test *Uniterm* del *College of Aeronautics*, Cranfield, en el Reino Unido. Sin embargo, en sentido estricto, esta tendencia se inaugura con el proyecto *Cranfield* (Cleverdon, 1997), dirigido por Cyril Cleverdon a partir de 1957, con la colaboración de la estadounidense *Fundación Nacional de la Ciencia (NSF)* y de la *Asociación Británica de Bibliotecas Especializadas y Gabinetes de Información (Aslib)*. Fue desarrollado en dos etapas (*Cranfield I* y *II*) y, más que a sus resultados, su trascendencia y repercusión se deben a que creó un modelo, una metodología y unas pautas para la evaluación de los modelos aún vigentes.

Cranfield I (1957-1962) comparó el funcionamiento de cuatro sistemas de indización: la *Clasificación decimal universal*, una lista alfabética de materias, un sistema de ordenación facetado y el modelo *Uniterm*. En este proyecto ya se comienzan a perfilar las herramientas que se utilizarían en años posteriores para la evaluación:

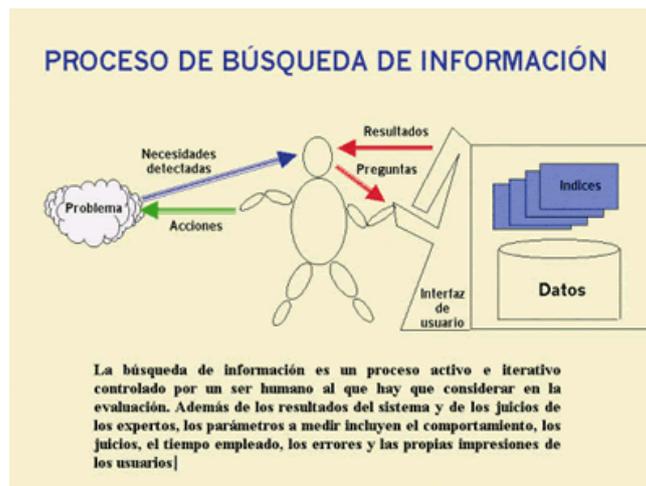
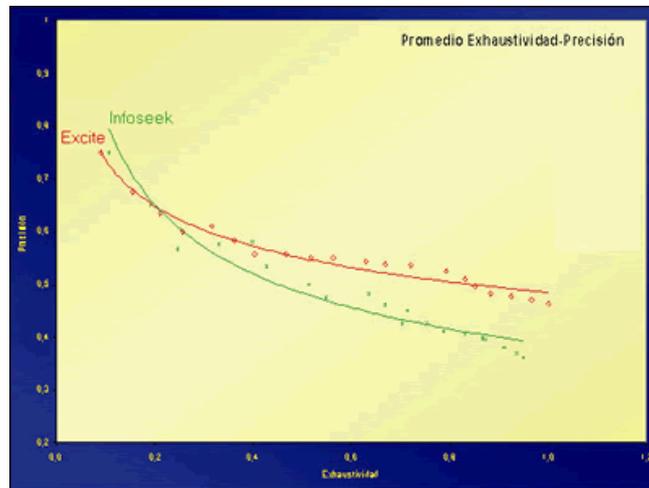
- Una colección de documentos (fuente) de los que se obtuvieron las preguntas que se le plantearían al sistema.
- Una serie de juicios de relevancia de los documentos recuperados y expresados en tres niveles: muy útil, útil, nada útil.
- El uso de las medidas de exhaustividad y precisión para analizar los resultados.

Este estudio dio paso a una segunda serie de pruebas que permitieron la consolidación de los denominados "instrumentos de *Cranfield*" antes mencionados. Así, *Cranfield II* (1963-1966) evaluó treinta y tres tipos de lenguajes de indización, diferenciados en terminología y estructura. En este caso, la colección de documentos y el número de preguntas era mucho mayor. Éstas se obtuvieron pidiendo a los autores de los artículos seleccionados (documentos base) que reconstruyeran las búsquedas que habían planteado originalmente al recopilar información para escribir sus trabajos.

Las consultas se realizaron usando estrategias basadas en diferentes niveles de coordinación de los términos. En estos estudios, centrados en la eficacia de esos lenguajes de indización, los autores determinaron la relevancia de los documentos en relación con las preguntas formuladas. Para ello se aplicó una escala de varios niveles, de uno a cuatro, y no un sistema de relevancia binaria (relevante-no relevante) (Cleverdon, 1997).

Este modelo tradicional implica que:

1. El usuario reconoce una necesidad de información.
2. Dirige a un SRI una consulta basada en ella.
3. El sistema la compara con las representaciones de los documentos de su base de datos.
4. Presenta al usuario el texto o textos que pueden interesarle.
5. El usuario los examina y juzga su relevancia.



La intención es que algunos o todos los documentos recuperados satisfagan total o parcialmente su necesidad de información.

Tanto *Cranfield I* como *II* fueron objeto de críticas (**Harter y Hert**, 1971; **Swanson**, 1965 y 1971) aunque, pese a todo, los componentes principales utilizados en los experimentos —una colección de documentos para la evaluación, un grupo de preguntas, una serie de juicios de relevancia y el uso de las medidas de exhaustividad y precisión— sentaron las bases del diseño de evaluaciones posteriores. Ésta parece ser la primera aparición de la relevancia como criterio de funcionamiento para la evaluación de la RI (**Ellis**, 1996).

«Con las pruebas de Cranfield se pasa de una aproximación especulativa en los diseños de los SRI a una concepción empírica y experimental»

Estos instrumentos siguieron utilizándose en gran parte de los estudios posteriores hasta nuestros días y, tras su incorporación a importantes proyectos como la renombrada *Trec* (*Text Retrieval Conference*) (**Harman**, 1993, 1995, 1997 y 1998), han cobrado nuevo impulso en este tipo de investigaciones. A este respecto, las pruebas de *Cranfield* constituyeron una aportación fundamental en el campo de la evaluación, ya que se pasa de una aproximación especulativa en los diseños de los SRI a una concepción empírica y experimental. Por otra parte, ofrecieron un marco teórico a partir del cual se desarrolló esta subdisciplina dentro las ciencias de la información, y los procedimientos adoptados en estas pruebas constituyeron la primera piedra de una tradición en la investigación de su diseño y su evaluación.

Esta aproximación algorítmica se utilizó en otros muchos estudios relacionados con teorías de clasificación, de indización y de cuestiones vinculadas a los lenguajes controlados y con la representación del lenguaje natural, así como para analizar métodos alternativos a la recuperación booleana, como la búsqueda *best-match* o por proximidad, la ordenación de resultados por relevancia, los sistemas probabilísticos o vectoriales de recuperación, etc.

Entre los primeros estudios cabe señalar el de *Medlars* (*Medical Literature Analysis and Retrieval System*), y el *Stairs* (*Storage and Retrieval System*). En los años 1966 y 1967 se llevó a cabo en la *Biblioteca Nacional de Medicina* (*NLM*) de los Estados Unidos un programa dirigido por **F. Wilfrid Lancaster** para evaluar el funcionamiento del primero. Fue diseñado para establecer las necesidades del usuario y determinar las capacidades de *Medlars* respecto a factores como: cobertura, exhaustividad, precisión, tiempo de respuesta, formato de los resultados y cantidad de esfuerzo del usuario. Concretamente,

se pretendía identificar las principales causas de los fallos en las búsquedas para mejorar su rendimiento (**Lancaster**, 1997).

«Muchos conceptos relativos a la RI son difíciles de definir, tanto operativa como conceptualmente»

En este proyecto se consideraron los motivos de los resultados insatisfactorios y se estudiaron todos los factores que incidían en éstos: el lenguaje documental utilizado, la indización efectuada, la interfaz de usuario, la búsqueda planteada, etc. Las medidas empleadas fueron, nuevamente, la precisión y la exhaustividad basadas en la relevancia. Este trabajo, que por su rigor y originalidad contó con una gran difusión y reconocimiento, constituyó un destacable primer intento de acercar y adaptar la evaluación de SRI a modelos reales en funcionamiento.

Aunque ya se había aplicado esta metodología en el *Smart* de **Gerard Salton** y el *Medlars* de **F. W. Lancaster**, se seguía planteando como principal objeción la dudosa aplicación de los pequeños tests de laboratorio a operaciones de recuperación reales realizadas en un entorno normal de usuario (**Salton**, 1992).

En 1985 **Blair** y **Maron** publicaron las conclusiones de un experimento a gran escala cuyo fin era evaluar la eficacia de un sistema de búsqueda y recuperación a texto completo. La base de datos de *Stairs* contenía aproximadamente 350.000 páginas de documentos que incluían informes de ingeniería, memorias de gestión, actas de reuniones, etc. Los resultados mostraron que la media de exhaustividad no era de más del 20%, con un grado de precisión medio del 79%. Estos datos, según los autores, no fueron sorprendentes puesto que es difícil obtener un gran nivel usando un modelo de recuperación a texto completo cuando la base de datos de documentos es muy grande.

Esto se explica porque, para una recuperación eficaz, el usuario debe ser capaz de predecir y usar en los términos de su pregunta aquellas palabras, frases y combinaciones que aparecen en la mayor parte de los documentos relevantes y que no aparecen en la mayoría de los que no lo son (**Blair** y **Maron**, 1990). Como en todos los proyectos realmente significativos, los resultados del experimento *Stairs* inspiraron un debate (**Salton**, 1986 y 1992) y sigue siendo objeto de estudio y mejora por parte de sus autores (**Blair**, 1996).

En la actualidad, los buscadores de la *www*, que cuentan con enormes y dinámicas bases de datos a texto completo, han llevado al límite este problema y la evaluación de grandes SRI en funcionamiento. Dada la ausencia de lenguajes documentales para la recuperación en internet y el enorme volumen de datos existente, más que exhaustividad los usuarios anhelan una alta tasa de precisión en sus consultas, si bien es éste un aspecto que las herramientas de búsqueda deben mejorar.

Los buscadores intentan facilitar la tarea al presentar los resultados según su relevancia a la pregunta planteada, mostrando los documentos supuestamente más importantes en la parte superior del *ranking*. En 1983, **Salton** y **McGill** habían propuesto un método para la evaluación del funcionamiento de los sistemas que ofrecen las respuestas ordenadas según este procedimiento. Así, los valores de exhaustividad y precisión dependerán del punto de corte, lo que **Blair** (1980) denomina "umbral de futilidad", es decir, el punto en el que el usuario cesa de examinar la lista de documentos recuperados. El par de valores exhaustividad-precisión se calcula para cada posición en la lista de resultados, para cada rango, usando éste como un nivel de recuperación (**Tague**, 1992). El método desarrollado por **Salton** y **McGill** goza de gran aceptación en la comunidad investigadora y, como se aprecia en la figura 1, puede aplicarse a la evaluación de los resultados ofrecidos por los buscadores de internet con los debidos ajustes que lo adapten a las particularidades de la web (**Olvera**, 1998).

Como puede comprobarse, el proyecto *Cranfield* marcó las pautas para el desarrollo de gran parte de los experimentos de evaluación posteriores. Sin embargo, desde finales de la década de los 70 y principios de los 80 esta situación comienza a cambiar y surge una actitud más ecléctica. En estos años aumentó la insatisfacción con la aproximación formalista imperante —basada en modelos formales matemáticos aplicados a datos preexistentes que constituyen la colección— si bien la investigación, en gran parte, continuó empleando estos diseños.

De esta manera *Cranfield* presenta algunas deficiencias desde la nueva orientación cognitiva. Las principales críticas a la aproximación algorítmica se centran en que se aprecia generalmente la ausencia de los usuarios y no se consideran sus estructuras cognitivas. Algunos autores señalan la tendencia de este enfoque al reduccionismo al obviar el objetivo fundamental: la transmisión eficaz del conocimiento, y caracterizarse por su "febril desconfianza en los fenómenos cognitivos y mentales o, dicho en términos técnicos, por su antimentalismo" (**García**, 1995).

Por otra parte, no se controlan muchos aspectos que son importantes en el proceso de recuperación como las interfaces, el acceso a las clasificaciones, etc.

Además, no se tiene en cuenta el gran número de estudios basados en el comportamiento de los usuarios en las búsquedas, por lo que se detecta una casi total ausencia de investigaciones del impacto de las técnicas algorítmicas sobre ellos en el contexto socio-organizativo (**Ingwersen y Willett**, 1995).

Finalmente, cabe señalar que se han detectado diversos problemas en este modelo experimental tradicional (**Harter y Hert**, 1997) en relación con:

1. La validez y fiabilidad: provocados por la omisión del usuario.
2. La generalización: falta de ejemplos tomados al azar, uso de pequeñas colecciones de evaluación y conclusiones contradictorias.
3. La utilidad: resultados de los experimentos no aplicables a sistemas reales.
4. Conflictos conceptuales: problemas de "solidez" que plantean los conceptos básicos en los que se apoya la evaluación de SRI.

En noviembre de 1992 tuvo lugar un acontecimiento clave: la primera conferencia *Trec*, celebrada en el estadounidense *Instituto Nacional de Estándares y Tecnología (Nist)*, auspiciada por éste y el *Arpa* (**Harman**, 1993).

<http://trec.nist.gov>

Se constituyó en una iniciativa sin precedentes que, en gran parte, hereda la filosofía de *Cranfield* y es la principal y más significativa contribución a la evaluación de la RI en varios sentidos. Reunió a investigadores de este campo para discutir los resultados que ofrecían cada uno de sus sistemas, tomando como base una nueva gran colección de evaluación, *Tipster* (**Harman**, 1998). En este primer encuentro participaron 22 grupos (entre universidades y empresas) de los Estados Unidos, el Reino Unido, Alemania y Australia.

«En el caso de los SRI interactivos, una evaluación realista debe ser multi-dimensional que contemple todas las facetas de la interacción con el usuario»

En la conferencia se estableció la creación de una colección constituida por un conjunto de documentos —que incluía el *Wall Street Journal*, *San Jose Mercury News*, artículos de informática de *Ziff*, informes del *Registro Federal*, patentes estadounidenses, resúmenes del *Dept. of Commerce*, etc.—, los temas o preguntas y un conjunto de valoraciones de relevancia. Lo novedoso del experimento fue su tamaño y variedad —dos gigabytes que, para la época, era una escala hasta entonces desconocida, y que desde entonces no ha dejado de crecer—. *Trec* está realizando un papel muy importante al ofrecer un paradigma de referencia sobre evaluación para la investigación en la materia.

Lo más significativo es su filosofía global (**Ellis**, 96) en cuanto que hay una vuelta al "realismo" en la evaluación. Además, a pesar de que el "concepto *Trec*" se basa en una colección ideal y emplea aproximaciones estadísticas y probabilísticas, la gama de técnicas exploradas es más amplia y refleja la creciente heterogeneidad de las aproximaciones en la investigación en RI.

La segunda conferencia se celebró en agosto de 1993, donde se incorporaron nueve grupos más y se analizaron varios métodos de recuperación. Se convirtió en una cita anual y la tercera, como ya se esperaba, se celebró en 1994 con 33 asociaciones participantes. Las sucesivas reuniones se centran en el mismo cometido, basado en los modelos clásicos de RI, por lo que la evaluación de *Trec* usó las medidas tradicionales de exhaustividad y precisión (**Harman**, 1998). La atención se enfocó sobre dos tareas que simulaban situaciones básicas de recuperación: la tarea de *routing* y la *ad hoc*.

En la primera se usan peticiones ya realizadas para localizar nuevos datos, como en un perfil de usuario para la difusión selectiva de la información (*DSI*). En la *ad hoc* se plantean nuevas preguntas a datos ya existentes, de la misma forma que un usuario se dirige a un sistema en línea para realizar una consulta. Así, en *Trec*, el *routing* se representa mediante temas conocidos y documentos relevantes a esos temas también conocidos. En la tarea *ad hoc* las preguntas se generan a partir de nuevas materias y se consultan los datos sin saber cuáles son los documentos más importantes de acuerdo con la petición.

Algunos de los 200 temas de evaluación (en *Trec-6*, celebrada en noviembre de 1997, eran ya 350) fueron creados por analistas de información que también ofrecían juicios de relevancia para los documentos recuperados. La valoración de la relevancia presentaba el problema principal. Con tantos documentos en la colección, el análisis completo para cada tema hubiese requerido cientos de millones de estos juicios. En vez de eso se optó por hacerlo sobre los primeros 200 recuperados por cada sistema, ya que los resultados se presentaban ordenados por relevancia, y éstos fueron agrupados (*pool*) para estimar la exhaustividad y precisión relativas. Esta aproximación era similar a la revisada por **Sparck Jones y Van Rijsbergen** en su consideración sobre la necesidad de una colección ideal (**Harman**, 1998).

«La aproximación por dominios se ha querido proponer como una tercera vía o tendencia»

La programación de las acciones llevadas a cabo en el marco de esta conferencia se desarrollan de febrero a noviembre de cada año. De febrero a

junio los participantes —es decir, universidades y empresas de diversos países que quieren evaluar sus SRI— indizan mediante sus sistemas la colección constituida por diversos tipos de documentos a texto completo. En junio, el *Nist* envía los temas de búsqueda a los participantes que deberán realizar las consultas hasta el mes de agosto. Después, les remiten sus resultados y se comparan las referencias recuperadas por los distintos sistemas con los documentos considerados relevantes a esas preguntas por los asesores externos y que habían servido, a modo de documentos fuente, para construir las preguntas. El *Nist* procede a la evaluación de los resultados y en octubre informa a los participantes. Las conclusiones finales son analizadas y discutidas por ambos hasta el mes de noviembre, fecha de la convocatoria anual de la conferencia.

Las medidas comparativas de funcionamiento basadas en la exhaustividad y la precisión se calculan para cada sistema y cada tema, de acuerdo con el método de **Salton y McGill** (1983), y se pueden representar según la curva exhaustividad-precisión que indica las medias obtenidas para los temas o bien para los sistemas analizados. Su uso, obviamente, asume que los resultados ofrecidos están ordenados según su relevancia respecto a la pregunta planteada.

Trec continúa avanzando con éxito a la vanguardia de la recuperación textual, ofreciendo un foro para el análisis de sistemas, utilizando datos y métodos comunes, actuando como un punto de referencia central de discusión sobre cuestiones metodológicas en la evaluación de la RI (**Harman**, 1998). Ha creado un amplio marco de colaboración en el que participan numerosos equipos de investigación y organizaciones interesadas en comparar los resultados de sus sistemas, con un diseño en continua evolución. Asimismo, permite difundir rápidamente sus descubrimientos, fomentar la investigación para grandes aplicaciones ofreciendo una colección importante y procedimientos de cálculo (*scoring*) uniformes. Actualmente es el principal esfuerzo experimental.

Los participantes han examinado una gran variedad de técnicas de recuperación, incluyendo métodos para utilización de tesoro automático, ponderación de términos sofisticada, técnicas de lenguaje natural, *feedback* de relevancia, etc. En cada nueva convocatoria se amplían las áreas de investigación. La última, *Trec-7* se celebró en Gaithersburg, Maryland, donde se encuentra la sede del *Nist*, en noviembre de 1998. *Trec-8* se celebra del 16 al 19 de noviembre de este año también en el *Nist*.

Sin embargo, estas reuniones también han estado sujetas a críticas. La más seria es que la evaluación se basa en juicios de relevancia (**Harter y Hert**, 1997). También se centran en la forma de generar los temas y en los juicios poco realistas para usuarios de verdad (**Beaulieu et al.**, 1996). **Sparck Jones** (1995) se fija en la ausencia del usuario final y en que documentos, temas y juicios de relevancia son “no naturales”. Otras inciden en el uso de la exhaustividad como medida de funcionamiento, que considera insatisfactoria en grandes bases de datos y señala que el uso de las gráficas exhaustividad-precisión ocultan el funcionamiento real de los sistemas al no reflejar sus posibilidades totales (**Belkin et al.**, 1993). *Trec* ha ido evolucionando para responder a las críticas planteadas y ya dispone de una inmensa colección de documentos, usa temas de búsqueda reales y un amplio abanico de juicios de relevancia.

Por otra parte, se ha introducido una prueba interactiva, a iniciativa de varios participantes, con la que concentran su esfuerzo en el proceso de recuperación dentro del marco general de *Trec* y siguen las orientaciones que tratarán de asegurar la comparabilidad de los resultados de diferentes sistemas (**Beaulieu et al.**, 1996). Además, se han incorporado nuevos campos de interés como las búsquedas en español, francés o recuperación de información en chino o de lenguaje oral —de la que se encarga, entre otros, **Sparck Jones**—. De esta forma, al igual que *Cranfield*, constituye un hito en la investigación en cuanto a la evaluación de la RI (**Harter y Hert**, 1997).

La aproximación cognitiva

El desarrollo de una teoría de la disciplina ha dado lugar a la aparición de diferentes modelos para explicar este fenómeno (**Codina**, 1995). No obstante, y expresado de forma muy simplista, se puede decir que la recuperación implica encontrar alguna información deseada en un depósito de información o base de datos (**Meadow**, 1992). Sin embargo, esta claridad sólo es aparente ya que muchos de los conceptos relativos a la RI son difíciles de definir, tanto operativa como conceptualmente.

Entre estos “conceptos inefables” (**Belkin**, 1981) se encuentran la información y la necesidad de información, el significado y la aproximación al tema, la relevancia, la satisfacción y la efectividad. En este sentido, **Swanson** (1988) cita varios “postulados de impotencia” y afirma que nuestra comprensión de los mismos sigue siendo deficiente.

La idea de información es protagonista de casi constantes reflexiones teóricas (**Tramullas**, 1997) en el campo de las ciencias de la información, y aunque se han ofrecido numerosas definiciones, ninguna es totalmente satisfactoria. La tendencia actual es que el estudio del fenómeno informacional se realice

mediante enfoques centrados en el usuario, en el individuo, lo que ha llevado a concebirla como algo "subjetivo, individualizado, que forma parte del proceso continuo que sigue cualquier persona en su relación con el entorno que le rodea" (**Fernández**, 1994).

La necesidad de información también es un concepto vago. Puede surgir cuando un individuo reconoce que su almacén actual de conocimiento es insuficiente para resolver la tarea en curso, para solucionar conflictos en un campo temático o para cubrir lagunas en cualquier área, y puede, o no, ser totalmente consciente. De hecho, pueden coexistir diferentes niveles de conciencia sobre la misma, es decir, en ciertas etapas un usuario puede incluso no saber su alcance.

En consecuencia, desde la perspectiva del sistema, una demanda articulada hecha por un usuario con frecuencia no representa forzosamente su necesidad real. Por lo tanto su satisfacción puede no resolver el problema (**Pao**, 1989). Constituye básicamente un estado cognitivo del individuo —si bien puede surgir por motivos de carácter físico o afectivo (**Wilson**, 1981)— y hay diversas posturas a la hora de definirla, según las cuales puede ser:

- a. Una incongruencia contextual en la que la estructura cognitiva de la persona no se adecúa a una tarea (**Ford**, 1980).
- b. El reconocimiento de algo erróneo en el estado del conocimiento (*ASK, anomalous state of knowledge*) del usuario, que desea resolver la anomalía (**Belkin**, 1980).
- c. El fin de la percepción interna de la coherencia de la información (**Dervin**, 1980).
- d. La conciencia de que la información disponible es insuficiente para rellenar lagunas o solventar inseguridades o conflictos en el área de conocimiento (**Horne**, 1983).
- e. Un defecto del modelo mental (**Marchionini**, 1989).

Éstos y otros planteamientos han generado diferentes líneas de estudio (**Dervin**, 1987) ya que, por su naturaleza compleja, pusieron de manifiesto que la investigación orientada únicamente a técnicas no podía ofrecer una panorámica completa del proceso en su integridad. Para lograr este objetivo es obligado incorporar las características del sistema —incluyendo los métodos de representación y de recuperación de las aproximaciones algorítmicas—, las características situacionales del usuario y los intermediarios imprescindibles, el más importante de los cuales es la interfaz, el mecanismo principal de enlace entre el sistema y el usuario (**Ingwersen y Willett**, 1995).

El término usuario, igualmente, es bastante ambiguo. De forma genérica pueden establecerse varios tipos (**Pao**, 1989): el potencial, el previsto y el beneficiario. En el contexto de una organización los primeros son los que aún no disfrutan de un acceso al servicio de información. Los previstos son los que sí lo tienen y planean hacer uso del mismo. En cuanto a los últimos, son los que ya han obtenido algún beneficio de los datos recuperados. Sin embargo, el éxito o el fracaso de la búsqueda puede depender de características personales muy dispares —experiencia previa en búsquedas, edad, personalidad, estatus académico— y el tipo de usuario. En consecuencia, la forma de enfrentarse a los procesos y sistemas de recuperación se ven afectados por numerosos factores.

«Blair denomina 'umbral de futilidad' al punto en el que el usuario cesa de examinar la lista de documentos recuperados»

Por otra parte, la interacción cobra un protagonismo fundamental. Puede definirse como aquellos procedimientos de comunicación en los que intervienen todos los agentes importantes: el usuario, el intermediario y el propio sistema, donde este último se entiende como la suma de la información potencial —sobre todo en forma de texto o de su representación— y de sus propias características, tales como la estructura de la base de datos y las técnicas de RI (**Ingwersen**, 1992).

El análisis del comportamiento de los usuarios es de gran importancia y debe tenerse muy en cuenta en el diseño de los sistemas (**Belkin**, 1990). Por ello su análisis ha de reflejar su capacidad para satisfacer al usuario (**Van Rijsbergen**, 1979). En el caso de los SRI interactivos, una evaluación realista debe ser multi-dimensional, porque estos modelos han creado la necesidad de nuevas medidas que contemplen todas las facetas de la interacción con el usuario.

No obstante, aún no se han terminado de consensuar cuáles han de ser las medidas de la eficacia de recuperación realmente aplicables a la interacción producida en los SRI. Con su uso es cuando surge un gran debate de fondo: conceptual, de interpretación, de aplicabilidad, etc. Éste constituye hoy uno de los temas que despiertan mayor interés.

La perspectiva de la RI orientada a usuarios se centra en la representación de los documentos y de los problemas de información, el comportamiento en las búsquedas y los componentes humanos de los sistemas en situaciones reales.

Se nutre principalmente de la psicología cognitiva y emplea métodos de las ciencias sociales. La principal diferencia entre la aproximación orientada al usuario y la algorítmica estriba en las distintas actitudes de los investigadores ante el concepto de información, la naturaleza de la necesidad de información, el entorno de investigación usado para la experimentación y los roles del intermediario, del usuario y de los SRI (**Ingwersen**, 1992).

Este acercamiento ha permitido iluminar aspectos sustanciales del comportamiento mental de los usuarios y de las características de sus consultas, tanto en relación con el individuo como con los contextos sociales y organizativos. Por ejemplo, sus necesidades pueden ser estables, pero también cambiar en el curso de la interacción con un SRI. Por otra parte, también hay que contar con la posibilidad de ser mal definidas debido a una falta de conocimientos previos apropiados. Los análisis muestran que es necesario contextualizarlas mediante información suplementaria sobre intenciones, propósitos y metas.

La búsqueda de información y la formación sobre su propia necesidad se contemplan como procesos cognitivos del usuario individual en donde el sistema de recuperación y los intermediarios funcionales (como la interfaz) son componentes cruciales del proceso de contextualización.

Una consecuencia inmediata es que la amplia gama de técnicas de representación y de búsqueda actuales se consideran estructuras complementarias de información de diferente naturaleza y origen cognitivos. Esto conduce a una teoría cognitiva, una empresa globalizadora que considera todos los componentes de la recuperación como representativos de estructuras cognitivas de diversa complejidad que cooperan en un proceso interactivo de comunicación (**Ingwersen**, 1996).

La aproximación orientada al usuario dio paso al acercamiento cognitivo propiamente dicho, cuya principal finalidad es mejorar la representación documental (**Izquierdo**, 1995) y diseñar y construir sistemas que le sean más cercanos (**Chen y Dhar**, 1991). Asimismo, se centra en las actividades mentales cognitivas, emocionales y de motivación en relación con todos los componentes del proceso. Por tanto, en este enfoque cobran especial importancia la semántica del texto y el estudio del lenguaje natural en el entorno de las necesidades de información del usuario (**Tramullas**, 1998). Sin embargo, el eje de esta tendencia es el análisis de sus estructuras cognitivas.

«Trec continúa avanzando con éxito a la vanguardia de la recuperación textual, ofreciendo un foro para el análisis de sistemas, utilizando datos y métodos comunes»

En este sentido una de las primeras aportaciones interesantes es la de **Belkin** (1980) y su *ASK*. El marco conceptual en el que se elabora es explícitamente cognitivo y estructural. Los modelos individuales del conocimiento son transformados en texto por el emisor con la intención de cambiar la organización de la imagen de algún receptor. Hay dos componentes separados, aunque relacionados, en este marco: uno se relaciona con la decisión de un autor o emisor de comunicar algún aspecto de su estado de conocimiento mediante un texto; el otro elemento se refiere a los factores que subyacen bajo una decisión personal de búsqueda de información y en la idea de que un documento concreto ofrece lo que se pedía.

El propósito de una persona de comunicar algún aspecto de su entendimiento del mundo supone aislar, de todo su estado de conocimiento, esa parte de la estructura individual que va a ser transmitida. Sin embargo antes de hacerlo es cambiada según las intenciones del autor y la percepción que tenga del probable estado de conocimiento de los destinatarios. Una vez modificada es convertida en un texto que se une a otros, a los que los receptores potenciales tendrán acceso, por ejemplo, gracias a una base de datos.

El concepto de *ASK* es introducido en relación con el segundo componente. Un receptor o usuario reconoce, o es consciente, de una deficiencia o anomalía en su estado del conocimiento y la convierte en una estructura comunicable, dirigiéndose después a un SRI. Ambas estructuras interactúan y será el receptor el que, al examinar los textos que son recuperados, decida si ha sido resuelta —en cuyo caso el sistema se cerraría— o no lo ha sido, por lo que el proceso continuaría hasta que la pregunta fuera satisfecha.

Estos modelos cognitivos sugieren que los procesos realizados durante la recuperación pueden verse como las interacciones entre diferentes tipos de estructuras cognitivas (**Ingwersen y Willett**, 1995), las cuales son consideradas generalmente como manifestaciones de la cognición humana, de la reflexión o de las ideas. En RI toman la forma de transformaciones generadas por diferentes actores humanos con gran variedad de orígenes: diseñadores y productores de sistemas, desarrolladores de técnicas de recuperación, constructores de reglas de indización, autores de texto e imágenes, diseñadores de mecanismos intermediarios, y usuarios en un contexto de un determinado dominio social u organizativo (**Ingwersen**, 1996).

Por tanto, se observa un desplazamiento de los centros de interés hacia los aspectos cualitativos del procedimiento, puesto que el requerimiento actual es

el acceso y utilización intelectual de la información (**Tramullas**, 1998). En este sentido **Allen** (1991), en su revisión de la literatura sobre investigaciones cognitivas, señala varias áreas de estudio relacionadas con la recuperación de información y la cognición:

- a. El conocimiento y los modelos cognitivos: el conocimiento del mundo de los usuarios, el del SRI, del tema, el grado de conciencia de los objetivos de búsqueda, etc.
- b. Procesos cognitivos generales y específicos: aquellos independientes de la RI y los desarrollados especialmente para aplicarlos en esta actividad.
- c. Capacidades y habilidades cognitivas, dado que las diferencias interpersonales afectan a los procedimientos utilizados.
- d. Estilos cognitivos y preferencias relativamente estables en el modo de pensar, aprender y resolver problemas.

Esta línea de investigación también ha recibido un abanico de críticas de entre las que cabe destacar las siguientes. No todas son válidas y, actualmente, las más populares han obviado algunos hechos fundamentales del pensamiento humano, como su dependencia de la interacción con otros sujetos del entorno de la persona estudiada, o el carácter interno del contexto.

Como quiera que las conexiones neuronales se realizan en grandes cantidades hasta los once años, para después descender y estabilizarse, se puede afirmar que la interacción y el entorno condicionan directamente el estilo y las capacidades cognitivas del usuario. No hay, pues, una distinción clara entre forma y fondo, entre el sistema —diseñado con un fin y un usuario determinado en mente— y el individuo que lo usa. En otras palabras, el cerebro y la mente no se deben estudiar disociados de su contexto sociocultural. En este sentido hay que recordar, por ejemplo, que las aproximaciones bibliométricas permiten aprehender tendencias sociales en el desarrollo del conocimiento.

Una de las principales críticas, sin embargo, se centra en la falta de rigor en sus planteamientos metodológicos. **Kuhn** (1990) señala como una de las condiciones indispensables para la existencia de un paradigma científico los ejemplares compartidos o *shared exemplars*, es decir, modelos de experimentación especialmente exitosos que han marcado un hito en el establecimiento de métodos para investigar un determinado problema. Estos ejemplares son los que permiten detectar empíricamente un paradigma y los que dotan de contenido y fuerza explicativa a una teoría científica.

A este respecto **Ellis** (1994) subraya que, mientras que en la aproximación algorítmica los tests de *Cranfield* cumplen esa función, es extremadamente difícil señalar una prueba o un estudio particulares como modelo de la aproximación cognitiva. Ciertamente se observa una clara tendencia dentro de este enfoque a superar los primeros trabajos, de carácter más filosófico.

De esta manera se están desarrollando numerosos estudios empíricos —sobre todo los relacionados con la teoría del *sense making* (**Jacobson**, 1991), que ha diseñado un modelo constructivista— que están contribuyendo a validar las propuestas cognitivas ya formuladas. Sin embargo este punto de vista aún no cuenta con un modelo de evaluación de SRI ampliamente aceptado y debe generalizar la aplicación de estrategias que incorporen el uso de las nuevas medidas y métodos que proponen de forma teórica.

Frohmann (1990) y **Hjørland** (1991), por su parte, señalan que las aproximaciones cognitivas asumen que el estudio del pensamiento o del comportamiento de los usuarios puede ayudar a descubrir leyes, principios o regularidades ocultas aplicables al diseño de SRI. Ambos opinan que el objetivo no debe ser descubrirlos, sino construirlos como instrumentos para optimizar ciertas prácticas sociales de naturaleza informativa.

Hacia la integración

Algunos autores perciben el enfoque cognitivo como integrador (**Fernández**, 1994; **Frohmann**, 1992; **García**, 1994), principalmente desde el punto de vista teórico ya que, según mantienen, abarca el proceso de comunicación completamente. Sin embargo, como se ha visto, no todos comparten esta opinión. **Harter** y **Hert** (1997) creen que las aproximaciones algorítmicas y las orientadas al usuario son complementarias por naturaleza para la descripción tanto de las bases de datos automatizadas como de los contextos de búsqueda y de uso de la información recuperada, por lo que su estudio intenta contribuir a minimizar la cesura entre los dos acercamientos al problema siempre vigente de la RI en bases de datos textuales (y su evaluación).

Aunque todo método es hijo del paradigma desde el que se construye, muchos se usan con independencia del sustento teórico original. Por ejemplo, en el caso de los protocolos de pensamiento en voz alta, una técnica de investigación psico-lingüística para estudiar la solución de problemas por humanos se ha aplicado al análisis de los procesos mentales de la traducción y la interpretación, a aquellos sobre la facilidad de uso de programas informáticos y a la misma RI.

Por otro lado, una teoría no debe permitir realizar generalizaciones abusivas. Éste es el caso de adoptar una aproximación estrictamente cognitiva, puesto

que usa abstracciones demasiado generales como para ser intersubjetivamente válidas. De ahí que algunos de los más renombrados estudios lleguen a conclusiones contradictorias.

La cuestión de las abstracciones no es vana. La aproximación por dominios — *domain-approach*, que se ha querido proponer como una tercera vía o tendencia— afirma que el horizonte más fructífero para las ciencias de la información es estudiar dominios de conocimiento como comunidades de pensamiento o de discurso, que son parte integrante de la división del trabajo en la sociedad. (Hjørland y Albrechtsen, 1995). A pesar de sus enormes aciertos —sobre todo en sus críticas a modelos anteriores— la propuesta no hace sino desplazar el problema de la abstracción del usuario al cuerpo social. En el primer caso siempre se puede llegar al individuo; en el segundo, jamás se alcanzará lo que, por definición, no existe más que en las mentes de quienes lo construyen y aceptan.

«El término usuario es bastante ambiguo. De forma genérica pueden establecerse varios tipos: el potencial, el previsto y el beneficiario»

Si bien desde una perspectiva pesimista puede decirse que la documentación sufre de un auténtico caos paradigmático, también es cierto que la complementariedad de los distintos paradigmas es evidente (García, 1995). Una nueva propuesta metodológica válida debería contar con un sustento teórico sincrético, pero no ecléctico. Es decir, una integración de elementos anteriores de un modo riguroso y coherente.

En principio, un modelo adecuado para la evaluación de la RI, como cualquier otro cuerpo riguroso de conocimientos, ha de cumplir con el *principium rationis sufficientis* de Leibnitz. En otras palabras, debe ser lo suficientemente poderoso como para aprehender todos los hechos relevantes del fenómeno estudiado. La premisa, pues, de todo estudio evaluativo ha de tener en cuenta la combinación de los tres elementos fundamentales que intervienen en la recuperación de información: el sistema, el usuario y la información. El primero debe estudiarse para verificar su idoneidad respecto a un usuario (o un grupo) y un tema (o conjunto de ellos) o tipos de información.

El comportamiento de los individuos, posiblemente clasificable en subtipos más o menos afinados y estables, es importante porque determina la funcionalidad del modelo. Finalmente, la información, su calidad y su misma naturaleza son relevantes porque no todos los temas permiten un tratamiento similar y su exclusión puede distorsionar los resultados de la evaluación. Estos aspectos se habían estudiado ya anteriormente y lo novedoso consistiría, precisamente, en el intento de integración coherente y rigurosa de la aproximación clásica en el análisis del sistema y la cognitiva en el estudio del usuario en cuanto a su comportamiento y estructuras cognitivas, así como del renovador *domain approach* de Hjørland y Albrechtsen (1995) para el análisis de la información en sí.

Para finalizar sólo queda mencionar algunos de los diferentes temas de interés en torno a la evaluación de la recuperación de información que se vienen observando (Harter y Hert, 1997) y que constituyen líneas de investigación actualmente en desarrollo. Éstos se relacionan con: el uso de métodos y dimensiones múltiples, la perspectiva de los mantenedores del sistema, los modelos interactivos, los diferentes niveles y estratos del contexto en el que se encuentran, la dimensión temporal más o menos prolongada del estudio o la denominada evaluación formativa —realizada durante el desarrollo del proceso o producto— frente a la sumativa, aplicada a un sistema o proyecto finalizado.

Agradecimientos

La autora desea dar las gracias a los referees de la revista por su interés y sus útiles sugerencias, las cuales han contribuido a la mejora de este trabajo.

Bibliografía

- Allen, B. "Cognitive research in information science: implications for design". En: *Annual review of information science and technology*, 1991, v. 26, pp. 3-37.
- Beaulieu, M.; Robertson, S.; Rasmussen, E. M. "Evaluating interactive systems in *Trec*". En: *Journal of the American Society for Information Science*, 1996, v. 47, n. 1, pp. 85-94.
- Belkin, N. J. "Anomalous states of knowledge as a basis for information retrieval". En: *Canadian journal of information science*, 1980, v. 5, pp. 133-143.
- Belkin, N. J. "Ineffable concepts in information retrieval". En: Sparck Jones, K. (ed.) *Information retrieval experiment*. London: Butterworths, 1981, pp. 44-58.
- Belkin, N. J. "The cognitive viewpoint in information science". En: *Journal of information science*, 1990, v. 16, n. 1, pp. 11-16.
- Belkin, N. J., [et al]. "The effect of multiple query representations on information retrieval system performance". En: *Sixteenth annual international*

ACM-Sigir conference, 1993, pp. 339-346.

Blair, D. C. "Searching biases in large interactive document retrieval systems". En: *Journal of the American Society for Information Science*, 1980, v. 31, n. 4, pp. 271-277.

Blair, D. C.; Maron, M. E. "An evaluation of retrieval effectiveness for a full-text document-retrieval system". En: *Communications of the ACM*, 1985, v. 28, n. 3, pp. 281-299.

Blair, D. C.; Maron, M. E. "Full text information retrieval: further analysis and clarification". En: *Information processing and management*, 1990, v. 26, n. 3, pp. 437-447.

Blair, D. C. "Stairs redux: thoughts on the Stairs evaluation, ten years after". En: *Journal of the American Society for Information Science*, 1996, v. 47, n. 1, pp. 4-22.

Chen, H.; Dhar, V. "Cognitive process as a basis for intelligent retrieval systems design". En: *Information processing and management*, 1991, v. 27, n. 5, pp. 405-532.

Cleverdon, C. "The Cranfield tests on index languages devices". En: **Sparck Jones, K.; Willett, P.** (ed.) *Readings in information retrieval*. San Francisco: Morgan Kaufmann, 1997. pp.47-59. Isbn 1558604545.

Codina, L. "Teoría de la recuperación de información: modelos fundamentales y aplicaciones a la gestión documental". En: *Information World en Español*, 1995, n. 38, pp. 18-22.

Dervin, B. "Communication gaps and inequities: moving toward a reconceptualization". En: **Dervin, B.; Voigt, M.** (ed. lit.). *Progress in communication sciences*, v. 2. Norwood: Ablex, 1980, pp. 73-112, Isbn 0893910600.

Dervin, B.; Nilan, M. "Information needs and uses". En: *Annual review of information science and technology*, 1986, v. 21, pp. 5-33.

Ellis, D. "Paradigms in information retrieval research". En: **Kent, A.** (ed.) *Encyclopedia of library and information science*, v. 54. New York: Marcel Dekker, 1994, pp. 275-291. Isbn 0824720547.

Ellis, D. "The dilemma of measurement in information retrieval research". En: *Journal of the American Society for Information Science*, 1996, v. 47, n. 1, pp. 23-36.

Fernández, J. C. "Enfoques objetivo y subjetivo del concepto de información". En: *Revista española de documentación científica*, 1994, v. 17, n. 3, pp. 320-331.

Ford, N. "Relating information needs to learner characteristics in higher education". En: *Journal of documentation*, 1980, v. 36, n. 2, pp. 99-114.

Frohman, B. "Rules of indexing: a critique of mentalism in information retrieval theory". En: *Journal of documentation*, 1990, v. 46, n. 2, pp. 81-101.

Frohman, B. "The power of images: a discourse analysis of the cognitive viewpoint". En: *Journal of documentation*, 1992, v. 48, n. 4, pp. 365-386.

García, F. J. "Paradigmas científicos en representación y recuperación de la información". En: *Organización del conocimiento en sistemas de información y documentación*, 1995, pp. 99-112.

Harman, D. K. "A special conference report: the first *text retrieval conference (Trec-1)*, Rockville, Md., USA, 4-6 november, 1992". En: *Information processing and management*, 1993, v. 29, n. 4, pp. 411-414.

Harman, D. K. "Overview of the second *text retrieval conference (Trec-2)*". En: *Information processing and management*, 1995, v. 31, n. 3, pp. 271-289.

Harman, D. K. "The *Trec* conferences". En: **Sparck Jones, K.; Willett, P.** (ed.) *Readings in information retrieval*. San Francisco: Morgan Kaufmann, 1997, pp. 247-256. Isbn 1558604545.

Harman, D. K. "*Text retrieval conferences (Trecs)*: providing a test-bed for information retrieval systems". En: *Bulletin of the American Society for Information Science*, 1998, abril-mayo, pp. 11-13.

Harter, S. P. "The *Cranfield II* relevance assessments: a critical evaluation". En: *Library quarterly*, 1971, v. 41, n. 3, pp. 229-243.

Harter, S. P.; Hert, C. A. "Evaluation of information retrieval systems". En: **Willian, M. E.** (ed.) *Annual review of information science and technology*, 1997, v. 32, pp. 3-94.

Hjørland, B. "The cognitive paradigm in library and information science". En: *Biblioteksarbejde*, 1991, v. 33, pp. 5-37.

- Hjørland, B.; Albrechtsen, H.** "Toward a new horizon in information science: domain-analysis". En: *Journal of the American Society of Information Science*, 1995, v. 46, n. 6, pp. 400-425.
- Horne, E.** "Question generation and formulation: an indication of information need". En: *Journal of the American Society of Information Science*, 1983, v. 34, n. 1, pp. 4-15.
- Ingwersen, P.** *Information retrieval interaction*. London: Taylor Graham, 1992. Isbn 0947568549.
- Ingwersen, P.; Willett, P.** "An introduction to algorithmic and cognitive approaches for information retrieval". En: *Libri*, 1995, v. 45, n. 3-4, pp. 160-177.
- Ingwersen, P.** "Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory". En: *Journal of documentation*, 1996, v. 52, n. 1, pp. 3-50.
- Izquierdo, J. M.** "Estructuras conceptuales para la representación documental". En: *Organización del conocimiento en sistemas de información y documentación*, 1995. pp 27-49.
- Kuhn, T. S.** *La estructura de las revoluciones científicas*. 1a. ed., 14ª. reimpr. México D. F.: Fondo de cultura económica, 1990. Isbn 843750046X.
- Lancaster, F. W.** "Medlars: report on the evaluation of its operating efficiency". En: **Sparck Jones, K.; Willett, P.** (ed.) *Readings in information retrieval*. San Francisco: Morgan Kaufmann, 1997, pp. 223-246. Isbn 1558604545.
- Lesk, M.** *Seven ages of information retrieval*. Consultado 27 junio 99. <http://community.bellcore.com/lesk>
- Marchionini, G.** "Information seeking strategies of novices using a full-text electronic encyclopedia". En: *Journal of the American Society for Information Science*, 1989, v. 43, n. 2, pp. 156-163.
- Meadow, C. T.** *Text information retrieval systems*. San Diego: Academic press, 1992. Isbn 012487410X.
- Olvera, M. D.** *Evaluación de la recuperación de información en internet: un modelo experimental*. Tesis doctoral. Universidad de Granada, marzo 1998.
- Pao, M. L.** *Concepts of information retrieval*. Englewood: Libraries, 1989. Isbn 0872874052.
- Salton, G.; McGill, J.** *Introduction to modern information retrieval*. New York: McGraw-Hill, 1983. Isbn 0070544840.
- Salton, G.** "Another look at automatic text-retrieval". En: *Communications of the ACM*, 1986, v. 29, n. 7, pp. 648-656.
- Salton, G.** "The state of retrieval system evaluation". En: *Information processing and management*, 1992, v. 28, n. 4, pp. 441-449.
- Sparck Jones, K.** "Reflections on *Trec*". En: *Information processing and management*, 1995, v. 31, n. 3, pp. 291-314.
- Swanson, D. R.** "The evidence underlying the *Cranfield* results". En: *Library quarterly*, 1965, v. 35, pp.1-20.
- Swanson, D. R.** "Some unexplained aspects of the *Cranfield* tests of indexing performance factors". En: *Library quarterly*, 1971, v. 41, n. 3, pp. 223-228.
- Swanson, D. R.** "Historical note: information retrieval and the future of an illusion". En: *Journal of the American Society for Information Science*, 1988, v. 39, pp. 92-98.
- Tague-Sutcliffe, J. M.** "The pragmatics of information retrieval experimentation, revisited". En: *Information processing and management*, 1992, v. 28, n. 4, pp. 467-490.
- Tague-Sutcliffe, J. M.** "Some perspectives on the evaluation of information retrieval systems". En: *Journal of the American Society for Information Science*, 1996, v. 47, n. 1, pp. 1-3.
- Tramullas, J.** "Los sistemas de información: una reflexión sobre información, sistema y documentación". En: *Revista general de información y documentación*, 1997, v. 7, n. 1, pp. 207-229.
- Tramullas, J.** "Una propuesta de concepto y definición para la disciplina "documentación automatizada". En: *Revista general de información y documentación*, 1998, v. 8, n. 1, pp. 263-282.
- Van Rijsbergen, C. J.** *Information retrieval*. 2nd ed. London: Butterworths, 1979. Isbn 0408709294.

Wilson, T. D. "On the user studies and information needs". En: *Journal of documentation*, 1981, v. 37, n. 1, pp. 3-15.

M^a Dolores Olvera Lobo. *Facultad de Documentación, Universidad de Granada.*

molvera ARROBA platon.ugr.es

Enlace del artículo:

http://www.elprofesionaldelainformacion.com/contenidos/1999/noviembre/evaluacion_de_sistemas_de_recuperacion_de_informacion_aproximaciones_y_nuevas_tenc