

La investigación en Recuperación de Información: Revisión de tendencias actuales y críticas

José Antonio Salvador Oliván,

Rosario Arquero Avilés

Resumen

La investigación en recuperación de información ha estado dominada por dos paradigmas, uno centrado en el diseño de algoritmos y sistemas de recuperación de información, y otro centrado en todo lo que rodea al usuario en la búsqueda de información. En este artículo se revisan brevemente las características y tendencias actuales de investigación en ambos paradigmas, así como las críticas que han recibido, derivadas de la diferente metodología y de la dificultad para definir la relevancia y medir los cambios que se producen en el estado mental del usuario.

Palabras clave

Recuperación de información; relevancia; evaluación.

Abstract

Two paradigms have marked research in Information Retrieval. The first paradigm focused in algorithms design and Information Retrieval Systems, and, the second, focused in all factors that have and effect on user in information search process.

This article revised characteristics and current trends in research in both paradigms, analyzing criticisms than these paradigms have received related to different methodology and difficulty to define relevance and to measure changes derived from mental state of user.

Keywords

Information Retrieval; Relevance; Evaluation.

INTRODUCCIÓN

El papel que la investigación tiene en el desarrollo de una ciencia es tan evidente e importante que podemos afirmar que «es a su vez causa y efecto del desarrollo de una disciplina, ya que para que la investigación tenga lugar es preciso un cierto desarrollo de la misma»¹. Naturalmente, la recuperación de información (RI) no es una excepción, y su desarrollo progresivo como disciplina ha venido marcada, para lo bueno y para lo malo, por los resultados de la investigación realizada.

La recuperación de información es un campo donde se realiza una actividad importante de práctica profesional y, también, de investigación científica. Por un lado, dio lugar a la aparición de la industria de la información, con sus bases de datos y sistemas de información, con un

marcado componente práctico y profesional centrado en el uso de las bases de datos para satisfacer las necesidades de información de los usuarios. Por otro lado, la investigación científica ha estado dirigida hacia el diseño de sistemas de recuperación de información (SRI) más eficaces, dando lugar a diversas teorías, modelos y experimentos en los que la evaluación ha ocupado un papel central. A lo largo de la historia no ha existido una relación muy estrecha entre ambos componentes, científico y profesional, al menos hasta en los últimos años en los que parece haberse producido un cierto acercamiento y una implementación de las teorías y modelos experimentales en las aplicaciones comerciales.

La investigación en recuperación de información ha favorecido sobre todo la estimulación de ideas; estas ideas, conforme se han ido explorando y comprobando en sistemas experimentales, se han convertido en teorías o modelos. Por lo tanto, estas ideas, teorías y modelos desarrollados han estado influenciados en gran parte por la comprensión y el conocimiento empírico obtenido a partir de experimentos en los que la evaluación ha estado omnipresente. Robertson reconoce que no hay una teoría general y global de recuperación de información, y puede que nunca la haya, aunque el descubrimiento de nuevos modelos y teorías contribuirán a una mejor comprensión².

Son muchos los experimentos, tests e investigaciones llevadas a cabo en el campo de la recuperación de información, y también muchas las críticas y discusiones sobre los resultados obtenidos derivadas, en su mayoría, por la diferente metodología utilizada, en parte debida a la gran variedad de problemas investigados y, en parte, debida también a la diversa procedencia académica de los investigadores. Ole Pors señala que esto último es a la vez una fortaleza y una debilidad³. Una fortaleza, porque supone un alto grado de importación de enfoques y teorías interesantes de otras disciplinas académicas. Una debilidad, porque hace muy difícil el desarrollo científico en el sentido de acumulación de teorías y conocimiento. El autor destaca la carencia de un lenguaje científico común que defina bien los conceptos fundamentales, algo que caracteriza a una disciplina científica madura, siendo uno de los principales problemas que caracterizan los experimentos en recuperación de la información.

CARACTERÍSTICAS GENERALES DE LA INVESTIGACIÓN EN RI

Contemplaremos la investigación en recuperación de la información desde una perspectiva global, teniendo en cuenta tres características generales que constituyen la esencia de su evolución y existencia:

1. La recuperación de información es un campo interdisciplinar. A pesar de que son muchas las disciplinas directamente relacionadas (Informática, Ciencia de la Información, Documentación, Lógica, Lingüística, Psicología, Inteligencia Artificial, etc.), todas comparten el mismo objetivo: facilitar la búsqueda y obtención de información relevante que satisfaga las necesidades de información de los usuarios. Es por ello deseable que la investigación en este campo sea realizada por grupos interdisciplinarios, o como mínimo, que la comunicación y la colaboración entre los diferentes equipos de investigación sea fluida, algo que no siempre ha ocurrido y ha sido motivo de discusión y controversia entre los propios investigadores.

2. La investigación en recuperación de la información ha estado muy influida por la evolución y los avances producidos en las tecnologías de la información que, necesariamente, han introducido nuevos enfoques, procedimientos y métodos en la organización, almacenamiento y acceso a la información, así como en el uso de nuevos sistemas de recuperación de información.

3. La recuperación de la información ha desempeñado un papel esencial en la evolución de la Sociedad de la Información. Hay que recordar que la recuperación de información surge para buscar soluciones y dar una respuesta al problema de la explosión de información científica; en la actualidad, la World Wide Web, como medio de acceso a la información más utilizado, juntamente con la facilidad para poder publicar en él ha provocado que uno de los principales problemas a los que se enfrenta cualquier persona es cómo localizar información pertinente ante el exceso de información existente. Por lo tanto, el problema en sí es el mismo que existía hace más de 50 años y la investigación debe ir dirigida a dar respuesta a este problema, aunque con los cambios lógicos que supone el uso de este medio, como por ejemplo, (a) que el acceso a la información se ha generalizado y es el propio usuario quien busca la información, por lo que han de diseñarse interfaces de búsqueda más amigables; (b) y que el concepto de información se amplía y ya no sólo se trata de documentos o texto sino también de imágenes y sonido, por lo que debe de investigarse cómo recuperar estas formas de información (recuperación multimedia).

La investigación en recuperación de información va a permitir el desarrollo de nuevas herramientas que faciliten un acceso más eficiente a la información, en cualquier formato en que esté, y de ello se aprovecharán tanto científicos, como profesionales de la documentación y cualquier usuario que necesiten información para sus investigaciones, su trabajo, la toma de decisiones, etc., y ello, indiscutiblemente, repercute en que la Sociedad esté más y mejor informada.

VARIABLES IMPLICADAS EN LOS EXPERIMENTOS EN RI

La investigación en recuperación de la información ha perseguido siempre mejorar el proceso de recuperación de información, y esto sólo es posible a través de estudios de evaluación que midan el grado con que los SRI satisfacen las necesidades de información de los usuarios y de estudios que determinen el efecto de los factores que influyen en el rendimiento del sistema. Estos tests pueden ser realizados en condiciones de laboratorio (experimentales) o en condiciones reales, pero en cualquier caso es necesario efectuar un control sobre las diferentes variables que pueden influir en la variabilidad de los resultados si se desean comparar con otros sistemas o inferir conclusiones válidas.

Independientemente del tipo de test, del grado de control a efectuar, de la metodología a utilizar, y de cualquier otro aspecto relacionado con el propio diseño de un estudio experimental y científico, las variables que se han investigado y se siguen investigando en el proceso de recuperación de información se pueden clasificar en los siguientes grupos⁴:

1. *Base de datos o colección de documentos*: Las variables de esta categoría están relacionadas con el tamaño, cobertura temática, forma, medio y uso de la base de datos.

2. *Representación de la información*: Se refiere a la estructura física y lógica de la información almacenada, y que puede representarse a través de un modelo. Para describir la estructura lógica de la información se han utilizado diferentes modelos, de los que caben destacar el booleano, vectorial y probabilístico. Los modelos utilizados para describir la estructura física incluyen árboles binarios, hashing, multilistas, etc.

También se incluyen en este apartado diferentes variables relacionadas con el proceso de indización: exhaustividad, especificidad, grado de control, exactitud, consistencia inter-indizador, grado de pre-coordinación, etc.

3. *Usuarios*: Ha sido uno de los principales focos de investigación a partir de los 80, especialmente por autores como Dervin, Belkin y Wilson. Las variables relacionadas con esta categoría son el tipo de usuarios, contexto en el que se va a utilizar la información, tipos de necesidades de información, urgencia de la necesidad de información, etc. También se han medido características personales como habilidades lógicas, modo de aprendizaje, experiencia en la búsqueda y nivel educativo.
4. *Consultas y sentencias de búsqueda*: La consulta es la expresión, escrita u oral, de la necesidad de información del usuario, y la sentencia de búsqueda es la traducción de la consulta al lenguaje del sistema. Las consultas pueden ser reales (representando las necesidades de información de usuarios) o artificiales (extraídas de los títulos u otras partes del documento), siendo éstas últimas las más utilizadas en los experimentos de laboratorio pero también uno de los aspectos más criticados precisamente por no representar la búsqueda de información en situaciones reales.

Las sentencias de búsqueda están formadas por términos del lenguaje natural y controlado, truncado, códigos de clasificación, operadores, y su estructura depende del modelo en que se basa el SRI; así, se pueden utilizar operadores booleanos y de proximidad, pesos que indiquen la importancia de cada término, o simplemente el lenguaje natural. Las sentencias de búsquedas se han medido en relación a su tamaño, exhaustividad y especificidad.

5. *Proceso de búsqueda*: Las búsquedas pueden variar en muchas maneras, y esta variación puede relacionarse con la interacción entre las personas (usuario e intermediario) y con la interacción entre las personas y los SRI. Estas fuentes de variación se pueden recoger, a un macro-nivel, en tres variables: (a) búsqueda delegada a un intermediario frente a búsqueda por el usuario final; (b) técnicas de búsqueda utilizadas; y (c) modos de acceso del buscador, sea a través de comandos, menús o por navegación mediante enlaces hipertexto.

A nivel individual, el proceso de búsqueda puede describirse por los tipos de interacción entre el usuario y el buscador cuando la búsqueda es delegada o por las actividades preliminares cuando no es delegada, y por los tipos de comandos y estrategias utilizadas durante el proceso de búsqueda.

6. *Evaluación de la recuperación*: Las medidas tradicionales para evaluar el rendimiento en la recuperación son la llamada y la precisión, que a pesar de sus numerosas críticas, todavía se siguen utilizando en la actualidad. Con el paso del tiempo, se han ido incorporando otras medidas como la eficiencia global del sistema, la llamada normalizada, pero también análisis coste-eficacia y coste-beneficio, razón de novedad, y otras medidas no cuantitativas que reflejan el grado de satisfacción del usuario.

En parecidos términos se expresa Niels Ole Pors al indicar que, como mínimo, los siguientes elementos pueden formar parte del proceso y cuya interacción forma el centro de la investigación en recuperación de información³: (1) la base de datos; (2) el usuario; (3) el lenguaje utilizado en la representación; (4) el SRI algorítmico; (5) el intermediario/interfaz; (6) el lenguaje del usuario; y (7) la necesidad de información/petición/consulta.

El nivel de detalle con que se puede analizar cada variable sería largo y extenso, pero sí podemos decir que los temas expuestos representan las principales variables que se han investigado en los experimentos realizados en recuperación de la información desde la perspectiva de la Documentación y de la Ciencia de la Información.

EVALUACIÓN Y RELEVANCIA: ASPECTOS ESENCIALES EN LA INVESTIGACIÓN

Mención especial merece la evaluación en la investigación en recuperación de información. Desde el inicio hasta la actualidad, la evaluación ha sido un tema central y omnipresente en la investigación y desarrollo de este campo, algo que no es de extrañar cuando el objetivo que se persigue es mejorar el rendimiento de los SRI o del propio proceso de recuperación.

La evaluación siempre ha estado dirigida a demostrar qué sistemas, modelos y técnicas ofrecen un mejor rendimiento y una mayor eficiencia en la recuperación de información, midiéndose a través de la relevancia de los resultados. Y es aquí donde se presenta uno de los principales problemas, ya que para medir una variable primero hay que definirla.

La relevancia ha sido uno de los más temas más importantes^[a], estudiados y que mayor volumen de literatura científica ha generado relacionada con la recuperación de información, pero también uno de los que mayor discusión y debate ha provocado, debido fundamentalmente a la falta de acuerdo sobre la naturaleza exacta y significado de la relevancia y a cómo se podría operativizar en el diseño y evaluación de los sistemas de información. En este sentido, Froehlich destaca una serie de cuestiones comunes en los artículos sobre la relevancia⁵: 1) la incapacidad para definir la relevancia; 2) lo inadecuado que resulta utilizar el tema de los documentos como base para valorar la relevancia; 3) la diversidad de criterios en torno al usuario y no relacionados con el tema que afectan al juicio de relevancia; 4) el carácter fluido y dinámico del comportamiento en la búsqueda de la información; y 5) la necesidad de modelos más robustos y complejos en el diseño y evaluación de sistemas.

La discusión y la falta de acuerdo viene determinada, entre otras razones, porque no existe una única relevancia. Es el concepto y la naturaleza de la relevancia, los diferentes tipos que existen y los numerosos factores que influyen en los juicios de relevancia, muchos de ellos

subjetivos, lo que impide una definición formal y una de las razones, para Mizzaro⁶, de que no exista una teoría global e integradora en recuperación de la información.

Cuando Schamber comentó que el papel de la relevancia en la evaluación de los SRI era “relativamente pequeño pero importante”⁷, Harter justificaba el término *importante* porque los juicios de relevancia forman los cimientos sobre los que se ha construido el modelo de evaluación experimental tradicional, y *relativamente pequeño* porque en la literatura apenas se ha estudiado la validez de estos juicios de relevancia⁸.

Desde los primeros experimentos realizados en la recuperación de información hasta los realizados en la actualidad, la eficacia de los SRI se ha medido a través de la llamada y de la precisión, medidas que se basan en el juicio que alguien hace (usuario o no usuario) sobre la relevancia del documento (o representación de éste) recuperado. También desde el principio se puso de manifiesto el desacuerdo entre los investigadores a la hora de considerar qué documentos eran o no relevantes, algo que no es de extrañar cuando estudios posteriores han demostrado el gran número de variables y factores que pueden afectar al juicio sobre la relevancia. Así, Schamber, en una revisión de la literatura existente hasta entonces, identifica hasta 80 factores diferentes, comentando además que no es una lista completa⁷. De todas estas variables, algunas son de tipo subjetivo y psicológico, que dependen del estado cognitivo del usuario y de las circunstancias o situación del momento, lo que añade todavía mayor complejidad y variabilidad a los juicios de relevancia que se puedan realizar.

Todos estos factores que pueden influir en la valoración de la relevancia, se han reflejado a lo largo de la historia de la experimentación en recuperación de información en una gran variabilidad de criterios, instrumentos de medida y metodología utilizada y, en consecuencia, han provocado resultados diferentes y difícilmente comparables⁹. Esta indefinición y confusión que rodea a la relevancia, llevó desde un principio a una importante crítica sobre el uso de la relevancia como base para evaluar la recuperación; así por ejemplo, Rees y Saracevic planteaban serias dudas sobre si la relevancia era un criterio adecuado para medir, si era medible, y bajo qué circunstancias se podía utilizar¹⁰; dudas que no se han resuelto a lo largo de los años, tal y como pone de manifiesto Schamber, basándose en los estudios realizados sobre el tema, cuando dice que los científicos de la información han fracasado en alcanzar un consenso sobre cómo debería medirse la relevancia⁷.

David Ellis reconoce que el problema de la relevancia ha permanecido intratable desde los inicios en la investigación en recuperación de información, y que el problema no reside en cómo medirla sino en utilizar la relevancia como criterio de medida, de ahí que en los últimos años el interés de los investigadores se haya centrado en el uso de métodos cualitativos relacionados con los aspectos cognitivos, afectivos y del comportamiento que se producen en el proceso interactivo de recuperación de información¹¹. También Harter ha propugnado la necesidad de desarrollar instrumentos de medida que sean sensibles a las variaciones individuales que se producen en los juicios de relevancia, de manera que la evaluación refleje el mundo real de los usuarios reales¹².

A pesar de que los investigadores reconocen en la literatura los problemas que plantea la relevancia y de que se hayan propuesto otras medidas, sean cualitativas o cuantitativas, no

dejan de ser simplemente propuestas teóricas que no ofrecen soluciones prácticas a cómo medir las variaciones individuales cognitivas y que terminan siempre con la coletilla “es necesaria más investigación para implementar estas propuestas”. La solución al problema no parece fácil. El propio Harter reconoce que las soluciones no son obvias ni vendrán rápida y fácilmente¹². Si el objetivo de la recuperación de información es proporcionar información relevante que satisfaga la necesidad de información del usuario, la eficacia debería basarse en una medida capaz de medir los cambios que produce la información recuperada en el estado cognitivo del usuario, y como dice Ellis, no está claro cómo se puede hacer si es que se puede¹¹. Quizá por ello, a lo largo de la historia, no ha habido un sustituto de la relevancia como criterio de medida y se siga utilizando para evaluar la eficacia de los SRI en los estudios de evaluación.

PRINCIPALES PARADIGMAS EN LA INVESTIGACIÓN EN RI

Los estudios de informetría que analizan los autores co-citados y los trabajos co-citados resultan de gran interés ya que pueden representar de una manera válida la investigación realizada en una determinada ciencia. En la ciencia de la Información y Documentación destacan dos trabajos: el de Persson¹³, y el de White y McCain¹⁴. Ambos coinciden en que la recuperación de información es el área de investigación más importante y numerosa. Persson identifica dos subgrupos: (1) uno que trabaja en el diseño de algoritmos, denominado también la parte “dura”, y (2) el que se centra en la relación usuario-sistemas, denominado parte “blanda”. Por su parte, White y McCain identifican tres grupos: (1) recuperación experimental, que se centra en el diseño y evaluación de SRI, (2) recuperación práctica, orientado al estudio de sistemas y bases de datos reales, y (3) teoría de usuarios, que representan el lado cognitivo. Estos dos últimos grupos corresponderían al grupo “blando” de Persson.

Aunque los resultados de estos estudios pueden presentar un sesgo derivado de los autores y revistas seleccionadas (resulta sorprendente que en ninguno se seleccionan artículos de revistas como Online, Online & Cd-Rom Review y Database, utilizadas por los profesionales de la búsqueda y recuperación de información), hay una coincidencia generalizada en que la investigación en recuperación de información ha estado dominada por dos enfoques o paradigmas: el orientado a los sistemas y el centrado en los usuarios, o según Ellis¹⁵, enfoque arquetípico y enfoque cognitivo. Ambos enfoques investigan en el campo de la recuperación de información, pero desde perspectivas diferentes y están representados por organizaciones diferentes, e incluso los canales de comunicación (revistas, conferencias) también lo son. El principal punto de discusión entre ambos enfoques en la evaluación de SRI ha sido claramente resumido por Robertson y Hancock-Beaulieu: «El conflicto entre los experimentos de laboratorio y operativos es esencialmente un conflicto entre, por una parte, el control de las variables experimentales, la observabilidad y repetibilidad, y por otra, el realismo»¹⁶.

Enfoque orientado a los Sistemas

La investigación centrada en los sistemas tiene su origen a finales de los 50 con los tests de Cranfield, seguidos de los diferentes proyectos SMART y con la continuación, hoy en día, de los experimentos TREC (Text Retrieval Conferences). La caracterización de este enfoque se puede resumir en que se centra en el diseño de sistemas y algoritmos, la evaluación es un

componente importante del experimento, y no tiene en cuenta a los usuarios ni su interacción con los sistemas.

Los elementos más importantes que intervienen en estos experimentos son: (1) un conjunto de documentos o representaciones de éstos; (2) algoritmos y procesos que intervienen en la recuperación de información, tanto los relacionados con la indización de las consultas como de los documentos en la base de datos; (3) un conjunto de consultas artificiales creadas por intermediarios, modificadas de los usuarios o derivadas de partes de los documentos; y (4) un conjunto de documentos recuperados, evaluados generalmente por expertos en base a su relevancia y utilizando dos medidas, la llamada y la precisión.

Las principales contribuciones de este enfoque han sido el desarrollo de algoritmos de representación como los vectoriales y probabilísticos, algoritmos de feedback por relevancia y algoritmos de proceso de consultas, y aunque sus resultados, en un principio, estaban confinados a sistemas experimentales, en los últimos años se han ido incorporando a los motores de búsqueda utilizados en la World Wide Web.

Este tipo de investigación era realizada, principalmente, por grupos pequeños de investigadores en diferentes lugares del mundo, lo que significaba que las colecciones, incluidas las consultas, los documentos y los juicios de relevancia eran relativamente pequeños en cobertura y en número, ya que estos grupos de investigación no disponían de suficientes recursos para construir grandes colecciones. Estos problemas se solucionaron con la celebración a partir de 1992 de las TREC, un foro que reúne anualmente a investigadores de universidades, empresas y del gobierno que utilizan la misma y gran colección de documentos para comparar sus sistemas de recuperación de información. En las TREC se proponen cada año nuevos tipos de tareas, llamadas pistas (tracks), donde cada una trata un problema diferente de recuperación de información y el investigador (individual o en grupo) elige aquella en la que quiere participar. Constituyen, por consiguiente, uno de los puntos de referencia más importantes sobre la investigación actual en este campo. Para el año 2005 se han propuesto las siguientes pistas¹⁷:

1. *Enterprise track*: Es una nueva pista que tratará de las necesidades de búsqueda de información dentro de una organización. Es una continuación de la pista *web track*.
2. *Spam track*: Es la otra pista nueva de este año, donde se estudiarán los problemas del spam en los correos electrónicos.
3. *Genomic track*: El objetivo es estudiar la recuperación de datos de genoma, entendiendo por éstos no sólo las secuencias de los genes sino también toda la documentación de apoyo como artículos o informes.
4. *HARD Track*: El objetivo es conseguir una gran precisión en la recuperación de documentos utilizando información sobre el buscador y/o sobre el contexto de la búsqueda a través de la interacción con el buscador.
5. *Question answering track*: Es una pista que se acerca más a la recuperación de información que a la recuperación de documentos. Para cada conjunto de preguntas, los sistemas deben de proporcionar la parte de texto con la respuesta

exacta a la pregunta y un Identificador del documento. Los sistemas sólo pueden devolver una respuesta por pregunta.

6. *Robust Retrieval Track*: Incluye una tarea de recuperación tradicional *ad hoc*, pero con la atención puesta en la eficacia individual en lugar de global.
7. *Terabyte track*: El objetivo es investigar si se puede (y cómo) escalar la evaluación tradicional de una colección a una colección mayor de documentos que los utilizados actualmente en las TREC. Se utilizará para ello una colección de páginas web estáticas de aproximadamente un terabyte, probablemente de dominio .gov.

Otras pistas utilizadas en años anteriores y que resumen la investigación realizada en este foro son las siguientes:

Cross-language track: El objetivo es investigar la capacidad de los SRI para encontrar documentos sobre un tema independientemente del idioma en el que estén escritos. Se ha investigado en documentos escritos en árabe, lenguajes europeos y lenguajes asiáticos. Esta actividad ha pasado a evaluarse en otros foros, concretamente en CLEF (Cross Language Evaluation Forum, (<http://clef.isti.cnr.it/>) para los europeos, y en los seminarios NTCIR (National Institute of Informatics-NACSIS Test Collection for IR Systems, (<http://research.nii.ac.jp/ntcir/workshop/>) para los asiáticos.

8. *Filtering track*: Las necesidades de información son estables y se conocen algunos documentos relevantes, pero continuamente se introducen nuevos documentos. Para cada documento, el sistema debe de tomar una decisión binaria sobre si el documento debería ser recuperado o no.
9. *Interactive track*: En esta pista se estudia la interacción del usuario con los SRI, siguiendo un protocolo con usuarios reales que utilizan la misma colección de documentos y un conjunto de consultas.
10. *Novelty track*: Se investigan las capacidades de los sistemas para encontrar nueva información.
11. *Video track*: Investiga la recuperación de vídeo digital basada en el contenido.
12. *Web track*: Se caracteriza por una serie de tareas de búsqueda (búsqueda de un item conocido y la síntesis de un tema) sobre un conjunto de documentos de la Web.

Los resultados de las conferencias TREC a lo largo de estos años ha llevado a refinar algoritmos que producen mejoras en la llamada y precisión y que funcionan bien en grandes colecciones de documentos, pero también ha ampliado la investigación en la evaluación en recuperación de información al incluir recuperación en múltiples lenguajes, recuperación de documentos escritos e información multimedia, e investigación del feedback por relevancia.

Otro de los foros internacionales más importantes de investigación en este enfoque corresponde al Special Interest Group on Information Retrieval (SIGIR) de la Association for Computing Machinery (ACM) (<http://www.acm.org/sigir/>) que, desde el año 1973, intermitentemente, y desde el año 1985, todos los años, celebra unas Conferencias sobre Investigación y Desarrollo en Recuperación de la Información. En la Conferencia del año 2005, las principales áreas de interés son las siguientes¹⁸:

- Modelos formales, modelos de lenguaje, fusión y combinación.
- Aprendizaje por ordenador, minería de datos, clustering y categorización de texto.
- Recuperación multi-lenguaje, traducción automática.
- Detección y rastreo de temas, tareas de filtrado y agentes.
- Recuperación de información en la web, análisis de citas y enlaces, XML y metadatos, bibliotecas digitales y recuperación de información distribuida.
- Acceso a imágenes y vídeo, recuperación de audio y habla, recuperación de música.
- Representación e indización de texto, sumarización y extracción de texto, proceso en lenguaje natural.
- Rendimiento, comprensión, escalabilidad, arquitecturas, y aplicaciones a teléfonos móviles.
- Interfaces, visualización, recuperación interactiva, estudios y modelos de usuarios.
- Evaluación, diseño de colecciones de test, diseño experimental y métrica.
- Aplicaciones especializadas en recuperación de información, incluidas la genómica, de estructuras químicas e ingeniería de software.

Por último, destacar también los Special Interest Groups de la American Society for Information Science and Technology (ASIST)¹⁹ relacionados con la investigación en el diseño, desarrollo y evaluación de bases de datos y sistemas de recuperación de información, algunos de los cuales tratan dichos temas, si bien sorprende que haya desaparecido el único grupo dedicado exclusivamente a estas tareas (Special Interest Group on Computerized Retrieval Systems).

Problemas y críticas al enfoque orientado a los sistemas: Son muchos los problemas y las críticas que ha recibido por parte de la comunidad científica este enfoque. Harter y Hert las resumen muy bien en cuatro categorías²⁰:

- Validez y fiabilidad: Una de las principales críticas, sobre todo por los autores del enfoque cognitivo, ha sido que no tiene en cuenta al usuario en el proceso de recuperación de información. También que los experimentos se realizan en un laboratorio y no en el mundo real, donde existen diferentes comportamientos de búsqueda y una interacción entre los usuarios o buscadores y los sistemas. Además, los juicios de relevancia utilizados en este enfoque no tienen en cuenta su naturaleza subjetiva, dinámica y cambiante.
- Generalización: Las principales críticas son consecuencia de la falta de aleatorización en la selección de muestras, tanto de los usuarios como de las necesidades de información, por lo que las consultas utilizadas en los experimentos no son representativas. También se ha criticado que las colecciones utilizadas en los experimentos son pequeñas y que tratan principalmente de temas de ciencia y tecnología.

- Utilidad: Las críticas se refieren a que los resultados de los tests experimentales no se pueden aplicar a los SRI reales.
- Nivel conceptual: No existe una base teórica firme en la que se basen las medidas y métricas utilizadas en la evaluación de la recuperación de información, ni está claro que la llamada y la precisión sean significativas para los que buscan información, o que las diferencias estadísticas de estas medidas entre los SRI sean significativas en contextos y situaciones reales.

Las TREC tampoco se han librado de críticas. Blair manifiesta que los resultados de estas conferencias deberían verse con cierta precaución debido, fundamentalmente, al uso de la llamada relativa como medida para evaluar la eficacia (no se utilizan todos los documentos de la base de datos) y a que no son extrapolables a los SRI que se utilizan en un marco real, en los que la interactividad es un aspecto importante de la recuperación y no se tiene en cuenta²¹. Este autor es replicado por dos cartas al editor de la revista enviadas por Tefko Saracevic²² y Sparck Jones²³ en las que rebaten los argumentos de Blair; simplemente, la discusión está ahí.

Enfoque orientado a los Usuarios

Aunque se ha reconocido desde siempre que la consulta de información es una representación de una necesidad de información subyacente, y que esta necesidad surge de un estado mental y/o cognitivo del usuario, no es hasta finales de los 70 y especialmente en la década de los 80, cuando comienza a desarrollarse una línea clara de investigación sobre los aspectos cognitivos, interactivos y contextuales que ocurren en el proceso de recuperación de información, centrándose, no ya sólo en los sistemas, sino en los usuarios y cómo éstos usan e interactúan con los sistemas.

Las primeras razones para una aproximación cognitiva fueron expuestas por De Mey al declarar que «cualquier proceso de información, sea perceptual o simbólico, está mediado por un sistema de categorías de conceptos que, para un mecanismo de proceso de información, son un modelo del mundo»²⁴, un punto de vista que favorece el desarrollo de otros modelos donde se concede mayor importancia al usuario. Esta perspectiva destaca el papel del estado actual del conocimiento en el proceso de la información, e implica que si un mensaje no puede ser mediado por el estado de conocimiento de un receptor, no tiene lugar ningún proceso de información.

Saracevic resume el cambio paradigmático que se produce hacia el modelo cognitivo como una reorientación de la investigación centrada en los sistemas hacia temas como las estructuras del conocimiento, la interacción hombre-máquina, la búsqueda de información, y el comportamiento humano en el proceso de información²⁵. La razón es que una mejor comprensión de estas áreas contribuirá al diseño de mejores sistemas de información y al diseño de elementos fundamentales orientados al usuario, como los interfaces inteligentes.

La importancia del enfoque centrado en el usuario queda reflejada en el siguiente párrafo: «Este nuevo enfoque reconoce que un elemento fundamental en la recuperación de información es el desarrollo de una necesidad de información, consecuencia de un estado de conocimiento inadecuado.....Para que la recuperación de información tenga éxito, esa necesidad de información debe representarse de manera adecuada, y los demás elementos

del sistema (representación de documentos, mecanismo de recuperación) se deben de representar o construir en base a esa representación»²⁶.

Considerar a los individuos como punto de partida en la investigación, implica una concentración sobre el significado, semántica y aspectos pragmáticos de la información, existiendo dos áreas principales de interés. La primera trata de estudiar el significado que los usuarios dan a la información en el contexto de búsqueda y uso de la información. La segunda trata de las implicaciones en el diseño de los sistemas de información.

Este tipo de estudios ha evolucionado hacia la investigación del comportamiento humano en la búsqueda, exploración, filtrado, uso y difusión de la información. Se centran en los usuarios reales y suelen utilizar medidas y métodos más cualitativos que el enfoque de sistemas. El principal propósito de la evaluación en este enfoque es determinar lo bien que interactúan el usuario, el mecanismo de recuperación y la base de datos, para extraer información en situaciones reales²⁷. Los juicios de relevancia, a diferencia del enfoque de sistemas, tienen que ser realizados por el usuario y basarse en su necesidad de información, que, lógicamente, puede cambiar a lo largo de la sesión de búsqueda.

Sin embargo, no ha sido hasta hace muy pocos años cuando este tipo de estudios han empezado a tener cierto impacto en el diseño de sistemas. Las razones pueden atribuirse a: (1) la explosión de la información electrónica que ha posibilitado un mayor acceso a la información y un mayor uso de los SRI; (2) la entrada en escena de un nuevo campo como es la gestión del conocimiento, más amplio que la gestión de la información, que requiere no sólo una comprensión de los datos sino también cómo se produce la comunicación y uso de estos datos²⁸.

Los foros más importantes de comunicación de la investigación llevada a cabo en esta área son las conferencias internacionales sobre Information Seeking in Context (ISIC), bianuales a partir de 1996, y que tratan siempre temas relacionados con las necesidades de información, búsqueda y uso de la información en factores contextuales. A destacar el grupo SIG-USE (Special Interest Group-Information Needs, Seeking and Use) de ASIST (American Society for Information Science & Technology), cuyas actividades están encaminadas al reconocimiento de las necesidades de información de los usuarios, la búsqueda de información para satisfacer estas necesidades, explorar las fuentes de información, recuperar información de las fuentes disponibles y colaborar con todos aquellos que se interesan por la comunicación entre las personas y la información. Otros grupos de investigación de ASIST relacionados con la recuperación de información son el SIG-HCI (Human Computer Interaction) y SIG-DL (Digital Libraries).

También en las TREC existe un acercamiento al enfoque cognitivo, como se deduce de la pista interactiva que trata de la interacción de los usuarios con los SRI, así como desde la ACM con los grupos SIG-CHI (Computer-Human Interaction) y SIG-IR (Information Retrieval), cuya conferencia anual incluye siempre un apartado sobre modelos y estudios de usuarios.

Los estudios de usuarios, en el campo de la recuperación de información, han contribuido principalmente a identificar el comportamiento humano en la búsqueda de información,

cubriendo el vacío entre las necesidades de información de los individuos y los SRI, y a influir en el diseño de nuevos tipos de SRI con interfaces gráficas más amigables y donde la interacción con los sistemas resulte de utilidad al usuario para encontrar información relevante (mediante técnicas de visualización, interacción con lenguaje natural, etc.). Además, los estudios de usuarios han ayudado a descubrir la naturaleza dinámica, cognitiva y situacional de la relevancia y, por consiguiente, a reconsiderar los criterios de evaluación utilizados en la recuperación de información.

Problemas y críticas al enfoque cognitivo: Tampoco este tipo de estudios está exento de críticas, especialmente las relacionadas con la generalización, utilidad y conceptualización.

- **Generalización:** Debido a que los estudios se centran en una población específica de usuarios y a que el número de usuarios seleccionado es pequeño, no se pueden extrapolar los resultados a la población en general. Para poder generalizar o inferir los resultados a la población, los estudios deberían replicarse en diferentes marcos y poblaciones, algo que no es fácil de hacer debido a que este tipo de estudios suelen llevar meses e incluso años en completarse.
- **Utilidad:** No es fácil ver la relación entre los resultados de los estudios de usuarios y el diseño y evaluación de los SRI. Ya que los investigadores de este enfoque son expertos en ciencias sociales y del comportamiento, y no necesariamente en tecnología y diseño de sistemas, la cuestión es cómo elaborar y traducir sus resultados para que otros puedan utilizarlos en la creación de nuevos sistemas. Sin embargo, es un hecho reconocido la falta de colaboración entre investigadores del enfoque cognitivo y del enfoque de sistemas.
- **Nivel conceptual:** No es fácil comparar o sintetizar resultados de estudios de usuarios debido, sobre todo, a que los métodos de análisis y los constructos conceptuales utilizados son diferentes. Actualmente se está trabajando en el desarrollo de teorías del comportamiento humano en la información que abarquen varios contextos y situaciones, y así poder identificar nuevas medidas y métodos de evaluación; sin embargo, los resultados de este tipo de investigaciones no aportan luz al tema y habrá que esperar para ver si ofrecen resultados positivos.

Una de las principales dificultades a la que se enfrenta la investigación en este enfoque es cómo describir los cambios que se producen en el estado mental o estructuras cognitivas del buscador durante la interacción con el sistema y con la lectura de los textos.

HACIA LA INTEGRACIÓN DE LOS DIFERENTES PARADIGMAS

Es algo conocido por todos que, en la sociedad actual, la información se ha convertido en un recurso esencial, pero también que la gran abundancia de información y su representación en diferentes formatos pueden hacer que las tareas de recuperación sean más complejas²⁹. Es necesaria una colaboración interdisciplinar entre los diferentes enfoques para demostrar su utilidad y eficacia. Saracevic lo resume muy bien cuando dice que hay que aprender una lección básica a partir de las experiencias acumuladas a lo largo de los años: «El éxito o fracaso de cualquier sistema interactivo y tecnología, depende del grado con que se atienden también

a los factores humanos y usuarios, tratados de la manera adecuada desde la teoría, conceptualización y diseño para el desarrollo, evaluación y provisión de servicios»³⁰.

Wu y Sonnenwald³¹ proponen un marco interesante para la evaluación de los SRI que sintetiza diferentes enfoques, utilizando para ello los atributos que Rogers³² identificó y que influyen en la adopción de innovaciones. A pesar de que no refleja todas las medidas, criterios y técnicas que se pueden utilizar, y de que finalizan su artículo con la típica frase *“further work is necessary to demonstrate the utility of framework”*, resulta de gran utilidad para sintetizar las direcciones en la investigación en evaluación de los nuevos SRI, antes de que éstos sean implementados y utilizados por los usuarios en general:

- **Ventaja relativa:** Representa el grado con el que un nuevo sistema sustituye al que se venía utilizando hasta ese momento. Se pueden utilizar criterios como relevancia del sistema, relevancia temática, velocidad o beneficio económico, medidos a través de la llamada y precisión, tipo y cobertura de las fuentes, tiempo de respuesta del sistema y análisis coste-beneficio.
- **Compatibilidad:** Se refiere al grado de consistencia de un nuevo sistema con los valores, experiencias y futuras necesidades de los usuarios. Este atributo se puede operativizar a través de tres criterios: relevancia motivacional, relevancia para la organización y relevancia social, que se pueden medir por lo bien que el sistema satisface las expectativas de los usuarios, de las organizaciones y de la sociedad.
- **Complejidad:** La dificultad de aprender a utilizar y comprender un nuevo sistema o tecnología. Se pueden usar criterios de usabilidad, relevancia cognitiva y situacional, medidas a través del tiempo que se tarda en completar una tarea, tasa de error y satisfacción del usuario en su contexto.
- **Capacidad de prueba:** Se refiere a la facilidad de experimentar con un nuevo sistema e incluye el nivel de esfuerzo necesario y el riesgo que supone participar en demostraciones del sistema a pequeña escala, así como los costes para dejar de utilizar el sistema. Este criterio se puede medir con variables como disponibilidad, tiempo de formación y otros costes iniciales.
- **Perceptibilidad:** Es el grado con el que los resultados del nuevo sistema son perceptibles, y puede medirse por el coste de esta percepción.

Estamos convencidos de que es necesaria una integración y colaboración interdisciplinar de los diferentes enfoques en la investigación en recuperación de información; sólo teniendo en cuenta las características y comportamiento de los usuarios es posible diseñar SRI más eficientes. Son muchas las líneas de investigación actuales que pueden atraer a ambos enfoques y establecer una colaboración multi e interdisciplinar, tales como la interacción hombre-máquina con el diseño de interfaces y la creación de SRI inteligentes, donde es necesario entender cómo los usuarios utilizan las estrategias y términos de búsqueda, qué resultados son juzgados relevantes por los usuarios para el mismo problema de búsqueda, etc. Aunque permanece como área importante de investigación el trabajo sobre cómo mejorar los algoritmos de búsqueda y otras técnicas automatizadas de representación y proceso de texto, desde hace pocos años existe un gran volumen de trabajo empírico para tratar temas relacionados con la interacción, especialmente la aplicación de nuevos enfoques y métodos

que soporten la búsqueda y recuperación interactiva; entre ellos, se incluye (1) el uso de técnicas de visualización para explotar la cobertura de colecciones de documentos, (2) la integración de técnicas de extracción o sumarización para explorar el contenido de conjuntos de documentos recuperados, y (3) la creación de estructuras de hipervínculos para encadenar documentos relacionados³³. En estas investigaciones es muy frecuente tratar cuestiones relacionadas con las interfaces, así como incluir evaluaciones del usuario como parte inherente de la investigación.

De igual manera, es importante destacar el interés creciente por propuestas híbridas de evaluación en los sistemas de recuperación de información interactivos (IIR) que combinen tanto el enfoque de sistemas como el enfoque de usuarios. El marco en el que se desarrollan estos experimentos engloba tres componentes básicos: (1) la implicación de usuarios potenciales, (2) la aplicación de necesidades de información individuales y dinámicas, y (3) el uso de juicios de relevancia multidimensionales y dinámicos. La aplicación de estos componentes experimentales cubre las demandas de los modelos de evaluación híbridos, al tener en cuenta tanto la relevancia como las revoluciones cognitivas e interactivas²⁷.

En este sentido, los profesionales de la Documentación, pueden y deben de participar activamente como miembros de grupos de investigación interdisciplinares, en cualquiera de las líneas de investigación mencionadas anteriormente, ya que su experiencia obtenida a lo largo de muchos años, tanto en su papel de usuarios expertos de los SRI reales como de intermediarios entre los usuarios y la información, puede ser aprovechada para el fin último del amplio campo de la recuperación de información, que no es otro que el usuario obtenga la información que necesita.

BIBLIOGRAFÍA

1. ORERA ORERA, L. "Reflexiones en torno a la investigación en Biblioteconomía". En: TRAMULLAS SAZ, J. (Editor). *Tendencias de Investigación en Documentación*. Zaragoza: Biblioteconomía y Documentación, Universidad de Zaragoza, 1996, p. 19-49.
2. ROBERTSON, S.E. "Computer retrieval". En: Vickery, B.C. (Ed). *Fifty Years of Information Progress. A Journal of Documentation Review*. London: Aslib, The Association for Information Management, 1994, p. 119-146.
3. OLE PORS, N. "Information retrieval, experimental models and statistical analysis". *Journal of Documentation*, 2000, vol. 56, nº 1, p. 55-70.
4. TAGUE-SUTCLIFFE, J. "The pragmatics of information retrieval experimentation, revisited". *Information Processing and Management*, 1992, vol. 28, nº 4, p. 467-490.
5. FROEHLICH, T.J. "Relevance reconsidered-Towards and agenda for the 21st Century: Introduction to special topic issue on relevance research". *Journal of the American Society for Information Science*, 1994, vol. 45, nº 3, p. 124-133.
6. MIZZARO, S. "Relevance: The whole history". *Journal of the American Society for Information Science*, 1997, vol. 48, nº 9, p. 810-832.

7. SCHAMBER, L. "Relevance and information behavior". En: M.E. Williams (Ed.). *Annual Review of Information Science and Technology*. Medford, NJ: American Society for Information Science, 1994, vol. 29, p. 3-48.
8. HARTER, S.P. "Variations in relevance assessments and the measurement of retrieval effectiveness". *Journal of the American Society for Information Science*, 1996, vol. 47, n° 1, p. 37-49.
9. GREISDORF, H. "Relevance: An interdisciplinary and information science perspective". *Informing Science* (Special Issue on Information Science Research), 2000, vol. 3, n° 2, p. 67-72.
10. REES, A.M.; SARACEVIC, T. "The measurability of relevance". *Proceedings of the American Documentation Institute. Annual Meeting, Santa Monica, California, October 3-7, 1966*. Woodland Hills, California: Adrienne Press, 1966, vol. 3, p. 225-234.
11. ELLIS, D. "The dilemma of measurement in information retrieval research". *Journal of the American Society for Information Science*, 1996, vol. 47, n° 1, p. 23-36.
12. HARTER, S.P. "Variations in relevance assessments and the measurement of retrieval effectiveness". *Journal of the American Society for Information Science*, 1996, vol. 47, n° 1, p. 37-49.
13. PERSSON, O. "The intellectual base and research fronts of JASIS 1986-1990". *Journal of the American Society for Information Science*, 1994, vol. 15, n° 1, p. 31-38.
14. WHITE, H.D.; McCAIN, K.W. "Visualizing a discipline: An author co-citation analysis of information science, 1972-1995". *Journal of the American Society for Information Science*, 1998, vol. 49, n° 4, p. 327-355.
15. ELLIS, D. *Progress and Problems in Information Retrieval*. London: Library Association Publishing, 1996.
16. ROBERTSON, S.E.; BEAULIEU, M.M. "On the evaluation of IR systems". *Information Processing and Management*, 1992, vol. 28, n° 4, p. 457-466.
17. NIST. TREC Tracks, 2005 <<http://trec.nist.gov/tracks.html>>. [Consulta: 21 de febrero de 2005].
18. ACM. The 28th Annual International ACM SIGIR. Conference on Research and Development in Information Retrieval, 2004. <<http://www.dcc.ufmg.br/eventos/sigir2005/callforpapers.php>>. [Consulta: 21 de febrero de 2005].
19. American Society for Information Science and Technology. ASIS&T Special Interests Groups (SIGs), 2005. <<http://www.asist.org/>> [Consulta: 21 de febrero de 2005].
20. HARTER, S.; HERT, C. "Evaluation of information retrieval systems: Approaches, issues and methods". En: M. Williams (Ed). *Annual Review of Information Science and Technology*. Medford, NJ: Information Today, 1997, vol. 32, p. 3-94.
21. BLAIR, D.C. "Some thoughts on the reported results of TREC". *Information Processing and Management*, 2002, vol. 38, n° 3, p. 445-451.

22. SARACEVIC, T.; VOORHEES, E.; HARMAN, D. "Letter to the Editor. *Information Processing and Management*, 2003, vol. 39, nº 1, p. 153-156.
23. SPARCK JONES, K. "Letter to the Editor. *Information Processing and Management*, 2003, vol. 39, nº 1, p. 156-159.
24. De MEY, M. "The relevance of the cognitive paradigm for information science". En: Harbo, O., Kajberg, L. (Eds). *Theory and Application of Information Research. Proceedings of the 2nd International Research Forum on Information Science*. London: Mansell, 1980, p. 49-61.
25. SARACEVIC, T. *Information Science revisited*. Rutgers, N.Y.: Rutgers University, School of Communication, Information and Library Studies, 1990.
26. BELKIN, N.J.; ODDY, R.N.; BROOKS, H.M. "ASK for information retrieval. Part I. Background and theory". *Journal of Documentation*, 1982, vol. 38, nº 2, p. 61-71.
27. BORLUND, P. "Experimental components for the evaluation of interactive information retrieval systems". *Journal of Documentation*, 2000, vol. 56, nº 1, p. 71-90.
28. OLUIC-VUKOVIC, V. "From information to knowledge: Some reflections on the origin of the current shifting towards knowledge processing and further perspective". *Journal of the American Society for Information Science and Technology*, 2001, vol. 52, nº 1, p. 54-61.
29. TAGUE-SUTCLIFFE, J. "Some perspectives on the evaluation of information retrieval systems". *Journal of the American Society for Information Science*, 1996, vol. 47, nº 1, p. 1-3.
30. SARACEVIC, T. "Information Science". *Journal of the American Society for Information Science*, 1999, vol. 50, nº 12, p. 1051-1063.
31. WU, M.; SONNENWALD, D.H. "Reflections on information retrieval evaluation". *Proceedings of the 1999 EBTI, ECAI, SEER & PNC Joint Meeting*. Academia Sinica, Taipei, 1999, p. 63-81. < <http://pnclink.org/annual/annual1999/1999pdf/wu-mm.pdf>>. [Consulta: 21 de febrero de 2005].
32. ROGERS, E. *Diffusion of Innovations, 4th Edition*. New York: The Free Press, 1995. <<http://nlnm.gov/pnr/eval/rogers.html>>. [Consulta: 21 de febrero de 2005].
33. BEAULIEU, M. "Interaction in information searching and retrieval". *Journal of Documentation*, 2000, vol. 56, nº 4, p. 431-439.

^[a] Un hecho significativo que demuestra su importancia es que el primer número monográfico de la revista JASIS (Journal of the American Society for Information Science) sobre temas especiales fue dedicado al tema de la relevancia: *JASIS*, 1994, vol. 45, nº 3.