

# Filtrado de información para la Búsqueda de Respuestas\*

Elisa Noguera, Fernando Llopis y Antonio Ferrández

Grupo de Investigación en Procesamiento del Lenguaje Natural y Sistemas de Información

Departamento de Sistemas y Lenguajes Informáticos

Universidad de Alicante

{elisa,llopis,antonio}@dlsi.ua.es

**Resumen:** La complejidad de los sistemas de Búsqueda de Respuestas (BR) dificulta que sean aplicados a grandes colecciones de documentos. Por este motivo, los sistemas de BR utilizan como paso previo un sistema de Recuperación de Información (RI), para detectar los documentos relevantes y disminuir de esta forma la información a procesar. En este artículo, presentamos un método para filtrar los documentos no relevantes en el contexto de BR. Concretamente, el método aplicado para filtrar información está basado en el teorema del valor medio. Hemos evaluado este método con tres medidas de similitud distintas: la medida del coseno, okapi y dfr. En los experimentos obtenemos una reducción de hasta el 89% de información sin apenas pérdida significativa en la precisión. Además, mediante este método, mejoramos un 7.05% la precisión global de un sistema de BR.

**Palabras clave:** Recuperación de Información, Búsqueda de Respuestas

**Abstract:** The complexity of the Question Answering (QA) systems difficults their application to large collections of documents. For this reason, QA systems use Information Retrieval (IR) as preliminary stage in order to detect the relevant documents and decrease the processing data. This paper studies the problem of filtering information in the context of QA. Specifically, we have applied the Mean-Value Theorem in order to filtering data without loss precision in the results. This method has been tested with three similarity measures: cosine, okapi and dfr. In the experiments, we achieved a 89% data reduction without significant data loss. Moreover, this method improves a 7.05% the precision of the QA system.

**Keywords:** Information Retrieval, Question Answering

## 1. Introducción

La Búsqueda de Respuestas (BR) puede definirse como la tarea que intenta encontrar respuestas concretas a preguntas en colecciones de texto. Esta tarea es muy útil cuando los usuarios no necesitan leer todo el documento o fragmento para obtener la información requerida.

Muchos de los sistemas de BR (ej. (Roger et al., 2005)(Narayanan y Harabagiu, 2004)) usan técnicas de procesamiento de lenguaje natural (PLN), las cuáles son muy costosas computacionalmente, y esto dificulta su aplicación a grandes colecciones de documentos. Por este motivo, la forma en la que los sistemas de BR afrontan esta limitación es aplicando Recuperación de Información (RI) (Baeza-Yates y Ribeiro-Neto, 1999) o Recu-

peración de Información basada en Pasajes (RP) (Kaszkiel, Zobel, y Sacks-Davis, 1999) a la colección de documentos y BR sólo al conjunto de documentos relevantes. La ventaja de usar PR con respecto a RI es que RP devuelve el pasaje más relevante en lugar del documento completo.

Aunque la mayoría de los autores se centran en el módulo de extracción de la respuesta, algunos ya han empezado a evaluar el impacto de este primer módulo en la BR (Monz, 2003)(Tiedemann, 2005).

El objetivo de este trabajo es mejorar RP en el contexto de BR. La salida de un sistema de RP es importante para BR porque si la respuesta no está en la salida, entonces el sistema de BR no encontrará la respuesta correcta (ver la figura 1). Además, RP debe reducir lo máximo posible la cantidad de texto a procesar por la BR.

Concretamente, se ha propuesto un méto-

\* Esta investigación ha sido financiada por el Ministerio de Ciencia y Tecnología (MCYT) con el proyecto CICYT número TIC2003-07158-C04-01.

do que está basado en el teorema del valor medio (Tadashi, 1999). Este método se ha usado para obtener el umbral de corte que obtiene los pasajes realmente relevantes. Además, el método devuelve un número diferente de pasajes para cada pregunta.

La organización del artículo es la siguiente: en la sección 2 se comentan brevemente los antecedentes del filtrado de información en BR, concretamente nos centramos en el método propuesto por el grupo QALC. En la sección 3 se describe nuestra propuesta basada en el método del valor medio. En la sección 4 se describen los experimentos realizados, los recursos utilizados y los criterios de evaluación. En la sección 5 se realiza un experimento adicional para evaluar nuestra propuesta dentro de un sistema de BR. Por último las secciones 6 y 7 muestran las conclusiones obtenidas y el trabajo futuro.

## 2. Estado del arte

La mayoría de los sistemas de BR (ej. (Fleischman, Hovy, y Echihabi, 2003)(Kise et al., 2004)) seleccionan de la salida del módulo de RI sólo la información realmente relevante para cada pregunta. Una aproximación común para el filtrado de información consiste en usar el siguiente procedimiento de dos pasos: primero se calcula la similitud entre el pasaje y la pregunta, y se le asigna una puntuación a ese pasaje para determinar el grado de relevancia de cada pasaje. Por último se hace un filtrado de pasajes, seleccionando aquellos cuya puntuación supere un determinado umbral.

Un método que sigue esta aproximación y el cual podemos encontrar en el estado del arte, es un algoritmo que fue propuesto por el grupo QALC (Ferret et al., 2001)(Ferret et al., 1999). Esta aproximación calcula el umbral de corte para extraer los documentos relevantes. Concretamente, este algoritmo detecta el decrecimiento relativo del peso de los documentos con respecto a su precedente.

## 3. Propuesta de un método para el filtrado de información

Se ha realizado una aproximación para el filtrado de información en un sistema de BR. Este aproximación se basa en la clasificación de los documentos/pasajes<sup>1</sup> en relevantes o no relevantes.

<sup>1</sup>Este método se puede aplicar a filtrado de documentos o filtrado de pasajes. Por motivos de claridad,

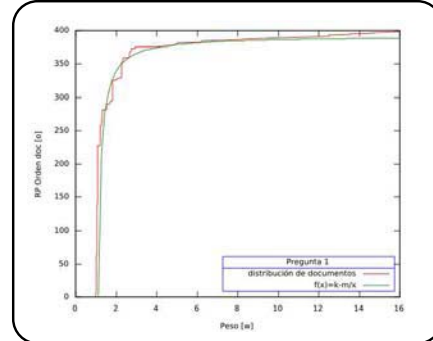


Figura 2: Pregunta 1 - Distribución de los pesos de los documentos

El método de filtrado usado sigue la estructura citada en 2 de dos pasos. En primer lugar obtenemos la lista de documentos que el sistema de RI devuelve y a partir de esta lista, seleccionamos los documentos que superan un determinado umbral.

El primer paso es inmediato, ya que simplemente tenemos que ejecutar nuestro sistema de RI. En consecuencia vamos a centrarnos en el segundo paso: seleccionar los documentos relevantes. Para cada pregunta, el sistema nos devuelve una lista de documentos con un peso asociado. Por lo tanto, la tarea a realizar es seleccionar un peso de corte para clasificar los documentos en relevantes o no relevantes.

Para determinar este umbral, observamos gráficamente la distribución que siguen estos pesos (veáse figura 2). Al representar los pesos de los documentos (obtenidos por la medida de similitud) observamos que siguen la distribución de una función racional (veáse 1).

$$f(x) = k - \frac{m}{x} \quad (1)$$

En cuanto a los documentos relevantes, en la figura 3 se puede observar que se distribuyen de forma más condensada en posiciones más altas y es a partir de un determinado punto que esta distribución es más dispersa. Nuestro objetivo es calcular dicho punto (que será nuestro punto de corte o umbral), ya que, como se observa en la figura 3 los documentos relevantes no son tan frecuentes.

Para calcular este punto de corte, usamos

en esta sección, se hace referencia únicamente a documentos

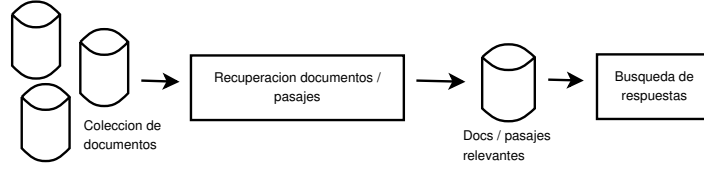


Figura 1: Proceso de filtrado en BR

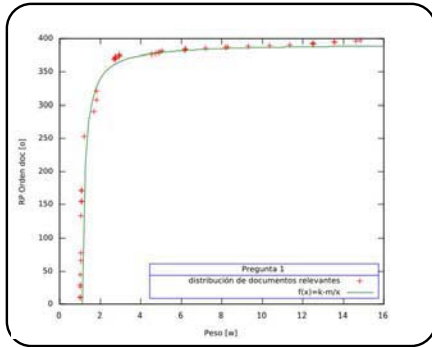


Figura 3: Pregunta 1 - Distribución de los pesos de los documentos relevantes

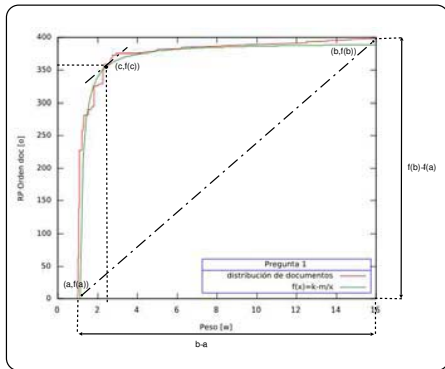


Figura 4: Pregunta 1 - Aplicación geométrica del teorema del valor medio

el teorema del valor medio (MVT) (ver teorema 1). Éste es un teorema de Cálculo que no ha sido utilizado previamente en RI y BR.

**Teorema 1:**

Si  $f(x)$  es una función que cumple que es continua en el intervalo  $[a, b]$  y diferenciable en  $(a, b)$ , entonces hay al menos un punto llamado  $c$  en el intervalo  $(a, b)$  (que cumple que  $a < c < b$ ) tal que:

$$f(c)' = \frac{f(b) - f(a)}{b - a}.$$

Como se ha mencionado, si consideramos la relación que siguen los pesos de los documentos como un intervalo de la función racional 1, entonces podemos aplicar el teorema 1 para calcular el punto de la función al que corta la recta paralela a dicho intervalo y en el que, el valor de la pendiente de dicha recta es igual a la pendiente media (veáse 4).

A continuación se detalla el método que se ha seguido para obtener el punto de corte  $(c, f(c))$  (veáse 4).

A partir del teorema 1, obtenemos el valor de la derivada  $f(c)'$ . Por lo tanto, si derivamos la función  $f(x)$  (1) y la igualamos a la anterior derivada podemos obtener el punto de corte.

En primer lugar derivamos la función  $f(x)$  (1):

$$f(x)' = \frac{m}{x^2} \quad (2)$$

Si consideramos  $x = c$  obtenemos  $f(c)'$  (3) a partir de  $f(x)'$  (2).

$$f(c)' = \frac{m}{c^2} \quad (3)$$

Despejamos  $c$  de la ecuación anterior (3).

$$c = \sqrt{\frac{m}{f(c)'}} \quad (4)$$

A continuación sustituimos  $f(c)'$  en la ecuación 5 a partir del teorema del valor medio.

$$c = \sqrt{\frac{m \cdot (b - a)}{f(b) - f(a)}} \quad (5)$$

Y por último, obtenemos  $f(c)$  (6).

$$f(c) = k - \frac{m}{c} \quad (6)$$

En conclusión, se ha considerado este punto  $(c, f(c))$  como el punto de corte para filtrar los documentos no relevantes, porque a partir de este punto los documentos relevantes se encuentran más dispersos. Se puede observar

en (2) que el 50% de los documentos relevantes se encuentran encima de dicho punto. La evaluación de este método se detalla en el siguiente capítulo.

#### 4. Experimentos y resultados

Con el fin de evaluar el método propuesto y compararlo con el resto de métodos hemos llevado a cabo una serie de experimentos. En esta sección primero se describen el conjunto de datos, los recursos y las medidas de evaluación utilizadas. A continuación, se muestran los experimentos realizados y sus resultados.

##### 4.1. Descripción de la evaluación

###### 4.1.1. Colección de datos

Para los experimentos, hemos usado las colecciones para Español QA@CLEF-2003, QA@CLEF-2004 and QA@CLEF-2005 (CLEF, 2005). Concretamente, las colecciones de documentos son EFE1994 y EFE1995, las cuales están formadas por 454.045 documentos (1086 Mb). Aunque la colección de preguntas está formada por un total de 600 preguntas, nosotros hemos evaluado sólo las preguntas para las cuáles existe al menos una respuesta en la colección. Por lo tanto, nuestra colección de preguntas está compuesta por 522 preguntas<sup>2</sup>.

###### 4.1.2. Recursos utilizados

Hemos usado el sistema IR-n (Llopis y Noguera, 2005), que es un sistema de PR que establece como baseline un número fijo de frases. Nosotros hemos considerado 8 frases como tamaño de pasaje en los experimentos, ya que es el tamaño de pasaje óptimo establecido en experimentos previos.

Además, como medidas de similitud hemos usado okapi (Roberston, Walker, y Beaulieu, 1998), la medida del coseno (Salton, 1989), dfr (Amati y Van Rijsbergen, 2002) porque son las medidas que empíricamente han obtenido los mejores resultados. Hemos adaptado estas medidas para un sistema de RP. Además, nos gustaría destacar que okapi está basada en probabilidad mientras que la medida del coseno es una medida estadística.

###### 4.1.3. Medidas de evaluación

Para la evaluación de nuestro sistema hemos usado tres medidas: “Mean Reciprocal Rank” (MRR) (TREC, 2005), Cobertura (C)

<sup>2</sup>El número de preguntas es 180 para QA@CLEF-2003 y QA@CLEF-2004 y 162 para la colección QA@CLEF-2005.

(Roberts y Gaizauskas, 2004) y Redundancia (R) (Roberts y Gaizauskas, 2004). Hemos considerado que estas medidas capturan los aspectos del rendimiento de RI que son relevantes para BR como la redundancia, el orden de las respuestas o la proporción de respuestas correctas que devuelve el sistema, y que por tanto son más apropiadas que las medidas tradicionales de precisión y cobertura.

MRR asigna a cada pregunta el valor inverso de la posición del primer pasaje que contiene la respuesta correcta o cero si la respuesta no se encuentra entre los pasajes. El valor final es la media de los valores para todas las preguntas. Esta medida es usada en BR y otorga una puntuación más alta a los pasajes que se encuentran en las primeras posiciones de la lista, ya que mide tanto la precisión como el orden de las respuestas correctas. Esta medida se obtiene con la siguiente fórmula (7).

C es la proporción de preguntas para las cuales al menos una respuesta correcta se encuentra entre los pasajes recuperados. Esto se obtiene con la fórmula (8).

R es la proporción de los pasajes recuperados por pregunta que contienen la respuesta correcta. Podemos obtenerla con la fórmula (9).

$$MRR = \frac{\sum_{i=1}^q 1/far(i)}{q} \quad (7)$$

$$C = \frac{\sum_{i=1}^q a_i}{q} \quad (8)$$

$$R = \frac{\sum_{i=1}^q pa_i/p_i}{q} \quad (9)$$

Donde

- $far(i)$  es la posición del primer pasaje que contiene la respuesta correcta para la pregunta  $i$ .
- $q$  es el número de preguntas.
- $1/far(i)$  será 0 si la respuesta no se encuentra en ningún pasaje.
- $a_i$  será 1 si la respuesta correcta para la pregunta  $i$  se encuentra en algún pasaje y 0 en otro caso.
- $p_i$  es el número de pasajes devueltos para la pregunta  $i$ .
- $pa_i$  es el número de pasajes devueltos para la pregunta  $i$  que contienen la respuesta correcta.

En todos los experimentos hemos usado el criterio de evaluación en el cual determinamos que un pasaje es relevante para una pregunta determinada, si contiene el patrón de la respuesta correcta para esa pregunta.

## 4.2. Experimentación

En los experimentos se ha evaluado nuestra propuesta (MVT) y se ha comparado con el método QALC y con un sistema de referencia (Baseline), el cual devuelve un número fijo de pasajes (400) para cada pregunta.

En las tablas 1, 2 y 3 se puede observar la evaluación del sistema base y los métodos MVT y QALC utilizando las medidas del coseno, okapi y dfr como medidas de similitud.

Estos métodos han sido evaluados con las colecciones QA@CLEF-2003, 2004 y 2005 (ver 4.1.1). Además, se ha realizado una evaluación global (llamada total) para evaluar el rendimiento medio de las propuestas (con las tres colecciones).

Las tablas 1, 2 y 3 muestran la media de pasajes que son devueltos y la reducción de datos para cada método. También se muestran las medidas de evaluación previamente descritas: MRR, Cobertura (C) y Redundancia (R) para los diferentes métodos (véase sección 4.1.3).

Analizando las tablas 1, 2 y 3 observamos que MRR es casi constante para los tres sistemas, porque esta medida da más valor a las respuestas que se devuelven en las primeras posiciones. Es importante resaltar que para el método MVT, R es mayor (0,32) que para los otros métodos (Baseline (0,2) y QALC (0,29)). Obviamente, C decrece en MVT con respecto al Baseline, porque como se ha indicado anteriormente, MVT devuelve un subconjunto de pasajes de los que son devueltos por el Baseline. Hay que destacar que las medidas MRR y C en MVT son mejores que en el sistema QALC teniendo en cuenta que MVT devuelve una menor cantidad de pasajes.

En resumen, podemos concluir que el método MVT es efectivo para las tres medidas de similitud. El método se comporta de la misma forma independientemente de la medida de similitud utilizada, reduciendo los pasajes en una media del 89 % con la medida del coseno, 88 % con la medida okapi y 86 % con la medida dfr, que viene a demostrar la robustez del método propuesto.

Coleccion	Metodo	MRR
2003	Baseline	0.4725
	MVT	0.5067
2004	Baseline	0.4283
	MVT	0.4478
2005	Baseline	0.4208
	MVT	0.4608
Total	Baseline	0.4406
	MVT	0.4717

Cuadro 4: Evaluación del sistema de BR

## 5. Evaluación de MVT en BR

Se ha evaluado la propuesta como módulo de un sistema de BR para probar la efectividad del método dentro de esta tarea. Para este fin, hemos usado el sistema de BR "Ali-QAn" (Roger et al., 2005) que es un sistema monolingüe (español) y de dominio abierto. Esta aproximación se basa fundamentalmente en el uso de patrones sintácticos para la detección de las posibles respuestas.

Hemos evaluado el sistema de BR con las tres colecciones de documentos descritas en 4.1.1 y con la medida de similitud okapi.

En la tabla 4 se pueden observar los resultados de evaluar tanto el sistema de BR con el Baseline (devolviendo 400 pasajes) y con el método de filtrado de pasajes.

El sistema de BR con el Baseline ha obtenido un MRR medio de 0.4406, mientras que con el método MVT, el sistema de BR obtiene un MRR medio de 0.4717.

Podemos concluir que el método MVT mejora los resultados con respecto al Baseline en el sistema de BR en un 7.05 %. Esta mejora se debe a la reducción de pasajes que obtiene el método MVT, ya que si se reduce el número de pasajes se introduce un menor número de pasajes no relevantes (ruido) en el sistema de BR.

Además, esta reducción de información supone una mejora en el tiempo de procesamiento del sistema de BR que podría ser adecuado para sistemas de BR en tiempo real o sistemas de BR interactivos, ya que estos sistemas requieren la condición de responder al usuario en el menor tiempo posible.

## 6. Conclusiones

En este trabajo hemos estudiado el problema de filtrado de información en RP aplicado a BR. Concretamente, hemos propuesto un

Coleccion	Metodo	Media Pasajes	Reducción de datos	MRR	C	R
2003	BASELINE	400	0	0,6320	1	0,19
	MVT	41	-89,75 %	0,6271	0,96	0,32
	QALC	174	-56,5 %	0,6226	0,93	0,25
2004	BASELINE	400	0	0,6044	0,97	0,15
	MVT	43	-89,25 %	0,5998	0,92	0,23
	QALC	112	-72 %	0,5969	0,85	0,24
2005	BASELINE	400	0	0,7469	0,97	0,26
	MVT	45	-88,75 %	0,7432	0,96	0,36
	QALC	117	-70,75 %	0,7423	0,92	0,38
Total	BASELINE	400	0	0,6611	0,98	0,2
	MVT	43	-89,25 %	0,6567	0,95	0,3
	QALC	134	-66,42 %	0,6539	0,90	0,29

Cuadro 1: Resultados de los experimentos con la medida de similitud del coseno

Coleccion	Metodo	Media Pasajes	Reducción de datos	MRR	C	R
2003	BASELINE	400	0	0,6558	1	0,2
	MVT	43	-89,25 %	0,6530	0,96	0,3
	QALC	198	-50,5 %	0,6499	0,92	0,26
2004	BASELINE	400	0	0,6249	0,97	0,15
	MVT	47	-88,25 %	0,6237	0,93	0,3
	QALC	137	-65,75 %	0,6171	0,88	0,27
2005	BASELINE	400	0	0,7864	0,98	0,26
	MVT	52	-87 %	0,7845	0,97	0,34
	QALC	160	-60 %	0,7842	0,96	0,36
Total	BASELINE	400	0	0,6890	0,98	0,2
	MVT	47	-88,17 %	0,6871	0,95	0,31
	QALC	165	-58,75 %	0,6837	0,92	0,3

Cuadro 2: Resultados de los experimentos con la medida de similitud okapi

Coleccion	Metodo	Media Pasajes	Reducción de datos	MRR	C	R
2003	BASELINE	400	0	0,6009	0,99	0,19
	MVT	54	-86,5 %	0,5968	0,93	0,35
	QALC	269	-32,75 %	0,5982	0,96	0,25
2004	BASELINE	400	0	0,6422	0,97	0,14
	MVT	50	-87,5 %	0,6441	0,92	0,28
	QALC	216	-46 %	0,6418	0,94	0,22
2005	BASELINE	400	0	0,7319	0,96	0,24
	MVT	63	-84,25 %	0,7319	0,95	0,48
	QALC	233	-41,75 %	0,7317	0,95	0,36
Total	BASELINE	400	0	0,6583	0,97	0,19
	MVT	55	-86,08 %	0,6576	0,93	0,37
	QALC	239	-40,17 %	0,6572	0,95	0,28

Cuadro 3: Resultados de los experimentos con la medida de similitud dfr

método para filtrar pasajes calculando el umbral de corte (en el peso de los pasajes) para determinar los pasajes relevantes (los pasajes que probablemente contendrán la respuesta). Este método está basado en el teorema del valor medio (MVT). Nuestra propuesta ha sido comparada con un sistema base (un número fijo de 400 pasajes), y con un método conocido para filtrar pasajes (QALC (Ferret et al., 2001)(Ferret et al., 1999)). Los experimentos se han realizado empleando las medidas de similitud del coseno, okapi y dfr, además han sido llevados a cabo con las colecciones QA@CLEF-2003, 2004, 2005, obteniendo una reducción significativa en el número de pasajes (hasta un 89%) y sin una reducción apreciable en el MRR. Además, la mejor Redundancia (R) se obtiene con nuestra propuesta. Por otro lado, hemos evaluado el rendimiento global del método dentro de un sistema de BR y hemos obtenido una mejora de un 7.05% con respecto al Baseline.

## 7. Trabajo futuro

Finalmente, las futuras direcciones que tenemos planeadas llevar a cabo son: por un lado mejorar este modelo, detectando no sólo los pasajes que son relevantes para cada pregunta, sino también las frases relevantes en cada pasaje. Y por otro lado, también se va a introducir en el sistema conocimiento lingüístico (generalmente usado en BR): reconocimiento de entidades, análisis morfológico, análisis sintáctico parcial,... con el fin de establecer un método de filtrado más robusto en base a restricciones lingüísticas. Un ejemplo de restricción lingüística sería seleccionar sólo aquellos documentos que contengan el tipo de entidad que se espera como respuesta.

## Bibliografía

- Amati, G. y C. J. Van Rijsbergen. 2002. Probabilistic Models of information retrieval based on measuring the divergence from randomness. *ACM TOIS*, 20(4):357–389.
- Baeza-Yates, Ricardo A. y Berthier A. Ribeiro-Neto. 1999. *Modern Information Retrieval*. ACM Press / Addison-Wesley.
- CLEF. 2005. Workshop of Cross-Language Evaluation Forum (CLEF) 2005. En *Workshop of Cross-Language Evaluation Forum (CLEF)*, Lecture notes in Computer Science. Springer-Verlag.
- Ferret, Olivier, Brigitte Grau, Martine Hurault-Plantet, Gabriel Illouz, y Christian Jacquemin. 2001. Document Selection Refinement based on linguistic features for QALC, a Question Answering System. En *Proceedings of Recent Advances on Natural Language Processing RANLP'01, Bulgaria*.
- Ferret, Olivier, Brigitte Grau, Gabriel Illouz, Christian Jacquemin, y Nicolas Masson. 1999. QALC - the Question-Answering program of the Language and Cognition group at LIMSI-CNRS. En *Proceedings of TREC*.
- Fleischman, Michael, Eduard H. Hovy, y Abdessamad Echihabi. 2003. Offline Strategies for Online Question Answering: Answering Questions Before They Are Asked. En *Proceedings of ACL*, páginas 1–7.
- Kaszkiel, Marcin, Justin Zobel, y Ron Sacks-Davis. 1999. Efficient passage ranking for document databases. *ACM Trans. Inf. Syst.*
- Kise, Koichi, Markus Junker, Andreas Dengel, y Keinosuke Matsumoto. 2004. Passage Retrieval Based on Density Distributions of Terms and Its Applications to Document Retrieval and Question Answering. En *Reading and Learning*, páginas 306–327.
- Llopis, Fernando y Elisa Noguera. 2005. Combining Passages in the Monolingual Task with the IR-n System. En *Proceedings of CLEF*.
- Monz, C. 2003. *From Document Retrieval to Question Answering*. Ph.D. tesis, University of Amsterdam.
- Narayanan, Srinu y Sanda Harabagiu. 2004. Question answering based on semantic structures. En *Proceedings of COLING 2004*.
- Roberston, S., S. Walker, y M. Beaulieu. 1998. OKAPi at TREC-7. En *Proceedings of Seventh Text RETrieval Conference, volume 500-242*, páginas 253–264. National Institute of Standard and Technology. Gaithersburg, USA.
- Roberts, Ian y Robert J. Gaizauskas. 2004. Evaluating Passage Retrieval Approaches for Question Answering. En *Proceedings of ECIR*, páginas 72–84.

- Roger, S., S. Ferrández, A. Ferrández, J. Peral, F. Llopis, A. Aguilar, y D. Tomás. 2005. AliQAn, Spanish QA System at CLEF-2005. En *Proceedings of Cross Language Evaluation Forum*.
- Salton, G. 1989. Automatic Text Processing: The Transformation, Analysis, and retrieval of Information by Computer.
- Tadashi, Tokieda. 1999. A Mean Value Theorem. *The american mathematical monthly*, 106(7):673.
- Tiedemann, Jörg. 2005. Improving Passage Retrieval in Question Answering Using NLP. En *Proceedings of EPIA*, páginas 634-646.
- TREC. 2005. Overview of the TREC 2005 question answering track. En Ellen M. Voorhees, editor, *In Proceedings of TREC 2005*.