

SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

TEMA 1. EJERCICIO 1

Pregunta completa	Pregunta breve	Considerar respuesta correcta si...
¿Cómo se hace una paella de marisco?	paella de marisco	<u>Instrucciones</u> para hacer una paella de <u>marisco</u>

1. Realiza la pregunta breve anterior en los motores de búsqueda Google y Teoma y completa la siguiente tabla.

	Google	Exalead
Número resultados encontrados	Más	Menos
Número de resultados correctos entre los 10 primeros	9	8
Si se realiza la pregunta breve con las palabras en otro orden ¿varía el resultado?	Si	No
¿Has obtenido algún resultado que diera un error al intentar acceder a la página (entre los 10 primeros)? Indica cuantos	No	No
¿Se pueden filtrar los resultados por país, idioma, dominio o tipo de documento?	Si	
¿Se pueden refinar las búsquedas? Es decir ¿se puede hacer una pregunta sobre los resultados de otra pregunta?	Si	
¿Se permite añadir operadores booleanos a la consulta?	Si	
¿Se permite el uso de paréntesis en una consulta con operadores booleanos?	no	

2. ¿Cuál de los dos motores de búsqueda anteriores te parece mejor? Determina un método para evaluar un motor de recuperación: ¿qué parámetros tendrías en cuenta y cuáles valorarías más?

- Posicionamiento: no importa tanto qué cantidad de páginas da como resultado si no las posiciona bien (no las vamos a mirar todas)
- Opciones de búsqueda: diferencia según orden de las palabras, operadores booleanos, filtrados, posibilidad de refinamiento
- Actualización

SISTEMAS AVANZADOS DE RECUPERACIÓN DE INFORMACIÓN

- d. Tamaño/cobertura: cuanto más amplia es la porción de la Red que abarcan, más probabilidades de encontrar buenas páginas. Aunque a partir de cierta cantidad de páginas deja de tener tanta importancia.
- e. Visualización de los resultados: breve resumen página, título, caché, cambio de formato a HTML. Ejemplo: buscadores que permiten visualizarla en un cuadro pequeño en la misma página de resultados. Otro ejemplo: Kartoo, que las muestra en forma de mapa.
- f. Tiempo de búsquedas: en general son todos instantáneos

3. Si se realiza una pregunta en su forma completa en lugar de breve generalmente los resultados empeoran ¿Por qué? ¿Mejorarían los resultados de un buscador respecto a otro?

Porque generalmente los buscadores no utilizan técnicas de procesamiento de lenguaje natural, por lo que la introducción de otras palabras no identificativas de lo que queremos saber introduce ruido.

Este criterio sólo serviría para comparar buscadores si alguno de éstos utilizara PLN. Ask Jeeves por ejemplo tiene algo de PLN.

4. Si se realiza una pregunta en inglés en lugar de en español en general los resultados mejoran (a menos que sea una pregunta específica de un lugar de habla no inglesa) ¿Por qué? ¿Mejorarían los resultados de un buscador respecto a otro?

Desde luego la paella es bastante específica de nuestro país y aún así, tras probarlo, obtuve mejores resultados en inglés (otra cosa es que la receta sea buena porque hablaban de microondas por todos lados..:)

Esto ocurre porque la cantidad de páginas en inglés es mucho mayor que en otros idiomas. Por tanto, por probabilidades, es más factible encontrar buenos resultados en inglés. Claro que esto es cierto para preguntas no relacionadas con un país o idioma específico: por ejemplo, si buscamos las farmacias que hay en una calle de Madrid, obviamente tendremos mejores resultados si la hacemos en español.

Este criterio nos serviría para comparar buscadores si la cobertura que tienen para diferentes idiomas varía mucho.

5. CONCLUSIONES: en general el modo de hacer una pregunta no influye en los resultados de un buscador respecto de otro, pero estos dos últimos aspectos sí podrían tenerse en cuenta **Diseña esquemáticamente una metodología para evaluar buscadores en Internet.**