

# RECUPERACIÓN DE INFORMACIÓN EN INTERNET

## P.O.P. Lengua y Usos Profesionales

Facultad de Filología, Universidad de A Coruña

Curso 2008/2009

### TRABAJO PRÁCTICO

El siguiente trabajo práctico consta de 3 ejercicios que debes resolver utilizando las técnicas de recuperación de información que hemos visto a lo largo del curso.

El trabajo es individual. Debes elaborar una breve memoria explicativa de los pasos que has seguido para resolver cada ejercicio (aprox. 2 o 3 páginas para cada ejercicio). Posteriormente, debes enviar dicha memoria a la dirección de correo electrónico del profesor Jorge Graña ([grana@udc.es](mailto:grana@udc.es)). La fecha límite de entrega es el viernes 6 de febrero de 2009.

Los profesores de este curso revisarán los trabajos presentados, y comunicarán las calificaciones a los alumnos, reservándose el derecho de convocarlos previamente a una defensa personal, en caso de que algún aspecto del trabajo no esté expuesto de manera clara.

Por supuesto, desde hoy hasta la fecha de entrega, puedes consultar a los profesores del curso las dudas relacionadas con la realización del trabajo, o bien por correo electrónico, o bien en persona durante las horas de tutorías que tenemos reservadas para ello.

Los ejercicios son los siguientes:

#### 1. Preparación de un viaje.

Queremos pasar una semana en Benidorm, del 6 al 12 de abril de 2009. Queremos alojarnos en un camping de primera categoría o, en su defecto, en uno de segunda categoría. Queremos viajar en tren desde A Coruña hasta Alicante y, una vez en Alicante, queremos llegar hasta Benidorm utilizando algún otro tipo de transporte público (autobús o tren de cercanías). Durante el tiempo que estemos alojados, para poder realizar alguna excursión, queremos alquilar un coche durante 2 o 3 días.

Averigua si todo esto es posible. Indica aquellas búsquedas que han sido fructíferas y aquéllas que no. Haz mención explícita de aquellas páginas que permiten directamente efectuar reservas y de aquéllas que no. En caso de que necesites más información, y que la única manera de obtenerla sea ponerse en contacto con alguna persona de alguna de las entidades consultadas, puedes hacerlo. Simplemente, hazlo constar y, por supuesto, no lo hagas por teléfono, sino por correo electrónico (la idea es no gastar absolutamente ningún dinero en los preparativos) y procura no molestar en exceso, ni ser demasiado insistente (ya que al final se trata de un ejercicio, no de efectuar reservas reales).

Si lo deseas, puedes cambiar el destino, las fechas, el tipo de alojamiento y el medio de transporte, por otros que te parezcan más atractivos o que te interesen más. Eso sí, intenta que la *complejidad* del viaje y de las actividades que se van a realizar durante los días de estancia sea similar a la del supuesto que te proponemos.

Al final, queremos también que nos transmitas tu opinión sobre si consideras o no que Internet y las herramientas de recuperación de información son verdaderamente útiles para preparar este tipo de actividad, y queremos también que nos indiques cuál es el tiempo

medio que consideras necesario para realizar toda la operación, desde que comienzas las primeras búsquedas, hasta que confirmas las reservas. Aunque puede ser muy variable, dependiendo de las fechas y del destino, este último dato te servirá para otra vez, para saber con cuánto tiempo de antelación se deben comenzar los preparativos.

## 2. Realización de una compra por Internet.

Queremos comprar por Internet una cámara fotográfica reflex digital marca Nikon, modelo D90. Su precio aproximado en las tiendas de A Coruña es de unos 950 euros, sólo el cuerpo, es decir, sin ningún objetivo. Se trata pues de conseguirla más barata, o incluso por un precio similar pero con algún objetivo básico, y averiguando además algunos detalles de cierta relevancia, como pueden ser los siguientes:

- ¿Cómo es el método de pago? ¿Es suficiente con tener una tarjeta de crédito, o debo darme de alta en algún sistema de verificación de pagos?
- ¿Están los gastos de envío incluidos en el precio? Si no es así, ¿a cuánto ascienden?
- ¿Quién envía la mercancía? ¿Algún servicio oficial de correos? ¿Alguna empresa de mensajería privada? En caso de que venga desde algún país de fuera de la comunidad europea, ¿me garantiza el vendedor de alguna manera que el producto no será parado en la aduana, y que por tanto podré librarme de pagar el IVA y los aranceles? Si es así, ¿cómo me lo garantiza? Y si al final ocurre eso, ¿dispongo de algún mecanismo de protesta?
- ¿Qué garantía tiene la cámara? En caso de que venga desde algún país de fuera de la comunidad europea, ¿es válida esa garantía para España? Si no es así, ¿puedo hacer una extensión de garantía? ¿A cuánto asciende esa extensión?
- En caso de que la mercancía me llegue defectuosa o dañada, ¿existe algún tipo de seguro que me garantice que puedo enviársela de vuelta al vendedor para que me la sustituya sin coste alguno por mi parte?

Al igual que en el ejercicio anterior, haz un breve resumen de las páginas que han sido útiles y de las que no, a la hora de recopilar información sobre el artículo y sobre todos estos detalles.

Como antes, si lo deseas, puedes cambiar también el artículo a comprar por otro que te parezca más atractivo o que te interese más. Como antes, intenta mantener el mismo nivel de *complejidad* en lo que se refiere a los detalles de la compra. Por último, procura no ponerte en el supuesto de comprar algún producto que ya tengas. Sería muy desagradable descubrir que podrías haberlo conseguido mucho más barato.

Igual que antes, nos interesa también conocer tu opinión sobre la verdadera versatilidad de Internet como medio de comercio electrónico.

## 3. Estudio comparativo del rendimiento de varios sistemas de recuperación de información.

Elige 3 buscadores. Elige una consulta relacionada con algún tema que conozcas bien. Lanza esa consulta en cada uno de los 3 buscadores. Fíjate en los 10 primeros resultados de cada buscador. Calcula la unión de esos documentos. Serán como mucho 30 documentos, aunque probablemente sean menos, ya que es de suponer que algún documento será recuperado por más de un buscador. Llamemos entonces  $D$  al número de documentos diferentes, siendo, como ya hemos dicho,  $D \leq 30$ . Imagina ahora que esos  $D$

documentos son toda la base de datos documental de la que disponemos. Accede a cada uno de esos  $D$  documentos, y decide si es relevante o no para la consulta que has hecho. En función de esas relevancias, calcula la precisión, el *recall* y la precisión no interpolada de cada buscador. Indícanos qué buscadores has utilizado, cuál ha sido la consulta, cómo son las tres listas sí/no de relevancias, cuál es el valor de cada una de las tres medidas de rendimiento, y qué conclusiones sacas al respecto.

Repite de nuevo el mismo experimento, pero esta vez con una consulta relacionada también con algún tema que conozcas, igual que antes, para que seas capaz de juzgar rápidamente la relevancia de cada documento recuperado, pero que sea de un tema completamente diferente. ¿Qué ha ocurrido esta vez? El buscador que más te satisface ahora, ¿es el mismo de antes o es otro? ¿Qué conclusiones sacas al respecto?

Por supuesto, en cada uno de los dos experimentos, tomando inicialmente un único buscador, puedes tantear con la consulta e intentar refinarla hasta que el buscador ofrezca resultados interesantes. Pero eso sí, una vez que encuentres la consulta que te gusta, lánzala tal cual en cada uno de los tres buscadores, para que todos ellos compitan al mismo nivel y ninguno se vea perjudicado respecto a los demás.

Miguel A. Alonso Pardo ([alonso@udc.es](mailto:alonso@udc.es))  
Jesús Vilares Ferro ([jvilaras@udc.es](mailto:jvilaras@udc.es))  
Jorge Graña Gil ([grana@udc.es](mailto:grana@udc.es))