

BUSCADORES INTERACTIVOS INTELIGENTES

31 de Julio de 2006

Ante la creciente necesidad de aplicaciones que faciliten el acceso y tratamiento de la ingente cantidad de información digital, la Universidad de Jaén ha centrado sus esfuerzos en el desarrollo y evaluación de sistemas de búsqueda de respuestas y recuperación de documentos en escenarios multilingües.

Ro Córdoba

La creciente información digitalizada y no estructurada que se ha producido en los últimos veinte años hace necesaria la investigación en sistemas de información textual que faciliten la localización, acceso y tratamiento de toda esta ingente cantidad de datos. A esto hay que añadir el crecimiento exponencial de información digital en diferentes lenguas, lo cual requiere sistemas que permitan buscar o recuperar información solicitada en el idioma no sólo origen (es decir en el que se formula la pregunta) sino también en el idioma destino (es decir en el que está escrita la pregunta) en grandes volúmenes de información digitalizada.

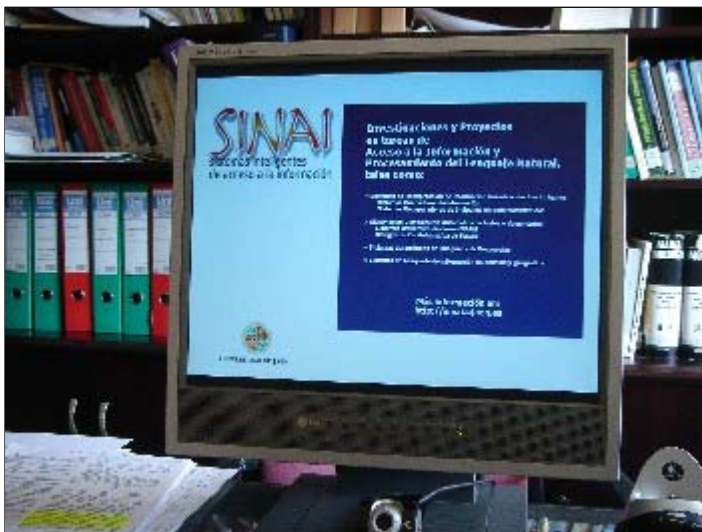
Con el fin de dar solución a esta problemática, la Universidad de Jaén participa en el macroproyecto 'Recuperación de Respuestas en Documentos Digitalizados', encargado por el Ministerio de Ciencia y Tecnología, con el que se pretende desarrollar nuevas técnicas y herramientas interactivas que sean capaces, no sólo de recuperar información específica en documentos digitales, sino que puedan localizar respuestas a preguntas concretas en grandes volúmenes de información digitalizada.

Se trata de un proyecto coordinado en el que participan, junto con la Universidad de Jaén (UJA), la Universidad Nacional a Distancia (UNED), la Universidad Politécnica de Valencia (UPV) y la Universidad de Alicante (UA). Los grupos de investigación de estas universidades son los de mayor relevancia en investigaciones en el acceso a la información.

El grupo de 'Sistemas Inteligentes de Acceso a la Información' de la Universidad de Jaén, a través de la puesta en marcha de uno de los cuatro subproyectos que componen este macroproyecto, titulado 'Desarrollo de sistemas de recuperación de información multilingües: técnicas y herramientas', trata de desarrollar un sistema de búsqueda de información multilingüe. Las premisas marcadas en el diseño son básicamente tres: en primer lugar, el sistema debe ser escalable, al menos, a ocho idiomas: español, inglés, francés, alemán, italiano, danés, finlandés y sueco. En segundo lugar, debe soportar modelos de recuperación de información basados en documentos, incluso, en imágenes. Por último, debe permitir una fácil integración en un sistema de preguntas y respuestas multilingüe.



Grupo 'Sistemas Inteligentes de Acceso a la Información' de la UJA



El sistema será evaluado en el forum internacional CLEF

CLEF (*Cross- Language Evaluation Forum*), siendo éste otro de los objetivos del desarrollo de infraestructura necesaria para la experimentación y evaluación de sistemas de recuperación de información que trabajen sobre las lenguas europeas en contextos monolingües y translingües, y en la creación de conjuntos de datos reutilizables por los sistemas desarrollados. Estas campañas son un excelente escaparate de las tecnologías desarrolladas, ya que en ellas participan los principales grupos internacionales relacionados con el tema.

Las campañas CLEF destacan por su carácter dinámico, dado que los investigadores pueden proponer nuevas tareas para ser evaluadas. Por otro lado, la participación en estas campañas supone para los grupos participantes, y en concreto para los grupos españoles, un gran reto y una gran valía cualitativa, ya que pueden probar y evaluar sus ideas además de la tecnología desarrollada.

Más información:

Luis Alfonso Ureña López

En cuanto a la aplicabilidad de los resultados, L. Alfonso Ureña resalta que se limitan a una posible transferencia tecnológica de las técnicas y recursos desarrollados, para la creación de un buscador en Internet que pudiera recuperar en lenguaje natural información en distintos idiomas. El sistema interactivo desarrollado se adaptará a consultas específicas, en las que la necesidad de información pueda satisfacerse con un solo documento, a la vez que esta búsqueda de respuestas pueda hacerse en idiomas desconocidos.

En definitiva, con este trabajo se pretende "aportar técnicas y herramientas interactivas que faciliten la accesibilidad a la información, y, sobre todo, que mejoren la calidad de la información digitalizada, incluso la existente en otros idiomas", según indica L. Alfonso Ureña, investigador responsable del subproyecto de Jaén.

Valoración internacional

Todos estos sistemas interactivos de búsqueda y recuperación de información multilingüe, desarrollados con la puesta en marcha de este ambicioso proyecto, se están evaluando a nivel internacional en la campaña internacional del macroproyecto. Este forum internacional se centra en el desarrollo de infraestructura necesaria para la experimentación y evaluación de sistemas de recuperación de información que trabajen sobre las lenguas europeas en contextos monolingües y translingües, y en la creación de conjuntos de datos reutilizables por los sistemas desarrollados. Estas campañas son un excelente escaparate de las tecnologías desarrolladas, ya que en ellas participan los principales grupos internacionales relacionados con el tema.

Universidad de Jaén
Tel.: 953 21 28 95

Email: laurena@ujaen.es

<http://sinai.ujaen.es/wiki/index.php/Proyectos>

[\[IMPRIMIR\]](#)

[\[CERRAR VENTANA\]](#)