



Está usted en: [Principal de la UGR](#) : [Noticias UGR](#) : Noticias Ciencia

[Noticias UGR](#)

[Noticias Ciencia UGR](#)

[Science News UGR](#)

[Nouvelles Scientifiques UGR](#)

 [English Version](#)

 [Version Française](#)

NOTICIAS CIENCIA

[Agroalimentación](#)

[Ciencias de la Vida](#)

[Ciencias Sociales, Económicas y Jurídicas](#)

[Ciencia y Tecnología de la Salud](#)

[Física, Química y Matemáticas](#)

[Humanidades](#)

[Recursos Naturales y Medio Ambiente](#)

[Tecnologías de la Información y las Comunicaciones](#)

[Tecnologías de la Producción](#)

[Genérico, otros](#)

Noticias por:

[Instancias de la UGR](#)

[Orden cronológico](#)

[Las 15 más consultadas](#)



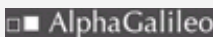
UCC



FECYT



SINC



Alpha Galileo



Eureka Alert!



Cordis

20/04/2006 - El Laboratorio Europeo de Física de Partículas emplea un nuevo sistema documental creado por un investigador de la UGR

- Arturo Montejo Ráez, del [Departamento de Lenguajes y Sistemas Informáticos](#), ha elaborado un método de clasificación multi-etiquetado automático para colecciones digitales.

- La base de datos de este centro de investigación, en el que participan veinte países de todo el mundo y que se ubica en Suiza, recibe cada semana más de mil nuevos artículos científicos.

C@MPUS DIGITAL La creciente digitalización de las bibliotecas tradicionales y el aumento en la producción científica en dominio, como ocurre en el ámbito de la Física de Altas Energías, ha llevado a considerar obsoletos los sistemas de indexación manual, por resultar inviables en su práctica. Y para muestra, un botón: cada semana, más de mil nuevos artículos llegan a las bases de datos del [Laboratorio Europeo de Física de Partículas](#) (CERN, por sus siglas en francés), ubicado en Suiza, cerca de Ginebra, y considerado hoy en día un modelo de colaboración científica internacional y uno de los centros de investigación más importantes en el mundo.

Este hecho ha motivado la búsqueda de nuevos sistemas de clasificación multi-etiquetado automáticos, que aceleren el proceso de documentación en bases de datos tan amplias como la del CERN.

Es el caso del sistema elaborado por **Arturo Montejo Ráez**, investigador del Departamento de Lenguajes y Sistemas Informáticos de la [Universidad de Granada](#) y profesor del departamento de Informática de la [Universidad de Jaén](#) cuyo trabajo ha sido objeto de una tesis doctoral dirigida por los profesores Luis Alfonso Ureña López y Ralf Steinberger.

Montejo propone una solución basada en [técnicas de Recuperación de Información y Aprendizaje Automático](#) para resolver el problema del multi-etiquetado en colecciones digitales. Su investigación se ha centrado en documentos del dominio de la Física de Altas Energías, creando un novedoso sistema de documentación, que podría ser aplicado a otras bibliotecas digitales.

El investigador se centró en la categorización y la clasificación de los textos, en el que palabras clave predefinidas son consideradas categorías que se asignan a los documentos en función del contenido semántico de los mismos. Durante el desarrollo de su trabajo, realizado principalmente en el Laboratorio Europeo para la Investigación Nuclear, la colección de documentos manejada desveló problemas no cubiertos con anterioridad por la literatura especializada.

“Clasificadores base binarios”

La asignación automática de palabras clave a los documentos abre nuevas posibilidades en la exploración documental, y su interés ha despertado en la



Noticias Ciencia



Science News



Nouvelles de la Science

comunidad científica internacional la búsqueda de soluciones. El sistema propuesto por el investigador de la Universidad de Granada es una estrategia de clasificación multi-etiquetado que pueda construirse a partir de algoritmos de aprendizaje automático conocidos como "clasificadores base binarios". Además., su trabajo de campo valida la hipótesis de que la integración de información bibliográfica disponible en las bibliotecas digitales mejora los sistemas de clasificación.

El algoritmo propuesto por Montejo está siendo aplicado por el CERN en su servidor documental (<http://cds.cern.ch>). Asimismo, otras grandes bibliotecas digitales de algunas organizaciones internacionales han mostrado su interés en el sistema para su uso e integración, debido a la gran cantidad de aplicaciones que los sistemas de multi-etiquetado automático ofrecen.

El Laboratorio Europeo de Física de Partículas fue fundado en 1954 por doce países europeos, y actualmente cuenta con 20 estados miembros. Aparte de los científicos de los estados miembros, científicos de 220 institutos y universidades de países no miembros usan sus instalaciones.

Los países miembros son Austria, Bélgica, Bulgaria, República Checa, Dinamarca, Finlandia, Francia, Alemania, Grecia, Hungría, Italia, Los Países Bajos, Noruega, Polonia, Portugal, Eslovaquia, España, Suecia, Suiza y el Reino Unido. Los países y organizaciones observadoras son la Comisión Europea, India, Israel, Japón, Rusia, Turquía, la UNESCO y Estados Unidos. La lista de países no miembros implicados en programas del CERN la completan Argelia, Argentina, Armenia, Australia, Azerbaiyán, Bielorrusia, Brasil, Canadá, República Popular China, Croacia, Chipre, Estonia, Georgia, Islandia, Irán, Irlanda, México, Marruecos, Pakistán, Perú, Rumanía, Serbia, Eslovenia, Sudáfrica, Corea del Sur, Taiwan y Ucrania.

Referencia:

Prof. Arturo Montejo Ráez. Dpto. de Lenguajes y Sistemas Informáticos de la Universidad de Granada.

Tlfs. 958 244 344 - 953 212 882.

Correo e.: amontejo@ujaen.es.

Web: <http://www.di.ujaen.es/~amontejo>

Buscar en Noticias Ciencia

[Contacto. Consultas de carácter general por correo electrónico](#)

© Servicio de Comunicación - Universidad de Granada

Resolución recomendada: 1024 x 768 - Conforme W3C: [WAI Doble A](#) - [Accesibilidad](#)