

Convocatoria pública de ayudas para la campaña competitiva: Medical Semantic indexing in Spanish BioASQ MESINESP Task

El Plan de Impulso de las Tecnologías del Lenguaje (Plan TL) tiene como objetivo fomentar el desarrollo del procesamiento del lenguaje natural, la traducción automática y los sistemas conversacionales en lengua española y lenguas cooficiales. En el marco de dicho plan, el Ministerio de Economía de Empresa, a través del Encargo realizado por la Secretaría de Estado para el Avance Digital (SEAD) al Consorcio Barcelona Supercomputing Center - Centro Nacional de Supercomputación (BSC-CNS) para la realización de actuaciones de apoyo técnico especializado en en el área de biomedicina entre otros, se anuncia la realización de convocatorias públicas para la divulgación y concurrencia de participantes en la campaña que se detalla a continuación.

De conformidad con lo previsto en los artículos 17.3.b y 20.8.a de la Ley 38/2003, de 17 de noviembre, General de Subvenciones, el artículo 6 de la ley 9/2017 de 8 de noviembre, de Contratos del Sector Público, las normas especiales aplicables a los concursos de proyectos para la entrega de premios establecidos en la misma y la asignación de recursos presupuestarios establecidos para el ejercicio 2019 mediante la financiación del encargo de gestión formalizado a cargo del crédito presupuestario 20.12.467I.640, se detalla a continuación las bases de la convocatoria de la 1a campaña a llevar a cabo.

Objeto y descripción de la campaña: Como parte de la iniciativa BioASQ, la OTG de Sanidad del Plan TL organiza la tarea: Indización semántica médica en español (BioASQ MESINESP Task), la primera tarea sobre indexación semántica de textos médicos en español.

La tarea se basará en el proceso estándar seguido por IBECS y LILACS para indexar los resúmenes de revistas en español. Se les pedirá a los participantes que clasifiquen los nuevos documentos de IBECS/LILACS, escritos en español. La evaluación de los sistemas participantes se realizará una sólo vez con un conjunto anotado manualmente, expresamente creado para la tarea.

Véase más información en siguiente enlace: <http://temu.bsc.es/mesinesp/>

Inscripción: Para inscribirse como participante en la tarea objeto de la actual convocatoria, los participantes deben rellenar el formulario de registro que se encuentra en el siguiente enlace:

<http://participants-area.bioasq.org/accounts/register/>

Plazos de presentación: La tarea se realizará en invierno del 2019-2020 (se anunciará el programa detallado). Los participantes, después de descargar los conjuntos de pruebas publicados, deberán enviar los resultados dentro de un período de tiempo limitado. Más información de las fechas de la campaña en el siguiente:

<http://temu.bsc.es/mesinesp/index.php/participation/#schedule>

Evaluación: La evaluación de los sistemas participantes se realizarán comparándolos contra un conjunto de datos anotado manualmente especialmente para la tarea. Los sistemas participantes serán evaluados en base a una medida plana: micro medida-F basada en etiquetas. Más información en el siguiente enlace:

<http://temu.bsc.es/mesinesp/index.php/evaluation/>

Comité organizador: Esta tarea ha sido organizada por la OTG de Sanidad del Plan TL en colaboración con el Centro Nacional de Investigación Oncológica (CNIO), el Centro Nacional de Investigación Científica Demokritos (Grecia) y la Biblioteca Nacional de Ciencias de la Salud de El Instituto de Salud Carlos III. Más información de los miembros del equipo organizador puede encontrarse en el siguiente enlace:

<http://temu.bsc.es/mesinesp/index.php/organising-committee/>

Comité científico evaluador:

<http://temu.bsc.es/mesinesp/index.php/scientific-committee/>

Selección de ganadores: La tarea será evaluada en base a una medida plana: micro medida-F basada en etiquetas. Los equipos participantes clasificados en los tres primeros lugares recibirán un premio. Para más detalles, consulte el siguiente enlace:

<http://temu.bsc.es/mesinesp/index.php/premios/>

Generación de un Corpus Silver Standard público: Los participantes deberán anotar 1 millón de documentos heterogéneos, llamado el Conjunto de fondo, que incluirá los 500 documentos que forman el Conjunto de Evaluación. Estas anotaciones generarán un Corpus Silver Standard público de documentos indexados semánticamente.

Presupuesto y cuantía de las ayudas: El presupuesto para esta convocatoria es de 3.000 euros. El equipo participantes clasificado en primera posición recibirá un premio de 1.500 euros, el equipo clasificado segundo, un premio de 1.000 euros, y el equipo clasificado en tercera posición, un premio de 500 euros.

Hay un premio adicional de 500 euros para cualquier equipo clasificado entre los cinco primero que publique su sistema en GitHub como componente de código abierto. La plantilla requerida para el archivo README se puede descargar de: <http://github.com/PlanTL-SANIDAD/MESINESP>

Información de contacto: PlanTL-Sanidad@bsc.es

Barcelona, 10 de diciembre de 2019