

[Inicio](#) > El BSC aprovecha los últimos avances en IA para desarrollar sistemas que permitan aprender terminologías en español para el proyecto TeresIA

[El BSC aprovecha los últimos avances en IA para desarrollar sistemas que permitan aprender terminologías en español para el proyecto TeresIA](#)

El desarrollo de terminologías médicas, esencial en el proceso de digitalización del sector de la sanidad, es uno de los escenarios de aplicación en los que participa el BSC



Hay un creciente interés en el uso de la inteligencia artificial (IA) y tecnologías de procesamiento del lenguaje para la creación de terminologías relacionadas con ámbitos temáticos de alto impacto como pueden ser la salud, investigación biomédica o el dominio legal. Recursos tales como diccionarios, ontologías o vocabularios controlados juegan un papel fundamental en la caracterización, clasificación y el aprovechamiento de los ingentes volúmenes de datos textuales que se generan a diario.

Las terminologías están integradas en una gran variedad de soluciones tecnológicas tales como herramientas avanzadas de búsqueda, sistemas de traducción automática, IA conversacional o sistemas de pregunta-respuesta.

Por lo tanto, es necesario impulsar el desarrollo de sistemas de IA que aprendan a reconocer automáticamente términos científicos o técnicos, así como detectar las relaciones existentes entre estos términos para contenido en español. Esto permitiría el desarrollo de recursos computacionales que sean capaces de enriquecer sistemáticamente terminologías existentes. Estos aspectos forman parte de los objetivos básicos del proyecto TeresIA - Terminologías en España y servicios de inteligencia artificial en la que participa la unidad de Procesamiento del lenguaje natural para el análisis de información biomédica (NLP4BIA) del BSC.

El proyecto TeresIA se engloba dentro de los ejes estratégicos de la Estrategia Nacional de Inteligencia Artificial (ENIA), que tiene como objetivo desarrollar plataformas de datos e infraestructuras tecnológicas que den soporte a la IA y contribuyan a impulsar el proceso de transformación digital del país. El proyecto TeresIA fue presentado también a la Comisión Europea el pasado diciembre, obteniendo una [dotación económica de 1,4 millones de euros de la Secretaría de Estado de Digitalización e Inteligencia Artificial del Ministerio de Transformación Digital](#). Además, TeresIA también forma parte de la [Estrategia de Inteligencia Artificial 2024](#) aprobada por el gobierno en el mes de mayo, y ha recibido un galardón en la [categoría de emprendimiento e investigación por los Premios Internet](#).

TeresIA, liderado por el grupo de Elea Giménez del Consejo Superior de Investigaciones Científicas (CSIC), cuenta además con la participación de una diversidad de grupos que incluyen el Instituto Cervantes, la Universidad Politécnica de Madrid, la Asociación Española de Terminología y la Dirección General de Traducción de la Comisión Europea.

El proyecto TeresIA está directamente vinculado al desarrollo de una plataforma pública de terminología para dar acceso unificado a toda la terminología del español mediante un potente metabuscador. El desarrollo de este portal contribuye, por tanto, a la apertura de terminologías y a su uso compartido por la comunidad científica en primera instancia, y adicionalmente del tejido productivo español. Para esto, TeresIA aprovecha modelos de lenguaje y sistemas de IA junto con el trabajo cualitativo de especialistas en distintos ámbitos temáticos, así como de lingüistas para implementar sistemas inteligentes de generación de recursos terminológicos. Para la implementación de tales sistemas basados en modelos de IA es esencial el acceso a recursos computacionales como los del MareNostrum 5, parte de las instalaciones del BSC.

La unidad de NLP4BIA del BSC participa en aspectos tecnológicos clave del proyecto TeresIA como son el acceso a contenidos, el desarrollo de conjuntos de datos para el entrenamiento de sistemas inteligentes y la implementación de herramientas que aprovechen los modelos de lenguaje para la extracción de términos y relaciones semánticas. El BSC también se encarga de la evaluación técnica de los algoritmos de extracción de términos analizando aspectos relacionados con la calidad, robustez, interoperabilidad y escalabilidad de las soluciones implementadas. La participación del BSC en TeresIA pasa por el desarrollo de KeyCARE, una librería diseñada para la extracción de términos clave, su clasificación y la extracción de relaciones entre términos. Esta librería ya ha sido publicada y en noviembre será presentada en el Congreso de la Sociedad Española de Ingeniería Biomédica 2024 en Sevilla.

Entre los escenarios de aplicación y dominios de uso contemplados por el proyecto TeresIA en los que participa el BSC se encuentran el desarrollo de terminologías médicas, que tienen un papel esencial en el proceso de digitalización del sector de la sanidad o el aprovechamiento de terminologías para sistemas de recuperación de información científica multilingüe.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 19 Nov 2024 - 14:51): <https://www.bsc.es/es/noticias/noticias-del-bsc/el-bsc-aprovecha-los-%C3%BAltimos-avances-en-ia-para-desarrollar-sistemas-que-permitan-aprender>