

[Inici](#) > El BSC aprofita els últims avenços en IA per desenvolupar sistemes que permetin aprendre terminologies en espanyol per al projecte TeresIA

[El BSC aprofita els últims avenços en IA per desenvolupar sistemes que permetin aprendre terminologies en espanyol per al projecte TeresIA](#)

El desenvolupament de terminologies mèdiques, essencial en el procés de digitalització del sector de la sanitat, és un dels escenaris d'aplicació en què participa el BSC



Hi ha un interès creixent en l'ús de la intel·ligència artificial (IA) i tecnologies de processament del llenguatge per a la creació de terminologies relacionades amb àmbits temàtics d'alt impacte com poden ser la salut, la investigació biomèdica o el domini legal. Recursos com ara diccionaris, ontologies o vocabularis controlats tenen un paper fonamental en la caracterització, classificació i aprofitament dels ingents volums de dades textuales que es generen diàriament.

Les terminologies estan integrades en una gran varietat de solucions tecnològiques com ara eines avançades de cerca, sistemes de traducció automàtica, IA conversacional o sistemes de pregunta-resposta.

Per tant, cal impulsar el desenvolupament de sistemes de IA que aprenguin a reconèixer automàticament termes científics o tècnics, així com detectar les relacions existents entre aquests termes per a contingut en espanyol. Això permetria el desenvolupament de recursos computacionals que siguin capaços d'enriquir sistemàticament terminologies existents. Aquests aspectes formen part dels objectius bàsics del projecte

TeresIA - Terminologies a Espanya i serveis d'intel·ligència artificial on participa la unitat de Processament del llenguatge natural per a l'anàlisi d'informació biomèdica (NLP4BIA) del BSC.

El projecte TeresIA s'engloba dins dels eixos estratègics de l'Estratègia Nacional d'Intel·ligència Artificial (ENIA), que té com a objectiu desenvolupar plataformes de dades i infraestructures tecnològiques que donin suport a la IA i contribueixin a impulsar el procés de transformació digital del país. El projecte TeresIA també va ser presentat a la Comissió Europea el desembre passat, obtenint [una dotació econòmica d'1,4 milions d'euros de la Secretaria d'Estat de Digitalització i Intel·ligència Artificial](#) del Ministeri de Transformació Digital. A més, TeresIA també forma part de l'[Estratègia d'Intel·ligència Artificial 2024](#) aprovada pel govern al mes de maig, i ha rebut un guardó a la [categoria d'emprenedoria i investigació pels Premis Internet](#).

TeresIA, liderat pel grup d'Elea Giménez del Consell Superior d'Investigacions Científiques (CSIC), compta a més a més amb la participació d'una diversitat de grups que inclouen l'Institut Cervantes, la Universitat Politècnica de Madrid, l'Associació Espanyola de Terminologia i la Direcció General de Traducció de la Comissió Europea.

El projecte TeresIA està directament vinculat al desenvolupament d'una plataforma pública de terminologia per donar accés unificat a tota la terminologia de l'espanyol mitjançant un metabuscador potent. El desenvolupament d'aquest portal contribueix, doncs, a l'obertura de terminologies i al seu ús compartit per la comunitat científica en primera instància, i addicionalment del teixit productiu espanyol. Per això TeresIA aprofita models de llenguatge i sistemes d'IA juntament amb el treball qualitatiu d'especialistes en diferents àmbits temàtics, així com de lingüistes per implementar sistemes intel·ligents de generació de recursos terminològics. Per a la implementació d'aquests sistemes basats en models d'IA és essencial accedir a recursos computacionals com els del MareNostrum 5, part de les instal·lacions del BSC.

La unitat de NLP4BIA del BSC participa en aspectes tecnològics clau del projecte TeresIA com són l'accés a continguts, el desenvolupament de conjunts de dades per a l'entrenament de sistemes intel·ligents i la implementació d'eines que aprofitin els models de llenguatge per a l'extracció de termes i relacions semàntiques. El BSC també s'encarrega de l'avaluació tècnica dels algorismes d'extracció de termes analitzant aspectes relacionats amb la qualitat, la robustesa, la interoperabilitat i l'escalabilitat de les solucions implementades. La participació del BSC a TeresIA passa pel desenvolupament de KeyCARE, una llibreria dissenyada per a l'extracció de termes clau, la seva classificació i l'extracció de relacions entre termes. Aquesta llibreria ja ha estat publicada i al novembre serà presentada al Congrés de la Societat Espanyola d'Enginyeria Biomèdica 2024 a Sevilla.

Entre els escenaris d'aplicació i dominis d'ús contemplats pel projecte TeresIA on participa el BSC hi ha el desenvolupament de terminologies mèdiques, que tenen un paper essencial en el procés de digitalització del sector de la sanitat o l'aprofitament de terminologies per a sistemes de recuperació d'informació científica multilingüe.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 18 nov 2024 - 16:44): <https://www.bsc.es/ca/noticies/noticies-del-bsc/el-bsc-aprofita-els-%C3%BAltims-aven%C3%A7os-en-ia-desenvolupar-sistemes-que-permetin-aprendre-terminologies>