

[El projecte MarIA del BSC, premi Archiletras de la Lengua a la innovació](#)

Es tracta d'un sistema d'intel·ligència artificial massiu i expert en comprendre i escriure en llengua espanyola, creat a partir del patrimoni documental digital de la Biblioteca Nacional d'España.



El projecte MarIA, el sistema de models de llengua creat al Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS) a partir dels arxius web de la Biblioteca Nacional d'España (BNE), i emmarcat i finançant amb el Pla de Tecnologies del Llenguatge de la Secretaria d'Estat de Digitalització i Intel·ligència Artificial (SEDIA), ha estat guardonat amb el Premi a la innovació a la primera edició dels premis Archiletras de la Lengua.

Els premis Archiletras, de caràcter anual i convocats per l'editorial Prensa y Servicios de la Lengua, reconeixen els mèrits en la promoció, suport, investigació i desenvolupament de la llengua espanyola o d'algunes de les altres llengües en contacte amb l'espanyol en qualsevol dels seus àmbits territorials.

El lliurament del guardó es va fer aquest dijous a la Casa Amèrica de Madrid, en un acte en què el BSC va estar representat per Marta Villegas, responsable del projecte i líder de la Unitat de Minería de Textos del BSC, que va rebre el premi de mans de Carme Artigas, Secretària d'Estat de Digitalització i Intel·ligència Artificial.

“Rebre el premi Archiletras ens fa especialment feliços. És un honor i un reconeixement a l'equip de

professionals entusiastes del BSC que, en col·laboració amb la BNE i la SEDIA, ha treballat perquè l'espanyol disposi de recursos lingüístics suficients i de qualitat”, ha assegurat Marta Villegas.

MarIA es va imposar a la votació final del jurat als altres dos finalistes, l'aplicació mòbil Dialectos del Español, dissenyada per detectar i predir trets generals i característics de tots els dialectes del món hispanoparlant, i Euskal Herriko Ahotsak (Veus del País Basc), projecte que recopila i difon el patrimoni cultural oral i dialectal basc.

MarIA situa la llengua espanyola entre els idiomes que disposen de models massius d'accés obert

El projecte MarIA és un sistema d'intel·ligència artificial massiu i expert a comprendre i escriure en llengua espanyola. Pel volum i les capacitats, ha situat la llengua espanyola entre el grup dels idiomes que disposen de models massius d'accés obert, després de l'anglès i el mandarí.

Un model de llenguatge és un sistema d'intel·ligència artificial format per una xarxa neuronal profunda entrenada per adquirir una comprensió de la llengua, el seu lèxic i els seus mecanismes per expressar el significat i escriure com un humà.

Aquests models estadístics complexos, que relacionen paraules en textos de manera sistemàtica i massiva, són capaços d'“entendre” no només conceptes abstractes, sinó també el context dels mateixos. Amb aquests models, els desenvolupadors de diferents aplicacions poden crear eines per a múltiples usos, com ara classificar documents o crear correctors o eines de traducció.

MarIA s'ha construït a partir del patrimoni documental digital de la BNE, que rastreja i arxiva les webs elaborades en espanyol i s'ha entrenat amb el superordinador MareNostrum 4 del BSC. Es publica en obert perquè els desenvolupadors d'aplicacions, companyies, grups de recerca i la societat en general puguin utilitzar-lo en infinitat d'usos.

Els darrers avenços de MarIA constitueixen una fita en la consecució d'objectius de l'Estratègia Nacional d'Intel·ligència Artificial i del Pla de Recuperació, Transformació i Resiliència, amb què Espanya pretén liderar a nivell mundial el desenvolupament d'eines, tecnologies i aplicacions per a la projecció i ús de la llengua espanyola als àmbits d'aplicació de la IA.

MarIA està igualment vinculat al Projecte estratègic per a la recuperació i transformació econòmica (PERTE) Nova economia de la llengua, plantejat com una oportunitat per aprofitar el potencial de l'espanyol i de les llengües cooficials com a factor de creixement econòmic i competitivitat internacional en àrees com la intel·ligència artificial, la traducció, l'aprenentatge, la divulgació cultural, la producció audiovisual, la investigació i la ciència.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 19 set 2024 - 07:14): <https://www.bsc.es/ca/noticies/noticies-del-bsc/el-projecte-maria-del-bsc-premi-archiletras-de-la-lengua-la-innovaci%C3%B3>