

Sobre Aina

Neix 'AINA', el projecte per garantir el català en l'era digital

- El projecte AINA generarà els recursos digitals i lingüístics necessaris per facilitar el desenvolupament d'aplicacions basades en la Intel·ligència Artificial i les Tecnologies de la Llengua com ara els assistents de veu, traductors automàtics o agents conversacionals en català
- L'objectiu últim és que la ciutadania pugui participar en català en el món digital al mateix nivell que els parlants d'una llengua global, com ara l'anglès, i evitar així l'extinció digital de la llengua
- El primer recurs generat és el 'corpus' del català per entrenar els algoritmes d'Intel·ligència Artificial (IA), el més gran creat fins al moment, amb 1.770 milions de metadades associades a paraules
- El proper pas serà generar els models de la llengua, models de la parla i models per a la traducció utilitzant xarxes neuronals multicapa perquè les empreses que creen aplicacions basades en intel·ligència artificial (IA), com ara assistents de veu, traductors automàtics, agents conversacionals, etc., puguin fer-ho fàcilment en català.

Pressupost i abast del projecte

El projecte AINA té un pressupost global de 13,5 milions d'euros per al període 2020-2024 i arrenca amb una fase inicial de 250.000€ del Departament de Polítiques Digitals ha assignat al BSC per ampliar els corpus de la llengua catalana i així obtenir models lingüístics que abastin les diferents variants i registres.

El projecte es beneficia de la col·laboració del Plan de Impulso de las Tecnologías del Lenguaje, de la Vicepresidència d'Afers Econòmics i Transformació Digital del Govern espanyol.

Ara, amb l'impuls del Departament de Polítiques Digitals, es crearan nous corpus per incorporar les diferents variants dialectals del català, diferents registres lingüístics (col·loquial, literari, administratiu, etc.) i arxius de veu i imatge. La Corporació Catalana de Mitjans Audiovisuals aportarà tot el seu repositori documental.

Tots els models que crearà el BSC estaran a disposició de totes aquelles empreses o entitats que les vulguin fer servir, ja que es publicaran en obert i amb llicències permissives.

El món digital, una oportunitat i un repte per al català

Això ha de permetre que el català faci un salt qualitatiu i quantitatiu en l'ecosistema digital. De fet, el món digital és avui una oportunitat i un repte per a la llengua catalana. Actualment, les tecnologies de veu i les aplicacions i interfícies de veu per a l'accés al món digital són estratègiques per al desenvolupament ple de la llengua en tots els sectors. La interacció entre les persones i la tecnologia ha entrat en una nova fase on cada vegada menys es fa a través de dispositius com el teclat, el ratolí o les pantalles tàctils, per donar pas a una nova forma d'interacció més natural a través de la veu i la parla. I aquest fet atorga especial rellevància a la llengua, que esdevé un dels principals vehicles d'interacció. Aquesta nova interacció ha de ser possible fer-

la, també, en català.

La IA i les Tecnologies del Llenguatge

Les Tecnologies del Llenguatge són aquelles que ja utilitzem en el nostre dia a dia quan corregim de forma automàtica un text al correu electrònic, utilitzem un cercador web a Internet, traduïm de forma automàtica una pàgina web, donem ordres de veu al telèfon mòbil, interactuem amb assistents virtuals o seguim les indicacions del navegador GPS, entre d'altres. I són les tecnologies que permetran dialogar amb ordinadors, aparells domèstics i fins i tot amb el nostre vehicle de manera natural.

Les noves tecnologies d'Intel·ligència Artificial i les Tecnologies del Llenguatge es basen en l'aplicació d'algorismes sobre grans conjunts de dades de qualitat, però els conjunts de dades sobre els quals s'entrenen els algorismes són específics per a cada llengua.



Generalitat de Catalunya
**Departament de Polítiques Digitals
i Administració Pública**



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

Plan TL

Plan de Impulso de las
Tecnologías del Lenguaje



Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 24 nov 2024 - 15:12): <https://www.bsc.es/ca/research-and-development/not-assigned-pages/sobre-aina>