

La normalització del català en l'àmbit digital pren cos

■ El projecte Aina enllesteix el corpus lingüístic per preparar la intel·ligència artificial que faran servir els assistents de veu ■ El Barcelona Supercomputing Center participa en el disseny científic i el càlcul

Xavi Aguilar
BARCELONA

L'estratègia de la Generalitat de garantir la supervivència i competitivitat del català en un entorn cada cop més globalitzat i digital comença a prendre forma. El conseller de Polítiques Digitals i Administració Pública, Jordi Puigneró, en va oferir ahir els primers resultats en la presentació del projecte Aina, un programa amb un pressupost de 13,5 milions d'euros per a quatre anys que vol dotar la llengua pròpia del país dels recursos digitals i lingüístics per tal que les empreses que creen aplicacions basades en intel·ligència artificial, com ara assistents de veu i traductors automàtics,

puguin fer-ho fàcilment en català.

El Barcelona Supercomputing Center (BSC) serà un agent cabdal en aquesta adaptació a l'entorn virtual, ja que hi aporta el coneixement científic i la capacitat de càlcul per aixecar el projecte. De moment, ja s'ha elaborat el corpus lingüístic textual, el més gran que s'ha fet mai de la llengua catalana, consistent en 1.770 milions de paraules reunides en 95 milions de frases. Per construir-lo s'han descarregat i processat textos del *Diari Oficial de la Generalitat de Catalunya* i els seus webs, l'Agència Catalana de Notícies, els 500 dominis .cat amb més trànsit, la Viquipèdia i altres corpus i bases de da-

Les frases

“Si la dictadura no va eliminar el català, tampoc no ho farà la digitalització. Alexa parlarà català abans que es pugui fer al Congrés o al Suprem”

des ja existents.

Tot plegat va generar un volum d'informació que el superordinador *Ma-reNostrum* va polir i netejar de duplicats amb 2.000 hores de processadors. Ara 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, que es nodriran del repositori documental de la Corporació Catalana de Mitjans Audiovisuals.

Jordi Puigneró
CONSELLER DE POLÍTQUES DIGITALS

Amb tota aquesta informació, el següent pas serà entrenar xarxes neuronals multicapa perquè “aprenquin el català” i generin models de la llengua, models de la parla i models per a la traducció. Es tracta d'eines molt costoses de crear, perquè necessiten

gran capacitat de càlcul, però que es posaran a disposició dels desenvolupadors per crear assistents de veu, predictors i correctors lingüístics, xatbots, aplicacions de resum automàtic, cerques intel·ligents, aplicacions per a l'anàlisi de sentiments o motors de traducció i subtitulació automàtica, entre altres.

Batejat en honor d'Aina Moll, filòloga menorquina que des de la direcció general de Política Lingüística de la Generalitat va ser figura central de la normalització del català, aquest projecte arrenca amb una inversió inicial de 250.000 euros del govern al BSC i espera rebre finançament dels fons europeus Next-GenerationEU. ■

SOCIETAT La Intersindical alerta de la docència amb Covid

La Intersindical-CSC ha denunciat que el sistema educatiu “està debilitat” i que “no es donen les condicions adequades per atendre els alumnes en el marc de la pandèmia”, alhora que ha fet un crida a la mobilització dels professionals del sector. El sindicat reconeix amb la campanya *Imprescindibles* tots els professionals de l'educació perquè “fan possible” que els centres educatius funcionin. ■ R.M.B.

SOCIETAT La taula del tercer sector reclama la llei

La taula del tercer sector confia que la llei del tercer sector social a Catalunya sigui “de les primeres iniciatives legislatives” que s'engeguin la legislatura vinent, un cop passades les eleccions del 14-F, segons va dir la presidenta, Francina Alsina, ahir en una jornada, segons l'ACN. L'entitat lamenta que el fet de no tenir un marc legal català els dificulta que la tasca “sigui reconeguda.” ■ REDACCIÓ

Vins
Caves
Destil·lats
Cerveses

Torrons
Neules
Olis
Gourmet

món la cata

VINITECA CENTRAL DEL MARESME ARENYS DE MUNT

Pernils
Formatges
Embotits
Foies

Lots
Regals
Cistelles
Estoigs