

ECIA: El Català i la IA

Description

El projecte té com a objectiu generar els recursos lingüístics necessaris per al català, que permetin i facilitin desenvolupar aplicacions en llenguatge natural com chat bots, traductors automàtics, aplicacions d'extracció d'informació o resum automàtics. La tecnologia de la llengua ha entrat de ple en el món de l'aprenentatge profund i el big data, per tant, cal posar a l'abast de la comunitat científica i de la indústria, models de llengua en català preentrenats sobre grans quantitats de dades lingüístiques de qualitat.

Es produirà també un bon nombre de dades anotades que serviran per entrenar i, posteriorment, avaluar models per a tasques específiques (com ara sistemes de pregunta resposta, de classificació semàntica i d'altres tasques que impliquen comprensió del llenguatge). La generació de dades manualment anotades és un procés crític i molt costós que aquest projecte pretén resoldre.

Finalment, es desenvoluparà una aplicació de benchmarking que permetrà l'avaluació contínua de models i sistemes utilitzant dades anotades. L'objectiu final de l'aplicació és impulsar la investigació en el desenvolupament de sistemes generals i robusts de comprensió del llenguatge natural.

El projecte tindrà un gran impacte, tant en la indústria de la intel·ligència artificial i del processament del llenguatge com en els ciutadans; per primera vegada, el català disposarà de recursos suficients per a poder ser integrat en aplicacions intel·ligents de manera fàcil i assequible i els ciutadans podran fer vida digital en català amb normalitat. El desenvolupament d'aquestes aplicacions es preveu com a part del projecte [AINA](#).

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

Source URL (retrieved on 19 Oct 2024 - 09:42): <https://www.bsc.es/es/research-and-development/projects/ecia-el-catal%C3%A0-i-la-ia>