

Euskaltzaindiak Lexikoaren Behatokia corpora eguneratu du

diariovasco.com/zabalik/euskaltzaindiak-lexikoaren-behatokia-20210416151108-nt.html

April 16, 2021



Aldian-aldian eguneratuta, jada ia ehun milioi hitzek osatzen dute, 2009an baino ia 25 aldiz gehiago



Andres Urrutia, artxiboko argazki batean. / Iñigo Arizmendi

Egunkarietako albisteak, aldizkari orokorretako testuak, herri-aldizkari batzuenak, literaturan eta saiakeran ondutako testuak, hezkuntzako materiala eta administrazio materiala baliatuta Euskaltzaindiak 'Lexikoaren Behatokia' corpora eguneratu du eta **jada 98.060.414 testu-hitz dauzka, 2009an baino ia 25 aldiz gehiago**. Orduan 4 milioi zituen eta EHUKo Donostiako Informatika Fakultateko IXA taldea, Elhuyar Fundazioa eta UZEI ditu bidelagun izanda «XXI. mendeko euskararen corpus zabala eta eguneratua» osatzea lortu da.

2008an mende berriaren beharrei erantzuteko jarri zuen abian Euskaltzaindiak Lexikoaren Behatokia «hizkuntzaren bilakaera hurbiletik jarraitzeko, hitz eta adierazmolde berriak ezagutzeko, arauak zenbateraino betetzen edo urratzen diren jakiteko eta, hori guztia ezagututa, arauak finkatzeko. bai eta hizkuntzaren erabilera

sakonago ezagutzeko, gramatika edo estilistika aldetik, erregistroen ezaugarriak aztertze edo ikuspegi soziolinguistikotik ikertze nahiz edukien ideologia, historia eta abar aztertze».

«Lan arauemailea egiteko»

Euskaltzaindiak «hiztegia aztertze edo lan arauemailea egiteko» baliatzen du corpusa eta une oro akademiaren webgunean kontsulta daiteke. Etiketatua eta automatikoki lematizatua da, hau da, hizkuntzaren erabilera hurbiletik jarraitzea ahalbidetzen du. Gainera, erabiltzaileak testu-hitzak bila ditzake era arrunt, aurreratu zein konparazio bidezkoan.

Egitasmoak aurrera dihardu aipatutako kideen laguntzarekin eta corpusa testuz elikatze Euskaltzaindiak hainbat argitaratzaileekin hitzarmenak sinatu ditu.

Temas

Zabalik

Comentarios