

Click!

Viernes, 15 de diciembre de 2006

Registro Hemeroteca BUSCAR en Internet

PORTADA ÚLTIMA HORA ECONOMÍA DEPORTES OCIO PARTICIPACIÓN

Servicios | Clasificados | Tienda

EDICIÓN IMPRESA

EDICIÓN IMPRESA

EL DIARIOVASCOGipuzkoa
San SebastiánEdiciones Al día
Opinión
Política
Deportes
Economía
Mundo
Cultura
Gente
TV
Contraportada
De un vistazo
Ejemérides

ALDIA

Elhuyar eta IXA-k 'Zientzia eta Teknologiaren Corpusa' sortu dute

Euskarazko lehen corpus espezializatua da eta dagoeneko sarean kontsulta daiteke

DONOSTIA. DV. Elhuyar Fundazioak eta Euskal Herriko Unibertsitateko IXA taldeak zientzia eta teknologiaren alorreko euskara ikertzeko hizkuntza- baliabide bat garatu dute: Zientzia eta Teknologiaren Corpusa. Euskarazko lehen corpus berezi edo espezializatua da.

 [Publicidad](#)Click!
Click!
Click!

Corpusa elikatzeke, 1990-2002 bitartean argitaratuko zientzia eta teknologiaren alorreko obrak hartu dira kontuan. Corpus egituratua eta linguistikoki etiketatua da, automatikoki prozesatua, eta 8 milioi hitz ditu, horietatik 1,6 milioi eskuz berrikusita eta zuzenduta daudela. Corpusa www.ztcorpusa.net web gunean kontsultagai dago.

Elhuyar Fundazioak eta EHUko IXA taldeak sortutako euskarazko lehen corpus espezializatu hau sarean dago dagoeneko. Zientzia eta teknologiaren alorreko euskarazko testuen bilduma egituratu eta etiketatua da, eta alor horietako euskararen erabilera ikertzeko baliabidea izatea du helburu nagusia.

Corpusa etiketatuta dago, bai testuaren egiturari eta formatuari dagokionez, bai linguistikoki, eta testuko hitz bakoitzaren lema eta kategoria/azpikategoria etiketatuta dira. Corpusaren lehen bertsio honetan, 8 milioi hitz daude, eta horietatik, 1,6 milioi hitz eskuz berrikusi, desanbiguatu eta zuzendu dira. Bera kontsultatzeko interfaze ahalsua antolatu da, eta erabiltzaileak era askotako bilaketa bakunak eta konplexuak egiteko aukera izango du, horretarako parametro-multzo zabala erabiliz: lema, testu-forma, kategoria, eremua, generoa eta corpus-atala (eskuz zuzendua/corpus osoa), besteak beste.

2007tik aurrera ELDAren (Hizkuntza Baliabideen Ebaluazio eta Banaketarako Agentzia Europarra) baliabideen artean egongo da corpusa, ustiapen komertzialerako eskuragarri, lizentzia bidez.

Hizking21 ikerketa estrategikoko proiektuaren barnean hasi zen egiten corpusa. Hizking21 proiektuak honako laguntza hauek jaso ditu: Eusko Jaurlaritzaren Industria Sailaren Etorrek programa (2002-2004) eta Gipuzkoako Foru Aldundiaren Gipuzkoako Zientzia, Teknologia eta Berrikuntza Sarea programa (2004).

Subir

vocento

© Copyright DIARIOVASCO.COM DIGITAL VASCA, S.A.
Camino de Portuetxe, 2. 20018 San Sebastián
CIF A-20677878RSS
Powered by
SARENET

Reservados todos los derechos. Queda prohibida la reproducción, distribución, comunicación pública y utilización, total o parcial, de los contenidos de esta web, en cualquier forma o modalidad, sin previa, expresa y escrita autorización, incluyendo, en particular, su mera reproducción y/o puesta a disposición como resúmenes, reseñas o revistas de prensa con fines comerciales o directa o indirectamente lucrativos, a la que se manifiesta oposición expresa.

[Contactar](#) | [Mapa web](#) | [Aviso legal](#) | [Política de privacidad](#) | [Publicidad](#) | [Master El Correo](#)