

Web mining o mineria de web

Web mining o mineria de web?

ÀREA TEMÀTICA: **Terminologia. Informàtica. Telecomunicacions**

TÍTOLS ASSOCIATS: **minería de web, web (minería de), minería de dades**

DATA D'ACTUALITZACIÓ: 04/07/2003

SOLUCIÓ PROVISIONAL: **Mineria de web**

ARGUMENTACIÓ:

En diversos textos elaborats per l'Idescat apareixien els termes *data mining* i *web mining*:

- *Els objectius es resumeixen en el desenvolupament d'una metodologia, basada en **web mining** i *data mining*, que permeti correlacionar les dades estadístiques de llars i de petites i mitjanes empreses amb les d'ús de webs*

El primer va passar en el seu moment pel Consell Supervisor del Termcat i es pot consultar a la base de dades Cercaterm, i a més se'n pot llegir el comentari lingüístic en el butlletí **IDESCAT LLENGUA** número 15 o a la mateixa web de llengua i terminologia de l'Institut d'Estadística de Catalunya <http://www.idescat.es/publicacions/llengua/idescatllengua.stm>. La *minería de dades* (castellà: *extracción de datos, minería de datos*; anglès: *data mining*) és una "tècnica informàtica que consisteix a analitzar un gran volum d'informació emmagatzemada en diferents bases de dades a fi de deduir patrons de coneixement que puguin generar aplicacions pràctiques".

La *minería de web* és una extensió de la minería de dades que consisteix a aplicar-ne les tècniques a documents i serveis del web. Es tracta, doncs, de descobrir la informació o el coneixement potencialment útils i prèviament desconeguts a partir de dades web. Totes les persones que visiten un lloc a Internet deixen empremtes digitals (adreces d'IP, navegador, galetes, etc.) que els servidors automàticament emmagatzemen en un diari d'accessos (*log*). Les eines de *minería de web* analitzen i processen aquests diaris per produir informació significativa.

Una vegada definit el terme, cal veure quina és la forma més adequada de representar-lo en català. Atès que el Consell Supervisor del Termcat ja va aprovar *minería de dades* com a terme equivalent de l'anglès *data mining*, sembla lògic mantenir la mateixa estructura amb aquest nou terme, sobretot perquè n'és una tècnica anàloga: *minería de web* (anglès: *web mining*). Només queda un dubte: si *web* ha de ser en plural o no. Per la definició o els comentaris fets per especialistes en la matèria (es tracta de fet de descobrir informació útil a partir de *dades web*, és a dir procedents d'un web), la solució més apropiada és en singular. A més, i a títol orientatiu, la forma *minería de web* es troba documentada en 75 webs i *minería web* en 57 webs (si més, segons la cerca feta en el Google). Caldrà veure, més endavant, quina és la solució definitiva que prendrà el Consell Supervisor del Centre de Terminologia Catalana per poder confirmar aquesta proposta de treball.