

Padicat, el patrimoni digital de Catalunya

DIGITAL

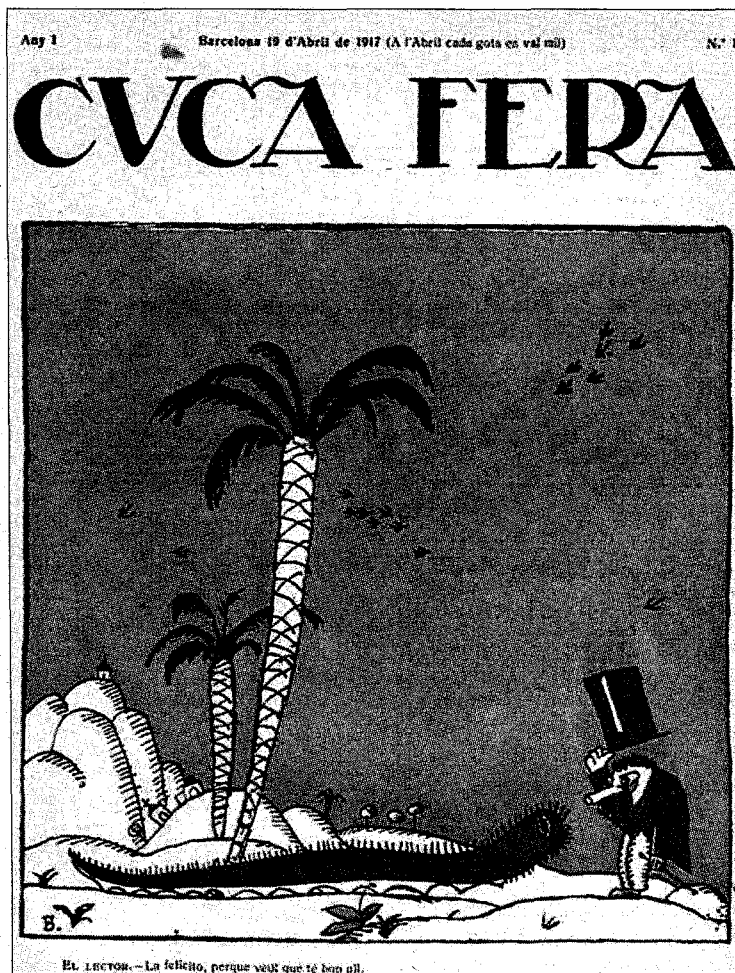
LAIA REVENTÓS

La memòria dels webs catalans s'anomena Padicat. Des del passat mes de juliol, aquest dipòsit de la Biblioteca de Catalunya es dedica a recopilar, conservar i difondre la producció cultural i científica catalana creada en format digital. L'objectiu és que Padicat tingui al 2009 una capacitat de 30 TB (Terabytes) i emmagatzemi 100.000 pàgines webs i 50 milions d'arxius, que corresponen aproximadament a 100.000 enciclopèdies catalanes. Aquesta informació estarà disponible gratuïtament a Internet, però no cal esperar dos anys per consultar el que arxiva: la pàgina ja està en marxa: www.padicat.cat.

Seguint l'exemple de l'Internet Archive (archive.org) que es dedica a guardar totes les pàgines webs que s'han creat al món des de l'any 2000, diversos països tracten de garantir, a través de les seves biblioteques nacionals, la pervivència del seu patrimoni digital per evitar que les generacions futures perdin la possibilitat de consultar aquests continguts en formats electrònics: un web té 44 dies de vida de mitjana.

Hi ha tres models diferents de conservació: l'exhaustiu, el selectiu i el mixt. Als països escandinaus conserven tot allò publicat en format digital dirigit al seu públic, sigui sota el domini nacional o els dominis '.com', '.net', '.org'... A Dinamarca, a més, també recopilen el que es diu a Internet sobre un esdeveniment concret de la vida social del país. Els britànics i australians prefereixen, en canvi, ser més selectius: un consell científic de l'arxiu web selecciona allò que consideren més interessant publicat en llengüa anglesa.

Padicat s'ha decantat pel model mixt: conservar-ho tot i alhora fer una selecció d'esdeveniments com, per exemple, el que s'ha dit a Internet durant la recent campanya a les eleccions al Parlament de Catalunya. Un robot és l'encarregat de rastrejar la xarxa en diferents freqüències temporals i, al cap d'un mes ja és a Internet per a la seva consulta gratuïta. Els servidors que guarden la informació són al Centre de Supercomputació de Catalunya (CESCA), soci tecnològic del projecte, dotat amb 800.000 euros.



Primer exemplar de la revista *Cuca Fera* que es pot consultar a Internet.

L'acord de fa uns 15 dies amb la Fundació Puntcat permetrà que els 18.000 webs sota aquest domini de la comunitat lingüística i cultural catalana també es conservin al dipòsit.

Si Padicat conserva el patrimoni digital, la Biblioteca de Catalun-

Publicacions antigues com 'Pèl i Ploma' ja són accessibles a la xarxa

ya també s'encarrega de fer el procés invers: convertir en format electrònic les revistes antigues. Es tracta d'un projecte cooperatiu de preservació digital en el qual participa el Consorci de Biblioteques Universitàries de Catalunya (CBUC), l'objectiu del qual és donar accés al ric fons bibliogràfic existent però dispers en diferents

espais, i garantir-ne la seva preservació.

Arca és el nom del dipòsit de les publicacions periòdiques ja tancades, representatives de la vida i societat catalana de finals del segle XIX i XX, com el setmanari satíric nacionalista *Cuca Fera*, *Pèl i Ploma*, *La Tralla*, *Luz* o *La carcajada*, entre d'altres.

Totes aquestes publicacions antigues es poden consultar per Internet —de cada publicació hi ha una ressenya sobre la seva història i les dates del primer i l'últim número— i també es poden descarregar a l'ordinador en format PDF. A més, com que s'han digitalitzat amb un programa de reconeixement de caràcters (OCR), es poden fer recerques internes de les dades i textos publicats a cada revista.

BIBLIOTECA NACIONAL DE CATALUNYA: www.bnc.cat