



Secció oberta

IV Col·loqui Aula d'Història del Periodisme Diari de Barcelona
Hemeroteques digitals: experiències

per Lluís Codina

Professor de Documentació dels Estudis de Periodisme
de la Universitat Pompeu Fabra

Introducció

La digitalització de la informació de premsa pot ser un dels principals aliats de la memòria històrica, en permetre l'accés a la informació amb facilitats inèdites fins ara, o pot ser una de les seves principals amenaces si no es preveuen al mateix temps condicions adequades d'identificació de les informacions i de preservació dels materials.

Amb aquestes preocupacions com un dels elements de discussió, l'Aula d'Història del Periodisme Diari de Barcelona —una realització conjunta dels Estudis de Periodisme de la Universitat Pompeu Fabra (UPF) i de l'Arxiu Històric de la Ciutat de l'Ajuntament de Barcelona— va organitzar, el 28 de maig de 2001, el seu IV Col·loqui, que en aquesta edició va dur per títol «Hemeroteques digitals: experiències».

Tot i que cada edició del Col·loqui és independent, el títol del tema expressava una clara voluntat de continuïtat amb l'anterior edició (maig de 2000) que va iniciar un debat sobre premsa, recerca històrica i digitalització de la cultura que continua inacabat.

Pel que fa a les ponències, en aquesta ocasió, el Col·loqui va comptar amb la col·laboració d'experts en documentació i arxius i de diverses empreses del sector, segons detallarem més endavant. En quant al suport organitzatiu, va comptar amb la col·laboració de l'Associació d'Arxivers de Catalunya, la Societat Catalana de Comunicació, el Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya i l'Observatori de la Comunicació Científica. Va presentar el Col·loqui Jaume Guillaumet, degà dels Estudis de Periodisme de la Universitat Pompeu Fabra i varen actuar com a moderadors Ramon Alberch, arxiver en cap de l'Ajuntament de Barcelona i Lluís Codina, professor de Documentació Periodística de la UPF.

Text o imatge

Ferran Agelet, responsable de projectes de l'Arxiu de l'Ajuntament de Barcelona, va dedicar la seva intervenció a presentar un estat de la qüestió en relació amb l'accés a col·leccions hemerogràfiques a través d'Internet i, en aquest sentit, va destacar la gran dicotomia actual que hi ha entre la presentació de la informació de premsa en forma textual i la presentació en forma gràfica.

De moment, és molt més habitual en Internet la presentació textual, que ell va denominar també *presentació de tipus documental*, per oposició a la presentació de tipus facsímil.

En la presentació textual, resulta relativament fàcil accedir a les informacions, principalment a través de les bases de dades documentals de

(cada cop més) diaris, però aquesta forma d'accés a la informació també suposa presentar-la de manera fora de context.

La descontextualització, segons va ressaltar Agelet, és múltiple: d'una banda es perd el format original de la notícia, ja que no només perdem la tipografia i la forma concreta de presentació original, sinó també les fotografies o altres materials no textuais que sovint acompanyen la notícia. D'altra banda, amb aquest tipus de presentació és impossible veure el context global de la publicació.

Aquesta presentació respon a una demanda real, de tipus documental, en la qual l'interès consisteix en saber *què* s'ha publicat en un o més diaris o revistes, sobre un tema determinat. A més, resulta la forma més fàcil de digitalització de la informació periodística. No és estrany, per tant, que sigui, pel moment, la fórmula més utilitzada.

Ara bé, aquesta manera de digitalitzar la informació periodística, com va assenyalar, pot anar en contra dels interessos de la recerca històrica i de la filosofia profunda del treball arxivístic. Aquesta última intenta no només conservar o preservar la informació, sinó preservar-la en el seu context i amb la màxima fidelitat possible a l'objecte original, sigui un còdex del segle xv o un diari del segle xx.

Per a garantir una digitalització que sigui el més fidel possible al document original, cal realitzar una digitalització del document que doni com a resultat una imatge facsímil del mateix. Aquest és un procés més problemàtic, ja que, com hem dit, d'una banda, significa emmagatzemar imatges, amb requeriments molt superiors d'espai de disc i temps de processament i de manipulació. D'altra banda, aquesta imatge no allibera de la necessitat d'emmagatzemar també el text, ja que sense aquest últim, les col·leccions només es podrien consultar de forma més o menys seqüencial.

Malgrat tot, a poc a poc, van sortint projectes relacionats amb aquest tractament integral de les col·leccions hemerogràfiques que contempla tant la imatge facsímil com el text. A banda d'experiències internacionals, com la del *Washington Post* i altres, al llarg del Col·loqui es van donar a conèixer les experiències de l'Ajuntament de Barcelona i de l'empresa Moebius Net amb el seu projecte *Prensa històrica*, de les que parlarem més endavant. Agelet també va proporcionar dades d'alguns recursos d'Internet relacionats amb el tema, que s'indiquen al final d'aquest article.

Experiències empresarials

Si en alguna cosa van estar d'acord les representants de les dues empreses que van intervenir a continuació, Begoña Aguilera (Delphi Infor-

mations Solutions) i Mayte Retana (Tecnodoc) va ser en els problemes que cal resoldre en els processos de digitalització de documents en paper. En primer lloc, estan l'elecció de l'equipament informàtic que cal utilitzar, segons va destacar Mayte Retana, ja que aquests processos requereixen un *hardware* molt específic. Així, les empreses com la seva es veuen obligades a ser especialment curoses amb les solucions informàtiques que recomanen als seus clients.

Un problema afegit, en aquesta ocasió de tipus cultural, segons Retana, és l'escassa predisposició de les empreses de premsa a invertir en el seu propi patrimoni, aspecte que no només es manifesta en la resistència a realitzar aquests tipus d'inversions, sinó en la destrucció dels seus propis fons per simple desinterès.

A més a més, Retana va explicar que la seva empresa és partidària de complementar els processos de digitalització d'imatges (facsimils) amb alguna forma d'indexació del text, bé per procediments automàtics —text complet— o bé per procediments intel·lectuals —descriptors—; tot i que ella es va decantar per la utilització de descriptors.

Les representants de les dues empreses van estar d'acord en que també hi ha encara incerteses sobre els procediments de conservació més adequats pel que fa als suports informàtics. Aguilera va fer èmfasi, no obstant això, que la preservació de materials informàtics no hauria de ser massa diferent de la filosofia general de preservació d'altres materials: també es fan malbé els documents en paper si no es renova l'atenció i la preocupació per la seva conservació adequada. Retana va assenyalar, per exemple, la paradoxa de que està molt ben conservada la premsa del segle XIX (i, per tant, també és més fàcil de digitalitzar) que la premsa dels anys trenta del segle XX.

En general, es va estar d'acord en què, pel moment, l'única opció realista en aquest sentit consisteix en preveure de forma sistemàtica el transvasament (*migració* és el terme en l'argot tècnic) de suport cada cert període de temps, per exemple, cada 5 anys, així com en utilitzar sempre els formats més oberts i més estandarditzats possibles. Això significa, per exemple, fer servir ara el CD-ROM com a suport i ben aviat el DVD, així com formats com el TIFF o el JPEG per a la imatge, codis ISO per al text, metadades, etc.

Aguilera (Delphi) va apostar de forma decidida pel màxim d'automatització possible en aquests casos, incloent l'ús de *software* de control d'escàners que sigui capaç d'interpretar de forma automàtica els diversos formats de pàgina de diversos diaris, de forma que la sortida sigui no només els facsimils de cada notícia i de cada pàgina completa, sinó també el text ben diferenciat per les zones principals: títol, autor, etc.

Segons Aguilera, l'objectiu principal és substituir el màxim de processos manuals per procediments automàtics. Va fer referència al projecte europeu Meta-e de digitalització de premsa històrica amb la utilització d'un estàndard de metadades (dades sobre el document incloses en el mateix arxiu informàtic que conté el document).

La digitalització de *El Mundo Deportivo*

Manuel Rovira, director de l'Arxiu Històric de la Ciutat (Ajuntament de Barcelona) va exposar l'experiència de la seva institució en la digitalització de la col·lecció *El Mundo Deportivo*, aparegut el 1906.

En el moment de la seva exposició, Rovira va explicar que s'havien digitalitzat més de 24.000 pàgines (en format TIFF i PDF), corresponents al període 1906-1938 de la publicació, però que això només era la primera fase del projecte, ja que l'objectiu era digitalitzar les més de 720.000 pàgines que, segons els seus càlculs, té la col·lecció entre 1906 i 1999.

El fet d'haver escollit aquesta publicació obeeix a una doble motivació: interès intrínsec extraordinari de la publicació com a font per a la recerca històrica i les facilitats proporcionades per l'empresa editora que posseeix la capçalera de la publicació.

Figura 1. Formulari de consulta de l'hemeroteca de l'Arxiu Històric de la Ciutat.

The image shows a screenshot of a web browser window displaying a search form for a digital archive. The browser's title bar reads "ABSYS WEB Arxiu Històric de la Ciutat - Microsoft Internet Explorer". The page header includes the "HEMEROTECA DIGITAL" title and the "Absys" logo. Below the header, there are navigation icons and a dropdown menu for "Bases de Dades disponibles:" with "Hemeroteca Digital" selected. The main form area contains several input fields with labels: "Qualsevol camp:", "Títol:", "Autor Personal:", "Autor Corporatiu:", "Títol col·lecció:", "Congrés:", "Editorial:", "Materies:", "Abans de<=:", and "Any Publicació>:". Each field has a small "Absys" logo icon to its right. A "Cercar" button is located to the right of the "Títol col·lecció:" field. At the bottom of the form, there are "Netejar" and "Ajuda" buttons. The browser's status bar at the bottom shows the address bar with "http://www.arxiu-historic.de.la.ciutat.cat/absys/web/absysweb.html" and the current page title "ABSYS WEB Arxiu Històric de la Ciutat".

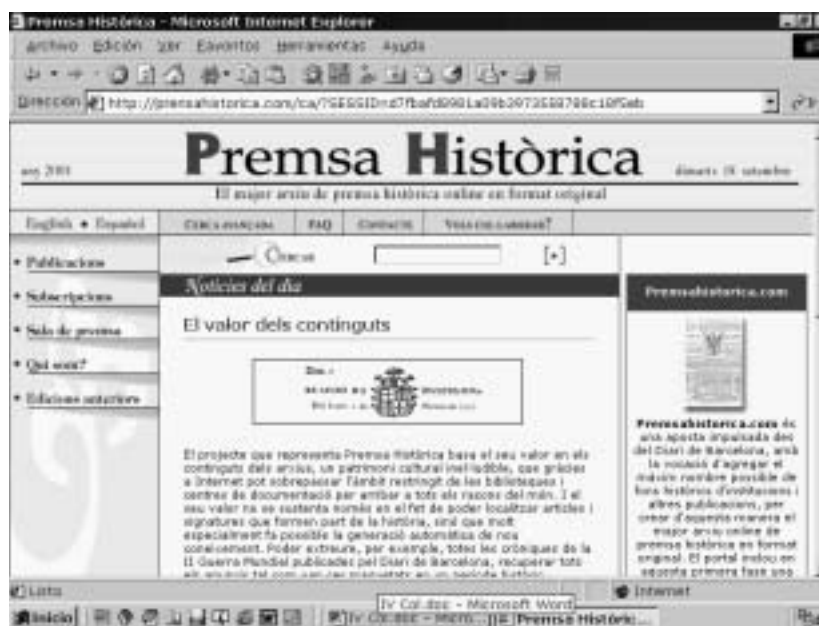
Rovira va fa èmfasi en que aquesta no és una operació aïllada, sinó que forma part d'un pla d'actuacions que des de fa anys es desenvolupa en l'Arxiu Històric i que té per objectiu la salvaguarda de les col·leccions de premsa històrica conservades a l'Hemeroteca de la Casa de l'Ardiaca.

També va explicar que en el seu pla d'actuacions no es contempla una oposició, que ell va considerar inadequada, entre digitalització i microfilmació, sinó una complementarietat. Va indicar que, al menys, durant uns anys, el que la seva institució contempla és tant la microfilmació com la digitalització de les col·leccions.

Projecte *Premsa històrica*

Cristina Ribas, directora del *Diari de Barcelona* (actualment, en edició per Internet) i de l'empresa Moebius Net va presentar el projecte *Premsa històrica*, de gran interès i que, en cas d'arribar a bon port, pot ser un dels més importants en aquest terreny. En tot cas, cal assenyalar que ja és, sens dubte, un dels més originals i apassionants.

Figura 2. Entrada al portal del projecte «Premsa Històrica».



Premsa històrica (*premsahistorica.com*) vol ser un «mitjà de referència a Internet per a la consulta de premsa històrica presentada en format original».

La seu web de *Prensa històrica* proporcionarà, al mateix temps, accés a una base de dades, eines de recerca i de visualització de la informació i continguts periodístics articulats en diversos canals temàtics.

Ribas va explicar que existeix una enorme reserva de coneixement inexplorada en els arxius de premsa històrica, però que el seu accés implica despeses molt elevades. Per aquest motiu —argumentà Ribas— es produeix una progressiva descapitalització d'una part del patrimoni cultural.

En aquest sentit, va explicar que la missió de *Prensa històrica* consisteix en contribuir a la capitalització del patrimoni cultural que resideix als arxius històrics de premsa de l'àmbit cultural català i espanyol, cosa que facilitarà l'accés des de qualsevol suport multimèdia.

La previsió de Moebius Net és que els usuaris de *Prensa històrica* vinguin tant de l'àmbit anomenat B2B (*business to business*) com de l'àmbit B2C (*business to consumers*). Un exemple del primer grup de futurs usuaris serien les xarxes de biblioteques universitàries i centres de documentació d'instituts de recerca i d'empreses del sector de la comunicació, la publicitat, etc. Un exemple del segon grup d'usuaris serien historiadors, investigadors, creatius i, en general, ciutadans amb inquietuds i capacitat de reflexió sobre la història.

Les fases del projecte contemplen la digitalització d'unes 30.000 pàgines del *Diari de Barcelona* i la posada en marxa del portal *Prensa històrica* en proves, seguit d'un procés de firma d'acords amb proveïdors de continguts de premsa històrica en català i en espanyol. Cap a finals de l'any 2001, el portal *Prensa històrica* hauria d'estar en ple funcionament i en condicions d'acceptar subscripcions tant de clients particulars com d'empreses. També està prevista la realització d'acords internacionals amb editors i arxius de premsa internacionals.

La intensa taula rodona que va seguir a les intervencions dels ponents va estar moderada per Ramon Alberch, arxiver en cap de l'Ajuntament de Barcelona, el qual, entre altres aspectes va reivindicar la necessitat d'un pla a escala autonòmica i estatal pel que fa a la digitalització del patrimoni. En la mateixa línia, algunes intervencions del públic, per exemple, la intervenció del periodista i historiador Josep M. Figueres, van reclamar una autèntica política hemerogràfica de Catalunya.

La web de l'Aula d'Història del Periodisme

Per últim, però no menys important, hem de destacar que en la presentació d'aquest IV Col·loqui, Jaume Guillamet, catedràtic d'universitat, degà dels Estudis de Periodisme i coordinador del grup de treball de l'Aula, va presentar la recentment creada pàgina web de l'Aula d'Història del Periodisme Diari de Barcelona (www.upf.es/periodis/auladb/index.htm).

El doctor Jaume Guillaumet va expressar el seu convenciment de la importància progressiva que, eines com Internet i la www poden aportar al treball dels historiadors en general i, en particular, als investigadors de la història del periodisme; així com va assenyalar els arxius de premsa com a font insubstituïble d'informació, al mateix temps que va animar als ponents i assistents a continuar la seva tasca en aquest camp.

Figura 3. Una part de la galeria d'imatges de la web de l'Aula d'Història del Periodisme.



La web de l'Aula d'Història del Periodisme té l'objectiu de proporcionar informació sobre les activitats principals de l'Aula i de temes d'història del periodisme.

Entre les seccions d'aquesta web hi ha una base de dades que permet consultar l'inventari de fonts de publicacions periòdiques de Catalunya, informació sobre els col·loquis organitzats fins ara, una galeria d'imatges de premsa històrica i una secció d'enllaços d'interès, així com novetats bibliogràfiques. Es preveu que publiqui també les intervencions dels ponents del Col·loqui.

El Col·loqui de l'Aula d'Història del Periodisme Diari de Barcelona es celebra anualment, des de fa 4 anys, a finals del mes de maig en l'Auditori de l'Edifici de Rambles de la UPF. Els/les interessats/des es poden adreçar a la Secretaria dels Estudis de Periodisme (*secretaria.periodisme@grup.upf.es*).

Webs relacionades amb el IV Col·loqui i les hemeroteques digitals

Aula d'Història del Periodisme Diari de Barcelona:
www.upf.es/periodis/auladb/index.htm

Arxiu Municipal de Barcelona:
www.bcn.es/arxiu/

Catàleg i Hemeroteca de l'Arxiu Històric:
www.bcn.es/cgi-absys/abweb/G0

The Washington Post Online Archives:
www.washingtonpost.com/wp-adv/archives/

Projecte LAURIN, projecte europeu de digitalització de diaris:
<http://laurin.uibk.ac.at/>

L'Hemeroteca Electrònica (UAB):
www.bib.uab.es/hemerot/internet.htm

Biblioteca de Catalunya:
www.gencat.es/bc/

Historical Newspapers Online:
historynews.chadwyck.com

News Directory:
www.newsdirectory.com

Premsa històrica:
www.premshistorica.com

Tecnodoc:
tecnodoc@landers.es

Delphi:
www.delphi.es