

Anàlisi d'enllaços cap a biblioteques i arxius digitals de patrimoni cultural des de la *Wikipedia* en espanyol i català

[\[Versión castellana\]](#)TOMÁS SAORÍN 

Professor de la Facultad de Comunicación y Documentación
Universidad de Murcia
tsp@um.es

EMILIO JOSÉ RODRÍGUEZ POSADA

Estudiante predoctoral. Escuela Superior de Ingeniería
Universidad de Cádiz
emiliojose.rodriguez@uca.es

Opciones

[Imprimir](#)[Recomandar](#)[Citació](#)[Estadístiques](#)[Metadades](#)[Similars](#)

Resum [\[Abstract\]](#) [\[Resumen\]](#)

Objectiu. Descriure i avaluar l'ús en la *Wikipedia* d'enllaços a les col·leccions digitalitzades en biblioteques, arxius i altres institucions culturals.

Metodologia. L'estudi es fa sobre la totalitat dels articles de les edicions en espanyol i català de la *Wikipedia*, utilitzant una eina d'anàlisi de wikis. Es fa una selecció àmplia de 81 col·leccions digitals espanyoles de diferent abast. També s'agafen dades d'altres projectes de digitalització per poder comparar els resultats.

Resultats. Es mostra una presència encara feble d'enllaços des de la *Wikipedia*, excepte per a la *Biblioteca Virtual Miguel de Cervantes*, les magnituds de la qual són sensiblement diferents. Algunes col·leccions especialitzades són més usades, però en general s'aprecia una falta d'atenció cap a aquestes col·leccions des del col·lectiu d'editors de la *Wikipedia*, la qual cosa s'hauria de tenir en compte en el desenvolupament dels projectes de digitalització tipus *Europeana*.

1 Introducció

L'*Agenda Digital para Europa* ha representat la inversió de quantitats de diners considerables per ampliar la disponibilitat digital del patrimoni cultural europeu. El seu màxim exponent és *Europeana*, que afegeix recursos de múltiples institucions culturals. A Espanya han crescut els continguts i repositoris de caràcter patrimonial, a l'empara de les convocatòries d'ajuts del Ministerio de Cultura, amb uns requisits exigents de qualitat de digitalització, publicació en accés obert i compliment d'estàndards de metadades. Espanya és un dels cinc països que més aporta a *Europeana* (el 8,31 % dels seus més de vint milions d'objectes).¹

No obstant això, *Europeana* concedeix cada vegada més importància en les seves estratègies a l'impacte en el mitjà digital, i no només a les dades d'ítems digitalitzats. La mera posada a disposició d'un gran catàleg no implica que sigui altament visible per a la ciutadania, i per tant s'intenta orientar el portal cap als usos i pràctiques digitals que connecten amb la realitat actual de la Xarxa. Les institucions culturals temen, a més, perdre visibilitat a l'espai digital, perquè els seus registres descriptius no siguin usats com a font de referència amb relació als objectes culturals que conserven, i es degradi la qualitat de la informació cultural (Ferrer-Sapena; Peset, 2012).

S'ha plantejat que s'han d'elaborar mètriques noves per mesurar l'impacte de les seves accions en

el mitjà digital, que en lloc de mesurar simplement el trànsit en els seus propis llocs web, tinguin en compte "la generació d'accessos en llocs ja freqüentats pels usuaris". Contemplen en un lloc prioritari la *Wikipedia*, de la qual destaquen que "Tota institució hauria de preocupar-se dels seus recursos a *Wikipedia*. Hauria de formar part del seu pla de negoci habitual" (Verwayen; Arnoldus; Kaufman, 2011).

L'interès del sector d'arxius, biblioteques i museus a la *Wikipedia* s'ha consolidat en un moviment conegut com a [GLAM](#). Des del febrer del 2012 la biblioteca digital *Europeana* inclou un *widget* per "[citar a la Wikipedia](#)", que busca facilitar que els seus continguts digitals siguin usats com a referència en articles de qualsevol edició de la *Wikipedia*. D'aquesta iniciativa cal deduir que en un futur proper es podria usar la presència a la *Wikipedia* com un dels indicadors d'impacte social digital, aplicable a repositoris de la mateixa naturalesa.

Els aspectes de col·leccions, polítiques, metadades i preservació de les col·leccions digitals patrimonials són tractats en altres estudis (Estivill Rius; Gascón García; Sulé Duesa, 2010), però hi ha menys informació sobre l'ús efectiu dels seus recursos. Aquest treball s'emmarca en el que diferents models d'avaluació de llocs web denominen *visibilitat* o *popularitat* (Codina, 2006; Guallar; Abadal, 2009), però analitzant una única font proveïdora d'enllaços d'impacte i ús considerables. Atès que el *pagerank* de Google combina l'abundància i diversitat d'enllaços amb la visibilitat de l'emissor de l'enllaç, els resultats d'aquest estudi també poden tenir interès en el camp del posicionament web (Codina; Marcos, 2005, p. 86). Dins de les investigacions centrades en la *Wikipedia*, hi ha una línia de recerca sobre l'anàlisi de l'ús i citació de fonts, que és un element bàsic de la qualitat i credibilitat dels articles de l'enciclopèdia (Nielsen, 2007; Luyt; Tan, 2010). Altres estudis han abordat l'augment del trànsit originat en la *Wikipedia* cap a les col·leccions digitals, quan s'han enriquit articles i inclòs enllaços sobre les temàtiques que recullen (Lallye; Dunford, 2007; Elder; Westbrook; Reilly, 2012).

Aquest treball té com a objectiu fer una primera aproximació a la presència d'enllaços de fonts contingudes en biblioteques digitals espanyoles de l'àmbit del patrimoni cultural. S'extrauran dades individuals i globals de cada repositori, que serviran per avaluar i comparar en el temps el seu impacte directe en la *Wikipedia* en espanyol i en català (*Viquipèdia*).

2 Metodologia de l'estudi

Per a l'estudi s'han seleccionat els repositoris i arxius digitals espanyols amb continguts de l'àrea de patrimoni cultural (patrimoni bibliogràfic i documental, patrimoni històric, patrimoni fotogràfic). Els documents de caràcter patrimonial conservats en arxius, biblioteques, museus i institucions culturals han estat de difícil accés fins a la seva digitalització i difusió recents a través de portals web.

Després d'haver-los identificat, s'analitzen enllaços que apunten cap als repositoris o objectes que contenen des de qualsevol article de les edicions en espanyol i català de la *Wikipedia*.² Les dades dels articles de la *Wikipedia* han estat processats amb una eina estadística especialment enfocada a l'anàlisi de wikis anomenada [WikiEvidens](#) (vegeu l'apèndix 3). Hi hem afegit les funcions necessàries per parametritzar la llista de fonts, els patrons de comparació i l'edició de la *Wikipedia* que s'analitza. WikiEvidens és programari lliure i es pot descarregar gratuïtament.

S'han identificat per a cada projecte totes les variants possibles de dominis i subdominis que resulten de totes les maneres d'enllaçar (a documents, fitxes de metadades, seccions, etc.).³

A l'hora d'establir el patró de vincles als documents podem trobar el problema de les "seleccions ad hoc". Quan la necessitat de promocionar un contingut representa "treure'l" del repositori, i presentar-lo a través del sistema d'arxius del gestor de continguts, pot afectar que el wikipedista esculli un enllaç extern al repositori, i per tant no sigui tingut en compte en analitzar els enllaços al repositori.⁴ També queden fora de l'estudi totes les digitalitzacions singulars publicades fora de repositoris identificables.

Un altre element que hem considerat que oferia resultats útils era comprovar la persistència dels enllaços, atès que en els tests preliminars detectem enllaços que havien deixat de funcionar (coneguts com a *enllaços trencats*). Per a això hem utilitzat un petit *crawler* (disponible entre els mòduls de WikiEvidens) per revisar sistemàticament els enllaços i detectar els que havien deixat d'estar accessibles.

Hem optat per fer una anàlisi panoràmica, a causa de la presència reduïda d'enllaços. En una segona aproximació es podria considerar la relació entre la mida del repositori i el nombre de citacions rebudes en la *Wikipedia*.

També hem considerat convenient centrar els resultats en la presència d'enllaços només en articles, de manera que hem eludit les pàgines de discussió i les de coordinació de la comunitat wikipedista. Tampoc no diferenciarem per a cada cas si l'enllaç és en la nota, en la bibliografia o en els enllaços externs, només presentarem dades globals. Quan les citacions a repositoris patrimonials siguin més abundants, es podran considerar aquestes anàlisis segmentades, i també d'altres sobre les categories dels articles o el nombre d'enllaços per article.

Un aspecte que no hem desenvolupat ha estat l'anàlisi de la destinació dels enllaços (fitxes de metadades; document digitalitzat, imatges, categories, col·leccions, recerques, exposicions digitals, etc.), la qual cosa podria aportar informació molt específica interessant per millorar l'experiència d'usuari i l'arquitectura de la informació d'aquestes biblioteques digitals. En alguns casos poden haver-se recollit enllaços a registres que no permeten l'accés al document complet, o enllaços genèrics al portal del repositori. Si les dades d'enllaços fossin més abundants s'estaria en condicions d'afinar més sobre la destinació preferida per a l'enllaç.

2.1 Selecció de biblioteques digitals

En aquest treball no es tenen en compte els nombrosos repositoris científics en accés obert, que contenen sobretot publicacions acadèmiques i de recerca.⁵ S'ha optat pels repositoris amb documents històrics i patrimonials (biblioteques i arxius digitals).

Hem seleccionat els principals projectes de biblioteques digitals espanyoles. S'hi inclouen també les biblioteques digitals regionals i els repositoris digitals d'universitats o arxius digitals d'altres institucions culturals que disposen d'un fons antic important. Cal tenir en compte que hi ha nombrosos projectes de digitalització monogràfics que no s'integren en repositoris digitals, i també digitalització de patrimoni publicada electrònicament de manera dispersa i, per tant, no identificats en aquest estudi.

També es recullen els projectes europeus en què hi ha una aportació significativa de continguts espanyols. Encara que l'abast és fonamentalment nacional, es tenen en compte algunes dades amb fonts de dades de col·leccions digitals importants com la de la Library of Congress, la *Biblioteca Digital Mundial* de la Unesco i *Memoria Chilena* de la Biblioteca Nacional Digital de Chile. Com a element de comparació, hi hem inclòs també les hemeroteques històriques de l'*ABC* i *La Vanguardia*, que comprenen tot el segle xx i són d'accés obert.

Per a l'elaboració de la llista de repositoris prenem com a punt de partida la selecció feta per la Biblioteca Nacional de España sota el títol "[Colecciones digitales españolas](#)", vinculades en la seva major part als ajuts de digitalització del Ministerio de Cultura dels darrers anys. Aquesta llista s'ha complementat amb una selecció d'altres projectes dels 561 del directori de col·leccions digitals [Hispana](#).

Les col·leccions històriques d'alguns repositoris universitaris presentaven problemes per a la seva inclusió, davant la impossibilitat de separar aquests registres de la resta de continguts científics i institucionals. Molts dels seus continguts estan disponibles a través d'*Hispana*, però la seva publicació indiferenciada fa que perdem aquestes dades.⁶ En algunes de les anàlisis hem aplicat un modulador als resultats segons el percentatge que representi la col·lecció històrica en el total del repositori, i ho hem indicat en nota. Així, per exemple, en el repositori *Roderic* de la Universitat de València, el fons històric representa la meitat dels continguts, mentre que a *Digitum*, de la Universidad de Murcia, és el 5 %.

S'ha fet atenció a la inclusió de col·leccions digitals catalanes, encara que fossin petites, per captar de la millor manera possible el comportament específic de l'edició en català de la *Wikipedia*.

També presenten problemes d'ajustament les dades que recollim d'*Hispana*, perquè es barregen repositoris de producció científica amb col·leccions patrimonials.

Per obtenir dades diferenciades segons l'abast de cada biblioteca digital, hem establert les categories següents: EU (europees), ES (àmbit nacional), REG (regionals o autonòmiques), LOC (locals o provincials), UNI (universitats), INS (altres institucions culturals). En la categoria TEST s'inclouen les col·leccions digitals usades per contextualitzar les xifres d'enllaços recollits.

3 Resultats

En total s'han seleccionat 81 biblioteques digitals (3 EU, 7 ES, 22 REG, 5 LOC, 28 UNI i 16 INS). S'han inclòs 6 col·leccions en la categoria TEST. La llista completa i els URL base per localitzar els enllaços es presenten en l'apèndix 0.

S'han recuperat els enllaços procedents de 372.286 articles de la *Wikipedia* en català (WP-CAT) i 885.935 de la *Wikipedia* en espanyol (WP-ES), a partir dels seus [dumps](#), amb data de 19 i 20 d'abril del 2012, respectivament.

La *Wikipedia* en espanyol disposa de 6.460 enllaços cap a les biblioteques que hem seleccionat, cosa que representa el 0,21 % del total d'enllaços existents en aquesta enciclopèdia. Aquests enllaços estan distribuïts per 4.393 articles, que constitueixen només el 0,50 % del total d'articles disponibles.

La *Wikipedia* en català té 1.666 enllaços distribuïts en 1.075 articles, la qual cosa representa un 0,11 % del total d'enllaços i un 0,29 % del total d'articles.

Edició de la <i>Wikipedia</i>	Enllaços a biblioteques digitals	Total enllaços	%
WP-ES	6.460	3.104.964	0,21 %
WP-CAT	1.666	1.516.015	0,11 %

El nombre d'enllaços a biblioteques digitals dins d'un mateix article presenta patrons similars en ambdues enciclopèdies: 1,47 enllaços per article en l'espanyola i 1,55 en la catalana.

Per poder posar en context aquestes xifres, hem obtingut les de les hemeroteques de l'*ABC* i *La Vanguardia*, dues grans hemeroteques de diaris espanyols del segle xx. La premsa és una gran proveïdora d'enllaços a la *Wikipedia*: *El País* ocupa el lloc 17 del rànquing d'enllaços de la [Wikipedia en espanyol](#) i la *BBC* el lloc 6 en la [Wikipedia en anglès](#). Per això, les magnituds d'aquest estudi poden contrastar-se amb els 4.148 enllaços des de la *Wikipedia* en espanyol cap a l'Hemeroteca de l'*ABC* distribuïts en 2.325 articles i els 1.203 enllaços cap a l'Hemeroteca de *La Vanguardia* en 722 articles. En la *Wikipedia* en català trobem 2.217 enllaços des de 942 articles per a l'Hemeroteca de *La Vanguardia* i només 160 enllaços cap a l'Hemeroteca de l'*ABC*, des de 137 articles.

També ens serà útil saber les dades d'altres col·leccions digitals estrangeres conegudes: *Google Books*, amb les excepcions anteriorment indicades, rep 18.497 enllaços en la WP-ES i 1.131 en la WP-CAT. Les col·leccions digitals de la Library of Congress dels Estats Units reben 1.052 enllaços en la WP-ES i 276 en la WP-CAT. El portal *Memoria de Chile* en rep 1.462 des de la WP-ES i 9 des de la WP-CAT. Finalment, la *Biblioteca Digital Mundial* de la Unesco amb prou feines genera 26 enllaços en la WP-ES i 3 en la WP-CAT.

Les tres taules següents recullen el nombre d'enllaços que apunten cap a les diferents biblioteques i els repositoris digitals seleccionats, en la WP-ES, la WP-CAT i de manera conjunta. Els llocs no enllaçats no han estat afegits en les taules (vegeu l'apèndix 2). Les dades de les biblioteques digitals mixtes (científiques i patrimonials) marcades amb (***) poden ser poc fiables per a col·leccions patrimonials. Per contra, les marcades amb (*) tenen una alta probabilitat d'ajustar-se a l'abast del treball.⁷

Biblioteca digital	Enllaços
<i>Biblioteca Virtual Miguel de Cervantes</i>	4.372
<i>Biblioteca Digital Hispánica y Hemeroteca Digital</i> de la Biblioteca Nacional de España	320
<i>Biblioteca Digital de Castilla y León</i>	226
<i>Biblioteca Digital del Real Jardín Botánico</i>	172
<i>Biblioteca Virtual de Derecho Aragonés</i>	152
<i>Dipòsit digital de documents de la UAB</i> (***)	87
<i>Biblioteca Digital Leonesa</i>	86
<i>Biblioteca Virtual de Prensa Histórica</i>	76

<i>Pares (Portal de Archivos Españoles)</i>	68
<i>Biblioteca Digital de Castilla-La Mancha</i>	63
<i>Fondo Documental Histórico de las Cortes de Aragón</i>	55
<i>Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico</i>	54
<i>e_Buah de la Universidad de Alcalá (***)</i>	48
<i>Biblioteca Virtual del Patrimonio Bibliográfico</i>	47
<i>Fondo antiguo de la Universidad de Sevilla</i>	47
<i>Biblioteca Digital Complutense</i>	42
<i>Biblioteca Digital del repositorio Gredos de la Universidad Salamanca (*)</i>	40
<i>Memoria Digital de Canarias</i>	39
<i>Red Digital de Colecciones de Museos de España</i>	38
<i>Memoria Digital Vasca</i>	38
<i>The European Library</i>	30
<i>Memòria Digital de Catalunya</i>	29
<i>Biblioteca Digital del Ateneo Madrid</i>	27
<i>Digitum. Fondo antiguo de la Universidad de Murcia (***)</i>	26
<i>Helvia: Repositorio Institucional de la Universidad de Huelva (**)</i>	22
<i>Arias Montano: repositorio institucional de la Universidad de Huelva (**)</i>	21
<i>Fondo Antiguo de la Universidad de Granada (*)</i>	19
<i>BI NaDI. Biblioteca Digital de Navarra</i>	18
<i>Repositorio da Universidade da Coruña (***)</i>	17
<i>Rodín. Universidad de Cádiz. Patrimonio bibliográfico (**)</i>	15
<i>Biblioteca Digital de Castellón (*)</i>	13
<i>Biblioteca Virtual Ignacio Larramendi</i>	12
<i>Memoria de Madrid</i>	11
<i>Jable: Archivo de Prensa Digital de la Universidad de las Palmas de Gran Canaria</i>	11
<i>BiValdi (Biblioteca Valenciana Digital)</i>	10
<i>Europeana</i>	9
<i>ANC (Arxiu Nacional Catalunya)</i>	8
<i>Biblioteca Digital de la Comunidad Madrid</i>	7
<i>Fototeca del Patrimonio Histórico</i>	6
<i>ARCA (Arxiu de Revistes Catalanes Antiques)</i>	6
<i>Biblioteca Digital de la Región de Murcia</i>	6
<i>Biblioteca Virtual del Principado de Asturias</i>	6
<i>Biblioteca Virtual de Andalucía</i>	6
<i>Fondo Histórico Digital de La Rábida de la Universidad Internacional de Andalucía</i>	6
<i>DUGi Fons Especials: Repositori Digital de la UdG</i>	5
<i>UVaDOC: repositorio documental de la Universidad de Valladolid (*)</i>	5
<i>Recursos de Investigación de la Alhambra</i>	4
<i>Fondo Antiguo Universidad Zaragoza (**)</i>	4
<i>RIUMA. Universidad de Málaga. Patrimonio (***)</i>	4
<i>Biblioteca Digital de la Real Academia de Historia</i>	3
<i>Liburutegi Digitala de la Biblioteca Foral de Bizkaia</i>	3
<i>Biblioteca Virtual de La Rioja</i>	3
<i>Minerva: Repositorio Institucional da Universidade de Santiago de Compostela (**)</i>	3
<i>Hispana (**)</i>	2
<i>Biblioteca Virtual de la Diputación de Zaragoza</i>	2
<i>Somni. Fons històric de la Universitat de València (*)</i>	2
<i>RUIdeRA: repositorio institucional de la UCLM (***)</i>	2
<i>Archivo del Ateneo Madrid</i>	1

<i>Iuris Digital</i> : Biblioteca Virtual de la Real Academia de Jurisprudencia y Legislación	1
<i>Teatro Siglo de Oro</i> de la Biblioteca Nacional de España	1
<i>Biblioteca Dixital de Galicia (Cidade de Cultura Galega)</i>	1
<i>Galiciana: Biblioteca Dixital de Galicia</i>	1
<i>Fons Sol-Torres</i> del Servei de Biblioteca i Documentació de la Universitat de Lleida	1
Biblioteca Digital del CEU (***)	1
Setze biblioteques digitals no tenen enllaços a l'edició en espanyol (vegeu l'apèndix 2)	

Taula 1. Enllaços en la *Wikipedia* en espanyol (WP-ES)

Biblioteca digital	Enllaços
<i>Biblioteca Virtual Miguel de Cervantes</i>	786
<i>Memòria Digital de Catalunya</i>	186
<i>Biblioteca Virtual de Derecho Aragonés</i>	174
<i>Dipòsit digital de documents de la UAB (***)</i>	147
<i>ARCA (Arxiu de Revistes Catalanes Antiques)</i>	95
<i>Biblioteca Digital del Real Jardín Botánico</i>	35
<i>ANC (Arxiu Nacional Catalunya)</i>	31
<i>Fondo Documental Histórico de las Cortes de Aragón</i>	31
<i>Biblioteca Virtual de Prensa Histórica</i>	24
<i>Biblioteca Digital Hispánica y Hemeroteca Digital</i> de la Biblioteca Nacional de España	18
<i>BiValdi (Biblioteca Valenciana Digital)</i>	13
<i>Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico</i>	12
<i>Biblioteca Digital de Castilla y León</i>	11
Fondo antiguo de la Universidad de Granada (*)	11
<i>Cartoteca Digital</i> de l'Institut Cartogràfic de Catalunya	10
<i>Biblioteca Digital de Castellón (*)</i>	9
<i>Biblioteca Digital Leonesa</i>	8
<i>Red Digital de Colecciones de Museos de España</i>	7
<i>Fons Sol-Torres</i> del Servei de Biblioteca i Documentació de la Universitat de Lleida	6
<i>Biblioteca Virtual del Patrimonio Bibliográfico</i>	5
<i>Pares (Portal de Archivos Españoles)</i>	4
<i>Biblioteca Digital del Ateneo Madrid</i>	4
<i>Memoria Digital de Canarias</i>	4
<i>DUGi Fons Especials: Repositori Digital de la UdG</i>	4
<i>Digitum</i> . Fondo antiguo de la Universidad de Murcia (***)	4
The European Library	3
<i>Biblioteca Virtual Ignacio Larramendi</i>	3
<i>Biblioteca Digital Complutense</i>	3
<i>Europeana</i>	2
<i>Biblioteca Virtual de Andalucía</i>	2
Fondo antiguo de la Universidad Sevilla	2
<i>Somni</i> . Fondo histórico de la Universitat de València (*)	2
<i>Biblioteca Digital</i> del repositorio Gredos de la Universidad Salamanca (*)	2
<i>Almirall: portal de pensament i cultura del segle XIX</i>	1
<i>Liburutegi Digitala</i> de la Biblioteca Foral de Bizkaia	1
<i>Biblioteca Virtual de La Rioja</i>	1
<i>Memoria Digital Vasca</i>	1
<i>Biblioteca Virtual del Principado de Asturias</i>	1
<i>BINaDI. Biblioteca Digital de Navarra</i>	1

<i>Helvia: Repositorio Institucional de la Universidad de Huelva</i> (**)	1
<i>Repositorio da Universidade da Coruña</i> (*)	1
Quaranta biblioteques digitals no tenen enllaços en l'edició catalana (vegeu l'apèndix 2)	

Taula 2. Enllaços en la *Wikipedia* en català (WP-CAT)

Biblioteca digital	Enllaços
<i>Biblioteca Virtual Miguel de Cervantes</i>	5.158
<i>Biblioteca Digital Hispánica y Hemeroteca Digital</i> de la Biblioteca Nacional de España	338
<i>Biblioteca Virtual de Derecho Aragonés</i>	326
<i>Biblioteca Digital de Castilla y León</i>	237
<i>Dipòsit digital de documents de la UAB</i> (***)	234
<i>Memòria Digital de Catalunya</i>	215
<i>Biblioteca Digital del Real Jardín Botánico</i>	207
<i>ARCA (Arxiu de Revistes Catalanes Antiques)</i>	101
<i>Biblioteca Virtual de Prensa Histórica</i>	100
<i>Biblioteca Digital Leonesa</i>	94
<i>Fondo Documental Histórico</i> de las Cortes de Aragón	86
<i>Pares (Portal de Archivos Españoles)</i>	72
<i>Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico</i>	66
<i>Biblioteca Digital de Castilla-La Mancha</i>	63
<i>Biblioteca Virtual del Patrimonio Bibliográfico</i>	52
Fondo antiguo de la Universidad Sevilla	49
<i>e_Buah</i> de la Universidad de Alcalá (***)	48
<i>Biblioteca Digital Complutense</i>	45
<i>Red Digital de Colecciones de Museos de España</i>	45
<i>Memoria Digital de Canarias</i>	43
<i>Biblioteca Digital</i> del repositorio <i>Gredos</i> de la Universidad Salamanca (*)	42
ANC (Arxiu Nacional Catalunya)	39
<i>Memoria Digital Vasca</i>	39
The European Library	33
<i>Biblioteca Digital del Ateneo Madrid</i>	31
<i>Digitum</i> . Fondo antiguo de la Universidad de Murcia (***)	30
Fondo antiguo de la Universidad de Granada (*)	30
<i>BiValdi (Biblioteca Valenciana Digital)</i>	23
<i>Helvia: Repositorio Institucional de la Universidad de Huelva</i> (**)	23
<i>Biblioteca Digital de Castellón</i> (*)	22
<i>Arias Montano</i> : repositorio institucional de la Universidad de Huelva (**)	21
<i>BiNaDI. Biblioteca Digital de Navarra</i>	19
<i>Repositorio da Universidade da Coruña</i> (*)	18
<i>Biblioteca Virtual Ignacio Larramendi</i>	15
<i>Rodin</i> . Universidad de Cádiz. Patrimonio bibliográfico (**)	15
<i>Europeana</i>	11
<i>Jable: Archivo de Prensa Digital</i> de la Universidad de las Palmas de Gran Canaria	11
<i>Memoria de Madrid</i>	11
<i>Cartoteca Digital</i> de l'Institut Cartogràfic de Catalunya	10
<i>DUGi Fons Especials: Repositori Digital de la UdG</i>	9
<i>Biblioteca Virtual de Andalucía</i>	8
<i>Biblioteca Digital de la Comunidad de Madrid</i>	7
<i>Biblioteca Virtual del Principado de Asturias</i>	7

<i>Fons Sol-Torres</i> del Servei de Biblioteca i Documentació de la Universitat de Lleida	7
<i>Biblioteca Digital de la Región de Murcia</i>	6
<i>Fototeca del Patrimonio Histórico</i>	6
<i>Fondo Histórico Digital de La Rábida</i> de la Universidad Internacional de Andalucía	6
<i>UVaDOC</i> : repositorio documental de la Universidad de Valladolid (*)	5
<i>Liburutegi Digitala</i> de la Biblioteca Foral de Bizkaia	4
<i>Biblioteca Virtual de La Rioja</i>	4
Fondo Antiguo Universidad de Zaragoza (**)	4
<i>Recursos de Investigación de la Alhambra</i>	4
<i>RIUMA</i> . Universidad de Málaga. Patrimonio (***)	4
<i>Somni</i> . Fondo histórico de la Universitat de València (*)	4
<i>Biblioteca Digital</i> de la Real Academia de Historia	3
<i>Minerva: Repositorio Institucional da Universidade de Santiago de Compostela</i> (**)	3
<i>Biblioteca Virtual de la Diputación de Zaragoza</i>	2
<i>Hispana</i>	2
<i>RUIdERA</i> : repositorio institucional de la UCLM (*)	2
<i>Almirall, portal del pensament i cultura del segle XIX</i>	1
Archivo del Ateneo de Madrid	1
<i>Biblioteca Dixital de Galicia (Cidade de Cultura Galega)</i>	1
Biblioteca Digital del CEU (*)	1
<i>Galiciana: Biblioteca Dixital de Galicia</i>	1
<i>Iuris Digital</i> : Biblioteca Virtual de la Real Academia de Jurisprudencia y Legislación	1
<i>Teatro Siglo de Oro</i> de la Biblioteca Nacional de España	1
Catorze biblioteques digitals no tenen enllaços a cap de les dues Wikipedia (vegeu l'apèndix 2)	

Taula 3. Enllaços en conjunt en la *Wikipedia* en espanyol i en català (WP-ES + WP-CAT)

La distribució de resultats mostra una disparitat clara en el cas de la *Biblioteca Virtual Miguel de Cervantes*, que acumula fins al 67 % dels enllaços en la WP-ES i el 47 % en la WP-CAT. En les taules d'estudi de la freqüència d'enllaços següents no inclourem les dades d'aquesta biblioteca digital, sobre la qual aprofundirem més endavant mitjançant anàlisis ad hoc. Això permet tenir una visió més ajustada de l'ús dels molts repositoris culturals que han proliferat en els últims anys.

La distribució d'enllaços per categories de biblioteques digitals, tenint en compte que la mostra no és exhaustiva per a col·leccions de caràcter local i institucions culturals, és la següent. Es prenen les dades conjuntes de la WP-ES i la WP-CAT.

Tipus de biblioteca digital	Enllaços	%
Regionals	899	35 %
Institucions	666	26 %
Estatals	615	24 %
Universitats ⁸	264	10 %
Locals	111	4 %
Europees	44	2 %

Taula 4. Distribució d'enllaços per tipus de biblioteca digital (WP-ES + WP-CAT)

En la WP-ES, 6 de les 80 biblioteques digitals estudiades concentren el 50 % dels enllaços recollits: *Biblioteca Digital Hispánica* i *Hemeroteca Digital* de la Biblioteca Nacional de España; *Biblioteca Digital de Castilla y León*; *Biblioteca Digital del Real Jardín Botánico*; *Biblioteca Virtual de Derecho Aragonés*; *Dipòsit digital de documents de la UAB* (***) i *Biblioteca Digital Leonesa*.

El 76 % s'assoleix amb només 17 repositoris: *Biblioteca Virtual de Prensa Histórica*; *Pares* (Portal

de Archivos Españoles); Biblioteca Digital de Castilla-La Mancha; Fondo Documental Histórico de las Cortes de Aragón; Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico; Fondo antiguo de la Universidad de Alcalá (***); Biblioteca Virtual del Patrimonio Bibliográfico; Fondo Antiguo y Archivo Histórico de la Universidad Sevilla; Colección Digital Complutense; Biblioteca Digital del repositorio Gredos de la Universidad Salamanca (*).

És a dir, en la WP-ES un 21 % de les biblioteques digitals reben el 76 % dels enllaços, la qual cosa representa valors propers a la distribució de Pareto. Un 80 % de les biblioteques digitals té una presència escassa.

Per a la WP-CAT la concentració d'enllaços és més gran, 6 repositoris acumulen el 76 % dels enllaços: *Memòria Digital de Catalunya*; *Biblioteca Virtual de Derecho Aragonés*; *Dipòsit digital de documents de la UAB* (***); *ARCA (Arxiu de Revistes Catalanes Antiques)*; *Biblioteca Digital del Real Jardín Botánico*; *Fondo Documental Histórico de las Cortes de Aragón*.

Per a la WP-CAT trobem que moltes de les deu primeres biblioteques digitals (excloent la *Biblioteca Virtual Miguel de Cervantes* i el *Dipòsit digital de documents de la UAB*) són de poc ús en la WP-ES, especialment la *Memòria Digital de Catalunya* i *ARCA (Arxiu de Revistes Catalanes Antiques)*, amb un 87 % i un 79 % d'enllaços originats des de l'enciclopèdia en català.

Biblioteca digital	% enllaços des de WP-CAT
<i>Memòria Digital de Catalunya</i>	87
<i>Biblioteca Virtual de Derecho Aragonés</i>	53
<i>ARCA (Arxiu de Revistes Catalanes Antiques)</i>	94
<i>Biblioteca Digital del Real Jardín Botánico</i>	17
<i>ANC (Arxiu Nacional Catalunya)</i>	79
<i>Fondo Documental Histórico de las Cortes de Aragón</i>	36
<i>Biblioteca Virtual de Prensa Histórica</i>	24
<i>Biblioteca Digital Hispánica y Hemeroteca Digital de la Biblioteca Nacional de España</i>	5
<i>BiValdi (Biblioteca Valenciana Digital)</i>	57
<i>Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico</i>	18

Taula 5. Distribució d'enllaços entre la WP-CAT i la WP-ES per al top ten d'enllaços en WP-CAT

Troblem algunes característiques en el format dels enllaços. A primera vista destaca l'abundància d'exemples de mala praxi per a URL nets i Cool URI (abundància de variables, controladors interns, etc.).⁹ L'extensió mitjana de les cadenes de l'URL, sense comptar el domini, és de 56 caràcters en la WP-ES i 55 en la WP-CAT. Gairebé no s'utilitza en la *Wikipedia* l'URL recomanat en els repositoris d'acord amb el sistema [Handle](#). Els URL Handle tenen 11 caràcters (una extensió similar a la usada per un servei massiu com Youtube). Alguns repositoris com la *Biblioteca Digital Hispánica* doblen el valor mitjà (111 caràcters), encara que la seva recomanació d'enllaç permanent és de 25. Molt pocs casos fan ús d'URL amigables.

Quant al nombre d'enllaços trencats, hem trobat que representen al voltant del 4 % en la *Wikipedia* en espanyol i el 12 % en la *Wikipedia* en català; els errors més freqüents són el de "pàgina no trobada" (error 404) i el de "temps expirat" (considerem 25 segons com a temps màxim d'espera). Les dades completes es presenten en l'apèndix 4. L'estabilitat dels enllaços és una cosa fonamental en la Xarxa —i poques vegades es compleix (la [vida mitjana d'una pàgina web en Internet](#) és de 77 dies)— i ineludible en els repositoris. Un 96 % dels enllaços accessibles en el cas de l'espanyola confirma l'estabilitat, però caldria estudiar l'origen dels errors.

Els enllaços es poden incloure en els articles sia com a nota final (referències), sia com a enllaços a secció independent (ja sigui en "Enllaços externs" o "Bibliografia"). Per fer el càlcul d'enllaços hem utilitzat dos *scripts* diferents. Encara que finalment el més eficient no permetia detectar la ubicació de l'enllaç dins de cada article, els càlculs fets amb l'altre indiquen que una mica més de la meitat dels enllaços apareixen dins de notes (amb l'etiqueta <ref></ref>), i referencien dades específiques de l'article (en la *Wikipedia* en català un 64,47 % i en la *Wikipedia* en espanyol un 56,28 %).

No podem deixar de destacar els elevats resultats obtinguts per les dues fonts que contenen

recursos d'Hispanoamèrica: *Biblioteca Virtual Miguel de Cervantes* i *Memoria de Chile*. D'altra banda, les dades dels dos grans agregadors analitzats, *Europeana* i *Hispana*, són molt baixes (onze i dos enllaços, respectivament). Tot i acumular gran quantitat de registres, són poc usats directament per enllaçar.

3.1 El cas de la *Biblioteca Virtual Miguel de Cervantes* (BVMdC)

Amb els 4.372 en la WP-ES, i 786 en la WP-CAT, aquesta biblioteca se situa entre els llocs webs amb més enllaços en la *Wikipedia*, en la posició 69. La diferència enorme en les xifres d'enllaços, com hem comentat anteriorment, ens permet avançar en una anàlisi més detallada sobre les pràctiques d'enllaçament dels editors de la *Wikipedia*. La BVMdC és, a més, una de les pioneres en digitalització a gran escala de textos complets en espanyol.

Els resultats relatius al total d'articles per a la WP-ES són més del doble que per a la WP-CAT (0,49 % enfront de 0,21 %). Farem servir només els enllaços de l'edició en espanyol per descriure les pautes d'enllaçament.

L'estabilitat dels URL és important en els repositoris, orientats a oferir un accés permanent i únic als recursos. No obstant això, és freqüent que els objectes digitals puguin ser enllaçats de diverses maneres diferents o que fins i tot els URL visibles per a l'usuari no siguin adequats per a l'enllaçament.

La majoria dels enllaços de la BVMdC usen el domini estàndard del web "www.cervantesvirtual.com". Però un significatiu 15 % fa servir la forma "descargas.cervantesvirtual.com" i un 4 % fa servir el subdomini "bib.cervantesvirtual.com".

Només 22 enllaços són genèrics a la pàgina principal de la BVMdC. La resta són vincles directes a documents i per tant suposem una precisió més gran per donar suport a la fiabilitat del contingut dels articles.

Es detecten enllaços antics que, tanmateix, funcionen correctament en ser transformats pel servidor d'acord amb la nomenclatura actual d'URL del lloc.¹⁰ Es percep l'evolució cap a URL amigables.

El terme *enllaç* no és exclusivament usat com a enllaç al document, i a més adopta diferents formes. El 4 % identifica en l'enllaç les biblioteques d'autor, mantenint en l'URL la referència a la forma d'organització de la col·lecció per autors.¹¹ El 6 % dels enllaços són a informació d'autors, en lloc d'informació a obres. Un 2 % són a continguts de les seccions específiques d'història.¹² Un 26 % enllaça directament a documents PDF i no a pàgines web convencionals—siguin fitxes informatives o textos font publicats en HTML.

La BVMdC s'organitza en portals temàtics i institucionals. No obstant això, la gran majoria d'enllaços usats en la *Wikipedia* no permeten identificar si el document està englobat en un d'ells. En enllaçar es perd part d'aquesta portalització que destaca el paper de les institucions que col·laboren amb els seus fons. Només el 6 % dels enllaços fan referència a "portal". Els més freqüents són el portal 0 (37 %), portal 33 (16 %) i portal de Chile (6 %).¹³

Es detecten fins a vint maneres diferents de construir URL. Això es pot deure tant a canvis en la tecnologia del servidor com a necessitats d'enllaç específiques dels usuaris (elecció entre obres, autors, matèries, pàgines informatives, enllaços directes al PDF).

```
/autor/  
bib_autor  
/buscador/  
Buscar.html  
/descargaPdf/  
/escaparate/  
FichaAutor.html  
FichaClasificacionMaterias  
FichaMateria  
FichaObra  
/obra/  
/obra-visor  
/obras-autor/
```

Hemeroteca
Historia
/portal/
BVMC_OBRAS
/servlet/Ficha
servlet/MuestraForo
servlet/SirveObra

Figura 1. Principals cadenes en els enllaços a la *Biblioteca Virtual Miguel de Cervantes*

4 Discussió dels resultats

Encara que és arriscat establir patrons explicatius per als resultats obtinguts, sembla que el valor dels documents pot tenir tanta importància com la quantitat d'ítems de cada repositori. Hi ha algunes qüestions que mereixerien aprofundir-hi més:

- Per què un repositori molt especialitzat com el del Real Jardín Botánico té resultats tan bons? La variable coherència temàtica de la col·lecció podria estar influïent en l'increment del seu ús i visibilitat.
- Quins patrons de conducta es podrien esperar en els editors d'articles de la *Wikipedia* per citar documents tan especialitzats com els de les biblioteques digitals? Si entenem la *Wikipedia* com una enciclopèdia generalista, els documents de les biblioteques digitals són molt menys adequats que manuals i obres de síntesi. Si l'entendem també com una enciclopèdia especialitzada que abraça cada vegada temes més delimitats, aquests documents tenen un valor potencial, i per tant seria raonable esperar trobar-los citats.
- Hi ha alguna relació entre diferents arquitectures de publicació digital de continguts i la freqüència de citacions en la *Wikipedia*? És difícil copsar en una investigació com una publicació digital amb una experiència d'usuari millor afectaria el seu ús des de la *Wikipedia*.
- Quina explicació podem trobar als baixos resultats obtinguts per agregadors com *Hispana* o *Europeana*? Això pot fer-nos reflexionar sobre el nivell de maduresa o com d'adequada és l'estratègia informativa dels projectes.

Si acceptem que la citació a la *Wikipedia* és un altre indicador d'ús dels continguts d'una biblioteca digital, seria útil poder conèixer millor quins factors externs al contingut mateix influeixen en l'ús. Sens dubte, aspectes com la qualitat de l'experiència d'usuari, l'organització mitjançant narratives digitals atractives o l'optimització per a cercadors, han de tenir la seva influència en l'ús dels repositoris digitals.

La diferència espectacular de resultats de la *Biblioteca Virtual Miguel de Cervantes* mereix una consideració a part. Potser és el projecte amb més antiguitat al nostre país, a més d'estar àmpliament vinculat amb la comunitat iberoamericana, la qual cosa n'amplia l'abast i les possibilitats. El factor hispanoamericà mereixeria atenció, atès que el projecte *Memoria Chilena* també presenta resultats molt superiors als valors mitjans de les biblioteques digitals espanyoles. A més d'un volum considerable de continguts, ha estat valorada com "un ejercicio permanente de curaduría digital" (Abbagliati, 2011). Totes dues han enfocat la seva organització de continguts amb arquitectures d'informació orientades a portals temàtics i no exclusivament al repositori.

Les biblioteques digitals regionals en conjunt tenen un bon comportament, cosa que podria indicar una certa preferència per una agrupació de continguts locals enfront dels grans projectes nacionals. També fa pensar l'ús escàs de la *Biblioteca Virtual de Prensa Histórica*, mentre que els arxius retrospectius de l'*ABC* i *La Vanguardia* són usats freqüentment. D'altra banda, no s'aprecien diferències a favor de la *Viquipèdia*, encara que ha fet regularment activitats de conscienciació i col·laboració GLAM.

5 Conclusions

Les dades d'ús d'enllaços a repositoris digitals de patrimoni cultural són, des de qualsevol punt de vista, molt baixos. Caldria preguntar-se si es tracta d'una limitació pròpia de la font (documents històrics), que en redueix el potencial d'ús a l'hora d'escriure articles enciclopèdics. Potser les fonts primàries documentals i bibliogràfiques siguin interessants per a la recerca, però menys per a

l'elaboració de continguts informatius com els de la *Wikipedia*. No obstant això, molts articles de la *Wikipedia* són homologables als d'una enciclopèdia especialitzada —fins i tot alguns articles tracten sobre algun dels documents presents en biblioteques digitals— on sí que es podria esperar l'enllaç a fonts primàries.

Aquestes dades es poden prendre com un indicador de falta de visibilitat real dels continguts patrimonials, i de connexió escassa amb una de les comunitats digitals més actives en la difusió i organització de coneixement. Els editors de la *Wikipedia* mostren un interès per sobre de la mitjana per fonts de qualitat per millorar els seus articles, i no obstant això sembla que coneixen o manegen poc les nombroses biblioteques digitals que, a priori, serien molt útils per enriquir els articles d'àmbit historicocultural.

Hi podria haver moltes i diferents accions per estimular l'ús de biblioteques digitals en la *Wikipedia*. Algunes són relatives al mateix lloc web de la biblioteca digital i d'altres es refereixen a la promoció d'aquestes biblioteques digitals entre els wikipedistes. Des del punt de vista de les biblioteques digitals podria ser útil:

- Millorar la construcció d'URL, facilitant la identificació de *cool URI* o *permalinks*, estables en el temps, senzills i curts.
- Incorporar un *widget* "citar en la *Wikipedia*" a semblança d'*Europeana*.
- Millorar els sistemes de promoció de continguts destacats, de manera que no sigui necessari extreure els documents del repositori per difondre'ls en els portals temàtics.
- Millorar l'experiència d'usuari i usabilitat del lloc.
- Millorar el posicionament en cercadors, especialment a Google.
- Integrar els repositoris en narratives digitals amb més potencial de connexió amb els usuaris. Per exemple, fent exposicions virtuals connectades amb els repositoris.

Aquest últim punt és present en les iniciatives posades en marxa a *Europeana* per integrar la informació aportada per usuaris (*user generated content*) i narratives digitals noves (*digital storytelling*). L'objectiu és connectar amb els usuaris digitals i impulsar l'ús creatiu dels seus recursos. La creació d'exposicions digitals utilitzant els objectes continguts en les biblioteques aprofita estratègies de tall museogràfic, didàctic o de comunicació social (Saorín, 2011).

La combinació de totes aquestes accions suggereix línies de recerca noves per establir els millors models per publicar biblioteques digitals que, a més dels requisits tècnics interns, incorporin la millora de l'experiència d'usuari a través de la funcionalitat, accessibilitat, *findability*, utilitat, estètica, credibilitat i usabilitat (Hassan Montero, 2006).

Des del punt de vista de la comunitat de la *Wikipedia*, seria interessant que el sector dels museus, arxius i biblioteques (GLAM) adoptés un paper més actiu, promovent l'ús de les fonts digitals en articles i preocupant-se pels articles i wikiprojectes més relacionats amb els seus temes d'interès. La figura del wikipedista en residència pot ser complementada incloent la *Wikipedia* en les rutines de treball, com a plataforma adequada per complir la seva missió de difusió del patrimoni.

En definitiva, aquest estudi té fronts molt interessants que poden ser abordats en el futur. En primer lloc, caldria ampliar la llista de repositoris amb els de la resta de països de la comunitat lingüística espanyola, i fins i tot d'altres d'internacionals. D'altra banda, és important incloure *Wikimedia Commons* i *Wikisource* per a aquest tipus d'anàlisi, ja que s'hi recullen fonts gràfiques i documentals. La naturalesa de molts dels recursos d'aquestes biblioteques digitals és dual: serveixen com a enllaç a la font i per il·lustrar l'article. Documents històrics, mapes, pintures, objectes culturals, etc. són sovint usats des del Banc d'imatges *Wikimedia Commons*. Volem destacar que les limitacions de llicència d'ús lliure que tenen en la seva majoria aquests repositoris —a més de marques d'aigua freqüents— exclouen radicalment la seva inclusió en Commons. Aquest aspecte podria estar minant les opcions de ser enllaçats, en no respondre a una necessitat gaire determinada dels editors de la *Wikipedia*.

Finalment, considerem que les dades d'aquest estudi poden servir per avaluar de manera periòdica els resultats de les accions que s'hagin emprès per millorar la difusió i estimular l'ús de les biblioteques digitals.

Bibliografia

Abbagliati, Enzo (2011). "Portal Memoria Chilena: un ejemplo de curaduría de contenidos digitales".

Hlb: revista de historia Iberoamericana, vol. 4, nº 2.

<http://revistahistoria.universia.cl/pdfs_revistas/articulo_166_1325114205359.pdf>. [Consulta: 24/05/2012].

Codina, Lluís (2006). "Metodología de análisis y evaluación de recursos digitales en línea: Parte II: Parámetros e Indicadores (v. 2006)". *Análisis y métodos en ciencias de la documentación*.

<<http://www.lluiscodina.com/metodos.htm>>. [Consulta: 24/05/2012].

Codina, Lluís; Marcos, Mari Carmen (2005). "Posicionamiento web: conceptos y herramientas". *El profesional de la información*, vol. 14, nº 2 (marzo-abril), p. 84-99.

<<http://www.elprofesionaldelainformacion.com/contenidos/2005/marzo/1.pdf>>. [Consulta: 24/05/2012].

Elder, Danielle; Westbrook, R. Niccole; Reilly, Michele (2012). "Wikipedia lover, not a hater: harnessing Wikipedia to increase the discoverability of library resources". *Journal of web librarianship*, vol. 6, issue 1, p. 32-44.

Estivill Rius, Assumpció; Gascón García, Jesús; Sulé Duesa, Andreu (2010). "Les col·leccions digitals patrimonials espanyoles: polítiques de col·lecció i presentació de la col·lecció". *BiD: textos universitaris de biblioteconomia i documentació*, núm. 25 (juny).

<<http://www.ub.edu/bid/25/estivill1.htm>>. [Consulta: 24/05/2012].

Ferrer-Sapena, Antonia; Peset, Fernanda (2012). "La reutilización de datos culturales". *Anuario ThinkEPI*, vol. 6, p. 193-196.

Guallar, Javier; Abadal, Ernest (2009). "Evaluación de hemerotecas de prensa digital: indicadores y ejemplos de buenas prácticas". *El profesional de la información*, vol. 18, nº 3 (mayo-junio), p. 255-269. <<http://hdl.handle.net/10760/13048>>. [Consulta: 24/05/2012].

Hassan Montero, Yusef (2006). "Factores del diseño web orientado a la satisfacción y no-frustración de uso". *Revista española de documentación científica*, vol. 29, nº 2, p. 239-257.

<<http://redc.revistas.csic.es/index.php/redc/article/view/291/353>>. [Consulta: 24/05/2012].

Huvila, Isto (2010). "Where does the information come from? Information source use patterns in Wikipedia". *Information research*, vol.15, no. 3 (September). <<http://informationr.net/ir/15-3/paper433.html>>. [Consulta: 24/05/2012].

Lallye, Anne M.; Dunford, Carolyn E. (2007). "Using Wikipedia to extend digital collections". *D-Lib magazine*, vol. 13, no. 5-6 (May-June). <<http://www.dlib.org/dlib/may07/lally05lally.html>>. [Consulta: 24/05/2012].

Luyt, Brendan; Tan, Daniel (2010). "Improving Wikipedia's credibility: references and citations in a sample of history articles". *Journal of the American Society for Information Science and Technology*, vol. 61, no. 4, p. 715-722. <<http://onlinelibrary.wiley.com/doi/10.1002/asi.21304/pdf>>. [Consulta: 24/05/2012].

Nielsen, Finn Årup (2007). "Scientific citations in Wikipedia". *First Monday*, vol. 12, no. 8 (August). <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1997/1872>>. [Consulta: 24/05/2012].

Saorín, Tomás (2011). "Exposiciones digitales y reutilización: aplicación del software libre Omeka para la publicación estructurada". *Métodos de información*, vol. 2, nº 2, p. 29-46.

<<http://www.metodosdeinformacion.es/mei/index.php/mei/article/viewFile/IIMEI2-N2-029046/740>>. [Consulta: 24/05/2012].

Smith-Yoshimura, Karen; Rose Holley (2012). *Social metadata for libraries, archives, and museums. Part 3: Recommendations and readings*. Dublin, Ohio: OCLC Research.

<<http://www.oclc.org/research/publications/library/2012/2012-01.pdf>>. [Consulta: 24/05/2012].

Verwayen, Harry; Arnoldus, Martijn; Kaufman, Peter B. (2011). *The problem of the yellow milkmaid: a business model perspective on open metadata*. *Europeana White Paper*, no. 2 (November).

<<http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5>>. [Consulta: 24/05/2012].

Notes

¹ *Europeana Facts & Figures*, April 2012: <<http://pro.europeana.eu/web/guest/about/facts-figures>>. [Consulta: 24/05/2012].

² *Dump* de la *Wikipedia* en espanyol del 20 d'abril del 2012 (més de 885.000 articles). *Dump* de la *Wikipedia* en català del 19 d'abril del 2012 (més de 372.000 articles).

³ Ens ha estat difícil identificar si s'han produït canvis en el format dels URL durant els últims anys, de manera que, si haguessin succeït, hi podria haver enllaços no identificats en el nostre estudi.

⁴ Un cas paradigmàtic podrien ser les Capitulaciones de Santa Fe, que el Ministerio de Cultura destaca a "Novedades Pares" en ser incloses per la Unesco com a "[Memoria del mundo](#)". Una citació a aquest document no comptabilitzaria en la nostra recollida de dades.

⁵ Que són la majoria dels recollits en [BuscaRepositorios](#) i recol·lectats en [Recolecta](#).

⁶ Per exemple, la col·lecció de 341 obres de la *Biblioteca Digital Floridablanca* dins el repositori [Digitum](#) de la Universidad de Murcia, o la col·lecció Somni del repositori [Roderic](#) de la Universitat de València.

⁷ Els repositoris mixtos estan marcats amb un asterisc (*) quan el percentatge de la col·lecció patrimonial és molt significatiu (més del 50 %), doble asterisc (**) quan és més gran del 25 % i triple (***) quan és inferior. Com que les xifres són baixes, fem servir la dada total d'enllaços. En l'apèndix 1 oferim els valors resultants d'aplicar la ponderació calculada per a cada repositori.

⁸ Fem servir l'estimació d'enllaços conseqüència dels criteris de ponderació de col·leccions patrimonials (vegeu l'apèndix 1).

⁹ Prenent les recomanacions de Cool URIs don't change: <<http://www.w3.org/Provider/Style/URI>>. [Consulta: 24/05/2012].

¹⁰ Per exemple: <<http://www.cervantesvirtual.com/FichaObra.html?portal=0&Ref=8183&video=1>> es transforma en un URL més descriptiu: <<http://www.cervantesvirtual.com/obra/conde-de-gondomar-en-la-real-biblioteca--0/>> [Consultes: 24/05/2012].

¹¹ Exemple: <http://bib.cervantesvirtual.com/bib_autor/alfonsoelsabio/>. [Consulta: 24/05/2012].

¹² Exemple: <<http://www.cervantesvirtual.com/historia/textos/medieval/>>. [Consulta: 24/05/2012].

¹³ Exemple: <<http://www.cervantesvirtual.com/FichaObra.html?Ref=1102&portal=157>>. [Consulta: 24/05/2012].

Apèndix 0. Biblioteques digitals seleccionades

- Àmbit europeu (EU):
 1. **Europeana**
<http://www.europeana.eu>
<http://exhibitions.europeana.eu>
<http://www.europeanaregia.eu>
 2. **Portal Europeo de Archivos**
<http://www.archivesportaleurope.eu>
 3. **The European Library**
<http://www.theeuropeanlibrary.org>
- Àmbit estatal espanyol (ES):
 1. **Biblioteca Digital Hispánica y Hemeroteca Digital de la Biblioteca Nacional de España**
<http://bdh.bne.es>
<http://bibliotecadigitalhispanica.bne.es>
<http://hemerotecadigital.bne.es>
 2. **Biblioteca Virtual de Prensa Histórica**

- <http://prensahistorica.mcu.es>
- 3. **Biblioteca Virtual del Patrimonio Bibliográfico**
<http://bvpb.mcu.es>
- 4. **Red Digital de Colecciones de Museos de España**
<http://ceres.mcu.es>
- 5. **Pares (Portal de Archivos Españoles)**
<http://pares.mcu.es>
- 6. **Fototeca del Patrimonio Histórico**
http://www.mcu.es/fototeca_patrimonio/
<http://ipce.mcu.es/documentacion/fototeca/fondos>
- 7. **Hispana**
<http://roai.mcu.es>
- Àmbit regional (REG):
 - 1. **Biblioteca Digital de la Región de Murcia**
<http://bibliotecadigital.carm.es>
 - 2. **Biblioteca Virtual de Andalucía**
<http://www.bibliotecavirtualdeandalucia.es>
 - 3. **Biblioteca Virtual de Aragón**
<http://bibliotecavirtual.aragon.es>
 - 4. **Fondo Documental Histórico de las Cortes de Aragón**
<http://www.cortesaragon.es/fileadmin/ESTATICAS/fondoAntiguo.htm>
 - 5. **Biblioteca Digital de Castilla y León**
<http://bibliotecadigital.jcyl.es>
 - 6. **Archivo de la imagen de Castilla-La Mancha**
http://clip.jccm.es/archivo_de_la_imagen
 - 7. **Biblioteca Digital de Castilla-La Mancha**
<http://clip.jccm.es/bidicam/es/>
 - 8. **Biblioteca Digital de la Comunidad Madrid**
<http://www.bibliotecavirtualmadrid.org>
 - 9. **Memòria Digital de Catalunya**
<http://mdc.cbuc.cat>
 - 10. **ARCA (Arxiu de Revistes Catalanes Antiques)**
<http://mdc2.cbuc.cat>
 - 11. **BiValdi (Biblioteca Valenciana Digital)**
<http://bv2.gva.es>
 - 12. **BINaDI. Biblioteca Digital de Navarra**
<https://administracionelectronica.navarra.es/binadi>
http://www.navarra.es/appsext/bnd/GN_Ficheros_PDF_Binadi.aspx
 - 13. **Liburuklik: Biblioteca Digital Vasca**
<http://www.liburuklik.euskadi.net>
 - 14. **Memoria Digital Vasca**
<http://www.memoriadigitalvasca.es>
 - 15. **Galiciana: Biblioteca Dixital de Galicia**
<http://galiciana.bibliotecadegalicia.xunta.es>
 - 16. **Biblioteca Dixital de Galicia (Cidade de Cultura Galega)**
<http://www.csbg.org/bibliotecadixital>
<http://www.csbg.org/prensagalega>
<http://www.cmg.xunta.es/visor/prensa/>
<http://www.cmg.xunta.es/mediateca/hemeroteca/>
http://bvg.centromultimedia.net/search*gag/fcmapa*/
<http://www.cmg.xunta.es/mediateca/cartografia/>
 - 17. **Memoria Digital de Canarias**
<http://mdc.ulpgc.es>
 - 18. **Archivos Históricos de la Región de Murcia. Proyecto Carmesí**
<http://www.regmurcia.com/servlet/s.SI?METHOD=FRMSENCILLA2&sit=c.373.m.139.serv.Carmesi>
 - 19. **Biblioteca Virtual del Principado de Asturias**
<http://www.bibliotecavirtual.asturias.es>
 - 20. **Biblioteca Virtual de La Rioja**
<http://bibliotecavirtual.larioja.org>
 - 21. **ANC (Arxiu Nacional Catalunya)**
<http://cultura.gencat.net/anc/>
 - 22. **Documentos y Archivos de Aragón**

Sense ànim de ser exhaustius, hem recollit unes quantes biblioteques digitals de caràcter local o provincial (LOC):

1. **Biblioteca Virtual de la Diputación de Zaragoza**
<http://www.bivizar.es>
2. **Biblioteca Digital de la ciudad de Murcia**
<http://www.murcia.es/bibliotecadigitaldemurcia/>
3. **Biblioteca Digital Leonesa**
<http://www.saber.es>
4. **Memoria de Madrid**
<http://www.memoriademadrid.es>
5. **Liburutegi Digitala de la Biblioteca Foral de Bizkaia**
http://www.bizkaia.net/kultura/foru_liburutegia/bibdigital_spi.asp?Tem_Codigo=2542&Idioma=EU

Hem seleccionat alguna de les col·leccions de fons antic més rellevants d'universitats espanyoles (UNI):

1. **Fondo Antiguo Universidad de Zaragoza (31 %)**
<http://zaguan.unizar.es>
2. **Digitum. Fondo antiguo de la Universidad de Murcia (5 %)**
<http://digitum.um.es>
<http://hdl.handle.net/10201/>
3. **Somni. Fons històric de la Universitat de València (51 %)**
<http://roderic.uv.es>
<http://hdl.handle.net/10550/>
4. **Fondo antiguo de la Universidad de Granada (62 %)**
<http://digibug.ugr.es>
<http://hdl.handle.net/10481/>
5. **Biblioteca Digital del repositorio Gredos de la Universidad de Salamanca (61 %)**
<http://gredos.usal.es>
<http://hdl.handle.net/10366/>
6. **Fondo antiguo de la Universidad Sevilla**
<http://fondosdigitales.us.es/fondos/>
7. **Col·leccions digitals de la Universitat de Barcelona**
<http://www.bib.ub.edu/recursos-informacio/colleccions/colleccions-digitals/>
8. **Biblioteca Digital Complutense**
<http://alfama.sim.ucm.es>
9. **Fondo Fotográfico de la Universidad de Navarra**
<http://coleccionfff.unav.es/byunav>
10. **Dipòsit digital de documents de la UAB (19 %)**
<http://ddd.uab.cat>
11. **Biblioteca Digital de la UIMP**
<http://www.bduimp.es/archivo>
12. **DUGi Fons Especials: Repositori Digital de la UdG**
<http://dugifonsespecials.udg.edu>
<http://hdl.handle.net/10256.2/>
13. **e_Buah de la Universidad de Alcalá (3 %)**
<http://dspace.uah.es/>
<http://hdl.handle.net/10017/>
14. **Helvia: Repositorio Institucional de la Universidad de Huelva (32 %)**
<http://helvia.uco.es>
<http://hdl.handle.net/10396/>
15. **Jable: Archivo de Prensa Digital de la Universidad de las Palmas de Gran Canaria**
<http://jable.ulpgc.es>
16. **Biblioteca Digital de Castellón (69 %)**
<http://repositori.uji.es>
<http://hdl.handle.net/10234/>
17. **Repositori Obert UdL. Fons Especials**
<http://repositori.udl.cat/handle/10459.2/>
<http://hdl.handle.net/10459.2/>
18. **Minerva: Repositorio Institucional da Universidade de Santiago de Compostela (34 %)**
<http://dspace.usc.es>

- <http://hdl.handle.net/10347/>
19. **Fondo Histórico Digital de La Rábida** de la Universidad Internacional de Andalucía (73 %)
<http://dspace.unia.es>
<http://dspace.unia.es/handle/10334/105>
 20. **RIUMA**. Universidad de Málaga. Patrimonio (14 %)
<http://riuma.uma.es>
<http://hdl.handle.net/10630/>
 21. **Rodin**. Universidad de Cádiz. Patrimonio bibliográfico (42 %)
<http://rodin.uca.es>
<http://hdl.handle.net/10498/>
 22. **Repositorio da Universidade da Coruña** (12 %)
<http://ruc.udc.es>
<http://hdl.handle.net/2183/>
 23. **RUIdeRA**: repositorio institucional de la UCLM (12 %)
<https://ruidera.uclm.es>
<http://hdl.handle.net/10578/>
 24. **UVaDOC**: repositorio documental de la Universidad de Valladolid (74 %)
<http://uvadoc.uva.es>
 25. **Arias Montano**: repositorio institucional de la Universidad de Huelva (25 %)
<http://rabida.uhu.es/dspace/>
<http://hdl.handle.net/10272/>
 26. **Biblioteca Digital del CEU** (8 %)
<http://dspace.ceu.es>
<http://hdl.handle.net/10637/>
 27. **Fons Sol-Torres** del Servei de Biblioteca i Documentació de la Universitat de Lleida
<http://soltorres.udl.cat>
<http://hdl.handle.net/10459/>
 28. **DigiBUO**: Biblioteca Digital de la Universidad de Oviedo
<http://digibuo.sheol.uniovi.es/>

Hem seleccionat un conjunt divers de col·leccions digitalitzades d'institucions culturals de diferent naturalesa i abast (INS):

1. **Biblioteca Virtual de Derecho Aragonés**
<http://www.derechoaragones.es/es/estaticos/contenido.cmd?pagina=estaticos/indice>
2. **Teatro Siglo de Oro** de la Biblioteca Nacional de España
<http://teatrosiglodeoro.bne.es/es/RecursosBibliograficos/indexEstudios.html>
3. **Biblioteca Virtual Miguel de Cervantes**
<http://www.cervantesvirtual.com>
4. **Biblioteca Digital Ateneo de Madrid**
http://www.ateneodemadrid.com/biblioteca_digital/
5. **Biblioteca Digital del Real Jardín Botánico**
<http://bibdigital.rjb.csic.es>
6. **Archivo del Ateneo de Madrid**
<http://archivo.ateneodemadrid.es>
7. **Biblioteca Virtual Ignacio Larramendi**
<http://www.larramendi.es/i18n/estaticos/contenido.cmd?pagina=estaticos/bibliotecall>
8. **Iuris Digital**: Biblioteca Virtual de la Real Academia de Jurisprudencia y Legislación
<http://bvrajyl.insde.es>
9. **Cartoteca Digital** de l'Institut Cartogràfic de Catalunya
<http://cartotecadigital.icc.cat>
10. **Almirall: portal de pensament i cultura del segle XIX**
<http://194.224.194.201:9080/Almirall/>
11. **Recursos de Investigación de la Alhambra**
<http://www.alhambra-patronato.es/ria/>
<http://hdl.handle.net/10514/>
12. **Repositorio institucional Arjona y Cubas** de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes
<http://repositorio.racordoba.es>
<http://hdl.handle.net/10853/>
13. **Biblioteca Virtual Sierra Pambley**
<http://bibliotecavirtualsierrapambley.org>
14. **Biblioteca Saavedra Fajardo de Pensamiento Político Hispánico**
<http://saavedrafajardo.um.es>

15. **Biblioteca Virtual de la Real Academia Nacional de Farmacia. España**
<http://bibliotecavirtual.ranf.com>
16. **Biblioteca Digital de la Real Academia de Historia**
<http://bibliotecadigital.rah.es>

Finalment, en la categoria TEST hem considerat les col·leccions següents:

1. **Hemeroteca de l'ABC**
<http://hemeroteca.abc.es>
2. **Hemeroteca de La Vanguardia**
<http://hemeroteca.lavanguardia.es>
3. **Memoria de Chile**
<http://www.memoriachilena.cl>
4. **Biblioteca Digital Mundial de la Unesco**
<http://www.wdl.org/es>
5. **Col·leccions digitals de la Library of Congress**
<http://chroniclingamerica.loc.gov>
<http://memory.loc.gov>
<http://lcweb2.loc.gov/diglib>
<http://www.loc.gov/pictures/collection>
<http://international.loc.gov>
6. **Google Books**
<http://books.google.com>

Apèndix 1. Dades ponderades per a repositoris mixtos

Dades conjuntes per a la WP-ES i la WP-CAT. Al costat del nom del repositori s'indica la ponderació aplicada, calculada a partir del percentatge de les col·leccions patrimonials sobre el total del repositori, i el nombre d'enllaços estimats. Les dades de les biblioteques digitals marcades amb (***) poden ser poc fiables per a col·leccions patrimonials. Per contra, les marcades amb (*) tenen una alta probabilitat de ser vàlides.

Repositoris mixtos		Total enllaços	% Pond.	Pond. enllaços
UVaDOC: repositorio documental de la Universidad de Valladolid	*	5	74	4
Biblioteca Digital de Castellón	*	22	69	15
Fondo antiguo de la Universidad de Granada	*	30	62	19
Biblioteca Digital del repositorio Gredos de la Universidad Salamanca	*	42	61	26
Somni. Fons històric de la Universitat de València	*	4	51	2
Rodín. Universidad de Cádiz. Patrimonio bibliográfico	**	15	42	6
Minerva: Repositorio Institucional da Universidade de Santiago de Compostela	**	3	34	1
Helvia: Repositorio Institucional de la Universidad de Huelva	**	23	32	7
Fondo Antiguo de la Universidad de Zaragoza	**	4	31	1
Arias Montano: repositorio institucional de la Universidad de Huelva	**	21	25	5
Dipòsit digital de documents de la UAB	***	234	19	44
RIUMA. Universidad de Málaga. Patrimonio	***	4	14	1
Repositorio da Universidade da Coruña	***	18	12	2
RUIdeRA: repositorio institucional de la UCLM	***	2	12	0
Biblioteca Digital del CEU	***	1	8	0
Digitum. Fondo antiguo de la Universidad de Murcia	***	30	5	2
e_Buah de la Universidad de Alcalá	***	48	3	1

Per al recol·lector *Hispana* hem fet una estimació del 30 % i només rep dos enllaços.

Apèndix 2. Biblioteques digitals sense enllaços des de la *Wikipedia*

Un total de catorze biblioteques digitals no tenen enllaços en cap de les dues edicions (WP-ES ni WP-CAT): *Biblioteca Virtual Sierra Pambley*; *Biblioteca Virtual de la Real Academia Nacional de Farmacia*. Espanya; *Repositorio institucional Arjona y Cubas* de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes; *Biblioteca Digital de la ciudad de Murcia*; *Archivo de la imagen de Castilla-La Mancha*; *Liburuklik: Biblioteca Digital Vasca*; *Biblioteca Virtual de Aragón*; *Documentos y Archivos de Aragón*; *Archivos Históricos de la Región de Murcia*. Proyecto Carmesí; *Biblioteca Digital de la UIMP*; *DigiBUO*: Biblioteca Digital de la Universidad de Oviedo; Col·leccions digitals de la Universitat de Barcelona; Fondo Fotográfico de la Universidad de Navarra; *Repositori Obert UdL*. Fons Especials.

Setze casos no tenen enllaços en la WP-ES: *Portal Europeo de Archivos*; *Almirall: portal de pensament i cultura del segle XIX*; *Repositorio institucional Arjona y Cubas* de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes; *Biblioteca Virtual de la Real Academia Nacional de Farmacia*. Espanya; *Biblioteca Virtual Sierra Pambley*; *Cartoteca Digital* de l'Institut Cartogràfic de Catalunya; *Biblioteca Digital de la ciudad de Murcia*; *Documentos y Archivos de Aragón*; *Biblioteca Virtual de Aragón*; *Archivo de la imagen de Castilla-La Mancha*; *Archivos Históricos de la Región de Murcia*. Proyecto Carmesí; *Liburuklik: Biblioteca Digital Vasca*; *DigiBUO*: Biblioteca Digital de la Universidad de Oviedo; Col·leccions digitals de la Universitat de Barcelona; *Repositori Obert UdL*. Fons Especials; Fondo Fotográfico de la Universidad de Navarra y *Biblioteca Digital de la UIMP*.

Quaranta biblioteques digitals no tenen enllaços en la WP-CAT: *Hispana*; *Fototeca del Patrimonio Histórico*; *Portal Europeo de Archivos*; Archivo del Ateneo de Madrid; *Iuris Digital*: Biblioteca Virtual de la Real Academia de Jurisprudencia y Legislación; *Recursos de Investigación de la Alhambra*; *Repositorio institucional Arjona y Cubas* de la Real Academia de Córdoba de Ciencias, Bellas Letras y Nobles Artes; *Teatro Siglo de Oro* de la Biblioteca Nacional de España; *Biblioteca Virtual de la Real Academia Nacional de Farmacia*. Espanya; *Biblioteca Virtual Sierra Pambley*; *Biblioteca Digital* de la Real Academia de Historia; *Memoria de Madrid*; *Biblioteca Virtual de la Diputación de Zaragoza*; *Biblioteca Digital de la ciudad de Murcia*; *Biblioteca Digital de la Comunidad Madrid*; *Documentos y Archivos de Aragón*; *Biblioteca Virtual de Aragón*; *Archivo de la imagen de Castilla-La Mancha*; *Biblioteca Dixital de Galicia (Cidade de Cultura Galega)*; *Biblioteca Digital de Castilla-La Mancha Archivos Históricos de la Región de Murcia*. Proyecto Carmesí; *Biblioteca Digital de la Región de Murcia*; *Liburuklik: Biblioteca Digital Vasca*; *Galiciana: Biblioteca Dixital de Galicia*; *RUIdeRA*: repositorio institucional de la UCLM; *Fondo Histórico Digital de La Rábida* de la Universidad Internacional de Andalucía; *Biblioteca Digital del CEU*; *Fondo Antiguo Universidad de Zaragoza*; *Minerva: Repositorio Institucional da Universidade de Santiago de Compostela*; *Rodin*. Universidad de Cádiz. Patrimonio bibliográfico; *DigiBUO*: Biblioteca Digital de la Universidad de Oviedo; Col·leccions digitals de la Universitat de Barcelona; *e_Buah* de la Universidad de Alcalá; *Repositori Obert UdL*. Fons Especials; Fondo Fotográfico de la Universidad de Navarra; *UVaDOC*: repositorio documental de la Universidad de Valladolid; *Biblioteca Digital de la UIMP*; *Arias Montano*: repositorio institucional de la Universidad de Huelva; *RIUMA*. Universidad de Málaga. Patrimonio; *Jable: Archivo de Prensa Digital* de la Universidad de las Palmas de Gran Canaria.

Apèndix 3. Detalls del procés de selecció d'enllaços

Reprement l'aspecte tècnic de l'anàlisi, hem utilitzat els [dumps](#) d'enllaços (*-externallinks.sql.gz) que proporciona la Fundació Wikimedia al seu web. Aquest fitxer inclou tots els enllaços a llocs web externs de la *Wikipedia*, inserits en qualsevol part dels articles. Inicialment considerem analitzar directament el text brut dels articles i extreure'ls d'aquí, però l'anàlisi amb el *dump* d'enllaços és més ràpida i exacta, si bé és cert que perdem informació sobre quan va ser inserit cada enllaç.

En l'anàlisi només s'han tingut en compte els enllaços que es troben en articles de l'enciclopèdia (conegut com a *espai de noms principal*), i s'han descartat per complet els situats en les seves pàgines de discussió o d'altres de manteniment.

Els resultats s'han exportat a fitxers CSV generant un [dataset](#) amb llicència lliure, amb l'ànim que els nostres resultats puguin ser analitzats i reutilitzats per qui vulgui.

Apèndix 4. Enllaços trencats en la WP-ES i la WP-CAT

Estat de l'enllaç	Wikipedia en espanyol	Wikipedia en català
Correcte (OK)	6.173	1.452
Pàgina no trobada (error 404)	119	40
error 500	69	37
Temps expirat (25 segons sense resposta)	68	108
Enllaços mal escrits	18	2
error 504	9	4
error 400	4	1
error 503	0	22
Total	Accessibles: 6.173 (95,56 %) Inaccessibles: 287 (4,44 %)	Accessibles: 1.452 (87,15 %) Inaccessibles: 214 (12,85 %)

