

LA DIFUSIÓ DELS ARXIUS I EL RECONeixEMENT DE CARÀCTERS. EL PROJECTE EINES DE L'ARXIU COMARCAL DEL BAIX LLOBREGAT I LES PERSPECTIVES DE FUTUR PER ALS ARXIUS

Enric Cobo Barri, cap del Servei de Coordinació General d'Arxius de la Generalitat de Catalunya
M. Luz Retuerta Jiménez, directora de l'Arxiu Comarcal del Baix Llobregat

En el marc de la societat del coneixement i l'impacte de les noves tecnologies en el món dels arxius, que han incidit en els diferents àmbits, com la protecció i la difusió del patrimoni documental, la gestió documental, l'administració electrònica o l'aplicació de les polítiques de transparència, accés a la informació i bon govern, s'han desenvolupat diferents projectes que de ben segur, en un futur immediat, esdevindran una línia més de treball dels arxius. Una d'aquestes línies que es consolidarà són les tecnologies de la visió per ordinador i el reconeixement massiu de caràcters aplicat als arxius. Com a exemple, l'experiència pionera desenvolupada a l'Arxiu Comarcal del Baix Llobregat amb l'ús d'aquestes tecnologies aplicades als padrons d'habitants de Sant Feliu de Llobregat, un projecte que obre un seguit de noves possibilitats per als arxius catalans.

Les tecnologies de la visió per ordinador són aquelles que permeten desenvolupar programes de reconeixement massiu de caràcters de lletra manuscrita i mecanoscrita per fer possible una transcripció automàtica o semiautomàtica dels documents digitalitzats. Simplificant molt, es tracta de fer que les màquines sàpiguen llegir documents mecanoscrits, manuscrits o d'altra tipologia. Aquesta transcripció permet posar a l'abast dels usuaris un gran volum d'informació i dades, a les quals, per la seva complexitat o laboriositat, no tothom pot tenir accés. Aquest tipus de tecnologies permet potenciar la recerca i ampliar la tipologia i el nombre de persones usuàries dels arxius. Els arxius, per tant, estan obligats a adaptar-se, tot superant les funcions exclusivament culturals i administratives, per transformar-se en un servei per al conjunt de la societat. I han de garantir l'abast universal i el caràcter democràtic i igualitari de l'accés dels ciutadans a la informació i la documentació.

1. LA DIFUSIÓ DELS ARXIS EN L'ENTORN DE LA SOCIETAT DEL CONEIXEMENT I L'ADMINISTRACIÓ ELECTRÒNICA

En les últimes dècades s'ha produït una gran transformació del paper que desenvolupen els arxius i la manera en què es presenten davant la societat i les administracions. Els arxius del present i del futur estan immersos en una societat que progressa ràpidament cap a un nou model de relació entre persones, entitats i administracions, fruit de les noves demandes i dels nous usuaris propis de la societat del coneixement i de l'administració electrònica. El ventall de persones que sol·liciten serveis als arxius i els tipus de demandes s'han ampliat i diversificat substancialment. Des del moment en què els arxius han deixat de ser serveis adreçats només a la recerca i a l'administració per dirigir-se al conjunt de la societat, han adquirit nous objectius, procediments i eines. Aquesta realitat representa un canvi en el paradigma de la comunicació dels arxius, tant pel que fa a les metodologies a aplicar com a la relació que aquesta ha de tenir amb les eines que cal utilitzar per garantir l'accés i la difusió eficient als fons i als documents. Som davant un nou escenari on els arxius han d'assolir els seus objectius en funció de la seva tipologia. Aquest fet pot provocar una especialització dels arxius, però en alguns casos l'arxiu és un únic servei que ha de vetllar per garantir la conservació, el tractament, l'accés i la difusió dels fons més antics, ahora que ha d'impulsar la implantació de sistemes de gestió documental, l'administració electrònica o, conjuntament amb altres unitats, l'aplicació de les polítiques de transparència, accés a la informació i bon govern.

A la llum de les transformacions tecnològiques i socials, els reptes tradicionals prenen una nova dimensió, alhora que s'hi afegeixen nous desafiaments. Entre els usos, els valors i les línies d'actuació dels arxius que s'han vist transformats, destaquem:

- › La **democratització de l'accés i la difusió dels continguts**, per mitjà de l'aprofundiment i la millora de les descripcions, que permet posar la documentació a l'abast dels usuaris de manera presencial i, sobretot, en línia. En aquest sentit, prenen rellevància els portals per accedir als documents i la proliferació de productes de difusió, tant en els suports tradicionals com en els nous, com ara visites, conferències, jornades, exposicions, materials didàctics, documentals, aplicacions de mòbils, etc.
- › Els **nous reptes del món educatiu**, entenent l'educació en un sentit ampli, com un procés d'aprenentatge que es desenvolupa al llarg de tota la vida. Els documents de l'arxiu poden ser un referent per a l'alumnat en totes les etapes educatives reglades, per a la ciutadania en general i per a col·lectius concrets (entitats, persones adultes, joves, associacions de dones, gent gran, professionals específics, etc.). Es tracta de desvetllar el potencial educatiu dels documents i de presentar-los i adaptar-los a les diferents necessitats dels humans. També s'ha de tenir en compte el repte de l'escola actual, que és engrescar l'alumnat a aprendre coneixements rellevants i significatius, motiu pel qual s'utilitzen metodologies diverses que responen a l'aprenentatge basat en problemes i projectes.¹ Darre-rament ha pres importància el disseny de les preguntes que possibiliten processos d'investigació i les activitats que ajuden a ensenyar a formular-les. La metodologia arxivística que ens apropa al context i la provi-nença, i que contribueix a interrogar-nos sobre l'origen i la funció dels documents, pot ser d'utilitat per a aquests nous plantejaments educatius.

Els arxius ofereixen més productes per a l'educació i progressivament incorporen les tecnologies de l'aprenentatge i la comunicació (TAC) com a instruments essencials dels processos d'aprenentatge. El portal **Patri-moni Cultural - Educació** és un exemple d'espai vinculat al patrimoni cultural en el qual es presenten activitats que, amb el desenvolupament de les tecnologies de visió per ordinador, podrien facilitar productes més consumibles, útils i atractius per als docents i per a l'alumnat.

- › La **dinamització del patrimoni per la necessitat d'enfortir i de crear identitats**. Els arxius són equipaments patrimonials bàsics i de referència en una societat global que necessita crear referents identitaris amb el territori i amb la seva gent. Els arxius són centres que gestionen patrimoni en un doble sentit, perquè els documents formen part del patrimoni documental d'una ciutat o d'un país, i perquè alhora són la font primària per informar d'un altre tipus de patrimoni, com és el patrimoni arquitectònic, natural, històric, etc.

- › Els documents d'arxiu són una font primària per a **la recerca i la construcció de nou coneixement**. El coneixement és la informació combinada amb l'experiència, el context, la interpretació i la reflexió. Les característiques específiques de la metodologia arxivística, vinculada a l'anàlisi, la classificació, la descripció del contingut i la contextualització dels documents, són molt útils a l'hora d'estimular les habilitats, especialment necessàries en el marc de la societat de la informació, ja que permeten ensenyar el camí per aconseguir destriar i donar significat a les dades necessàries per construir coneixement. D'altra banda, els arxius impulsen la recerca a través de l'assessorament als usuaris, creant beques i premis i donant suport a aquelles institucions encarregades de promoure la recerca (universitats, centres d'estudis, etc.).

- › Les funcions dels arxius no poden ser alienes a l'entorn en el qual actuen. Els arxius tenen responsabilitats envers la societat, el que es vincula al concepte de **responsabilitat social corporativa**, que intenta integrar les preocupacions socials i ambientals en les empreses i que, en el cas de l'Administració, implica anar més enllà de les obligacions jurídiques de la funció pública per tal de buscar una manera de fer i uns resultats d'allò que fem. L'objectiu final és incrementar el retorn de continguts i valors a la societat a través de:
 - Fomentar diversos usos de la documentació (formatius, lúdics, reivindicatius, de gènere, etc.) que contribueixin a fer una societat més equitativa (formació ocupacional, plans d'ocupació, etc.).
 - Fer accessible la informació perquè sigui reutilitzada per empreses, entitats o persones per elaborar productes culturals i socials propis o per obrir noves línies de negoci (*big data i open data*).
 - Possibilitar una memòria més plural per mitjà d'activitats adreçades a col·lectius i agents que actuen en el territori (sindicats, associacions, mitjans de comunicació, etc.).

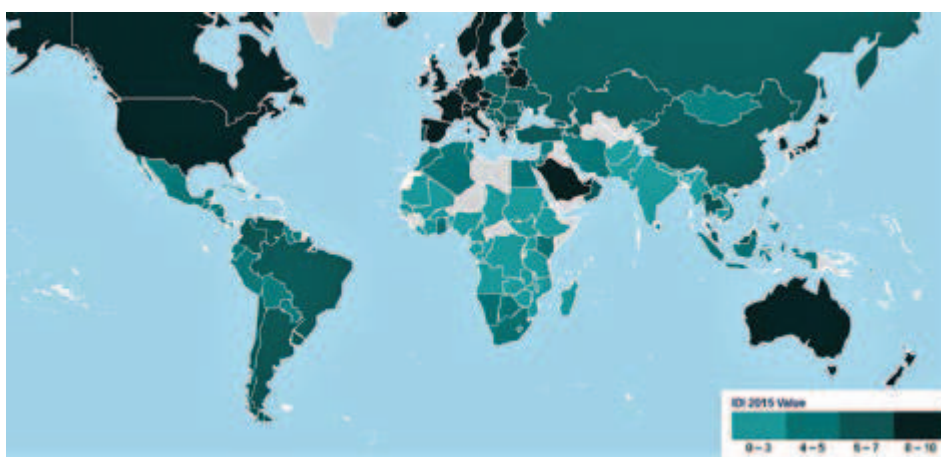
- › Les funcions de gestió documental i arxiu s'emmarquen en la progressiva necessitat de **transparència administrativa i bon govern**, que promou l'accés lliure a la documentació per enriquir el coneixement de la societat, impulsar la democràcia, protegir els drets de la ciutadania i millorar-ne la qualitat de vida.
- › En relació amb **l'administració electrònica i el treball multidisciplinari dels nous arxivers i gestors de documents**, la promulgació de nova legislació en matèria de procediment administratiu,² que ha entrat en vigor el passat mes d'octubre del 2016, obliga els arxius a reconfigurar estratègies, metodologies i procediments en la seva activitat ordinària. Aquestes noves estratègies requereixen coordinar-se constantment amb altres unitats de l'organització, alhora que els nous canals de l'administració electrònica i d'Internet comporten nous hàbits i noves fórmules de comunicació entre els ciutadans i les administracions.

Per tant, l'impacte dels arxius en la societat és molt alt i imprescindible per al bon funcionament de les Administracions Públiques, però també per garantir un accés a la informació més universal, ràpid i de qualitat. En aquest sentit els arxius han de modificar els seus hàbits de funcionament respecte dels seus usuaris.

1.1 L'ENTORN DE LA SOCIETAT DEL CONEIXEMENT EN EL MARC DELS ARXIS I ELS SERVEIS DE GESTIÓ: NOUS CANALS DE COMUNICACIÓ

En el *Mobile World Congress*, celebrat a Barcelona al març del 2015, va aparèixer una referència a la frase *internet de les coses* en què situa Internet en el cor del nou model de la humanitat dels propers anys, fins a arribar a la postulació extrema de dir que Internet és tan vital que qualsevol ciutadà del món havia de tenir-hi accés, per sobre de disposar d'aliments, condicions sanitàries o treball digne. Però més enllà d'aquest debat, el missatge és clar i llança una associació entre el que es pot crear, comprar o utilitzar en el món desenvolupat i Internet. En el cas del món desenvolupat, els usuaris d'Internet se situen per sobre del 65%, a Espanya superen el 76% i a Catalunya són el 79% dels ciutadans. Aquest aspecte, però, situa Internet en el punt de mira de qualsevol negoci o activitat i, en conseqüència, els productes i els serveis que desenvolupen les administracions i les empreses han de tenir molt present aquesta orientació. Hi ha sectors com el bancari, el comercial o el turístic que ja fa temps que han

incorporat Internet en les seves estratègies comercials. En l'entorn de l'Administració, tenim l'exemple de l'actuació de l'Agència Tributària del Govern d'Espanya, que fa anys que promou una interrelació amb els ciutadans a través de l'ús intensiu dels seus serveis per Internet. Per tant, els arxius tenen en Internet un canal extraordinàriament potent per oferir als ciutadans els seus productes i serveis i fer-los participants de les seves activitats i del seu funcionament, i per interactuar amb usuaris de tot el món. Aquesta interacció multicanal té l'objectiu de promoure la difusió del patrimoni cultural, la implantació de sistemes de gestió documental, administració electrònica o preservació digital i ser un referent en polítiques de transparència, accés a la informació i bon govern.



Imatge 1.
Mapa de països amb accés a Internet. Unió Internacional de Telecomunicacions. Informe sobre el desenvolupament del Mundial de las Telecomunicacions. <http://www.itu.int/net4/ITU-D/idi/2015/#idi2015map-tab>

Si el futur ens proposa un canal de comunicació i interrelació d'abast mundial, els arxius han d'oferir a la societat els seus principals valors: els fons i els documents. La porta, d'entrada i sortida, és Internet i requereix que els arxius adaptin els seus continguts a aquest canal, que té les seves regles i comportaments. Això vol dir que els arxius han de transformar o adaptar els seus productes i els seus serveis a aquesta nova realitat i fer-los visibles a través de les pàgines web pròpies o d'aquelles organitzacions que puguin valorar-los i implicar-los activament en les xarxes socials. L'impacte de l'ús de Facebook, Twitter, YouTube o Instagram per a un arxiu és molt alt i representa una excel·lent oportunitat per difondre els seus continguts, productes i serveis. Això explica que el rànquing dels arxius amb més piulades no es correspongui amb els que tenen fons més rellevants.³ Fer gran el petit o, millor dit, conegut el menys conegut (per raó de

la seva rellevància o volum documental), és una oportunitat competitiva que ofereix Internet. Per tant, els arxius han d'adaptar-se a aquesta situació i revisar el *modus operandi* en tots els nivells. En aquest sentit, s'han de plantejar modificacions importants en les estratègies corporatives o del mateix arxiu, en la metodologia dels tractaments arxivístics, en la presentació de la informació o documentació als ciutadans i en el rol de l'arxiu en el suport a les Administracions públiques.



Imatge 2.

Relació d'arxius i entitats arxivístiques amb més presència a twitter.

Font: Cruz, Adrià. Els Arxius a Twitter: 10 piladors que cal seguir, a El Núvol Digital, 2015.
<http://www.nuvol.com/opinio/els-arxius-a-twitter-10-piuladors-que-cal-seguir/>

Són molts els arxius que ja apliquen criteris d'anàlisi de necessitats i valoren les demandes dels usuaris perquè han detectat que són les tendències, els interessos i les necessitats de la societat els que han de fixar, entre altres, els objectius de l'arxiu. Utilitzant una terminologia poc arxivística, els serveis i els productes que ofereix l'arxiu no poden obviar la demanda del mercat. Aquesta posició estratègica dels arxius només s'assolirà si es modifiquen, es milloren i s'adapten algunes estratègies, continguts i procediments. Per assolir aquest objectiu, hi ha algunes recomanacions, entre les quals destaquem:

- › *Comprensió*. Els fons i els documents han de ser comprensibles per als ciutadans. L'arxiu ha d'elaborar els continguts pensant que els canals d'accés i difusió tradicionals es complementen i, progressivament, seran substituïts pels d'Internet.

- › *Accés a la documentació.* L'aplicació de polítiques de transparència, accés a la informació i bon govern obliguen a facilitar l'accés universal de gran part de la documentació dels arxius. Això comporta un nou disseny dels instruments d'accés i difusió orientat a canals de comunicació per Internet (web, plataformes, portals, Twitter, Facebook, YouTube, Instagram, etc.). La documentació es trasllada del seu hàbitat tradicional de la sala de consulta a l'espai d'Internet a través d'eines que faciliten un accés fàcil, ràpid i intuïtiu. Per això s'han de promoure plans de màrqueting i comunicació adequats a aquest nou escenari i que projectin els serveis i els productes dels arxius a la societat i a les administracions.
- › *Reutilització de dades.* La informació dels arxius esdevé, en gran part, reutilitzable, especialment la informació pública, tot aplicant criteris normalitzats que garanteixin el compliment de la legislació vigent com la Directiva 2003/98/CE del Parlament Europeu i del Consell, relativa a la reutilització de la informació del sector públic, o la Llei 37/2007, de 16 de novembre, sobre reutilització de la informació del sector públic, sense oblidar la part d'accés que preveu la Llei 10/2001, de 13 de juliol, d'arxius i documents; per tant, accions per facilitar la consulta de la informació amb l'ús de mitjans informàtics en formats comprensibles que permetin la reutilització de les dades i els documents. En aquesta línia, la disponibilitat de dades obertes esdevé un apartat altament requerit en els portals dels arxius.
- › *Interoperabilitat.* La gestió de les administracions i de les activitats que desenvolupen amb les persones i les entitats, públiques i privades, implica interoperabilitat. Les administracions han de garantir que la informació és compartida per facilitar una gestió eficient de les interaccions amb ciutadans, administracions i entitats. Per assolir aquest objectiu es requereix una adaptació de programaris, plataformes i recursos tecnològics de les administracions i els arxius. Es necessiten protocols d'interoperabilitat com els que planteja l'Esquema Nacional d'Interoperabilitat.

El nou escenari es basa en:

- › Un nou entorn en relació amb la metodologia del tractament massiu de dades (classificació, descripció, accés i difusió) i captura (digitalització, ingesta de documents, registre, avaluació, etc.).

- › La incorporació de nous sistemes informàtics i plataformes que garanteixin els principis de gestió dels documents electrònics sota el concepte d'arxiu únic i, no menys important, la preservació digital.
- › L'ús d'eines de participació ciutadana amb l'objectiu de recollir informació dels ciutadans, tant en relació amb els continguts dels fons i els documents dels arxius com amb el seu funcionament.
- › La participació en propostes d'interoperabilitat entre administracions que garanteixin una gestió documental correcta i una gestió eficient de l'activitat administrativa.
- › El desenvolupament i la implantació dels nous canals de comunicació entre l'arxiu i els seus usuaris (investigadors, genealogistes, historiadors, ciutadans en la consulta d'expedients, estudiants, etc.) o entitats (administracions, associacions, empreses, centres culturals, etc.).
- › La publicació de dades obertes a disposició de persones i entitats, públiques i privades, a partir de la legislació en matèria de transparència, accés a la informació i bon govern.

2. L'ARXIU COMARCAL DEL BAIX LLOBREGAT I EL PROJECTE EINES DE RECONeixEMENT DE CARÀCTERS

2.1. EL PROJECTE

El reconeixement massiu de caràcters aplicat als arxius es va experimentar per primera vegada a l'Arxiu Històric de Girona l'any 2006 amb els expedients de frontera de l'Arxiu Històric de Girona. En aquell projecte s'inicià el que podríem denominar primers «tests d'estrès» dels arxius respecte a la captura digital de documents i el seu posterior reconeixement massiu de caràcters. El valor d'aquestes primeres actuacions va raure més en la demostració del potencial dels arxius i en la validació de cara al futur de les tecnologies, les metodologies i els procediments de treball, que no pas en els resultats concrets. D'alguna manera, aquests primers anys de proves (assaig-error-solució) van facilitar, anys després, en un context diferent, la col·laboració entre el Centre de Visió per

Computador (CVC), el Centre d'Estudis Demogràfics (CED) i l'Arquebisbat de Barcelona. En aquest cas, la iniciativa va ser la digitalització i la transcripció dels llibres d'esposalles de l'Arxiu de la Catedral de Barcelona, un projecte finançat per la Unió Europea, batejat sota el nom de Five Centuries of Marriages.⁴ La tercera experiència és el projecte amb l'Arxiu Comarcal del Baix Llobregat (ACBL) sobre els padrons municipals d'habitants de Sant Feliu de Llobregat, que es va començar a definir a principi de l'any 2014, i que més endavant exposarem.

En aquest context, cal remarcar la sessió de treball a la seu del CVC, a principis de l'any 2015, en la qual van participar tècnics de la Subdirecció General d'Arxius i Museus i arxivers comarcals i municipals amb l'objectiu d'explorar possibles projectes d'impacte de la utilització de les tecnologies de visió per ordinador en els arxius. Arran d'aquesta sessió, i amb la voluntat de formalitzar un espai de treball, es va redactar un conveni de col·laboració entre el Departament de Cultura i el CVC per a l'elaboració de projectes de tractament i explotació de dades dels fons documentals dels arxius. Aquest conveni, que es signarà en els primers mesos del 2017, està orientat al treball amb els fons de l'Arxiu Nacional de Catalunya, la Xarxa d'Arxius Comarcals (XAC) i els arxius municipals relacionats amb la XAC, però darrere hi ha la voluntat d'ampliar-lo als arxius del Sistema d'Arxius de Catalunya (SAC). Finalment, cal destacar el conveni entre el CVC i l'Associació d'Arxivers-Gestors de Documents, de l'any 2015, adreçat a millorar la capacitat professional dels experts en arxivística i gestió documental.

A principis de l'any 2014 Josep Lladós Canet, director del CVC, va contactar amb el personal de l'ACBL per explicar que, conjuntament amb el CED, estaven elaborant una proposta de recerca relacionada amb documentació de l'ACBL, amb la qual volia intentar aconseguir finançament. El CVC és un centre de recerca de la Universitat Autònoma de Barcelona (UAB) dedicat a la investigació i el desenvolupament en el camp de la visió per ordinador. Duu a terme una recerca bàsica i investigació aplicada al més alt nivell internacional i fomenta la transferència de coneixements a la indústria i la societat en general. Aquest centre comparteix projectes de recerca amb altres centres, com el CED, pertanyent també a la Universitat Autònoma de Barcelona. D'altra banda, el CED és un centre dirigit aleshores per la demògrafa Anna Cabré, que s'encarrega de fer recerca sobre les estructures i les dinàmiques de la població. L'integren diversos equips de treball constituïts en funció de les propostes presentades a les convocatòries públiques d'R+D o de les demandes externes. Una de les seves línies de recerca és la de «Treball, família i gènere». Aquesta iniciativa té com a objec-

tiu la utilització dels padrons municipals d'habitants de Sant Feliu de Llobregat des del punt de vista demogràfic i també l'experimentació amb la tecnologia de la lectura d'imatges a través de la visió per ordinador. L'objectiu era: «...desenvolupar instruments i procediments que facilitin la informatització massiva de les fonts demogràfiques del passat orientada a la construcció de bancs de dades d'ús públic i a la millora i la consulta de documents dels arxius. Secundàriament, ens proposem també construir algunes eines d'anàlisi d'aquestes dades de cara a la generació de genealogies, l'establiment de trajectòries individuals i familiars i la localització especial de xarxes.»⁵

Després del projecte dels llibres d'esposalles, al CVC i al CED els interessava treballar amb els padrons municipals d'habitants perquè és una font present en gairebé tots els municipis de Catalunya i Espanya, que alhora té equivalències en la majoria de països europeus. Això permet que les dades siguin escalables a l'àmbit europeu. D'altra banda, el tipus d'estructura d'aquests documents, a manera de formulari, facilita poder fer estudis comparatius entre diversos padrons d'habitants, tant del municipi com d'altres poblacions.⁶

L'elecció dels padrons municipals d'habitants de Sant Feliu de Llobregat del període de 1828 a 1955 la van fer per les característiques demogràfiques d'aquest municipi i perquè els padrons ja estaven digitalitzats i accessibles a Internet. Gràcies a la consulta de les imatges a través del cercador d'Arxius en Línia, van comprovar que Sant Feliu era un municipi amb unes dades padronals molt completes i amb unes característiques demogràfiques molt adequades al projecte de recerca que volien fer. Aquest municipi durant aquest període no era excessivament gran (la població va oscil·lar entre els 1.585 h. del 1826 i els 8.529 h. del 1955); tenia una economia mixta, amb població activa dedicada a l'agricultura i a la indústria (hi havia fàbriques des de l'últim terç del segle XIX), i va experimentar totes les onades migratòries típiques de Catalunya durant els segles XIX i XX (amb les migracions de l'interior de Catalunya al llarg del segle XIX i principis del segle XX, de la zona de Múrcia i Almeria durant els anys vint i trenta del segle passat, i de la resta de l'Estat a partir dels anys cinquanta).

L'ACBL ha potenciat la digitalització de documentació municipal i ha fet possible la consulta, a través del cercador corporatiu, de les actes d'11 ajuntaments del Baix Llobregat (Castellví de Rosanes, Collbató, Corbera de Llobregat, Begues, Esplugues de Llobregat, Pallejà, el Papiol, Sant Feliu de Llobregat, Santa Coloma de Cervelló, Torrelles de Llobregat i Vallirana) i dels padrons d'habitants de Sant Feliu de Llobregat i Corbera i, properament, els de Castellví de Rosanes,

Collbató, Begues, Papiol, Santa Coloma de Cervelló i Torrelles de Llobregat, que ja han estat digitalitzats. Donar accés per Internet al major nombre possible de documents dels arxius és la millor manera de donar-los a conèixer i de fomentar la recerca, l'estudi i la divulgació dels seus continguts. És oferir un ventall de possibilitats, tant per a la col·lectivitat més pròxima com per a la comunitat científica que fa estudis de més abast i projecció. Mesos després, al març del 2015, el projecte Eines va ser una de les propostes seleccionades entre les 221 iniciatives presentades al programa RecerCaixa 2014.

Aquest projecte inclou, des del vessant de la demografia històrica, la creació d'un banc de dades informatitzat de 127 anys del municipi de Sant Feliu de Llobregat i, des de l'òptica informàtica, el desenvolupament de programes de reconeixement de lletra manuscrita per fer possible una transcripció semiautomàtica de les imatges dels documents digitalitzats. Es tracta de traslladar les tecnologies d'OCR (de l'anglès *optical character recognition*), utilitzades amb èxit amb les lletres d'impremta i les mecanografiades, a les lletres manuscrites. Això permetria posar a l'abast dels usuaris una gran quantitat d'informació a la qual no sempre té accés tothom, i facilitaria i potenciaria molt la recerca. Simplificant el tema, podríem apropar-nos una mica als resultats aconseguits amb la indexació i l'accés a la lletra d'impremta o feta a màquina. Tanmateix, l'aplicació del reconeixement de caràcters a les lletres manuscrites no és prou madura i, en el cas d'aquestes lletres, els autors del projecte proposen combinar la transcripció manual i el reconeixement amb l'ajut de màquines.⁷

El projecte té un impacte internacional dins del programa marc Horitzó 2020 i, en particular, en el repte «Europe in a changing world - Inclusive, innovative and reflective societies». El treball utilitza tècniques nominatives de *record linkage* o vinculació de documents, que relaciona dades de la mateixa entitat (domicili, persona, etc.) de diferents padrons d'habitants, cosa que permet seguir trajectòries personals al llarg del temps, construir genealogies i localitzar les xarxes familiars en l'espai.⁸

El projecte disposa de personal propi del CED i del CVC, amb la col·laboració del personal de l'ACBL i del personal voluntari. Utilitza una plataforma col·laborativa (*crowdsourcing*) per Internet per transcriure els padrons municipals, tasques que són a càrrec de la comunitat, en aquest cas bàsicament de la gent de Sant Feliu.⁹ Els voluntaris fan transcripcions assistides, sota la direcció i l'avaluació dels encarregats del projecte. En les transcripcions es fan servir el reconeixement

ment dels termes més freqüents i l'arrossegament cronològic de la informació, la qual cosa permet crear una estadística de la dinàmica de les llars.

La iniciativa vol contribuir a dinamitzar el patrimoni històric entre la comunitat com un instrument que ajudi a divulgar el coneixement i la comprensió del passat. S'intenta aconseguir també una projecció social en el procés de la recerca i en els resultats. També es vol aprofitar el coneixement col·lectiu de la població local per incorporar-lo al projecte.

Es tracta d'una iniciativa molt participativa que busca contribuir a popularitzar els arxius a través de centrar el reclam en l'activitat de cerca dels avantpassats. La captació de voluntaris es fa a títol individual i a través d'un apropament al món associatiu. Fins ara s'ha incidit directament en les entitats vinculades al patrimoni, com el Centre d'Estudis Comarcals del Baix Llobregat i Arrels Locals, però hem previst contactar amb altres tipus de col·lectius (gent gran, grups de dones, etc.). Una altra iniciativa que pot sumar esforços i que hem començat a explorar són les accions encaminades a donar a conèixer el projecte entre l'alumnat dels últims cursos d'ESO i de batxillerat dels centres de secundària de Sant Feliu de Llobregat.

El projecte preveu estendre's a altres municipis de Catalunya. En aquest sentit, l'Arxiu Comarcal del Baix Llobregat ha començat a donar suport a l'Arxiu Municipal de Molins de Rei per tal que s'afegeixi al projecte Eines. Ho ha fet oferint el maquinari i les aplicacions necessàries perquè personal adscrit a aquest arxiu pogués digitalitzar els seus padrons. Entre els anys 2015 i 2016 han fotografiat tots els padrons de Molins de Rei des del segle XIX fins a l'any 1955.

2.2. ACTIVITATS REALITZADES

El 16 d'abril del 2015 es va presentar el projecte a la seu de l'ACBL. Des d'aleshores s'han organitzat dos cursos de formació per aconseguir captar persones voluntàries per fer transcripcions, un l'11 de juny del 2015 i l'altre el 3 d'octubre del 2015 (en la segona sessió, a més de posar en comú dubtes sobre les transcripcions, es va fer formació a nous usuaris). A la formació s'ensenya l'aplicació (com accedir-hi i com introduir-hi dades), es proporcionen les normes de transcripció i es resolen els dubtes més freqüents. Fins ara s'hi han inscrit 30 persones voluntàries, 10 homes i 20 dones.



Imatge 3. Presentació resultats del projecte EINES a l'Arxiu Comarcal del Baix Llobregat. Abril de 2015. Foto: Marc Rius

Per tal de donar a conèixer l'activitat s'han fet diverses xerrades sobre el projecte i sobre la demografia de Sant Feliu de Llobregat, a càrrec de la demògrafa Anna Cabré i del director del CVC, Josep Lladós. Amb la finalitat d'arribar a més públic, algunes d'aquestes conferències s'han inserit dins d'una programació més àmplia com, per exemple, la conferència «La gent de Sant Feliu de Llobregat dedicada als oficis i a la indústria a través dels padrons municipals d'habitants», impartida per la professora Anna Cabré el 27 d'octubre de l'any passat en el marc de la celebració de la II Setmana de Patrimoni «El Sant Feliu de les xemeneies».



Imatge 4. Primera formació dels voluntaris a l'Arxiu Comarcal del Baix Llobregat. Juny de 2015.



Imatge 5. Cercador de Padrons del municipi de Sant Feliu de Llobregat.

El 25 de maig del 2016 es va presentar el cercador de la base de dades del projecte «Eines i procediments per a la informatització massiva dels padrons d'habitants de Sant Feliu de 1828-1955», al palau Falguera de Sant Feliu de Llobregat. En l'acte van participar diversos voluntaris de l'institut Olorda i de l'equip de voluntaris que habitualment hi col·labora. En aquell moment s'havien transcrit els padrons del 1881, el 1889 i el 1940. Amb aquest cercador els usuaris poden consultar qui estava censat a Sant Feliu en un determinat carrer o casa o amb un determinat cognom, els oficis més freqüents, la distribució de la població per gènere i edat, etc.



Imatge 6. Cercador de Padrons del municipi de Sant Feliu de Llobregat.

S'ha fet una divulgació de les activitats a través de la web de l'Arxiu Comarcal i la de l'Ajuntament de Sant Feliu de Llobregat, les xarxes social (Facebook i Twitter) i comunicats de premsa de l'oficina de premsa de la UAB i el Parc de Recerca de la UAB, del Departament de Cultura i de l'Ajuntament de Sant Feliu de Llobregat.

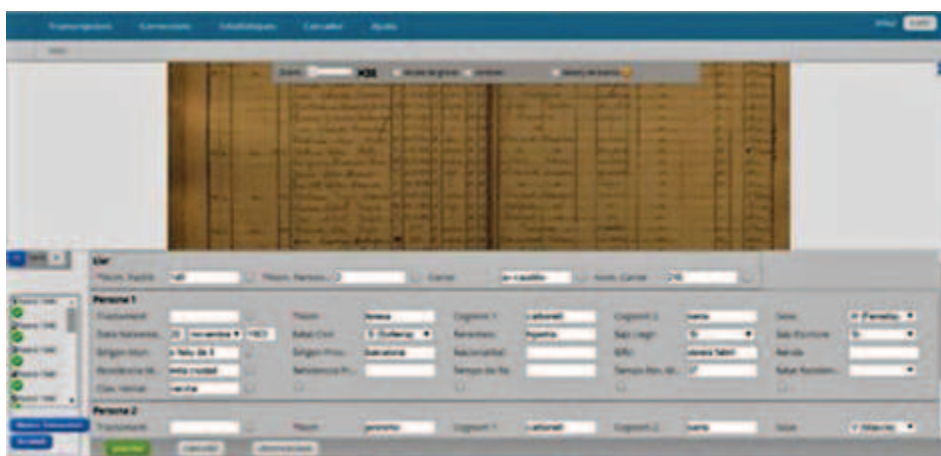
- › Accions als *mitjans de comunicació*. S'han publicat setze articles en diversos diaris i revistes, en suport paper i digital, i en diferents mitjans, tant d'abast estatal i català (*Ara, El Mundo, La Vanguardia*) com local i comarcal (*El Llobregat, El Far, Fet a Sant Feliu*), entre altres. També s'ha participat en diverses emissores d'àmbit estatal, català i local, com Ràdio Barcelona, Radio Nacional de España i Cadena Ser, així com ETV (Esplugues Televisió) i Ràdio Sant Feliu.

TV3 va fer un reportatge titulat «Projecte EINES: tecnologia punta per als antics padrons», a «Som el que fem», un espai de màxima audiència emès divendres, 14 d'octubre de 2016, després del Telenotícies del cap de setmana en el qual es van enregistrar imatges al CVC i a l'Arxiu. El reportatge va mostrar el treball d'un voluntari fent transcripcions, així com algun usuari demanant informació, alhora que va explicar la repercussió que té per als arxius l'aplicació d'eines de reconeixement de caràcters.

- › *Presentacions científiques*. En l'àmbit científic, l'equip del CVC i el CED han presentat les comunicacions següents a congressos:
 - Presentació de la comunicació «Herramientas de transcripción asistida colaborativa de censos históricos» a càrrec d'Alicia Fornés, Joana M. Pujades-Mora, Oriol Ramos i Anna Cabré al II Congreso Internacional de Humanidades Digitales Hispánicas «Innovación, globalización e impacto», celebrat a Madrid els dies 7 i 8 d'octubre del 2015.
 - Presentació de la comunicació «Construcción de bases de datos y participación ciudadana. Tres experiencias en curso» a càrrec d'Anna Cabré, Joana M. Pujades i Antonio David Cámara al XI Congreso de la Asociación de Demografía Histórica, celebrat a Cadis els dies 21 i 24 de juny del 2016.

2.3. ESTAT DE LES TRANSCRIPCIONS

S'han transcrit totalment els padrons del 1833, el 1840, el 1881 i el 1940 i, estan pràcticament acabats els del 1828 i el 1857. En total, des de l'11 de juny (data de la primera formació) fins setembre del 2016, s'ha fet un 19,7 % de les imatges transcrites. Ara (a principis de desembre de 2016) s'ha superat el 30 %. De les 30 persones inscrites, quatre s'han donat de baixa per manca de temps i cinc han estat donades de baixa per inactivitat.



Imatge 7. Pantalla de l'entrada de dades de la plataforma col·laborativa

Estat de les tasques de transcripció (16 de setembre de 2016)

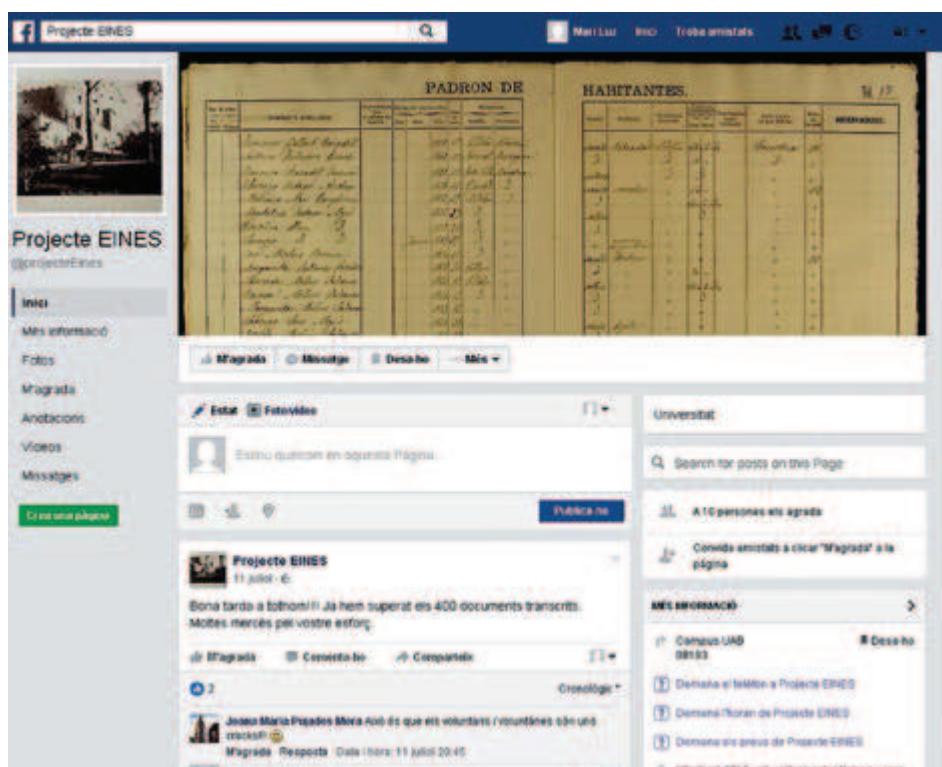
Any del padró	Nombre d'imatges	Transcrites	Percentatge de transcripció
1828	24	20	83,3%
1833-1840	140	140	100,0%
1857	44	40	90,9%
1878	108	0	0,0%
1881	97	97	100,0%
1886	100	0	0,0%
1889	102	0	0,0%
1906	115	0	0,0%
1910	114	0	0,0%
1915	130	23	17,7%
1920	133	0	0,0%
1924	172	0	0,0%
1930	201	0	0,0%

Any del padró	Nombre d'imatges	Transcrites	Percentatge de transcripció
1936*	210	0	0,0%
1940	206	206	100,0%
1945	279	0	0,0%
1950	224	0	0,0%
1955	272	0	0,0%
Total	2.671	526	19,7%

* La dada del 1936 és aproximada, ja que no hi ha el llibre del padró sinó 1.473 cèdules d'empadronament.

Valoracions dels resultats i perspectives de futur del projecte.

El treball amb el voluntariat té la vàlua especial d'implicar la comunitat local en tot el procés del projecte, però demana un seguiment profund i dinamització. El personal investigador del CVC i el CED fa el seguiment tècnic de les feines de transcripció i dóna suport per crear l'equip de participants, però cal personal que s'encarregui d'estimular la participació i de fer les revisions de les transcripcions. Des dels arxius podem acollir aquest projecte, però hem de disposar d'alguna persona que es dediqui a coordinar i dinamitzar el voluntariat.



Imatge 8. Facebook del projecte que promou la participació ciutadana a partir del model de crowdsourcing i facilita la presentació de resultats.

En el cas de Sant Feliu de Llobregat, una part molt important de la població és originària de les migracions dels anys seixanta i setanta i, per tant, pot no sentir-se vinculada amb uns avantpassats que no són els seus, al contrari que en altres municipis més petits, en els quals les migracions de la segona meitat del segle xx van ser menys fortes per això senten com a més pròpies aquestes genealogies. Això ens ha requerit i ens continua obligant a fer un esforç per establir vincles i empaties amb possibles persones de la comunitat local que puguin participar en el projecte.

El projecte Eines finalitzarà l'any 2017 i, com hem vist, ens falta el 80% de les transcripcions per acabar la feina. Per això ara estem intentant relançar el projecte amb un seguit d'actuacions encaminades a incrementar el nombre de voluntaris i, sobretot, el temps que cadascun d'ells pot dedicar a fer transcripcions. Hem d'aconseguir més implicació i continuïtat en el treball de les persones que hi participen. Ara ens estem adreçant a col·lectius específics perquè es comprometin a la transcripció de padrons concrets com, per exemple, persones jubilades o estudiants de secundària. En aquest sentit, ja a final del curs passat, vam buscar la participació d'un grup de la classe de quart de batxillerat de l'institut Olorda; tanmateix hem de tornar a aquesta iniciativa perquè l'acabament del curs va impedir que es materialitzés aquesta col·laboració. Ara s'ha previst que els responsables del CVC facin més demostracions en els centres de secundària de Sant Feliu de Llobregat. També estrenyerem els vincles amb les entitats locals perquè assoleixin més compromisos per tirar endavant Eines. En aquests moments és més fàcil aconseguir-ho perquè ja tenim un producte visible per Internet: la base de dades amb el buidatge de quatre padrons complets.

També buscarem incentivar la participació amb petits regals de recompensa: llibres d'història local i comarcal, reproduccions de fotografies antigues de Sant Feliu de Llobregat, etc. I, sobretot, hi destinarem una persona adscrita a l'Arxiu que es dedicarà a mitja jornada a impulsar el projecte fins al seu acabament.

Perquè progressin a bon ritme els projectes participatius d'aquestes característiques, cal disposar de personal que es dediqui de manera professional a incentivar la participació i a fer les revisions de les transcripcions, i que això es reculli en els pressupostos dels projectes. També seria convenient que el personal que faci les transcripcions sigui mixt, és a dir, que hi hagi personal voluntari i personal contractat. Els arxius tenen poc personal i recursos per fer-se càrrec en solitari d'aquesta dinamització i això pot fer que les tasques s'eternitzen i es trigui a tancar el projecte. També és important incloure el finançament de

petites recompenses en funció del treball aportat. I, sobretot, cal disposar de recursos econòmics perquè la recerca es pugui presentar després en suports de divulgació diferents, adreçats al públic en general, al món educatiu, a col·lectius concrets, etc.

En relació amb aquest últim aspecte, a través de l'Obra Social de La Caixa podrem començar a impulsar la incardinació d'Eines amb el món educatiu, ja que aquest projecte ha estat seleccionat per incloure'l en el programa Investiga amb RecerCaixa, un programa que proposa un projecte educatiu d'investigació per a les escoles, adreçat als nens i les nenes de cinquè i sisè de primària. Al llarg de tot el curs vinent, les escoles que ho sol·licitin podran escollir el projecte Eines i els responsables del CVC i del CED en faran l'acompanyament.

Les humanitats digitals és un camp emergent i multidisciplinari que està cridat a tenir un paper destacat en els arxius. Gaudir d'aquests recursos pot permetre fer més atractius els continguts que es custodien als arxius. Les noves tecnologies de visió per ordinador obren un ventall de grans possibilitats per a aquests serveis; però, perquè se'n puguin beneficiar els arxius, cal que vagin acompanyades dels recursos econòmics necessaris.

2.4 PERSPECTIVES DE FUTUR

Els beneficis de la captura massiva i semiautomàtica de dades en el món dels arxius són evidents i permetran una millor difusió dels continguts dels arxius, tant en el vessant de difusió cultural i social com a favor de la gestió documental. Des del punt de vista de l'anàlisi de costos, probablement amb el temps s'assolirà un escenari de reducció de costos de tractament, gestió, accés i difusió de la documentació/informació, però no podem obviar el benefici social i cultural que implica la incorporació de la participació ciutadana a través de projectes cooperatius de *crowdsourcing*.

Un altre aspecte estratègic en el desenvolupament d'aquests projectes és la promoció de les xarxes d'arxius. Aquells projectes que impulsin un treball de col·laboració entre arxius disposaran de més garanties d'èxit en el finançament i en l'impacte en la societat. Les xarxes permeten utilitzar tecnologia compartida, normes comunes i homologades, reduir els costos de les inversions i millorar la projecció i una difusió conjunta del projecte i dels arxius que hi participen. De fet, si es dissenyen projectes de gran impacte que apleguin el conjunt d'arxius

del Sistema d'Arxius de Catalunya (SAC), tant públics com privats, generaran resultats de gran ressò, tant des de la perspectiva del volum de dades com de la seva diversitat territorial. En aquest punt, seria molt recomanable que alguns dels projectes impulsats des de l'òrgan director del SAC es fessin extensibles al màxim nombre d'arxius del SAC, amb l'objectiu d'elaborar, així, projectes el més universals possible.

3. PERSPECTIVES D'INNOVACIÓ, RECERCA I DESENVOLUPAMENT DE TECNOLOGIA APLICADA EN MATÈRIA D'ARXIUS I GESTIÓ DOCUMENTAL

Els arxius estan davant un moment singular i paradigmàtic en la seva funció respecte de les administracions i els ciutadans. És una excel·lent oportunitat per consolidar-se com a centres estratègics del coneixement i de la gestió de les administracions i com a garants dels drets d'accés i impulsors de la transparència. La seva doble funció d'arxiu històric i arxiu administratiu, que actualment quasi podríem qualificar de multifunció per l'ampli abast de la seva acció, situa els arxius en un lloc estratègic en les organitzacions, però també els obliga a introduir importants canvis en la seva mentalitat, en el seu funcionament i en la coordinació amb altres àmbits de l'organització o, en alguns casos, de fora de la pròpia organització.

Aquesta visió dels arxius del futur és, alhora, una necessitat competitiva que requereix un replantejament de les eines i els procediments utilitzats tradicionalment. Per assolir aquest objectiu, els arxius han de tenir no només el suport de l'Administració, sinó el del conjunt d'agents implicats en el desenvolupament tecnològic, que són els que poden aportar un valor afegit als continguts dels arxius. En aquest sentit, la col·laboració amb universitats, empreses i administracions és estratègica i imprescindible. Sense aquesta retroalimentació, els arxius no podran desenvolupar productes ni serveis competitius.



Imatge 9. Arxiu Comarcal del Baix Empordà

Per tant, hi ha tres factors clau per a l'èxit en matèria d'arxius i gestió documental:

- › Les aliances estratègiques amb institucions que impulsen la innovació i la recerca vinculades a universitats i sector productiu (públic i privat), perquè a partir d'aquests projectes es generin productes o serveis de mercat útils per al conjunt de la societat i les administracions.
- › L'impuls de les administracions, especialment de l'òrgan director del SAC, tant per potenciar la innovació i la recerca en l'àmbit dels arxius i la gestió documental com per buscar aliances amb altres sectors que aportin valor afegit. I, també, promoure el suport al conjunt d'arxius del SAC en la seva aplicació.
- › El convenciment dels professionals, els arxivers i els gestors de documents perquè introdueixin projectes d'aquest tipus per millorar el servei que ofereixen als ciutadans i les entitats.

3.1 LÍNIES DE RECERCA, INNOVACIÓ I INVESTIGACIÓ: PROJECTES RELACIONATS AMB ARXIUS I GESTIÓ DOCUMENTAL

Amb la voluntat de treballar conjuntament amb institucions capdavanteres en projectes d'R+D+I, des de la Subdirecció General d'Arxius i Museus s'ha elaborat un primer esborrany de línies de treball. L'objectiu és iniciar un pla de col·laboració que incorpori de manera progressiva noves eines i procediments en l'explotació i la difusió dels documents dels arxius. En el marc de la col·laboració amb el Parc de Recerca de la UAB¹⁰ i, també, en el marc de la col·laboració consolidada amb el CVC, s'han dissenyat algunes línies de treball:

- › **Extracció massiva, automàtica o semiautomàtica de dades dels documents.** Els arxius disposen d'innombrables sèries documentals de temàtica molt diversa relacionada amb la producció de les administracions, associacions, empreses, persones, etc. En conjunt, un gran potencial de matèria primera (dades, informació, documents, digitalitzacions, descripcions, etc.) per ser explotada amb sistemes automatitzats o semiautomatitzats de captura de dades. Tot i que hi ha un gran camí a recórrer, els darrers projectes demostren que progressivament les automatitzacions garanteixen un percentatge més alt de fiabilitat dels resultats i, en conseqüència, una reducció del temps de dedicació dels humans a la seva validació. Aquests projectes, a més, situen la participació ciutadana i associativa com un element essencial, perquè aporten un valor social i d'identificació indiscutible. Tanmateix, el seu valor també pot ser el seu punt crític. És a dir, l'automatització que garanteixi un 100% de fiabilitat dels processos d'extracció massiva de dades només s'obté en aquells documents mecanoscrits amb perfectes condicions de reconeixement. En el cas dels documents manuscrits o mecanoscrits amb complexitat de digitalització (problemes de tinta, qualitat del paper, enquadernació, etc.), és imprescindible la figura del «validador». Normalment, aquest tipus de projectes disposen d'una campanya de *crowd-sourcing* que busca, tot fent participar els ciutadans, les associacions, els centres d'estudis, etc., implicar el màxim de persones en la revisió dels treballs automatitzats per a la seva definitiva validació. Per tant, ens trobem davant d'una proposta de participació ciutadana en els projectes d'arxiu que en ocasions esdevé la clau de l'èxit del projecte. Actualment, sense un percentatge important de participació, aquests projectes tenen el risc de no finalitzar-los o prolongar-se excessivament en el temps.

Aquesta línia incorporaria algunes propostes com les següents:

- Captura automatitzada de dades i informació de documents mecanoscrits dels segles XIX i XX. La facilitat de captura permet un percentatge alt de reutilització directa de les dades.
- Captura de documentació manuscrita en diferents llengües i tipologies d'escriptura. Els manuscrits medievals, amb complexitat de transcripció i comprensió per als usuaris, presents en la majoria dels arxius, serien un exemple clar del potencial d'una captura per socialitzar-ne i difondre'n el contingut. També es pot aplicar en documentació contemporània com els padrons municipals, el cadastre (inclouria plànols), la correspondència, etc. En el marc de la Xarxa d'Arxius Comarcals estem explorant la viabilitat del projecte d'extracció i explotació de les dades dels índexs notariais de tot Catalunya.
- Captura d'imatges a partir, especialment, de cartells i fotografies, com elements visuals. Però també amb la lectura dels textos inclosos en els documents icònics (cartells, indicadors...) i en els documents fotogràfics (cartells de botigues, noms de pobles, carreteres, etc., presents en les imatges).
- Captura de documents icònics amb l'objectiu de millorar la indexació i la descripció. En són exemples casos de textos o imatges procedents de la publicitat (premsa), goigs, còmics, cartells, etc.
- Captura de documents audiovisuals a partir del reconeixement de la veu i l'extracció de les dades per facilitar la comprensió del document i la seva difusió.
- Captura de documentació amb una estructura complexa com poden ser partitures musicals. La catalogació i els estudis de digitalització dels arxius amb continguts musicals permeten facilitar el tractament especial que aquests fons documentals requereixen.

- › **Extracció massiva, automàtica o semiautomàtica de dades dels expedients i els documents electrònics.** En relació amb la documentació administrativa (en suport paper i electrònic), la captura permet indexar i classificar automàticament expedients administratius contemporanis que representen un volum molt considerable de la documentació custodiada per l'arxiu i que poden tenir una alta demanda. L'estructura homogènia dels documents que formen els expedients permeten recuperar la informació a partir d'una anàlisi d'aquestes estructures. D'aquesta manera,

l'extracció massiva de dades a partir de l'estructura del document permet captures eficients i aprofitar dades del contingut i del context per descriure'ls o per generar les pròpies metadades. En relació amb la documentació electrònica, les possibilitats són extremadament àmplies a partir del que representa la captura i el tractament massiu de dades o metadades. El fet que els documents disposin d'unes estructures idèntiques en una organització és un gran avantatge que garanteix una reutilització massiva de les dades.

- › **Extracció massiva de dades sobre continguts de premsa local i comarcal.** Alguns arxius disposen de fons hemerogràfics que per a la consulta són un complement perfecte dels fons documentals. Un exemple són les més de 950.000 pàgines digitalitzades de XAC_Prensa de la Xarxa d'Arxius Comarcals i de molts arxius municipals. Per tant, si la potencialitat de la documentació hemerogràfica és molt alta i la seva captura massiva aporta resultats de qualitat elevats, és un camp clarament òptim per dedicar esforços en el desenvolupament de productes i serveis.
- › **Conversió d'imatges històriques en 2D a 3D** amb obtenció d'aparells estereoscòpics per ser visionats en anàglif. Aquesta conversió facilita una millor comprensió de la imatge i el seu entorn i pot ser especialment indicada en el món educatiu.
- › **Big data / open data.** Amb l'entrada en vigor de la Llei 19/2014, de transparència, accés a la informació i bon govern, els arxius esdevenen uns centres amb un enorme potencial de dades. De fet, les dades ja són als arxius des de fa temps, però la novetat és la necessitat incipient de difondre aquestes dades de manera massiva, comprensible i reutilitzable per les persones i les entitats. Aquesta transformació o adequació de les dades per fer-les obertes i reutilitzables és un dels reptes dels arxius moderns. Només fixant la mirada en l'article 35.a de la Llei 10/2001, d'arxius i documents, observem l'obligació dels arxius de fer públics i fàcilment accessibles instruments que permeten als ciutadans i els investigadors localitzar la documentació de cada administració. Per tant, un dels camps de màxima expansió en relació amb les dades obertes és la reutilització dels instruments de descripció dels arxius (guies, inventaris o catàlegs). En el cas de l'òrgan director del Sistema d'Arxius de Catalunya, també s'ha d'afegir l'accés i la reutilització de les dades del Sistema (registre, cens, directoris, consultes i usuaris, tipus d'activitats desenvolupades,

etc.). Finalment, un aspecte essencial de la introducció de la funcionalitat de descàrrega de dades en els programaris és l'aplicació de diferents formats que en garanteixin la reutilització (XML, Excel, CSV, JSON, etc.).

- › **Realitat augmentada.** Si entenem la realitat augmentada com la visió directa o indirecta d'un entorn real, en el qual s'afegeix informació virtual a la informació física ja existent, els arxius són centres de gran potencial. Amb l'ajuda de la tecnologia (visió per ordinador i reconeixement d'objectes), la informació del món real que envolta l'usuari es converteix en interactiva i, fins i tot, permet sumar continguts a la pròpia informació real. Un exemple que des del Departament de Cultura s'està explorant és el projecte [Google Arts & Culture](#), que proposa un accés virtual a l'edifici de l'arxiu i als seus continguts, els documents digitalitzats, les guies del centre, els catàlegs en línia, les exposicions virtuals, els materials didàctics i la pròpia singularitat de l'edifici en molts casos. Actualment s'està explorant la possibilitat d'endegar un pla pilot amb tres arxius. Més enllà d'aquest projecte, el potencial dels documents en el món de la realitat augmentada és molt alt i dibuixa un escenari de creixement important a partir de l'elaboració de productes per als arxius i els ciutadans.

- › **Modelització en 3D.** A partir dels documents es poden modelitzar en 3D representacions desconegudes. Un dels factors d'èxit de la reconstrucció del Born va ser aprofitar el contingut dels protocols notariaus que identifiquen de manera precisa la forma de vida de les persones en un determinat període de temps. A més, la impressió en 3D de documents a partir d'objectes en 2D pot ser un sector en creixement, sobretot perquè la impressió en 3D comença a ser considerada d'ús habitual en molts àmbits. En aquesta línia, els documents ofereixen diferents possibilitats que es poden explorar a partir de la transformació en 3D de fotografies històriques, plànols d'arquitectura, mapes, etc.

- › **Indústria audiovisual i sistemes d'explotació.** Només la Xarxa d'Arxius Comarcals té dipositades i gestiona més de 7 milions de fotografies, 45.000 registres sonors i 16.000 vídeos (molts de televisions locals i documentals locals i comarcals). El potencial d'informació i dades dels arxius en conjunt és molt elevat, però requereix un tractament específic.

- › **Metodologia i tractament de dades.** Els documents dipositats en els arxius tenen un gran potencial i són la matèria primera de l'elaboració dels estudis, per això la utilització d'una metodologia adequada en permetria una millor explotació. En aquesta línia, les experiències dels projectes Esposalles o Eines del padrons de Sant Feliu, impulsat pel CED, ho demostren.
- › **Reconstrucció del passat o ludificació.** Aquest és un dels camps amb més potencial atesa la singularitat dels documents que permeten, a partir del seu contingut, reconstruir el passat i, en el cas de la societat actual, promoure els principis democràtics i de transparència. Tanmateix, aquesta potencialitat s'ha de veure consolidada amb primeres experiències en el marc dels arxius catalans per valorar el possible impacte real d'aquesta tecnologia en els arxius.
- › **Ontologies i cerques intel·ligents d'informació.** El gran volum de dades contingudes en els documents, els fons i els arxius poden esdevenir un veritable mur en la cerca de la informació. El desenvolupament de llenguatges basats en ontologies poden ser la clau per ajustar la resposta als usuaris i facilitar un retorn qualitatiu de la informació desitjada.



Imatge 10. Fotografia 3D a la sala de consulta de l'Arxiu Comarcal de l'Alt Urgell. Al fons, ontologia sobre la Norma de Descripció Arxivística de Catalunya

Probablement, serà necessari superar la tradició dels *thesaurus*, siguin sobre el conjunt del patrimoni cultural o específics, i apostar per les ontologies i la web semàntica per tal de ser més eficaços en aquestes cerques i per aconseguir el retorn de la informació requerida.

- › **Aplicacions per a dispositius mòbils.** Els arxius gestionen una gran quantitat d'informació que és un actiu per a qualsevol desenvolupament que s'integri en dispositius mòbils. A més, els arxius, en qualsevol dels seus àmbits d'acció, tenen una interacció directa amb els ciutadans, les entitats i les administracions, aspecte pel qual, com s'ha exposat anteriorment, s'han d'utilitzar els canals adequats de comunicació. La recent transformació del portal de la Xarxa d'Arxius Comarcals de Catalunya al «model responsiu» respon a aquesta necessitat d'estar alineat amb les necessitats dels ciutadans; l'objectiu és posar al seu abast els continguts a través dels diferents canals i dispositius mòbils.

- › **Sistemes d'informació geogràfica i tractament de la informació georeferenciada.** La diversitat d'ubicació dels fons i els documents dels arxius del Sistema requereix una plataforma senzilla per als usuaris que permeti una identificació georeferenciada. La localització en una mapa dels continguts documentals de Catalunya i l'aportació de les característiques d'un model georeferencial serien una línia de desenvolupament important. Les possibilitats en aquest apartat no estan relacionades només amb una tipologia documental, sinó que haurien de permetre interconnectar documents a partir d'aquesta informació geogràfica, amb la finalitat de sumar continguts o dades de diferents documents, estiguin o no en el mateix arxiu físic. A més, permetria incorporar continguts documentals a altres àmbits del patrimoni i aportar una informació de context i contingut, per exemple, a un castell, una fàbrica del segle XIX o qualsevol altre element patrimonial. Aquesta figura milloraria propostes actualment en funcionament que aporten informació parcial d'un element patrimonial i que, amb el treball a partir de la georeferenciació, esdevindrien molt més útils i atractives per al ciutadà. També, en l'àmbit estrictament documental, aquest tipus d'informació geogràfica té un recorregut ampli, per geolocalitzar tant els arxius com un tipus de fons o documents.

- › **Didàctica de les ciències socials i de l'estudi del medi.** L'aplicació d'aquestes tecnologies pot permetre potenciar línies de recerca en matèria de didàctica de la història, les ciències naturals, l'evolució tecnològica,

l'educació per a la ciutadania, entre altres temes, en el marc de les ciències de l'aprenentatge.

Finalment, cal valorar l'experiència Library Living Lab per traslladar-la al món dels arxius a través d'un **Archives & Record Management Living Lab**, és a dir, una espècie de laboratori tecnològic, social, cultural i administratiu en matèria d'arxius i gestió documental. Disposar d'un espai similar permetria:

- › Aplicar una anàlisi de les necessitats de la societat en matèria d'arxius i gestió documental en la presa de decisions de les administracions.
- › Analitzar les tendències de consum i ús, així com les expectatives dels ciutadans, les entitats i les administracions en matèria d'arxius i gestió documental, però també en el marc de l'administració electrònica, la preservació digital, l'aplicació de polítiques, les accions de transparència, accés a la informació i bon govern.
- › Garantir una participació ciutadana en els productes i els serveis que ofereixen els arxius a partir d'una interrelació amb la societat en un espai físic i un espai virtual.
- › Explorar els nous productes i serveis, i fins i tot els procediments, que a partir de les anàlisis preliminars es considerin potencialment útils.

Aquestes propostes generen unes línies d'actuació que s'han de contrastar i valorar amb els diferents sectors implicats, però que apriorísticament semblen que tenen un potencial en el seu desenvolupament en el món dels arxius i la gestió documental. Són línies que les administracions i els arxius han d'explorar i analitzar en el marc dels serveis i els productes que ofereixen en el context de la societat de la informació i l'administració electrònica.

4. CONCLUSIONS

En el nou escenari tecnològic i d'impacte d'Internet, els arxius tenen un gran repte i, alhora, una oportunitat per posicionar-se millor en la societat en general i en les administracions en particular. Aquesta situació hauria de provocar una metamorfosi en la manera de funcionar dels arxius per poder sobreviure i estar

a l'alçada de les noves demandes. L'arxiu del futur, i ja del present, ha de ser un centre especialitzat en la protecció, la preservació, la gestió, l'accés i la difusió del patrimoni cultural, així com un referent en la gestió documental de l'organització, a través de l'aplicació de les polítiques que garanteixin la implantació de sistemes de gestió documental, l'administració electrònica i la transparència i el bon govern de l'Administració.

Les tecnologies de la visió per ordinador poden contribuir a posar a l'abast de la societat una gran quantitat de documents de diferent tipologia —textuals (mecanografiat i manuscrits), cartogràfics, icònic, tecnogràfics, etc.— que poden transformar radicalment l'accés que fins ara tenien els usuaris als arxius. Per mitjà d'aquestes tecnologies ja s'ha aconseguit la indexació amb garanties d'èxit d'una part important dels documents dels arxius (els escrits en lletres impreses o mecanografiades): per a les lletres manuscrites, el procés no és automàtic, cal combinar la transcripció manual i el reconeixement amb l'ajut de màquines. Això demana fer partícip la comunitat en les tasques de transcripció a través de plataformes col·laboratives. L'execució d'aquests projectes exigeix un important abocament de recursos als arxius (per a la digitalització de documents, per a la contractació de personal per dinamitzar la participació col·lectiva, per fer transcripcions, per adaptar els programaris dels arxius a la lectura automàtica de documents i per a l'elaboració dels productes que es poden fer; sense recursos difícilment es poden tirar endavant els projectes).

La demanda de recursos ha d'anar vinculada necessàriament a més vertebració dels arxius. Cal actuar a partir dels diferents vincles entre els arxius: orgànics, territorials o documentals per aconseguir el treball en xarxa. Això pot garantir més consistència en el missatge i en l'oferta de serveis adreçada a la societat, les administracions o el sector empresarial.

En els últims anys, des de la Direcció General d'Arxius, Biblioteques, Museus i Patrimoni s'ha col·laborat amb les universitats en programes d'R+D+I, amb el Centre de Visió per Computador, amb l'Escola Superior d'Arxivística i Gestió de Documents i, més recentment, amb el Parc de Recerca de la UAB. Pel que fa a aquesta última institució, hem d'esmentar la propera participació en el concurs d'idees de negoci entorn de les indústries culturals i les indústries basades en l'experiència, adreçat a investigadors de la UAB i a centres del campus. Finalment, cal destacar el projecte Eines, ja explicat, que exemplifica un cas real de captura massiva de dades d'una determinada sèrie documental, amb resultats tangibles per a la lectura d'imatges semiautomàtica i per als estudis demogràfics.



Imatge 11. Parc de Recerca de la Universitat Autònoma de Barcelona

Els arxius han d'assumir aquests reptes amb fórmules de col·laboració imaginatives a l'hora que sòlides (entre els mateixos arxius, amb les institucions de recerca, amb la participació de la comunitat local, amb fórmules mixtes de personal voluntari i contractat); si no, perdran l'oportunitat per compartir els lideratges de la transformació i per aconseguir els recursos que finalment ho facin possible.

I en aquesta línia de col·laboració i de generació de sinèrgies dirigides a la societat digital i la eficiència dels recursos, hem de valorar positivament la proposta del recent Pacte Nacional per a la Societat Digital, liderat pel Govern, juntament amb el món local, que promou la construcció d'una societat digital més enllà de la simple suma aritmètica de les diverses iniciatives que es duen a terme per part de les diverses administracions. Aquest és un marc excel·lent en el qual, el Sistema d'Arxius de Catalunya (SAC) pot aportar la seva experiència i coneixement en relació al món digital. Tant des de l'òrgan director del SAC com des de cada arxiu, no es pot obviar un pacte que té, entre els seus principis fundacionals la voluntat de treballar conjuntament i de forma consensuada amb tots aquests agents.

Agents, els arxius, que impulsen l'administració electrònica, i que també són estratègics en el món digital en tant en quan desenvolupen programes de digi-

talització, gestionen coordinadament amb els agents responsables les infraestructures tecnològiques, faciliten l'accés digital als milions de document i bilions de dades i promouen la seva difusió mitjançant eines digitals.

Esperem, en aquest sentit, que aquest pacte sigui prou coherent i incorpori en els respectius grups de treball aquells agents que estant treballant amb documents, ja siguin nascuts digitals o digitalitzats, però que són un actiu necessari per qualsevol política relacionada amb la societat digital del present i del futur.

NOTES

1. Sanmartí, Neus; Márquez Bargalló, Conxita. «Enseñar a plantear preguntas investigables». *Alambique. Didáctica de las ciencias experimentales*, núm. 70, gener, febrer i març de 2012.
2. Llei 39/2015, d'1 d'octubre, del procediment administratiu comú de les Administracions públiques, i Llei 40/2015, d'1 d'octubre, de règim jurídic del sector públic.
3. Cruz Espinosa, Adrià. «Els arxius a Twitter: 10 piuladors que cal seguir». Arxipèlag (recuperat el 20 de setembre del 2016 a: <http://www.nuvol.com/opinio/els-arxius-a-twitter-10-piuladors-que-cal-seguir>).
4. Villavicencio, Francisco; Jordà, Joan Pau; Pujadas-Mora, Joana M. «Reconstructing Lifespans Through Historical Marriage Records of Barcelona from the Sixteenth and Seventeenth Centuries». Bloothoof, Gerrit; Christen, Peter; Mandemakers, Kees; Schraagen, Marijn (ed.). *Population Reconstruction*. Suïssa: Springer International Publishing, 23 de juliol de 2015, p. 199-216 (ISBN: 978-3-319-19883-5).
5. Lladós i Canet, Josep (CVC), Cabrer i Pla, Anna (CED). «EINES: Eines i procediments per a la informatització massiva de les fonts històriques de població. Proposta de projecte per a la convocatòria d'ajuts a la recerca RecerCaixa 2014», 2014, p. 5.
6. Lladós i Canet, Josep (CVC), Cabrer i Pla, Anna (CED). «EINES...», p. 6.
7. Darrerament diversos equips europeus han estat col·laborant en la implementació de projectes de transcripció semiautomàtica de lletres manuscrites a través de projectes concrets, per exemple l'equip noruec que ha treballat amb els fulls del cens de Noruega i el català que ho ha estat fent amb els llibres d'esposalles de l'Arxiu de la Catedral de Barcelona. Vegeu: Thorvaldsen, Gunnar; Pujadas-Mora, Joana M.; Andersen, Trygve; Eikvil, Line; Lladós, Josep; Fornés, Alicia; Cabré, Anna. «A Tale of Two Transcriptions: Machine-Assisted Transcription of Historical Source», *Historical Life Course Studies*, 2015, vol. 2, p. 1-19.
8. Lladós i Canet, Josep (CVC); Cabrer i Pla, Anna (CED). «EINES...», p. 7.
9. Sobre experiències similars dutes a terme per aquest equip, vegeu: Fornés, Alicia; Lladós, Josep; Mas, Joan; Pujadas-Mora, Joana M.; Cabré, Anna. «A Bimodal Crowdsourcing Platform for Demographic Historical Manuscripts». *DATeCH (Digital Access to Textual Cultural Heritage)*, 2014, Madrid, 19-20 de maig de 2014, p. 103-108.
10. El Parc de Recerca de la UAB té la missió de facilitar i promoure la transferència de coneixement, la innovació i les capacitats d'R+D+I dels grups de recerca per atendre les necessitats del sector social i econòmic.

RESUM

Aquest article analitza les possibilitats que ofereixen les tecnologies de visió per ordinador i reconeixement massiu de caràcters aplicades a la documentació. Des de la perspectiva del potencial dels arxius, la seva accessibilitat i difusió per Internet, es proposen línies de desenvolupament en relació amb els nous canals i registres de comunicació dels arxius. S'incideix especialment en l'experiència pionera que s'ha desenvolupat amb els padrons d'habitants de l'Ajuntament de Sant Feliu de Llobregat dipositats a l'Arxiu Comarcal del Baix Llobregat, tot remarcant la relació entre tecnologia, arxius i participació ciutadana. Aquest projecte s'està desenvolupant a partir de l'aliança entre el Centre de Visió per Computador, el Centre d'Estudis Demogràfics, l'Arxiu Comarcal del Baix Llobregat i l'Ajuntament de Sant Feliu de Llobregat. Finalment, s'exploren les perspectives d'innovació, recerca i desenvolupament de tecnologia aplicada en matèria d'arxius i gestió documental en relació amb els arxius catalans, arxius amb un gran potencial com demostren les línies de projectes de la Xarxa d'Arxius Comarcals, fàcilment extrapolables al conjunt d'arxius del Sistema d'Arxius de Catalunya.

RESUMEN

Este artículo analiza las posibilidades que ofrecen las tecnologías de visión por ordenador y reconocimiento masivo de caracteres aplicadas a la documentación. Desde la perspectiva del potencial de los archivos, su accesibilidad y difusión por internet, se proponen líneas de desarrollo en relación con los nuevos canales y registros de comunicación de los archivos. Se incide especialmente en la experiencia pionera que se ha desarrollado con los padrones de habitantes del Ayuntamiento de Sant Feliu de Llobregat depositados en el Archivo Comarcal de El Baix Llobregat, remarcando la relación entre tecnología, archivos y participación ciudadana. Este proyecto se está desarrollando a partir de la alianza entre el Centro de Visión por Computador, el Centro de Estudios Demográficos, el Archivo Comarcal de El Baix Llobregat y el Ayuntamiento de Sant Feliu de Llobregat. Finalmente, se exploran las perspectivas de innovación, investigación y desarrollo de tecnología aplicada en materia de archivos y gestión documental en relación con los archivos catalanes, que tienen un gran potencial como demuestran las líneas de proyectos de la Red de Archivos Comarcals, fácilmente extrapolables al conjunto de archivos del Sistema de Archivos de Cataluña.

RÉSUMÉ

Cet article analyse les possibilités offertes par les technologies de vision par ordinateur et de reconnaissance massive de caractères appliquées aux documents. Des lignes de développement sont proposées à partir du potentiel des archives, de leur accessibilité et de leur diffusion sur Internet par rapport aux nouveaux canaux et registres de communication des archives. L'accent est notamment mis sur l'expérience pionnière développée avec les recensements d'habitants de la mairie de Sant Feliu de Llobregat déposés aux archives de la région d'El Baix Llobregat, tout en insistant sur la relation entre technologie, archives et participation citoyenne. Ce projet a été développé dans le cadre du partenariat entre le Centre de vision par ordinateur, le Centre d'études démographiques, les archives de la région d'El Baix Llobregat et la mairie de Sant Feliu de Llobregat. Les perspectives en termes d'innovation, de recherche et développement de la technologie appliquée en matière d'archives et de gestion documentaire en lien avec les archives catalanes, des archives à fort potentiel, sont explorées, ainsi que le montrent les lignes des projets du Réseau des archives régionales, facilement extrapolables à l'ensemble des archives du système des archives de la Catalogne.

ABSTRACT

This article analyses the possibilities of computer vision and massive character recognition technologies applied to documentation. Areas of development are proposed from the point of view of archives' potential, accessibility and dissemination via the Internet, in relation to the new channels and communication registers of archives. Particular attention is paid to the pioneering experience developed with the population registers of Sant Feliu de Llobregat Town Council deposited in the Regional Archive of El Baix Llobregat, with emphasis on the relationship between technology, archives and public participation. This project is being developed under an alliance between the Computer Vision Center, the Centre for Demographic Studies, the Regional Archive of El Baix Llobregat and Sant Feliu de Llobregat Town Council. Finally, the prospects for innovation, research and development of technology applied to archives and records management are examined in regard to Catalan archives, archives with great potential, as shown by the Regional Archive Network projects, which are easily extrapolated to all the archives in the Catalan Archives System.