

Geolocalització i dades obertes.

**Un projecte per a la
visibilització de la dona
com a productora de fons
documentals d'arxiu**

Paula Gomila Moll

2019



Aquest llibre està basat en el Treball Final de Màster que l'autora va elaborar i presentar al curs 2018-2019 al Màster en Arxivística i Gestió de Documents de l'Escola Superior d'Arxivística i Gestió de Documents (ESAGED) de la Universitat Autònoma de Barcelona.

Direcció del treball: Dra. Remei Perpinyà Morera i Dr. Remo Lucio Suppi Boldrito.

El treball va ser guardonat amb el Premi al Millor Treball Final de Màster del curs 2018-2019 que atorguen l'ESAGED i l'AAC-GD.

ISBN: 978-84-09-19731-6

Data: Abril 2020

Com citar aquest article:

Gomila Moll, Paula (2020); *Geolocalització i dades obertes. Un projecte per a la visibilització de la dona com a productora de fons documentals d'arxiu*. Bellaterra (Cerdanyola del Vallès): Escola Superior d'Arxivística i Gestió de Documents i Associació d'Arxivers i Gestors de Documents de Catalunya. ISBN: 978-84-09-19731-6. [Http://...](#) (consultat el ...)



Aquesta obra està subjecta a llicència Creative Commons Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya (<http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>). Es permet la reproducció total o parcial i la comunicació pública de l'obra, sempre que no sigui amb finalitats comercials, i sempre que es reconegui l'autoria de l'obra original. No es permet la creació d'obres derivades.

Títol: Geolocalització i dades obertes. Un projecte per a la visibilització de la dona com a productora de fons documentals d'arxiu

3



Resum

El present projecte explora i analitza oportunitats de difusió dels instruments de descripció dels fons i documents dels arxius públics catalans, concretament de l'Arxiu Nacional de Catalunya i de la Xarxa d'Arxius Comarcals.

S'elabora un model de dades on es defineixen els elements de descripció i les metadades mínimes a incloure en un projecte de publicació i obertura dels instruments de descripció amb perspectiva de gènere de fons familiars i patrimonials, fons personals i col·leccions. S'estudien les possibilitats de visibilització i difusió que ofereix la combinació de les dades obertes amb recursos de geolocalització i visualització de dades.

Finalment, es resumeixen les principals conclusions a les quals s'ha arribat, quant a la representació de la dona en arxius públics, a través de la publicació i anàlisi de les dades contingudes als instruments de descripció.

Paraules clau:

[dades obertes, geolocalització, perspectiva de gènere, dona, instruments de descripció, metadades, model de dades, reutilització, fons personals, CKAN]

Título: Geolocalización y datos abiertos. Un proyecto para la visibilización de la mujer como productora de fondos documentales de archivo

4



Resumen

Palabras clave:

El presente proyecto explora y analiza oportunidades de difusión de los instrumentos de descripción de los fondos y documentos de los archivos públicos catalanes, concretamente del Archivo Nacional de Cataluña y la Red de Archivos Comarcales.

[datos abiertos, geolocalización, perspectiva de género, mujer, instrumentos de descripción, metadatos, modelo de datos, reutilización, fondos personales, CKAN]

Se elabora un modelo de datos donde se definen los elementos de descripción y los metadatos mínimos a incluir en un proyecto de publicación y apertura de los instrumentos de descripción con perspectiva de género de fondos familiares y patrimoniales, fondos personales y colecciones. Se estudian las posibilidades de visibilización y difusión que ofrece la combinación de los datos abiertos con recursos de geolocalización y visualización de datos.

Finalmente, se resumen las principales conclusiones a las que se ha llegado en cuanto a la representación de la mujer en archivos públicos, a través de la publicación y el análisis de los datos contenidos en los instrumentos de descripción.

Title: Geolocation and Open Data. A project for the visibility of the woman as a creator of archival fonds

5



Abstract

Keywords:

[open data, geolocation, gender perspective, woman, finding aid, metadata, data model, reuse, personal fonds, CKAN]

The present project explores and analyses dissemination opportunities for the finding aids of fonds and documents of catalan public archives, specifically The National Archive of Catalonia and the Network of County Archives.

A data model is made and in it are defined the description elements and the minimum metadata to be included in an open data publication project of the finding aids with gender perspective for family fonds, personal fonds and collections. It also studies the possibilities of visibility and dissemination that offers the combination of open data with geolocation and data visualizations resources.

Finally, the main conclusions reached are summarized, regarding the representation of women in public archives, through the publications and analysis of the finding aid data.

Índex de continguts

Agraïments.....	8	3. Estudi de cas: la dona com a productora de fons documentals.....	30
1. Introducció	9	3. 1. El perquè de la dona productora	30
1. 1. Context i justificació	9	3. 1. 1. Els orígens dels arxius de dones	30
1. 2. Hipòtesi de treball i objectius	10	3. 1. 2. Els arxius de dones en l'era de la post custòdia arxivística	31
1. 3. Metodologia i estructura del treball	11	3. 1. 3. Tendències i projectes actuals.....	34
2. Marc teòric.....	12	3. 2. Abast del projecte	37
2. 1. Dades obertes.....	12	3. 2. 1. Els arxius	38
2. 1. 1. El concepte. Característiques i principis	12	3. 2. 2. Els fons	40
2. 1. 2. Per què és important obrir dades del sector públic.....	14	3. 2. 3. Els usuaris potencials.....	45
2. 1. 3. Legislació relativa a dades obertes i reutilització	20	4. Disseny del <i>dataset</i>	47
2. 1. 4. Arxius que ofereixen dades obertes	22	4. 1. Recopilació de les dades	48
2. 1. 5. Iniciatives d'obertura de dades amb perspectiva de gènere	23	4. 1. 1. Sol·licitud de les dades en formats reutilitzables	50
2. 2. Geolocalització	24	4. 2. Tractament de les dades	51
2. 2. 1. El concepte.....	24	4. 2. 1. Anàlisi de l'estat inicial de les dades.....	51
2. 2. 2. Arxius i projectes de geolocalització	24	4. 2. 2. Primer tractament.....	51
2. 2. 3. Com aporta valor afegit la georeferenciació al projecte.....	28	4. 3. Selecció dels elements de descripció	57
		4. 3. 1. Segon tractament.....	66
		4. 4. Els estàndards de descripció.....	67
		4. 5. Estructura del model de dades.....	68

9. Índex de taules i figures110

4. 5. 1. Selecció del model de dades	68
4. 5. 2. Contingut de les taules	69
4. 5. 3. Les relacions entre les taules	70
4. 6. Actualització de la base de dades.....	73
5. Publicació de les dades.....	74
5. 1. Definició dels termes de publicació de les dades	74
5. 1. 1. Els formats	74
5. 1. 2. Llicències d'ús	74
5. 1. 3. Plataforma de publicació.....	76
5. 2. Prova de concepte	76
5. 2. 1. El procés de publicació a CKAN i la creació de visualitzacions.....	76
6. Conclusions	92
7. Bibliografia.....	95
8. Annexes	98
8. 1. Annex 1. Quantitat d'elements NODAC informats als fons de dones	98
8. 2. Annex 2. Distribució dels elements descriptius al model de dades.....	99
8. 3. Annex 3. Esquema complet de les relacions definides al model de dades.....	104
8. 4. Annex 4. Vista prèvia de les graelles i metadades de cada <i>dataset</i>	105

Agraïments

Voldria donar les gràcies a tothom que, d'una manera o altra, ha contribuït a fer que aquest treball sigui una realitat:

Als directors, la Maria Remei Perpinyà i el Remo Suppi, pel seu acompanyament i suport, especialment en el plantejament inicial del projecte.

A l'Àngels Bernal Cercós, de l'Arxiu Nacional de Catalunya, i a l'Enric Cobo i Barri i l'Aniol Maria Vallès, de la Subdirecció General d'Arxius i Museus, els quals han mostrat un gran interès i han contribuït amb les exportacions de dades descriptives dels fons i les unitats documentals, sense les quals aquest projecte no s'hauria pogut materialitzar.

A la Glòria Gimeno Vázquez, per la seva comprensió i per animar-me a seguir endavant, amb valoracions i consells, quan més ho he necessitat.

A la Ingrid Gómez Rodríguez i al Francesc Olivé i Ollé, des de l'admiració i l'estima, per la confiança dipositada en mi i per donar-me a conèixer la professió, ara ja fa uns anys.

Al professorat del Màster en Arxivística i Gestió de Documents, pels coneixements aportats al llarg dels darrers dos anys. Així com a les companyes Anna Sancho, Klara

Gorosabel i Paula Pérez, que m'han animat en tot moment en l'elaboració d'aquest projecte.

Finalment, donar les gràcies a la meva família, pel seu suport incondicional sense el qual avui no seria qui soc.

1. Introducció

1. 1. Context i justificació

Reutilització, dades obertes, georeferenciació, visualització i difusió de dades, governs oberts, participació ciutadana i accés a la informació del sector públic, són conceptes que han rebut un impuls exponencial en les darreres dècades i actualment són clau en entorns de gestió de dades i informació. En aquest treball, dits conceptes es relacionen directament amb la perspectiva de gènere i la igualtat, així com amb la inclusió i la millora de la representació de la dona als arxius públics catalans.

Aquest projecte té dos punts de partida. D'una banda, parteix de l'evidència i constatació del fet que ens trobem en un context tecnològic que ofereix importants oportunitats de visibilització dels arxius, del contingut dels seus fons i de la feina que es porta a terme. Els professionals de l'arxivística no ens podem permetre restar al marge d'aquestes tendències si no volem quedar-nos enrere pel que fa a les noves formes de gestió i usos de la informació. Hem de buscar maneres d'aprofitar les oportunitats emergents.

La difusió de dades en obert i l'ús de georeferenciació i visualitzacions de dades, és un dels camps que presenta majors ocasions de treball per tal que els arxius entrin de ple en aquesta revolució de la informació. No només és interessant difondre dades en obert de tipus estadístic del funcionament

dels mateixos centres d'arxiu, sinó que aquests disposen d'infinitat d'informació susceptible de ser reutilitzada i explotada amb múltiples i impensables finalitats. Els instruments de descripció són eines de gran valor que, en la gran majoria, ara per ara, són consultables a Internet. Així i tot, és just en els darrers anys que s'han començat a explorar els possibles nous usos que poden tenir, especialment en diferents projectes acadèmics. Pel que fa al cas català, s'han elaborat des de mapes de fons notariais, passant per fons municipals fins a l'obertura de descripcions d'unitats documentals a l'Arxiu Nacional de Catalunya (Meinhardt Llopis, 2014; Munté Van Enkhui-zen, 2017; Pastrana Roman, 2018; Terradellas Prat, 2016).

El segon punt de partida es troba en la voluntat i l'interès personal per contribuir des de la professió arxivística a la reducció del biaix de gènere. D'aquesta manera, es planteja un projecte que vol donar visibilitat a la dona com a productora de fons documentals, així com donar a conèixer les activitats i experiències que ha desenvolupat al llarg de la seva trajectòria vital. A través d'aquest treball es dona una visió general del context i la situació actual de la dona als fons documentals dels arxius i se'n reivindica la seva rellevància. Igualment, es reafirma la importància de millorar l'accessibilitat, i facilitar la consulta i la reutilització de la informació.

Així doncs, es parteix de la combinació de dues idees: d'una banda, que les dades obertes faciliten l'accés a la

informació i la reutilització, generant valor afegit als arxius i als seus fons, i d'una altra banda, que la professió arxivística pot contribuir a reduir el biaix de gènere.

1. 2. Hipòtesi de treball i objectius

El principal problema detectat és la poca quantitat de fons de dones que hi ha als arxius públics, així com el tractament desigual que han rebut una bona part d'aquests. Així, a partir de la detecció d'aquesta problemàtica i de la reflexió feta al voltant de les constatacions inicials de les quals parteix el projecte, s'ha definit la hipòtesi de treball, que és la següent:

Les dades obertes permeten múltiples usos i, juntament amb visualitzacions de gràfics i mapes, milloren l'accessibilitat i la comprensió de dades per tant, l'obertura i la geolocalització dels fons produïts per dones contribueix directament a la visibilització de les mateixes com a productores de fons documentals.

Amb aquest projecte es vol donar evidència de la manca de representació de la dona als arxius, de tal manera que sigui un revulsiu que incentivi als professionals de l'arxivística, arreu de Catalunya a ser proactius en la captació de fons i impulsin estratègies inclusives per als col·lectius menys o gens representats a les institucions. En aquest cas concret, la dona, en tota la seva diversitat d'experiències, com a subjecte productor.

De la detecció del problema i el plantejament de la hipòtesi de treball parteix l'objectiu principal del projecte, l'elaboració d'un repositori de dades obertes que centralitzi els inventaris de fons produïts per dones i permeti posar els instruments de descripció a disposició dels usuaris en formats oberts, reutilitzables i explotables. Una eina que ofereixi una porta d'entrada als arxius diferent de la tradicional.

Tot i l'ambició inicial, donades les característiques del present treball ha estat necessari acotar la recerca. Pel que fa als arxius, es treballen els inventaris de la Xarxa d'Arxius Comarcals i de l'Arxiu Nacional de Catalunya, en tant que centres que custodien un volum important de fons documentals, tenen presència a tot el territori, i una part rellevant, tot i que definitivament insuficient, són de dones. Pel que fa a la tipologia de fons, es treballaran especialment els fons personals, però també s'han tingut en compte els fons patrimonials i familiars i les col·leccions, en tant que són susceptibles de contenir documentació que hagi estat produïda per una dona.

En resum, amb aquest projecte es vol impulsar l'ús de formats oberts i reutilitzables per presentar les dades als usuaris dels arxius; es vol conscienciar sobre la necessitat de fer pedagogia i divulgació per tal que el ciutadà conegui l'existència d'aquests fons, en aquest cas a través de les seves descripcions, i sàpiga què hi pot trobar; es vol facilitar aquesta informació utilitzant recursos de georeferenciació

i visualització de dades per tal d'afavorir la comprensió d'aquestes, i donar referència sobre com es pot explotar la informació descriptiva dels fons. Amb tot, també vol ser una mena de creació artística on es mostri poèticament allò que manca.

1. 3. Metodologia i estructura del treball

La investigació s'estructura en quatre apartats principals. Al primer apartat, s'elabora el marc teòric del projecte centrant la recerca principalment en dades obertes i geolocalització. A continuació, s'aborda l'anàlisi de l'estudi de cas i el seu context específic, tot emmarcant l'origen dels fons de dones, les seves característiques en el paradigma postmodern i les tendències de treball actuals. Es delimita l'abast del projecte tot definint els arxius i els fons a treballar i, al mateix temps, s'apunten algunes qüestions relatives al principi de provinença i al concepte de productor, les quals afecten directament a la visibilitat de la dona a través dels instruments de descripció dels fons que ofereixen els arxius, així com les dificultats que esdevenen a l'hora d'identificar els fons en els quals la dona n'és productora, en la seva totalitat o en part.

Un cop s'ha elaborat l'estat de la qüestió de la situació actual dels arxius quant a la difusió de dades obertes i, concretament des d'una perspectiva de gènere, s'ha procedit a la part més pràctica del projecte, la qual ocupa dos grans apartats: d'una banda, el disseny del *dataset* amb la seva

estructura i contingut, i d'una altra banda, la prova de concepte de publicació de les dades amb la creació de mapes i gràfics que faciliten la seva comprensió i mostren els principals resultats obtinguts amb l'estudi de cas.

A l'apartat 4. *Disseny del dataset*, en primer lloc, s'ha exposat el procediment per a recollir les dades, de quina informació es disposa, en quins formats i com està estructurada. En segon lloc, s'han analitzat les dades i s'ha fet un primer tractament per tal de tenir un conjunt possibilista i adreçat a tota mena de públic, és a dir, que sigui senzill i fàcil d'interpretar. En tercer lloc, s'han analitzat les normes i estàndards de descripció actuals per acabar seleccionant els elements idonis a incloure al *dataset*. En quart lloc, es defineix el model de dades, així com l'estructura i les relacions entre les diferents taules que el conformen.

Finalment, a l'apartat 5. *Publicació de les dades* s'explica el procediment seguit per posar a disposició dels usuaris les dades treballades. Es publiquen diferents *datasets* obtinguts a partir de la combinació de les taules definides al model de dades. Estan pensats per mostrar exemples de publicació de dades a diferents nivells descriptius. Alguns s'acompanyen amb visualitzacions concretes que exemplifiquen alguns dels usos i anàlisi que se'n poden fer.

2. 1. Dades obertes

2. 1. 1. El concepte. Característiques i principis

Dades obertes (*Open Data* en anglès) és un concepte que es refereix als conjunts de dades produïts o recopilats per les administracions públiques, així com els organismes vinculats o dependents d'aquestes, i que són posats a disposició del ciutadà per tal que aquest pugui utilitzar-los de la manera més lliure i senzilla possible.

Aquest terme parteix de la base de què la informació pública és de tots i que, per tant s'ha de fer accessible a tothom. Obrir dades millora la transparència de les administracions públiques i fa possible la seva reutilització per part d'empreses, entitats, associacions i ciutadania en general, ja que permet l'elaboració de nous productes i serveis que aporten valor afegit, innovació, coneixement i oportunitats de negoci.

L'*Open Knowledge Foundation*¹, és una organització internacional sense ànim de lucre que promou les dades obertes fomentant l'ús i distribució d'informació entre governs i administracions públiques, i dona a conèixer les possibilitats i oportunitats que ofereixen aquestes dades a la societat civil. A *The Open Data Handbook*, una guia elaborada per aquesta organització, es parla sobre els aspectes legals, socials i tecnològics de les dades obertes, i es defineix el

concepte com a “*datos que pueden ser utilizados, reutilizados y redistribuidos libremente por cualquier persona, y que se encuentran sujetos, cuando más, al requerimiento de atribución y de compartirse de la misma manera en que aparecen*”². Es detallen a continuació les característiques principals que han de complir les dades obertes:

- **Disponibilitat i accés.** Les dades han de ser accessibles com a tot, a un cost raonable i preferiblement s'ha de poder consultar i descarregar en formats oberts.
- **Reutilització i redistribució.** S'han de publicar les dades amb llicències d'ús que permetin la seva reutilització i redistribució, i han de preveure la possibilitat d'integrar conjunts de dades diferents.
- **Participació universal.** Tothom ha de poder utilitzar, reutilitzar i redistribuir les dades publicades.

Al manual s'apunten també algunes de les àrees en les quals fins al moment s'ha creat valor gràcies als governs oberts, entre les quals destaquen: la transparència i control democràtic, la participació ciutadana, l'apoderament, la millora i la creació de nous productes i serveis, innovació, la mesura de l'impacte de les polítiques públiques o la creació de nou coneixement a partir de la combinació de grans volums de dades.

Hem de tenir present que el concepte de dades obertes engloba tant les del sector públic com el privat, de tal manera que és més correcte parlar d'obertura de dades del sector públic o *open government data*. Les dades obertes són el component bàsic i essencial per poder parlar de governs oberts, i això es dona quan són útils, usables i usades. La base de la importància de remarcar tots aquests paràmetres per garantir que la publicació d'informació es pugui considerar com a obertura de dades és la interoperabilitat. L'essència de les dades obertes és que es puguin relacionar i barrejar *datasets*. Enfocar-se a publicar dades que compleixin estàndards d'interoperabilitat ha de servir també per evitar que hi hagi conjunts de dades que tinguin poques o cap possibilitat de ser combinats amb sistemes més grans, és a dir, allà on se'n podria extreure veritable valor.

L'any 2007, després d'una trobada a Califòrnia de trenta governs oberts, es van definir els 8 principis bàsics que defineixen i permeten parlar d'*open government data*. Amb la implementació d'aquests principis es reforça i promou l'ús de dades del sector públic. L'any 2010, la *Sunlight Foundation*, que també havia estat present a Califòrnia, va ampliar fins a 10 aquests principis³, i són els següents:

1. Integritat. Les dades publicades han de ser completes i degudament anonimitzades. Han d'incorporar metadades i informació de context de les dades en brut

(*raw data*), així com fórmules i explicació de com s'han recopilat. Això ha de permetre als usuaris entendre millor la informació disponible i treballar les dades amb major detall.

- 2. Ús de dades primàries.** Les administracions han de publicar les seves pròpies dades, i incloure-hi informació sobre la recopilació d'aquestes. Això ha de permetre conèixer la seva precisió i veracitat.
- 3. Disponibilitat.** Publicació de dades en el període de temps més breu possible des de la seva recopilació, tot donant prioritat a informació sensible al pas del temps. L'actualització de la informació en temps real potencia la utilitat de les dades.
- 4. Accessibilitat.** Facilitat amb la qual es pot obtenir la informació. Aquesta facilitat d'accés fa referència, tant a eliminació de barreres físiques com electròniques.
- 5. Llegible per màquines.** Prioritzar l'ús de formats que siguin fàcilment interpretables per màquines. Evitar especialment OCR i PDF. Si es presenta en aquests formats, també hauria d'estar disponible en altres fàcilment interpretables per màquines.
- 6. No discriminació.** Evitar barreres d'ús com registres o formularis per accedir i descarregar les dades. No hi ha necessitat d'haver de justificar l'accés.

7. **Ús de formats oberts.** Evitar formats propietaris. Prioritzar formats alternatius de programari lliure que permeten accedir al contingut sense necessitat de disposar d'un *software* amb llicència. Minimitzant o eliminant aquesta barrera s'arriba a un públic molt més ampli d'usuaris potencials.
8. **Lliure de llicències.** Per a una obertura completa, s'han d'oferir les dades lliures de restriccions d'ús o requisits d'atribució.
9. **Permanència.** Disponibilitat al llarg del temps. Cal fer un seguiment i establir criteris adequats per a l'emmagatzemament de les dades. Notificar als usuaris les possibles actualitzacions o modificacions que hagin patit els conjunts de dades en línia.
10. **Cost d'ús.** Reduir o eliminar les taxes d'accés i ús de la informació.

L'any 2015 es va redactar la Carta Internacional de Dades Obertes⁴, la qual defineix el terme de la següent manera: *“son datos digitales que son puestos a disposición con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar”*, una definició bastant semblant a la que ja s'havia formulat al *The Open Data Handbook*. Aquesta carta determina sis principis, alguns similars als ja esmentats:

1. Obertura de dades per defecte
2. Oportunes i exhaustives
3. Accessibles i usables
4. Comparables i interoperables
5. Per a la millora de la governança i la participació ciutadana
6. Per al desenvolupament inclusiu i la innovació

El Govern de la Generalitat de Catalunya es va adherir als compromisos que estableix aquesta carta el 17 de juliol de 2018⁵. A més, en el marc d'aquesta adhesió també en va resultar l'*ACORD GOV/154/2018*⁶, en el qual es defineixen actuacions per a la millora i la promoció del Portal de dades obertes de la Generalitat, i per al foment de l'ús i l'exploració de les dades que publiquen. Tot plegat, mitjançant polítiques de gestió integral dels sistemes informàtics que faciliten el tractament de dades amb finalitats de transparència i reutilització.

2. 1. 2. Per què és important obrir dades del sector públic

La informació de les administracions en molts de casos és pública per llei i, en conseqüència, pot ser oberta sense restriccions -amb excepció d'aquella que conté dades especialment protegides o que per interessos públics no poden ser revelades-. La informació del sector públic (ISP, *PSI* en anglès) conté informació molt valuosa per al

desenvolupament econòmic i social, de tal manera que la seva obertura aportarà valor i creixement.

Segons l'*Open Knowledge Foundation* cal obrir dades, principalment, per tres raons:

- **Transparència.** L'obertura de dades afavoreix la transparència i contribueix al fet que els ciutadans de les societats democràtiques tinguin coneixement de què fan els seus governs. Per això, cal que puguin accedir lliurement a les dades i a la informació governamental. Així i tot, apunten que transparència no és únicament tenir accés a les dades, sinó que sigui possible compartir-les i reutilitzar-les.
- **Valor social i comercial.** Les dades obertes potencien i promouen la creació de negocis i serveis innovadors que aporten valor social i comercial.
- **Participació i compromís ciutadà.** Permet anar molt més enllà de la mateixa transparència, ja que els ciutadans no només saben què fan els seus governants, sinó que estan millor informats i, per tant, poden prendre decisions fonamentades i contribuir al desenvolupament de les seves ciutats i estats.

En el mateix sentit, Nagore de los Ríos (2013) entén el govern obert com una forma de fer política, i comparteix els tres pilars sobre els quals es basa: transparència, partici-

pació i col·laboració. I remarca que ha de ser en aquest ordre, en tant que sense transparència no es pot participar amb garanties. Per ella, obrir dades *“provee transparencia en su estado más puro, porque en los datos está o debería estar el origen de la toma de decisiones, por lo tanto son básicos para contribuir a la transparencia y también para explicar los porqués de las decisiones. (...) deben ser fáciles de encontrar y de entender.”* (de los Ríos, 2013, 145).

Obrir dades no només és interessant i té impacte pel que fa a la innovació, les oportunitats de negoci i el creixement econòmic, sinó que també les mateixes administracions públiques poden millorar els seus serveis i generar nou coneixement a partir de la seva anàlisi i ús.

Com obrir dades

Fins al moment s'ha vist què poden fer les administracions públiques per esdevenir més obertes i democràtiques, és a dir, obrir les seves dades i posar-les a disposició dels ciutadans, per tal que en puguin fer ús lliure fàcilment. Com es pot intuir però, no només és important obrir les dades, sinó que perquè aquestes siguin útils, usables i usades, s'han de publicar i difondre seguint uns passos, criteris i condicions que així ho afavoreixin.

Per donar resposta al *com obrir les dades*, el *The Open Data Handbook* dedica un apartat a aquesta qüestió⁷. En primer lloc, explica i resumeix les principals regles a seguir per començar a obrir dades, que són:

- **Simplicitat.** Començar amb un conjunt de dades petit, que sigui senzill i ràpid d'obrir.
- **Rapidesa i freqüència** en les comunicacions amb l'usuari. Publicar les dades tan aviat com sigui possible i mantenir-les actualitzades per demostrar fiabilitat i crear confiança amb el ciutadà.
- **Donar solució als temors i malentesos** més habituals. Identificar els més habituals i abordar-los des de l'inici.

Per a la creació de les dades proposen quatre passes principals:

1. **Escollir les dades.** Aquesta elecció pot basar-se en la necessitat dels usuaris potencials (cal preguntar quins són els seus interessos) per conèixer la demanda; el valor de les dades; la possibilitat real que hi ha de poder publicar les dades (que no entri en contradicció amb la legislació vigent de protecció de dades).
2. Utilitzar, i per tant escollir, una **llicència oberta**.
3. **Fer disponible la informació**, utilitzant un format útil i adient al conjunt de dades. Que estiguin disponibles en un format llegible per màquina per facilitar la reutilització.

4. **Fer visibles les dades publicades.** Es poden utilitzar eines especialment dissenyades per a tal finalitat, com és el cas de Datahub⁸, un catàleg i repositori de conjunts de dades d'arreu del món.

Per la seva banda, el *World Wide Web Consortium* (W3C) el 2009 va publicar una guia d'ús⁹ per ajudar a les organitzacions governamentals d'arreu del món a alliberar les seves dades, i explica les passes a seguir per a la seva publicació:

- Publicar les **dades en brut**, en tant que és la manera més ràpida i senzilla de fer-les disponibles a Internet. Recomanen que, tot i això, estiguin estructurades. Evitar formats de dades que només permeten lectura i no exportació.
- Crear un **catàleg en línia** amb les dades en brut. Es considera que aquest és un bon punt de partida perquè reflecteix el següent estadi de l'evolució d'Internet, és a dir, una Web com a servidor de fitxers.
- Que les dades siguin **llegibles tant per humans com per màquines**. Recomanen enriquir els *datasets* amb XHTML (*Extensible HyperText Markup Language*).

Els formats habituals per presentar dades obertes

El format en què es presenta la informació és un factor molt important per tal de poder parlar de dades obertes. Recordem que han de ser reutilitzables, redistribuïbles,

llegibles per màquines i lliures de llicències. Per obrir dades, els formats més habituals són:

- **HTML** (*HyperText Markup Language*). Llenguatge desenvolupat per W3C. Es basa en l'ús d'etiquetes per estructurar la informació i desenvolupar la descripció del contingut, el qual es mostra com a text. Utilitza l'estructura i el codi per definir el contingut d'una pàgina web, un text, imatges, vídeos i altres objectes digitals.
- **XML** (*eXtensible Markup Language*). És un format estàndard obert, flexible i àmpliament utilitzat per emmagatzemar, publicar i intercanviar dades estructurades.
- **KML** (*Keyhole Markup Language*). Llenguatge de marques basat en XML utilitzat per representar dades geogràfiques i el seu contingut relacionat. Especifica característiques com marques de llocs, imatges, polígons o descripcions textuais. Conté: títol, descripció bàsica del lloc, longitud, latitud i, opcionalment, altitud. És un format que es pot obrir amb múltiples aplicacions gratuïtes.
- **RDF** (*Resource Description Framework*). Llenguatge desenvolupat pel W3C. S'utilitza per representar la informació a Internet aportant significat als continguts que s'ofereixen al Web. Format complex que pretén donar sortida al paradigma de la Web semàntica.

- **JSON** (*JavaScript Object Notation*). Format estàndard obert basat en text per a l'intercanvi de dades estructurades. És fàcil de processar i de ser llegit per màquines, però també senzill d'entendre i interpretar per humans.
- **CSV** (*Comma-Separated Values*). Format obert i senzill per representar dades en forma de taula, en què les columnes se separen per comes i les files per salts de línia. Els camps que continguin una coma, un salt de línia o una cometa doble han de ser tancats entre cometes dobles. Habitualment es fa necessari identificar a què fa referència cada fila o columna.
- **ODF** (*Open Document Format*). Formats oberts basats en XML per a la creació de documents d'ofimàtica. Existeixen diferents extensions ODF en funció del tipus de dades, per exemple: text (.odt), full de càlcul (.ods), presentacions (.odp) o bases de dades (.odb). És un format basat en estàndard obert que pot ser utilitzat per qualsevol persona lliurement. Un dels programaris més coneguts que utilitza aquest format és Open Office.

A la guia de bones pràctiques publicada el 2009 pel W3C, ja esmentada anteriorment, s'explica que el format principal per a dades que siguin llegibles per l'ésser humà és XHTML. A més, aquest consorci és pioner en XML i RDF, formats als quals es pot accedir com si es tractés de bases de dades, utilitzant SPARQL, XQuery, JavaScript i altres llenguatges informàtics. La recomanació que fa W3C a l'hora

d'escollir el format amb el qual publicar les nostres dades és que s'adeqüi a les característiques del nostre *dataset*, de tal manera que en permeti, per un costat, una producció i publicació fàcil i eficient, i per un altre costat, que es puguin enllaçar les dades.

Per la seva banda, el Portal Europeu de Dades elabora recursos, breus i concisos, sobre l'ús de les dades obertes. En un d'aquests recursos¹⁰, s'explica la importància del format de presentació de les dades, l'estructura que fa possible que humans i màquines les puguem entendre i interpretar, i es presenta una taula amb els diferents formats i categories:

Data format category	Example data formats	Extensions	Adjustable?
Documents	PDF, RDF, text document, Word document	PDF, RDF, txt, doc	PDF and RDF are not adjustable Text and word documents are adjustable
Spreadsheets	Microsoft Excel documents, XML, spreadsheet, OpenDocument spreadsheet	csv, tsv, xlsx, xls, xml, ODS, JSON, Atom feed	All spreadsheet documents are adjustable
Figures	Pictures, online map viewer	jpeg, png, viz, invoke	Figures are not adjustable
Data to make figures with	Publish data online (usually maps)	wms, wfs, wcs, shp	Usually the data is not adjustable
Compressed files	Compressed gzip file, compressed folder with multiple documents	GZ, zip	It depends on what is in the compressed file

Taula 1. Categorització dels formats de fitxers per a dades obertes. (Font: <https://www.europeandataportal.eu/>)

Linked Open Data i l'esquema de les 5 estrelles de les dades obertes

Assolir la culminació de les dades obertes és arribar a les Dades Enllaçades o *Linked Open Data*, mètode iniciat per Tim Berners-Lee, també inventor del Web i un dels pares d'Internet¹¹. Mentre que l'obertura de dades pretén l'alliberació d'aquestes per a l'ús i reutilització generalitzat, les dades obertes enllaçades van més enllà.

Linked Data és una tecnologia basada en l'intercanvi de dades estructurades mitjançant l'ús dels estàndards Web actuals (HTTP, RDF i URIs) que permeten interconnectar-les de tal manera que puguin ser consultades des de diferents fonts. Per fer-ho, cal definir i descriure les rela-

cions entre les dades (recursos) al Web¹² amb models RDF, que permeten estructurar i enllaçar documents i dades independentment del lloc on estan allotjats. Per facilitar la recuperació de les dades hi ha eines per a la consulta d'informació com SPARQL que permeten llegir les descripcions d'aquestes relacions. En general es busca garantir la interoperabilitat.

El mateix Berners-Lee ens proposa un esquema de cinc estrelles¹³, el qual defineix els diferents nivells d'obertura de dades i permet mesurar el grau d'integració d'aquestes al Web¹⁴. Analitza l'accessibilitat i la usabilitat tècnica dels *datasets* disponibles en línia. El significat i escala de l'esquema d'estrelles és el següent:

1. Ús de llicència oberta
2. Ús de format reutilitzable
3. Ús de format obert
4. Ús d'identificadors oberts¹⁵
5. Ús de dades enllaçades

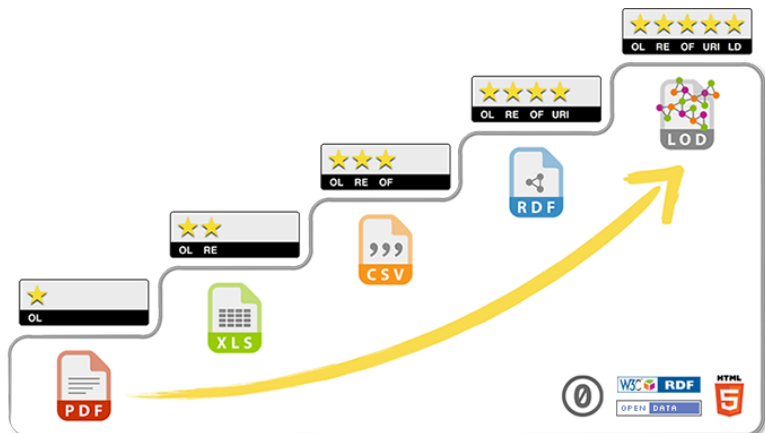


Figura 1. Esquema de cinc estrelles per a la publicació de dades obertes proposat per Tim Berners-Lee. (Font: <https://5stardata.info/es/>)

2. 1. 3. Legislació relativa a dades obertes i reutilització

Les administracions públiques no només disposen de la informació produïda per elles mateixes, sinó tot el que es coneix com a *informació del sector públic*, és a dir, inclou també tots aquells documents i dades que les persones físiques o jurídiques privades aporten a l'administració en el decurs d'un tràmit o gestió.

A nivell europeu, l'any 2003 s'aprova la Directiva 2003/98/CE¹⁶ estableix el conjunt mínim de normes i criteris homogenis per al tractament de la informació susceptible de ser reutilitzada. Deu anys després va ser modificada per la Directiva 2013/37/UE¹⁷, la qual fa extensiva la seva aplicació a “*las bibliotecas, incluidas las universitarias, los museos y los archivos*” argumentant que “*contienen una considerable cantidad de valiosos recursos de información del sector público, dado en particular que los proyectos de digitalización han multiplicado la cantidad de material digital de dominio público*”. Aquesta nova directiva afegeix entre d'altres, els punts 6 i 7 a l'article 2, on es defineix què és un *format llegible per màquina* i un *format obert* i en promou el seu ús. L'objectiu és facilitar la creació de productes i serveis que es basin en la reutilització d'informació del sector públic.

El 20 de juny de 2019 s'aprova la Directiva (UE) 2019/1024¹⁸, que introdueix noves actualitzacions a la ja modificada

Directiva 2003/98/CE. Una reforma que es considera necessària donat l'augment exponencial de dades que s'ha produït en aquesta última dècada i la constant evolució de la tecnologia amb la introducció de la intel·ligència artificial, l'Internet de les coses i la creació de nous serveis i aplicacions basats en l'anàlisi, explotació, agregació i combinació de dades que generen competitivitat i innovació.

Les principals novetats que s'incorporen són relatives a: prestació de dades dinàmiques en temps reals a través de mitjans tècnics adequats; augment del subministrament de dades de qualitat per a ser reutilitzades; ampliació de l'àmbit d'aplicació a dades d'empreses públiques de sectors com aigua, energia, transports i serveis postals; promoció de la creació de dades que siguin obertes des del disseny i per defecte de tots aquells documents als quals és d'aplicació la directiva per tal de millorar les condicions de reutilització; possibilitat de cobrar taxes superiors als costos marginals especialment per a les biblioteques, incloses les universitàries, els museus i els arxius, i organismes públics que han de generar ingressos per cobrir part dels costos derivats del seu exercici de servei públic.

A nivell estatal, la Llei 37/2007¹⁹ relativa a la reutilització de la informació del sector públic, estableix el marc normatiu de reutilització per tal de garantir-ne l'accés en igualtat de condicions. Aquesta llei serà modificada per la Llei 18/2015²⁰. La principal novetat és l'obligació d'autoritzar la reutilització de les dades (exceptuant aquelles en què

incorri alguna restricció) i s'hi actualitzen l'àmbit d'aplicació i els objectius pel que fa a l'àmbit d'arxius d'acord amb el que marcà en el seu moment la directiva europea, quelcom que queda definit al preàmbul i a l'article 3.3.h. A més, a l'article 3.1 es defineix reutilització com *“uso de documentos que obran en poder de las Administraciones y organismos del sector público, por personas físicas o jurídicas, con fines comerciales o no comerciales, siempre que dicho uso no constituya una actividad administrativa pública”*. I seguidament, l'article 4.5 de la mateixa llei disposa que les administracions crearan sistemes de gestió documental que permetin l'adequada recuperació de la informació. Han de facilitar eines que permetin la cerca de documents reutilitzables i accessibles, en els formats apropiats, en xarxa i llegibles per màquina.

Tot i això, el Govern de l'Estat Espanyol té fins al 17 de juliol de 2021 per actualitzar aquesta legislació i transposar el que disposa la recentment aprovada Directiva europea 2019/1024.

Altra legislació vinculada a la reutilització d'informació del sector públic és la Llei 19/2013²¹ de Transparència, Accés a la Informació Pública i Bon govern i la seva homònima a nivell autonòmic català, la Llei 19/2014²², les quals són molt similars en el seu contingut. Les dues preveuen la publicació i difusió d'actes administratius, documents i informació de les administracions públiques, en promouen la publicitat i l'accés, exigeixen més transparència i afavo-

reixen més drets pels ciutadans. La llei catalana elabora un llistat extens i enumerat de totes les obligacions de publicitat activa de l'administració. És rellevant destacar que l'article 5.4 de la legislació estatal disposa que les dades publicades hauran d'estar estructurades i preferiblement es presentaran en formats reutilitzables.

El Reial Decret 4/2010²³ regula l'Esquema Nacional d'Interoperabilitat en l'àmbit de l'Administració Electrònica i defineix les condicions necessàries a complir per garantir la interoperabilitat tècnica, semàntica i organitzativa dels sistemes i aplicacions utilitzats a les administracions públiques per tal de donar compliment a l'accés electrònic dels ciutadans als serveis públics. Posteriorment, la Resolució de 19 de febrer de 2013²⁴, de la Secretaria de l'Estat d'Administracions Públiques, aprova la Norma Tècnica d'Interoperabilitat de Reutilització de recursos de la informació, en la que s'estableixen *“selección, identificación, descripción, formato, condiciones de uso y puesta a disposición de los documentos y recursos de información elaborados o custodiados por el sector público”*.

Dret d'accés als instruments de descripció

La Llei 20/2015²⁵ de modificació de la Llei 10/2001 d'arxius i documents, modifica l'article 35 referent a publicitat activa i transparència, disposant que el Sistema d'Arxius de Catalunya (SAC) ha de complir amb les obligacions de transparència. Concretament es diu que els arxius han de fer públics els seus instruments de descripció, el registre

d'eliminació de documents i les limitacions de consulta de documentació. De la mateixa manera, també ha d'informar a l'usuari del seu dret a reclamar i quin és el procediment a seguir en el cas que se li hagi denegat l'accés. Tot això s'ha de poder consultar des del portal de transparència i a la seu electrònica de l'organisme titular de cada servei d'arxiu.

La legislació específica d'arxius no detalla que la publicació d'aquestes dades s'hagi de fer en formats oberts i reutilitzables, però hem de tenir present que la llei de transparència sí que diu que la publicitat s'ha de fer principalment en formats reutilitzables i, per tant, s'entén que es pot entendre extensiva la seva aplicació als arxius.

La normativa revisada tampoc defineix com s'han d'estructurar les dades que s'han de publicar. Això comporta que, donada la disparitat de recursos de què disposen els arxius, els instruments de descripció elaborats en cada cas són molt heterogenis. Així doncs, encara que es faci l'esforç d'obrir les dades, si aquestes no segueixen una estructura i uns estàndards d'aplicació, no es podrà garantir la seva interoperabilitat i serà complex creuar diferents conjunts de dades per generar valor afegit.

2. 1. 4. Arxius que ofereixen dades obertes

Pel que fa a experiències d'arxius que estan treballant amb dades obertes cal destacar, en l'àmbit internacio-

nal, els *The National Archives of America* (NARA). Al catàleg del Portal de Dades Obertes dels Estats Units d'Amèrica²⁶ els NARA hi han publicat un total 50 *datasets*, dels quals alguns només són enllaços HTML que redirigeixen a altres recursos. El que sí que és interessant és la possibilitat de descarregar, en format JSON, les metadades dels diferents conjunts publicats, inclús quan la informació que contenen no és accessible. Per altra banda, al web dels NARA podem descarregar en XML les descripcions arxivístiques²⁷ de la documentació federal que custodien.

Remarcable en l'àmbit internacional també és el Catàleg dels *The National Archives* del Regne Unit (TNA)²⁸, el qual permet exportar en format CSV, HTML i XML, les cerques que fa l'usuari. I el que és més interessant per aquest projecte, permet fer cerca segons tipus de productor i, en funció de la selecció que es faci, en alguns casos apareix un segon filtre que permet escollir entre diferents categories. En el cas de fons personals, aquest segon filtre permet seleccionar el gènere del productor. En resum, disposen d'un cercador bastant complet i molt interessant que permet obtenir dades en obert de les descripcions dels seus documents i incorpora la perspectiva de gènere en la manera d'oferir les dades.

En l'àmbit espanyol el cas més emblemàtic és l'Arxiu Municipal d'Arganda del Rey²⁹ el qual treballa amb el protocol OAI-PMH (*Open Archives Initiative – Protocol for Metadata Harvesting*)³⁰. Aquest sistema permet la recol·lecció automatitzada en fitxers

XML i amb esquemes normalitzats, quelcom que facilita la reutilització de les dades, així com el seu intercanvi, encreuament i explotació. El fet de tenir les dades estructurades i estandarditzades, fa que l'arxiu es converteixi en proveïdor de contingut per a altres portals com Hispana, Europea o el Portal Europeu d'Arxius.

Quant a dades obertes d'arxius en l'àmbit autonòmic català destaquen les iniciatives de l'Arxiu Nacional de Catalunya (ANC) i l'Arxiu Municipal de Girona. D'una banda, l'ANC, en compliment de la Llei 11/2007, publica al Portal de Dades Obertes de la Generalitat de Catalunya la *Llista de reparació jurídica de víctimes del franquisme (1938-1978)*³¹. A data de 27 de juny de 2019 aquest llistat inclou un total de 69.812 registres de víctimes i es pot descarregar en formats PDF, XLSX i CSV i, igual que en tots els *datasets* del catàleg de la Generalitat de Catalunya, es pot utilitzar l'API per incorporar les dades en eines com pot ser Instamaps³², per crear visualitzacions de dades, en aquest cas sobre mapes.

D'una altra banda, el Servei de Gestió Documental, Arxius i Publicacions de l'Ajuntament de Girona publica en format CSV els seus inventaris i catàlegs³³, així com dades bàsiques dels fons. De la mateixa manera, disposen d'informació sobre com utilitzar-les. Cal destacar la feina feta amb el *Llibre del Sindicat de Remença 1448*, segons únic document català inclòs en el Registre de la Memòria del Món de la UNESCO, del qual n'ofereixen diferents *datasets* en

obert, un dels quals relaciona les dones³⁴ que apareixen en l'esmentat document.

2. 1. 5. Iniciatives d'obertura de dades amb perspectiva de gènere

Existeixen nombroses iniciatives que s'han impulsat en els darrers anys que treballen per promoure l'obertura de dades amb perspectiva de gènere. A Argentina, des de 2017 es treballa per incloure la perspectiva de gènere en el *Sistema Estadístico de la Ciudad Autónoma de Buenos Aires*. Aquesta feina s'ha materialitzat en el Sistema d'Indicadors de Gènere³⁵, una agrupació de dades estadístiques i de gestió amb perspectiva de gènere que és consultable en línia i es pot descarregar en format XLSX. Aquestes dades permeten veure la situació diferenciada entre homes i dones a la ciutat.

Altres projectes destacables són *Datos 2x*³⁶, *Open Heroines*³⁷ i l'*Open Government Partnership*. Aquest darrer exemple es tracta d'una associació que treballa amb la societat civil per tal de determinar àrees a prioritzar en les polítiques públiques. Una de les principals estratègies en les quals s'enfoca és el gènere i la igualtat d'oportunitats. Col·labora en tres projectes dins aquesta línia de treball: *Break the Roles*, *Why Gender, Why Now* i *Addressing gender equity challenges in open government co-creation processes*³⁸. Finalment, esmentar també l'Institut Europeu per la Igualtat de Gènere el qual publica al Portal Europeu de

dades obertes, en format XLS i amb visualitzacions associades, dades referents a l'Índex d'Igualtat de Gènere³⁹ dels països membres de la Unió.

Tot i que hi ha exemples, i que es promou i impulsa l'obertura de dades públiques, la segregació per gènere encara no és tan habitual com caldria esperar.

2. 2. Geolocalització

2. 2. 1. El concepte

De què parlem, quan parlem de geolocalització? Pastana Roman (2018) constata en el seu treball amb una consulta a l'Institut Cartogràfic i Geològic de Catalunya (ICGC) que l'ús del terme geolocalització s'ha estès i és usat per a pràctiques que serien més aviat de georeferenciació i altres tècniques derivades.

L'ICGC disposa d'un diccionari⁴⁰ al seu web on defineix els diferents termes:

- **Geolocalització** és *“la determinació de la posició geogràfica d'un aparell connectat a una xarxa de comunicacions. Inclou la determinació de la posició geogràfica d'un dispositiu mòbil (...) i també la determinació de la posició geogràfica d'un ordinador connectat a Internet. (...) La geolocalització utilitza tecnologies, tècniques i mètodes de posicionament molt diversos i permet ob-*

tenir posicions amb diferents nivells de precisió (...). La geolocalització pot proporcionar posicions en forma de coordenades o simplement com una referència territorial (país, província, ciutat o barri)”.

- **Georeferenciació** és *“la representació de la posició dels fenòmens o entitats mitjançant un sistema de referència que permeti expressar posicions geogràfiques; és a dir, posicions sobre la superfície terrestres”.* Per tant, georeferenciar consisteix a ubicar una direcció en un mapa associant el punt a la coordenada.
- **Geocodificació**, es refereix al *“conjunt de mètodes de georeferenciació que expressen la posició geogràfica mitjançant referències a entitats geogràfiques de posició coneguda, en lloc de coordenades”,* o bé com a operació, en tant que *“procés d'obtenir la posició de les coordenades georeferenciades, ja siguin geogràfiques o projectades, a partir d'altres formes d'expressar la posició”.* Les coordenades poden ser utilitzades per localitzar el punt del mapa en un sistema d'informació geogràfica.

2. 2. 2. Arxius i projectes de geolocalització

D'uns anys cap aquí ha augmentat l'interès per l'ús de dades geogràfiques, per tal d'aportar valor a la informació del sector públic. Com s'ha vist, la Generalitat de Catalunya s'adhereix a la Carta Internacional de Dades

Obertes el desembre de 2018 i aprova l'estratègia de dades obertes a seguir pel Govern. D'aquesta estratègia és interessant destacar, pel que fa al present projecte, l'apartat 21, el qual fa referència a l'estandardització de dades, en tant que es diu que *“es prioritzarà la incorporació de la geolocalització i referències temporals i de gènere en les dades obertes, així com la incorporació de categories que fomentin l'anàlisi de les dades amb perspectiva social”*. Es reconeix doncs, la importància d'unir aquests dos conceptes per aportar valor i promoure el seu ús facilitant la comprensió de les dades i donant a conèixer exemples de què es pot arribar a fer amb elles.

La literatura respecte a l'ús de geolocalització des de l'àmbit arxivístic és més aviat escassa, però a poc a poc va creixent l'interès i s'estan veient les possibilitats que presenta. Especialment des de l'entorn acadèmic s'han començat a impulsar iniciatives que posen en valor l'ús de geolocalització per donar a conèixer els arxius i fer difusió del seu contingut.

Així, Carla Meinhardt el 2014 va presentar el projecte ArxiMaps, un treball pioner en aquesta línia d'investigació que analitza la realitat dels fons notariais de Catalunya i planteja el disseny d'una base de dades amb els instruments de descripció dels fons, informació geogràfica i un cercador que permet localitzar fàcilment la documentació arxivística a través d'un mapa. L'autora remarca la importància d'explotar al màxim les possibilitats que ofereixen

les noves tecnologies per donar a conèixer els actius d'informació dels arxius, així com mirar cap a les necessitats i els interessos concrets dels usuaris per tal d'oferir productes dissenyats expressament per a ells, i agafar certa distància respecte als termes més tècnics i propis de la professió a l'hora de projectar-se cap a fora.

Dos anys després, Enric Terradellas (2016) fa una proposta d'obertura de dades d'instruments de descripció de l'ANC, en aquest cas d'unitats documentals, analitzant el context i el punt de partida de l'arxiu, així com l'estructura de gestió interna, per tal de treballar les dades dels fons seleccionats, i finalment fer-ne una proposta de geolocalització. Segueix el treball de Karen Munté (2017), obrint el catàleg d'expedients de llicències d'obres particulars de l'Ajuntament de Santa Perpètua de la Mogoda i experimentant amb la creació de visualitzacions d'aquestes llicències sobre mapes a partir de diferents programes i analitzant les diferents possibilitats i problemàtiques que presenta cada cas. I finalment, el més recent presentat per Valeri Pastrana (2018), qui elabora una proposta de geolocalització d'expedients del fons de la Comissió Provincial de Regions Devastades i Reparacions de Girona, a partir dels catàlegs disponibles del fons.

En l'àmbit institucional, també comença a haver-hi algunes propostes destacables. D'una banda, la ja esmentada *Llista de reparació jurídica de víctimes del franquisme* publicada per l'ANC, s'acompanya amb dos mapes geolo-

calitzats, els quals s'han creat a partir de la mateixa llista, però un mostra l'origen de les persones executades i l'altre el lloc de residència d'aquestes. La georeferenciació dels dos mapes està feta a partir de les coordenades dels municipis, de tal manera que allà on hi hagi més d'una persona -registre- apareixerà un grup de punts o clúster, quelcom que pot dificultar la recerca, ja que caldrà clicar a sobre de cadascun per veure'n la informació i, si n'hi ha molts, és fàcil deixar-se'n algun. Així i tot, no es considera que

sigui un fet negatiu, doncs compleixen els objectius de la visualització. A més, els mapes no presenten la totalitat del *dataset*, sinó que s'han seleccionat i filtrat prèviament aquells camps útils per a la difusió d'aquest mapa concret -únicament mostra registres de les persones executades, és a dir, una part de tot el conjunt- de tal manera que es fa necessari anar a la taula original per accedir amb més detall a tota la informació.

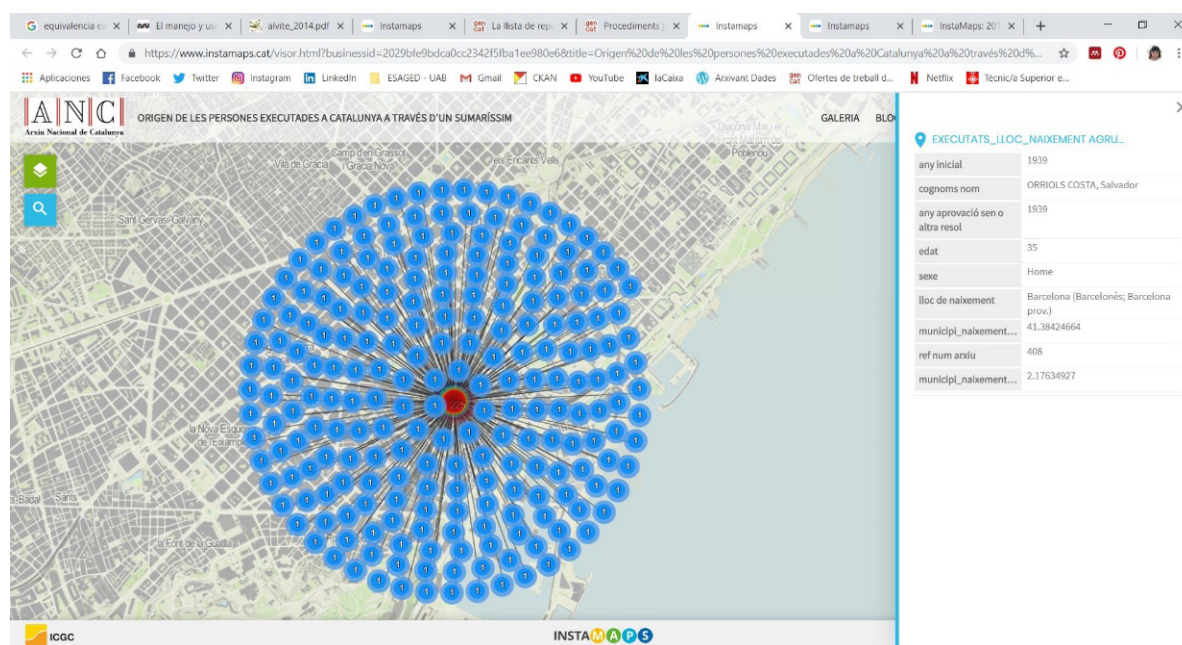


Figura 2. Detall del clúster de Barcelona. [Mapa Origen de les persones executades a Catalunya a través d'un sumaríssim](#). (Font: *Instamaps*. ICGC. Mapa elaborat per l'ANC)

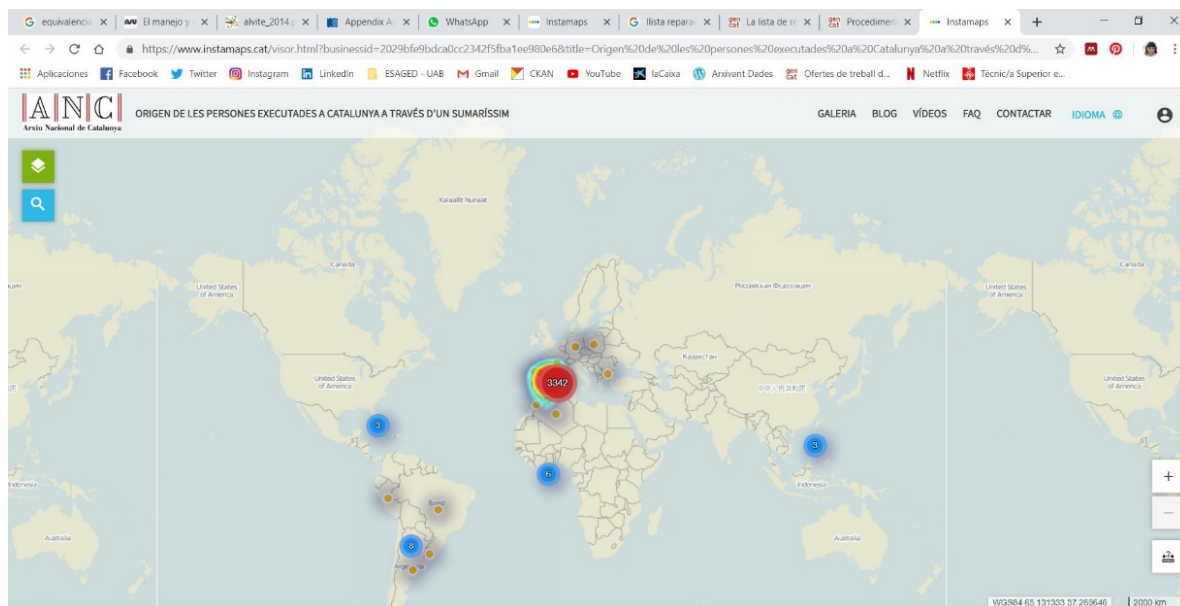


Figura 3. Visualització del mapa [Origen de les persones executades a Catalunya a través d'un sumaríssim](https://www.instamaps.cat/visor.html?businessid=2029bfe9bdca0cc2342f5fba1ee980e6&title=Origen%20de%20les%20persones%20executades%20a%20Catalunya%20a%20través%20d%20un%20sumaríssim). (Font: *Instamaps*. ICGC. Mapa elaborat per l'ANC)

L'altra característica interessant d'aquesta visualització geogràfica és l'ús de les capes d'agrupació i concentració, tot aplicant el mapa de calor per mostrar molt ràpidament la procedència de les persones que van ser executades.

Al Web de l'Archivo Histórico de Protocolos de Gipuzkoa es poden consultar dos mapes, un mostra on es localitzen els documents notariaus de Gipuzkoa⁴¹ i l'altre les comptadores d'hipoteques⁴² de la província. Ambdós presenten les mateixes característiques, s'han elaborat amb Google Maps, localitzen amb punts senzills i utilitzen colors per distingir els quatre districtes. Clicant sobre un dels punts

es mostren les dades relacionades com el nom, la descripció, les dates extremes i enllaços a descripcions del seu catàleg en línia.

L'Archivo de Arganda del Rey també ha elaborat diversos mapes, entre els quals destaca el que fa referència a plànols parcel·laris del municipi⁴³. Fet també amb Google Maps, permet filtrar segons tres anys diferents (1878, 1904 i 1945). Emmarca i divideix amb color les diferents parcel·les, i clicant a sobre apareix una finestra emergent en miniatura des d'on es pot accedir a la digitalització de la cèdula cadastral original.

2. 2. 3. Com aporta valor afegit la georeferenciació al projecte

Com es pot comprovar, no hi ha mapes sense dades. Res no existeix si no recopilem i preparem prèviament la informació. Els arxius que utilitzen mapes per fer difusió, ho poden fer perquè abans han plantejat i portat a terme estratègies d'obtenció i tractament de dades. Val a dir però, que no tots els mapes estan fets sobre dades obertes, però sí que faciliten la difusió i permeten als respectius arxius arribar a públics més amplis i diversos.

Els mapes que s'han consultat utilitzen eines i recursos diferents. Els objectius que busquen cadascuna de les propostes també són diferents i fan usos diversos dels estils de visualització. Tots però, tenen una voluntat comuna i és la de mostrar-se i obrir-se als usuaris digitals, així com ampliar el seu abast d'actuació.

En resum, podem dir que l'elaboració de mapes geolocalitzats, en termes generals però concretament als arxius, permet:

- Mostrar de manera diferent els seus instruments de descripció i el seu contingut.
- Establir relacions visuals entre arxius i/o fons i/o unitats documentals.
- Ús de diferents estils de visualització, com per exemple, l'ús de colors per establir diferències entre dades o categoritzar la informació. En els casos analitzats hem vist dos tipus d'ús en aquest sentit. D'una banda, els mapes de calor complementaris a la visualització de la volumetria, doncs varia el color en les zones on hi ha més o menys densitat de dades. D'una altra banda, les paletes de color, que generalment s'utilitzen per distingir categories de dades a partir de punts simples.
- Ús de finestres emergents per mostrar les dades mínimes del punt georeferenciat sobre el qual s'ha disposat el ratolí. Moltes de les plataformes per a la creació de mapes permeten que el propietari de les dades decideixi si vol publicar-les o no, de tal manera que es poden oferir en obert les dades utilitzades per crear-los.
- Establir diferents nivells de precisió geogràfica. Tot i que s'acostuma a ser el més precís possible a l'hora de georeferenciar una ubicació, les diferents possibilitats de geocodificació d'un punt permeten que l'exactitud sigui més o menys precisa segons els nostres interessos.
- Utilitzar les visualitzacions per demostrar i/o justificar les bones o males praxis, i enfocar-se a la millora de la gestió del servei, tant a escala interna com externa.
- Visualitzar d'un cop d'ull, volums o tipus de fons i quins arxius els custodien, el nivell de tractament dels fons,

les modalitats d'ingrés per les quals aquests fons han arribat als arxius, o qualsevol altra dada disponible al *dataset*.

- Plantejar-se noves estratègies d'actuació en adquisició de fons.
- Satisfer les necessitats dels usuaris a través d'un recurs que, juntament amb les dades en obert, permetrà que puguin donar diferents usos als instruments de descripció i trobar noves maneres d'explotar la informació que contenen.
- Diluir la necessitat dels arxivers d'exercir d'intermediaris entre els instruments de descripció i els usuaris dels arxius. Els pot fer més comprensibles i l'ús de mapes intuïtius permetrà que l'usuari tingui més autonomia i facilitat per comprendre el contingut de les descripcions.

3. Estudi de cas: la dona com a productora de fons documentals

30



3. 1. El perquè de la dona productora

3. 1. 1. Els orígens dels arxius de dones

Durant les dècades dels anys 1930 i 1940 les institucions acadèmiques dels Estats Units van augmentar els esforços per recuperar i preservar documentació no publicada. Aquesta dedicació per recopilar materials originals rarament va incorporar documents i fonts primàries per a la història de les dones, ja que durant molt de temps va ser una tasca que es va deixar quasi exclusivament en mans de feministes (Voss-hubbard, 1995).

El feminisme és un moviment que, en essència, busca motivar la dona a prendre consciència de si mateixa i del valor de la seva pròpia existència. Aquesta idea afavorirà, entre moltes altres coses, la salvaguarda de fons i col·leccions de dones, majoritàriament vinculades al moviment, i n'impulsa la seva conservació i preservació. Així és que durant aquestes primeres dècades, malgrat la poca consciència i predisposició de les institucions acadèmiques, van sorgir iniciatives principalment de dones actives dins el moviment feminista, que van treballar per a fer possible la creació de centres específics per a la recuperació de materials per a la història de les dones. El 1935 es van crear

el *The World Center For Women's Archives* (WCWA) a Nova York, i l'*International Archives for the Women's Movement* (IAV) a Amsterdam, avui conegut com a *ATRIA Institute on Gender Equality and Women's History*. Els paral·lelismes entre aquests dos centres, i altres dedicats a tasques similars, són clars. El rol d'arxiu que van desenvolupar va ser com "*a repository of documents and producer of new scholarly publications*", però també van realitzar la tasca de "*collection and production of feminist documents*" (Eichhorn, 2013, 38). En definitiva, centres la voluntat i l'objectiu principal dels quals va ser la recopilació de documents que no només fossin representatius per a la història de la dona, sinó que haguessin estat produïts per elles mateixes.

Segons Eichhorn (2013), hi ha similituds destacables entre WCWA, IAV i altres arxius i col·leccions relacionades, els quals es basen principalment en les idees amb les quals estaven compromeses aquestes organitzacions. D'una banda, la preservació de documents i altres materials relacionats amb el moviment feminista, així com la idea de ser uns espais d'educació i de foment per a la mobilització a partir de la producció i difusió de nous materials. D'altra banda, la conscienciació inicial vers la necessitat de mantenir estratègies i vincles amb institucions ja existents, així com amb les organitzacions més influents i poderoses.

Més enllà d'aquests primers arxius, no serà fins a la dècada de 1970 quan proliferarà veritablement la creació de centres dedicats a documentar la vida i activitat de les dones. Aquestes organitzacions es basen en les primeres experiències de les dècades de 1930 i 1940. L'objectiu principal de la fundació de tots ells és cobrir la inadequada representació de la dona que hi ha en els arxius tradicionals (Mason i Zanish-Belcher, 2007) taking as case studies two women's archives. Drawing on their experiences building the collections of the Archives of Women in Science and Engineering (Iowa State University i insistir en la conservació de la documentació generada per les dones per fer-la accessible als investigadors. La dècada dels 70 del segle XX, és un període en que està en auge la Segona Onada Feminista, quelcom que afavoreix encara més la presa de consciència i la necessitat de disposar de documents originals per poder estudiar la seva pròpia història i empoderar-se com a individus i subjectes amb identitat pròpia.

3. 1. 2. Els arxius de dones en l'era de la post custòdia arxivística

El canvi de paradigma postmodern, de la custòdia a la post custòdia arxivística, impulsa la presa de consciència respecte a la importància que té l'arxiver en la configuració de la memòria que tindran les generacions futures. Al voltant d'aquest debat, autors com Cook i Hernández Olivera (2007), apunten el fet que l'arxiver no és objectiu, i en conseqüència té la capacitat de decidir qui és visible i

qui no, i encara més, de decidir qui serà recordat i qui serà oblidat. En la mateixa línia, Eric Ketelaar (2001) afirma que arxivar implica també seleccionar què s'ha de preservar i què no, i que hem de ser conscients que d'aquesta decisió en dependrà directament la narració que creem a la nostra organització i la memòria que perviurà en la societat en què ens emmarquem.

Aquest nou paradigma implica que com a arxivers hem de ser conscients de la nostra subjectivitat i el biaix de representació que aquesta pot comportar (Mason i Zanish-Belcher, 2007). Qualsevol temptativa d'imparcialitat comportarà sempre, malgrat tot, una presa de decisió, sigui aquesta conscient o inconscient, que pot afectar en gran manera quins subjectes es troben o no representats als arxius. Així doncs, parlar de *women's collections* o *women's archives* i de recuperació de documents originals que siguin representatius de la vida i experiències de les dones, en aquest nou context també implica certes dificultats. La representació de la dona als arxius no està exempta d'aquest biaix de subjectivitats dels i les professionals del sector, i com a arxivers hem de buscar la manera de ser més inclusius, especialment amb les minories i col·lectius infra-representats i poc valorats socialment.

Algunes autores han reflexionat sobre diverses qüestions, entre les quals en destaca: si realment hauria de ser necessari separar els fons de dones dels fons generals o, almenys, fer-ne distinció dels d'homes (Moseley, 1973); la figura de

l'arxiver activista i la importància de la proactivitat en la recuperació dels fons i col·leccions per tal d'assegurar que la veu dels grups no dominants també és present als arxius (De Haan i Mevis, 2008; Jornet Benito i Tuset Páez, 2016; Mason i Zanish-Belcher, 2007; Römkens i Wiersma, 2017) i la dificultat que implica preservar-los donada la gran diversitat de suports, formats i tipologies (Jornet Benito i Tuset Páez, 2016); la necessitat de repensar els instruments de descripció per evitar amagar les dones a través dels processos arxivístics (Henningham, Evans i Morgan, 2017); fins i tot, i potser el més important, la necessitat de dissenyar polítiques d'adquisició que compleixin amb les expectatives de representació de grups minoritaris de la societat (De Haan i Mevis, 2008; Jornet Benito i Tuset Páez, 2016), quelcom que requereix estudis i anàlisis dels fons actuals per detectar els buits.

En la majoria de referències consultades, es considera cabdal buscar noves formes de promoure la donació de fons personals de dones que puguin reflectir les múltiples experiències de les mateixes: què fan, què pensen, com es relacionen, quins són els seus interessos, quines activitats i professions han desenvolupat i exercit o com han viscut. Visibilitzar la dona a través dels seus propis documents i arxius és una manera de donar importància i valor a les seves vivències.

Autores com Mason i Zanish-Belcher (2007) i Moseley (1973) coincideixen en posar sobre la taula que els forats de representació de la dona als arxius, no només es de-

uen a la manca de proactivitat o al biaix personal dels i les professionals de l'arxivística, sinó al fet que la majoria de les dones no tenen present ni són conscients que la documentació que han generat té valor per a la història i forma part de la memòria col·lectiva, i els han tractat en conseqüència, desvirtuant-los o, senzillament, no els han conservat. Encara més, Mason i Zanish-Belcher (2007) introdueixen en aquest debat el nivell educatiu com a factor clau en l'autovaloració de la documentació, molt més important inclús que la raça o ètnia. Només aquelles persones que han rebut una certa educació han estat conscients del valor d'allò que han produït i, per tant, han tingut major tendència a conservar la seva documentació. Anys abans, les mateixes autores ja consideraven que el vertader repte és *“to persuade women that their reminiscences, ephemera, and oral histories have value and interest”* (Mason i Zanish-Belcher, 1999, 45). És per això que, augmentar les donacions i ampliar la representació de la dona és, al mateix temps, una forma de potenciar la seva pròpia valoració com a individus i, per tant, afavorir la conservació d'allò que les fa ser conscients de la seva importància com a tal. Igualment, la feina de recopilació i difusió d'aquests fons pot contribuir a retroalimentar les donacions, en tant que es pot encoratjar a dones que mai s'havien pensat valuoses a si mateixes, a considerar-se mereixedores de ser recordades (Mason i Zanish-Belcher, 1999).

Ara bé, no sempre l'adquisició de fons és possible i no és l'únic mitjà de què disposen els arxius per reduir els biaixos

de representació. Cal valorar quina és la manera més adient d'omplir les mancances de representació, així com tenir present que és molt probable que les dones que no han tingut l'oportunitat de rebre educació, no hagin deixat cap mena de llegat documental escrit o que sigui poc representatiu en el seu conjunt. Això implica que hi haurà casos en que els forats de representació no podran ser coberts amb documentació escrita, pel simple fet que aquesta no existeix. Per aquest motiu, s'impulsen també molts projectes d'història oral, la importància dels quals destaquen moltes de les autores que han teoritzat sobre gènere i arxius (Caswell i Cifor, 2016; De Haan i Mevis, 2008; Henningham, N., Evans, J., i Morgan, H., 2017; Jornet Benito i Tuset Páez, 2016; Mason i Zanish-Belcher, 1999, 2007; Römkens i Wiersma, 2017). Aquests projectes permeten, a les dones que hi participen, deixar constància de les seves vivències amb la seva pròpia veu i amb les seves pròpies paraules.

Una preocupació comuna a bona part de la bibliografia consultada té a veure amb l'accés a la documentació i les dades recopilades. No n'hi ha prou amb l'adquisició dels fons o la posada en marxa de projectes de recollida de dades, com els d'història oral als que s'acaba de fer esment. És necessari que se'n faci difusió però, sobretot, és cabdal que aquesta documentació rebi un tractament adient per tal que sigui accessible. En aquest sentit, es detecten dificultats pel que fa a la identificació i descripció. Tal com apunten Mason i Zanish-Belcher (1999, 46), sovint l'accés a aquests fons es complica, ja que d'entrada tro-

bem dificultats a l'hora de localitzar-los perquè "*women's materials were often not identified as such, or were 'hidden' in the papers of male family members or colleagues or in organizational records*". I no només això, en moltes ocasions únicament es pot saber si es tracta o no de documentació produïda per una dona si l'organització que la conserva així ho ha identificat expressament⁴⁴.

Wernimont i Flanders (2010) encara han anat més enllà i s'han qüestionant si el fet de facilitar l'accés a aquests fons és realment una solució per a la visibilització de la dona i, sobretot, si és una manera de treballar activament per la inclusió social del col·lectiu i de fer front a l'exclusió sistemàtica. En relació amb la inclusió, De Haan i Mevis (2008) consideren que no sempre és possible aconseguir-la, encara que les polítiques d'adquisició de fons estiguin vertaderament dissenyades per aquest motiu.

Si bé és cert, Wernimont i Flanders (2010) també destaquen que, abordar la visibilització dels fons de dones a partir de la creació de mapes geolocalitzats i altres tipus de visualitzacions que contextualitzin els fons i documents, així com l'ús d'eines que faciliten la recerca, és una tasca important i necessària per a donar-los a conèixer. No té sentit facilitar l'accés a la informació si qui pot utilitzar i explotar les dades en desconeix la seva existència, localització i possibilitats d'ús.

Finalment, un altre aspecte que s'ha de tenir present a l'hora de tractar aquests tipus de fons és la tendència actual al biaix feminista. Cal ser-ne conscients i procurar evitar-lo per afavorir la inclusivitat. En cas contrari, correm el risc de no incloure documents de grup de dones que no comparteixen o s'oposen als valors dels feminismes, tals com les organitzacions de dretes o aquelles contràries a l'avortament (Mason i Zanish-Belcher, 2007).

Diverses autores (Carreño Corchete, 2017; Codina-Canet i San Segundo Manuel, 2016; Jornet Benito i Tuset Páez, 2016) que han treballat els arxius des del punt de vista feminista han arribat a la mateixa conclusió. Segons elles, és necessari constituir un arxiu del feminisme per tal de fer front a la dispersió de fons i col·leccions, i així donar visibilitat al moviment i a les dones que hi han participat.

Tot i ser una proposta interessant, no es considera que sigui la millor opció, doncs cal definir molt bé què s'arxivaran en aquests centres, quines dones s'hi arxivaran i com es tractarà la documentació que s'hi ingressi. I no només això, la representació de la dona va molt més enllà de la seva identificació i sensibilització (o no) amb els feminismes. Les dones que hagin participat i contribuït activament al moviment feminista, és probable que també hagin destacat en altres àrees tant en l'àmbit acadèmic, professional com personal. Per tant, en relació amb els objectius que es proposen en aquest projecte, el plantejament d'aquestes autores no dona resposta a una visibilització objectiva

del paper de la dona com a productora de documents més enllà del feminisme.

A banda del que s'acaba d'apuntar, en el context postmodern, els arxivers i arxiveres estem cridats a ser capaços de conservar fons de persones, físiques o jurídiques, que potser no voldríem perquè no encaixen amb les nostres creences o la nostra manera de pensar. Segons Cifor i Wood (2017), els apropaments feministes a la pràctica i estudi de l'arxivística s'han centrat a cobrir els buits de representació a través dels fons en sí mateixos, tot creant espais per conservar-los adequadament com a tal, però s'ha d'anar més enllà. El feminisme crític que proposen les autores, entès "*as an intersectional political philosophy committed to the dismantling of heteronormative, capitalists, racist patriarchy*" (Cifor i Wood, 2017, 2), té a veure també amb una intervenció i transformació dels arxius des de dintre. Hem de ser capaços de construir i mantenir una memòria el més diversa, completa i rigorosa possible, i això ho permet l'apropament feminista crític, el qual va molt més enllà inclús de buscar una millor representació de la dona.

3. 1. 3. Tendències i projectes actuals

Iniciatives a Catalunya

Durant l'última dècada han sorgit alguns projectes i iniciatives de diversa naturalesa, impulsats des de diferents àmbits, per a la incorporació de la perspectiva de gènere i la visibilització de la dona a Catalunya. De la mateixa

manera, s'han publicat articles que remarquen algunes de les línies estratègiques que s'haurien de treballar des dels arxius per reduir el biaix de representació.

El 2010 la Generalitat de Catalunya i el Govern de les Illes Balears van iniciar un projecte que, tot i que ha quedat inacabat, és interessant tenir present. Es tracta del *Diccionari Biogràfic de Dones*⁴⁵, un diccionari en línia, d'accés públic i gratuït que actualment compta amb un total de 654 biografies de dones, i que vol posar en relleu les aportacions d'aquestes en esferes polítiques, socials i culturals, dels territoris de parla catalana. La part més interessant d'aquest projecte és que les biografies es vinculen a les fonts, d'arxius i biblioteques, que recullen informació sobre cadascuna d'elles.

El 2016 Jornet Benito i Tuset Páez van publicar l'article *Construint la memòria dels feminismes: arxius, biblioteques i centres de documentació. Una mirada al passat, una reflexió per al futur*, en el qual es fa un recorregut pels diferents centres d'Europa que han treballat la memòria de la dona, majoritàriament des del feminisme. La principal conclusió que extreuen d'aquesta anàlisi, especialment pel que fa a la part de documentació d'arxius, és detallar

“les línies essencials que haurien de perfilar una política de la col·lecció: activar estratègies de recuperació i captació d'arxius personals, de dones vinculades al feminisme, d'associacions o entitats que s'hi relacio-

nin (...); sistematitzar models d'acord i convenis amb els titulars de fons; promoure la donació i sistematitzar-ne les vies; incorporar les fonts orals amb la gravació d'històries i entrevistes personals, i activar ingressos i conservació de nous materials.” (Jornet Benito i Tuset Páez, 2016)

De la mateixa manera, i més vinculat amb el present projecte, les autores consideren essencial la creació de “*guies de centres que, de manera virtual, apleguin els fons i el material dispers i ofereixin un “mapa” que doni sentit i reculli el nombre, la ubicació i la significació d'aquest material, com també tots els centres i actors implicats.*” (Jornet Benito i Tuset Páez, 2016).

Cal destacar igualment la publicació d'alguns articles en revistes digitals que han alertat sobre l'alarmant invisibilització de les dones en els arxius catalans. La Dra. Remei Perpinyà escrivia el juny de 2016 a *Núvol, el digital de cultura* un article sobre la presència invisible de les dones als arxius⁴⁶, en el qual oferia una relació de percentatges de la quantitat de fons personals de dones conservats als centres d'arxius més rellevants de Catalunya en comparació al volum total custodiat per aquests. Al mateix article ja es reclama més implicació per part de la comunitat arxivística en la recuperació i conservació de la memòria de les dones, indicat que “*cal establir els mecanismes que encoratgin tant dones a títol individual com organitzacions femenines a dipositar els seus documents en arxius i biblioteques, i crear i mantenir un registre de fons*

documentals de dones i del feminisme". Es planteja la necessitat de donar impuls a la "incorporació de fons de dones als arxius públics i tenir present que la història de les dones no és només la història del feminisme", un fet que ja s'ha remarcat anteriorment, i que és objectiu d'aquest projecte, el de visibilitzar la diversitat de dones molt més enllà dels feminismes. Més recentment, en un nou article⁴⁷ ha reclamat novament més atenció a la memòria de les dones per part dels arxius catalans, remarcant que des del feminisme sí que s'està fent aquesta feina de conservació de la documentació generada, tot advocant per incloure conceptes com *representativitat*, *diversitat* i *inclusió* per a valorar els criteris d'adquisició de fons als nostres arxius.

Cristian Rubio, historiador i arxiver, arran d'aquest últim article publicat per Perpinyà reafirma que "*confondre els arxius de dones amb els arxius del feminisme és un error i recollir en un centre tots els fons de dones és combatiu i comprensible però sembla més un recurs desesperat que una solució a llarg termini*"⁴⁸.

L'Arxiu Històric de Terrassa i l'Arxiu Comarcal del Vallès Oriental han publicat recentment una guia que defineix el seu compromís en la integració de la perspectiva de gènere. En aquest document s'exposen les principals línies estratègiques per fer-ho, entre les quals es troba "*representar la diversitat de dones de la ciutat*" i "*integrar la perspectiva de gènere en la política d'adquisicions i ingressos de fons i fer polítiques proactives de cerca de fons documen-*

tals produïts per dones i/o que donin testimoni de les activitats de les activitats de les dones a la nostra societat"⁴⁹. Novament, una crida a dissenyar polítiques d'adquisició de fons amb perspectiva de gènere per reduir el biaix de representació actual.

Per acabar aquest breu recorregut, cal destacar el projecte #Cuéntalo⁵⁰, impulsat des de l'Associació d'Arxivers i Gestors de Documents de Catalunya, pel que fa a la recuperació de les dades, i treballat en col·laboració amb *Barcelona Supercomputing Center* per al tractament estadístic, la categorització i visualització d'aquestes. Es tracta d'un projecte que dona visibilitat a les experiències de dones víctimes de violència masclista, unes vivències que elles mateixes van explicar a la xarxa social Twitter, sota el *hashtag* que dona nom al projecte, durant les dues setmanes posteriors a fer-se pública la sentència de 'La manada' l'abril de 2018. Es recopilaren els tuits i totes aquelles dades que els contextualitzen, en formats JSON, per poder fer posteriorment el seu tractament. És un projecte innovador i ambiciós que posa sobre la taula qüestions a debatre al voltant del tractament de fons d'origen social, debat encara incipient i que pot afectar molt directament a la memòria de minories o col·lectius infra-representats a la nostra societat com és la dona, especialment des del punt de vista de la interseccionalitat.

Un exemple internacional de bones pràctiques: el *Register* australià

En l'àmbit internacional és interessant destacar el projecte *The Australian Women's Register*⁵¹, una base de dades en línia que enllaça amb les històries de les dones australianes i que incorpora informació dels seus documents, publicacions i altres recursos. A 25 de juny de 2019, segons la informació referenciada a la pàgina d'inici del web, conté un total de “7017 entries with references to 4250 archival resources, 8868 published resources and 1626 digital resources”.

Es tracta d'un cercador en el qual es poden buscar dones i les seves organitzacions, i incorpora enllaços a arxius i biblioteques que custodien els seus documents, així com altra informació com publicacions i recursos digitals. Permet fer cerques per diferents *inputs*: cerca lliure, alfabètica o per classificació funcional, els quals són intuïtius i estan basats en les cerques que faria qualsevol usuari a Google o altres cercadors.

La part més interessant d'aquest projecte és que, d'una banda, a partir d'aquesta base de dades es van poder detectar forats de representació que s'han anat emplenant de manera proactiva a partir de la creació de nous documents, generalment a partir de projectes d'història oral. I d'una altra banda que, arxivísticament parlant, és un projecte innovador, en tant que no té responsabilitat de custòdia de documents, però sí de la conservació de la informació continguda en aquests. És un clar exemple

de la importància de l'automatització de la recopilació de dades OAI-PMH de registres segons EAD (*Encoded Archival Description*). Henningham, Evans i Morgan (2017) afirmen que la recopilació de dades de forma automatitzada fa possible augmentar la productivitat de tots els agents implicats en la creació, gestió i ús del material per a investigació històrica, de tal manera que no només s'afavoreix la recopilació en si, sinó el fet que les dades obtingudes siguin de qualitat, estructurades i reutilitzables. Es facilita l'ús, la recerca i la recuperació d'informació.

3. 2. Abast del projecte

Per tot el que s'ha exposat fins al moment, aquest projecte vol ser un model de bones pràctiques per a l'obertura de dades amb perspectiva de gènere, presentant un model mínim de recollida de dades que permeti fer difusió i donar a conèixer els fons documentals dels arxius. Així, s'analitzaran els instruments de descripció per tal de visibilitzar la dona com a productora, així com donar resposta a la manca de representació, en tota la seva diversitat d'expressions, actituds i interessos, en els fons custodiats pels arxius a Catalunya.

Així doncs, l'inici del projecte ha consistit a definir l'abast d'aquest. Per fer-ho, en primer lloc s'ha delimitat territorialment. En aquest cas, l'àmbit geogràfic d'actuació seran els arxius de Catalunya i, en principi, sense fer distincions segons si l'arxiu forma part del Sistema d'Arxius de Catalu-

nya (SAC) o no, en tant que és força probable que hi hagi una part important de fons de dones que no es trobin en arxius públics o inclosos al SAC. En segon lloc, s'ha valorat especialment quins arxius i quines tipologies de fons contenen dades de major interès per a l'estudi de cas proposat.

La selecció final té la voluntat de demostrar les constatacions plantejades a la hipòtesi de treball. De la mateixa manera, per un costat, es persegueix posar sobre la taula la falta de representació de la dona als arxius i, per un altre costat, es vol evidenciar la dificultat de tractament que suposa l'estat actual de les dades per a la seva obertura, i per poder garantir la seva reutilització.

Les dades que s'han utilitzat són aquelles que actualment estan disponibles a Internet, però com es veurà, els formats i la qualitat d'aquestes, no sempre són aptes per a la reutilització, i encara menys es poden obrir de forma pràctica i ràpida. Hi ha maneres més o menys senzilles d'automatitzar alguns dels processos, però la desagregació de les dades és complexa. L'actualització dels *datasets* generats per al prototip d'obertura que aquí es presenta no és possible, ja que les dades s'han treballat en local i, per tant, els resultats són estàtics.

Tenint present que l'objecte de l'estudi de cas és absolutament transversal, i que les possibilitats i estratègies que presenta són diverses, s'ha acotat molt la recerca. Per sobre de tot, s'ha enfocat d'acord amb termes quantitatius i

qualitatius de les dades que es podien obtenir. Així doncs, queda obert a futures intervencions l'ampliació, adaptació i millora dels resultats obtinguts.

3. 2. 1. Els arxius

Un cop definit l'abast territorial, s'ha procedit a delimitar quins arxius incorporar en l'anàlisi i l'execució del projecte. S'ha partit des de l'ideal d'actuació en aquest territori, i a continuació s'ha acotat fins a una situació realista, possibilista i viable per a fer una proposta adequada.

Així doncs, idealment s'haurien treballat tots els serveis d'arxiu que conformen el SAC, juntament amb altres centres i equipaments, com poden ser biblioteques, centres de documentació i associacions feministes, susceptibles de custodiar fons del nostre interès.

En un primer moment, es fa una recerca genèrica per tenir una visió global de quins centres poden tenir fons documentals de dones. Aquesta primera exploració es fa a les webs i cercadors de: la Xarxa d'Arxius Comarcals (XAC), l'Arxiu Nacional de Catalunya i els Arxius Històrics Provincials de Girona, Tarragona i Lleida; els arxius de les diputacions provincials; els arxius universitaris; alguns arxius municipals⁵²; algunes associacions feministes, amb el Centre de Documentació Ca la Dona com a entitat més emblemàtica; institucions públiques com el Centre de Documentació Joaquina Alemany i Roca de l'Institut Català de la Dona

i l'Institut de Cultura de la Dona (Espai Bonnemaison); la Biblioteca Nacional de Catalunya; i arxius diocesans i capitulars de l'Església Catòlica, buscant concretament els arxius de congregacions religioses femenines.

Els resultats d'aquesta primera recerca han evidenciat, en primer lloc, la dificultat per disposar fàcilment dels instruments de descripció en formats oberts i reutilitzables. La majoria d'arxius donen accés i difonen els seus inventaris i altres instruments de descripció, principalment en HTML o PDF, així com amb estructures i formats ben diversos. Només l'Arxiu Municipal de Girona ofereix dades en CSV, tal com ja s'ha exposat en el marc teòric. La realitat d'aquesta situació és que es dificulta moltíssim l'explotació de dades, i n'impossibilita la seva obertura per a la reutilització si prèviament no hi ha un tractament exhaustiu. Tal com apunta Yakel (2005, 118) *“accessibilitat no és sinònim d'accés”, i “mentre l'accés als arxius ha crescut de manera exponencial, l'accessibilitat ha quedat enrere”*.

En segon lloc, s'ha vist que els arxius i centres que custodien documentació de potencial interès per aquest projecte són diversos. Això implica que el tractament descriptiu que han rebut els fons, segons l'organisme que els conserva, és diferent i, per tant, els recursos resultants són heterogenis. No només pel que fa a si han estat descrits o no, sinó al fet que no tots els centres (especialment els que no són pròpiament arxius), no acostumen a utilitzar les normes de descripció arxivística. Com a exemple d'això

que s'acaba d'exposar es pot esmentar la Biblioteca Nacional de Catalunya (BNC), un organisme que conserva una quantitat important de fons documentals, entre els quals podem trobar fons de dones. Només en els casos que es descriu el fons personal en el seu nivell superior, la descripció acostuma a seguir la *Norma de Descripció Arxivística de Catalunya* (NODAC). No és així a les unitats documentals dels exemples consultats. Aquestes acostumen a estar separades segons matèries i no totes elles disposen d'inventari publicat a Internet. A més, en tots els casos valorats, les guies de fons estan en format PDF i alguns inclús incorporen enllaços dins el mateix document que fan encara més complexa l'exportació de les dades per a la seva reutilització posterior.

En tercer lloc, la complexitat de recerca que suposa conèixer i navegar pels portals web de cada un dels arxius, centres i equipaments. La dona com a subjecte productor de documents és transversal a qualsevol temàtica, matèria o àmbit, de tal manera que la dispersió pot ser enorme. Si a aquesta casuística s'hi afegeix que els centres que tenen fons no sempre disposen de guies, inventaris o catàlegs dels mateixos en formats digitals, localitzar-los és, ara per ara, una tasca molt difícil que requereix temps i recursos. Aquest és un dels motius pels quals es considera interessant proposar un model de dades que permeti centralitzar les descripcions seguint un esquema de metadades mínimes, per tal de difondre els fons amb perspectiva de gènere. No només se'n facilitaria l'obertura i la reutilitza-

ció de dades, sinó que serviria com una mena de guia que permetria filtrar, agrupar i centralitzar els fons de dones de manera ràpida i senzilla, contribuint a facilitar la localització de fons i documents.

Analitzada la situació de partida, finalment es decideix fer la proposta a partir dels instruments de descripció de fons del XAC i l'ANC, en tant que arxius destacables del SAC, que tenen presència a tot el territori català i en els quals s'ha vist que seria possible identificar i recopilar un volum interessant de dades, tant quantitativament com qualitativament parlant, que permetin desenvolupar amb èxit la prova de concepte de concepte de dades proposada pel projecte.

Aquests arxius centralitzen les seves descripcions a cercador Arxius en Línia⁵³. Els arxius del XAC i l'ANC hi migren les descripcions des dels seus gestors documentals, GIAC i GANC respectivament. Es tracta d'un cercador de fons i documents dels arxius del Departament de Cultura de la Generalitat de Catalunya, i alguns municipals com, per exemple, el de Castelló d'Empúries, Palafrugell o Sant Feliu de Guíxols. El cercador permet fer cerca simple, avançada o per inventari, i retorna la documentació independentment del centre d'arxiu on es custodii, si no s'hi aplica el filtratge per centre. Permet veure els resultats de la cerca a través d'un llistat, o en una galeria d'imatges, sempre que aquestes hi estiguin associades, seguint les regles i principis de la descripció multinivell.

3. 2. 2. Els fons

Una vegada definits quins arxius treballar en aquest projecte, s'ha procedit a seleccionar els fons a analitzar i incorporar a la proposta final.

D'aquesta manera, d'entrada s'ha valorat quines de les catorze agrupacions que conformen el Quadre de fons de la Generalitat de Catalunya (Departament de Cultura, 1994a, 1994b), poden contenir documentació en la producció del qual la dona hi tingui un paper actiu.

Dit quadre de fons contempla els següents tipus: (1) Fons de la Generalitat de Catalunya, (2) Fons de l'Administració Local, (3) Fons de l'Administració perifèrica de l'Estat, (4) Fons de l'Administració Reial i Senyorial, (5) Fons notariais, (6) Fons judicials, (7) Fons registrals, (8) Fons d'institucions, (9) Fons religiosos, (10) Fons d'associacions i fundacions, (11) Fons comercials i d'empreses, (12) Fons patrimonials i familiars, (13) Fons personals i (14) Col·leccions.

En un primer moment, en tant que la seva funció és vetllar per la ciutadania i no tenen definició de gènere, es decideix descartar tots aquells fons que són produïts per organismes públics perquè, d'entrada es considera que el paper que pugui tenir-hi la dona com a individu és més aviat mínim. Si bé és cert que podríem tenir en compte aquelles institucions que es dediquen a la defensa de la igualtat i l'erradicació de les violències masclistes, estari-

em anant més enllà de l'objectiu principal del projecte i, per tant, no s'han inclòs.

De la mateixa manera, tot i que es considera que seria interessant analitzar els fons religiosos de congregacions femenines, es proposa descartar-los, en tant que l'abast del projecte és limitat, i per tenir una idea del volum i contingut d'aquests fons seria més interessant fer un mapatge i recerca directament a arxius eclesiàstics, no únicament a Arxius en Línia. Així doncs, en un primer moment es descarten els fons dels tipus 1 al 9 del quadre, i es proposa fer una primera recerca de fons d'associacions i fundacions, comercials i d'empreses, patrimonials i familiars, personals i col·leccions, per tal de tenir una visió global de què es pot trobar en cada una de les tipologies.

El resultat més destacable d'aquesta primera aproximació és que, amb l'única i evident excepció dels fons personals, els quals generalment adopten el nom de la persona que els ha produït, és pràcticament impossible identificar els fons en els quals la dona pot haver-hi tingut un paper clau en la seva producció. Per fer-ho, cal entrar a la fitxa de cada fons i revisar si a la *Història del(s) productor(s)* s'identifica amb noms i cognoms la persona que ha fundat, presidit o dirigit l'organització -quelcom que no sempre és així-, o si les activitats que porta a terme estan directament relacionades amb la defensa dels drets de la dona o similars, per tal de poder valorar la seva adequació a l'estudi de cas.

El concepte de productor

En tant que punt bàsic d'accés, cal apuntar algunes qüestions respecte a la identificació de fons i el concepte de productor.

Segons la NODAC, un *fons* és un “conjunt de documents, de qualsevol tipologia o suport, produït orgànicament i/o reunit i utilitzat per una persona, una família o un organisme en l'exercici d'activitats i funcions que li són pròpies” i, per tant, és el resultat natural de qui l'ha produït. Pel que fa al concepte de *productor*, es defineixen com a “qualsevol entitat (institució, família o persona) que ha produït, reunit i/o conservat els documents en el desenvolupament de la seva activitat personal o institucional”.

Es tracta doncs, de dos conceptes indestruïbles directament lligats al principi de provinença i, en conseqüència, als subprincipis de respecte a l'estructura i l'ordre original dels fons. És sota aquests principis que s'articula tota la disciplina arxivística de l'actualitat. La NODAC defineix el principi de provinença com la “relació entre els documents d'arxiu i les organitzacions o les persones que els han produït, reunit i/o conservat i utilitzat en l'exercici de la seva activitat personal o corporativa”. La idea general és que no s'han de barrejar, de cap manera, documents que provenguin de productors diferents. Així doncs, per identificar i delimitar un fons primer és necessari identificar el seu productor. El significat teòric d'aquests principis doncs, *a priori*, sembla senzill. Malgrat tot, la seva aplicació pràctica resulta no ser

tan simple i és per aquest motiu que des de la disciplina arxivística s'ha teoritzat incansablement aquest concepte.

En relació amb els fons d'organitzacions, i únicament a tall d'exemple, es destaca el text de Duchein (1983) *Theoretical Principles and Practical Problems of Respect des fonds in Archival Science*⁵⁴, on s'exposen algunes de les principals dificultats d'aplicació pràctica d'aquest principi, i on es proposen unes solucions que dues dècades després adoptarà també l'ANC en la identificació dels seus fons de l'administració pública. El mètode que segueixen per identificar el productor és definir el seu grau d'autonomia seguint uns criteris concrets. Tant Duchein (1983) com Bernal i Duran (2001) enfoquen els seus treballs en l'aplicació del principi de provenença a fons el productor dels quals és una persona jurídica o organisme públic, en tant que els criteris que adopten per determinar aquest grau d'autonomia es basen en qüestions jurídiques o legals.

Quan el productor és una persona física o jurídica privada, la limitació dels fons esdevé molt més complexa. Un fons comercial o empresarial contindrà majoritàriament documentació comptable i relativa a la seva activitat econòmica. El fons d'una associació o fundació, encara que es regeixi per legislació diferent de la de les empreses, la documentació que contindrà serà bastant similar a l'empresarial. I, tant en un cas com en l'altre, les activitats que poden desenvolupar són molt diverses. Ara bé, especialment en els fons d'associacions, però no només, es donen

casos en què la documentació ingressada no és única-ment relativa a les activitats de l'associació, sinó que conté documents que formarien part del fons personal d'un particular, el qual acostuma a tenir relació directa amb l'entitat que desenvolupa l'activitat. En aquests supòsits, es fa necessari distingir clarament els dos productors i no sempre és feina senzilla.

Si no es procedeix a delimitar i separar els dos fons i es manté el nom de l'associació, fundació o empresa, com a punt clau d'accés i recuperació de la informació, correm el risc d'estar amagant la part de documentació personal. El nom del productor té un pes tan elevat a l'hora de localitzar la informació que, encara que les descripcions incloquin anotacions referents a la part personal dels fons, la seva recuperació serà complexa.

Aquest fet és encara més evident en els fons patrimonials i familiars, els quals es defineixen per contenir *documentació produïda per la família o llinatge i que són el resultat de les activitats personals i familiars dels seus membres, de l'administració dels seus béns i de l'exercici de les seves potestats jurisdiccionals*. Malgrat tot, en molts casos es desdibuixa el límit entre aquests fons i els personals, aquests últims definits com *aquells que contenen documentació generada per l'activitat pública, professional i associativa d'una persona, i en destaca especialment la seva obra creativa*.

A partir d'aquesta revisió es pot intuir que, molt probablement, hi haurà una part important de documentació que hagi estat produïda per una dona que podrà estar inclosa en altres tipus fons, especialment familiars, d'associacions o, fins i tot, en fons personals dels seus marits. Malgrat tot, s'entén que es trobarà representada majoritàriament en els seus propis fons personals. Per aquest motiu, i en tant que el projecte busca visibilitzar les seves experiències, s'ha fet una breu revisió de bibliografia que teoritza al voltant de l'aplicació dels principis de provenença i de respecte a l'ordre originals, concretament en aquests tipus de fons. Amb això, a més, s'evidencia que l'aplicació d'aquests principis no és tan clara i definida i ha generat força debat.

En aquest àmbit, autores com Jennifer Douglas (2018), Catherin Hobbs (2001) o Jennifer Meehan (2010) han enfocat els seus estudis en l'anàlisi de fons personals i, en termes generals, han reclamat més atenció al concepte de productor d'aquests fons. D'entrada, Hobbs (2001, 128) destaca que aquests fons necessiten un enfocament d'avaluació diferent que aquells que contenen documents administratius o governamentals. Justifica perquè els fons personals, igual que els d'organitzacions, documenten les seves activitats com a individus, però en tant que aquestes no són fruit d'una llei, un estatut o política, s'ha de tenir present que *“what is also recorded is the personal, the idiosyncratic, the singular views of people as they go about doing the things that they do and commenting on them”*. Els fons personals contenen el punt de vista més subjectiu

de les experiències dels individus, no només contenen les transaccions i evidències de les activitats que hagin portat a terme, sinó *“the whole experience of writing itself”*, una part molt íntima que no té un origen orgànic.

Meehan (2010) considera que la gran limitació del principi de provenença i respecte a l'ordre original dels fons, és que no especifica què fer en els casos en què la documentació és inconsistent, de tal manera que és pràcticament impossible discernir l'ordre original, quelcom que succeeix habitualment en fons personals. Inclús afirma que en molts casos, en aquests tipus de fons, si seguim el principi de provenença, no podem arribar a suposar que les relacions entre documents siguin realment significatives i aportin informació del context de la producció dels documents. Proposa repensar el concepte per fer-lo flexible i permetre'n l'estudi i anàlisi de les múltiples relacions possibles entre documents i activitats, de tal manera que el resultat sigui una aproximació real al context de producció del fons.

Per la seva banda, Douglas (2018) suggereix prestar més atenció al rol actiu del productor en la creació del seu propi fons. Entén que el concepte de producció no és una mera acumulació de documents i que, per tant, el productor no és passiu. Especialment en els fons personals, on la producció orgànica és difosa i no respon a processos ni procediments. L'autora arriba a contemplar la idea que la producció documental no és cosa d'un únic productor, sinó que també qui custòdia els documents i els arxivers

tenim un paper actiu en la seva conformació, fins i tot els usuaris cada vegada que en fan ús.

En una altra línia, afirma que es reconeix i s'accepta que els documents a l'interior d'un fons poden tenir autors diferents del seu productor. Així i tot, és més habitual, seguint el principi de provinença, remarcar el rol principal i la perspectiva del productor del fons. Essent això així, és possible que si la descripció no es desenvolupa suficientment i no se'n fa un ús correcte de descriptors, el rol dels autors no serà reconegut i quedaran ocults.

Es tracta d'un aspecte importantíssim per a l'estudi de cas que es planteja en aquest projecte. Un fons destacable per exemplificar el que apunta Douglas, i que molt probablement és excepcional, el trobem en el fons *Consultorio para la mujer Elena Francis*⁵⁵ conservat a l'Arxiu Comarcal del Baix Llobregat. Un fons empresarial que conté un volum important de correspondència escrita majoritàriament per dones. Aquí les dones no són les productores del fons, sinó les autores de les cartes. En tant que es tracten temes personals i sovint íntims, aquestes s'han anonimitzat i es pot accedir a les digitalitzacions a la sala de consulta de l'arxiu. En aquest cas, en realitat el més important no és conèixer el nom de l'autor/a de la carta, sinó contextualitzar el seu contingut i posar-lo a disposició amb dades desagregades per gènere, edat, classe social, motiu de consulta, i fins i tot, de la resposta obtinguda, en els casos que s'ha conservat. Això permetria tenir una idea molt més

rica del context de creació de la documentació, una visió més clara de les motivacions i preocupacions d'aquestes dones, així com de les relacions establertes entre autors i productor del fons. Lamentablement però, encara que aquí la dona no sigui estrictament la productora, si el tractament no es fa amb una mirada més àmplia, es desvirtua la rellevància que poden tenir aquestes cartes per a la investigació amb perspectiva de gènere del context concret de creació a tots els nivells de descripció.

Una opinió similar l'expressen Henningham, Evans i Morgan (2017, 104), en tant que consideren que els instruments de descripció arxivístics són excessivament rígids, doncs expressen allò que es coneix en el moment de fer el tractament, on s'assigna un únic i singular productor i un conjunt limitat de relacions. Les autores consideren que *"these archival processes maintain the silencing of women whose records remain undiscoverable in the collections of other"*.

Així doncs, i per tot el que s'acaba d'exposar, els fons que finalment s'incorporaran al projecte són fons personals i col·leccions i, si és possible la seva identificació, s'inclouran també fons familiars i patrimonials. Aquesta decisió es basa en dos arguments: d'una banda, la idoneïtat d'incloure la perspectiva de gènere segons la naturalesa del productor, i d'una altra banda, la idiosincràsia de la documentació generada en cadascun dels tipus de fons. Com hem vist, quan parlem de fons el productor dels quals és una persona jurídica, si bé és cert que en alguns casos po-

dria ser interessant tenir present el gènere de la persona hagi fundat o presidit l'organisme, es considera que no són els que millor s'adeqüen als objectius d'aquest projecte. Igualment, la documentació generada en fons generalment fruit d'activitats administratives, llevat de casos concrets, no aportaran context pel que fa a la producció documental en clau de gènere.

3. 2. 3. Els usuaris potencials

La principal característica de les dades obertes és que han de ser fàcilment interpretables i usables per qualsevol persona. Tot i això, tot producte té el seu públic objectiu i donades les característiques de les dades que es treballen en aquest projecte, es considera necessari elaborar un breu perfil del *target* al qual es dirigeix. Tot plegat, ha de servir per poder definir i consignar amb major rigor els camps a obrir, quelcom que és primordial i essencial per a millorar l'accessibilitat.

A continuació es defineix doncs, no el perfil dels usuaris actuals dels arxius, sinó el perfil bàsic del principal públic potencial al qual es vol dirigir el model de dades, i que pot estar interessat en la reutilització de la informació per a diferents qüestions com la investigació o l'anàlisi de dades. Es caracteritza principalment per:

- Dona
- Nascuda i/o resident a Catalunya

- Té interès per investigar i conèixer sobre la història de les dones
- És conscient i té sensibilitat vers les desigualtats socials i de gènere
- Nivell d'ofimàtica mitjà-baix
- Sap navegar per Internet i disposa de connexió a la xarxa
- És capaç d'interpretar conjunts de dades estructurades i senzilles
- Forma part o participa esporàdicament en associacions o grups de dones

Ara bé, cal tenir present que no és l'únic públic al qual es dirigeix el projecte. Hi ha un altre públic potencial a valorar: els mateixos serveis d'arxiu, o centres de diversa tipologia, que conserven documentació personal produïda per dones a Catalunya.

En aquest cas, per a l'usuari el més important no és tant la reutilització de dades, com fer-se més visible i obtenir reconeixement per la tasca que desenvolupa dia a dia. D'aquesta manera, l'anàlisi i valoració del funcionament del model per a cada organisme a nivell intern, ha de servir per replantejar la forma i qualitat en què s'ofereixen les dades actualment -en cas que es faci- i analitzar la seva idoneïtat i les possibilitats d'implantació.

D'aquesta manera, com més serveis d'arxiu utilitzin el model de dades, més es facilitarà no només la interoperabilitat, sinó també l'actualització de la informació, de tal

manera que el projecte de centralització dels instruments de descripció serà més potent i interessant, i augmentarà l'ús per part dels usuaris.

4. Disseny del *dataset*

L'objectiu principal d'aquest projecte és crear un model de dades que permeti centralitzar els instruments de descripció dels fons patrimonials i familiars, personals i col·leccions dels arxius de Catalunya. Més concretament, el prototip que es presentarà busca visibilitzar la infra-representació de la dona en els mateixos arxius amb un abast concret, ja definit al llarg de l'apartat anterior. A més, els conjunts de dades han de permetre assolir altres objectius com, per exemple, oferir una visió global de la realitat arxivística actual i veure el nivell de tractament dels fons o els volums de fons de dones per arxius.

No es tracta doncs únicament d'elaborar una taula amb dades descriptives dels diferents fons i pujar-la a un portal de dades. Vol ser una proposta i un model de recol·lecció, de tal manera que l'obertura dels inventaris i catàlegs es plantegi de forma desagregada i estructurada des de l'inici, facilitant l'encreuament i la reutilització.

Es plantejaran quines són les metadades mínimes a recollir, principalment d'acord amb la informació de què es disposa actualment i als objectius que es plantegen. En aquest projecte es busca treballar específicament fons on la dona és la protagonista, però analitzar dades en clau de gènere implica que necessàriament s'han de fer comparacions. Per tant, el model s'ha estructurat pensant en la

globalitat dels tipus de fons treballats i haurà de complir una sèrie de requisits, els quals es detallen a continuació:

En primer lloc, ha de respectar la descripció multinivell i ha de tenir present les normes i estàndards internacionals de descripció. Ara bé, per tal que els instruments siguin reutilitzables, interoperables i explotables, es considera que és necessari agafar una mica de distància i deixar de buscar les extenses descripcions a les quals es tendeix actualment -encara que sense pretendre deixar-les de banda completament-, i desagregar les dades en camps d'informació més petits, fàcilment compressibles i usables. D'aquesta manera se'ls podrà dotar de nous valors i noves dimensions d'ús.

En segon lloc, ha de contenir informació geogràfica, principalment dels arxius. Igualment, ha de contemplar la possibilitat d'incloure'n també tant a nivell de fons, d'unitats documentals com per als productors, amb la finalitat d'ampliar al màxim les opcions de visualització a partir de mapes.

En tercer lloc, els *datasets* publicats han de tenir una estructura senzilla, estar en formats oberts i lliures de llicències, per tal que els usuaris puguin utilitzar-los sense restriccions.

Finalment, es considera que, donat que les dades que s'obrin no presentaran tota la informació, ni ho faran de la mateixa manera que les descripcions de les quals es par-teix, en aquest cas seria interessant tenir una relació ràpi-da entre les dades obertes i les fitxes de descripció actuals disponibles a Arxius en Línia per tal de no perdre ni desvir-tuar el context. Tot i això, cal tenir present que incorporar un enllaç comporta certs riscos, doncs és una dada amb altes probabilitats de quedar obsoleta, amb la qual cosa es reduiria la qualitat del *dataset* i podria motivar que usu-aris potencials no l'utilitzin.

4. 1. Recopilació de les dades

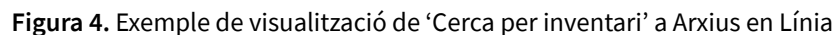
Una vegada definit sobre quins centres d'arxiu i quins tipus de fons centrar la recerca, aquesta s'ha fet des del buscador d'Arxius en Línia. Tal com s'ha esmentat, aquest cercador centralitza totes les descripcions de fons donats d'alta tant del XAC com de l'ANC i, per tant, per-met fer tota la recerca necessària per acomplir amb l'abast d'aquest projecte.

Des d'Arxius en Línia es dona accés als instruments de des-cripció dels fons, i a les digitalitzacions de les unitats do-cumentals, sempre que hi estiguin associades. Tot i la seva utilitat en aquest sentit, les possibilitats de cerca que ofe-reix són limitades. Tal com apunta Terradellas Prat (2016, 63) algunes de les mancances destacables del cercador, i que podrien esdevenir objectius de millora d'aquest, són

que: no disposa d'un espai personal per guardar les cer-ques, filtrar i ordenar resultats; no permet exportar ni des-carregar els resultats en cap format obert ni reutilitzable; tampoc ofereix la possibilitat de compartir les cerques a xarxes socials o a través del correu electrònic.

Per identificar els fons a incorporar en aquest projecte, el procediment ha estat el següent:

- A la pàgina d'inici del cercador d'Arxius en Línia anem a 'Cerca per inventari'.
- Seleccionem *documentació textual* i *documentació no textual* per tal que ens retorni fons que continguin qualsevol classe de documentació.
- Seleccionem el tipus de fons que volem consultar.



mirant, un per un, el nom del fons. Tal com s'ha predit a la breu reflexió feta al voltant del concepte de productor com a punt d'accés bàsic a la informació, per identificar els fons patrimonials i familiars es fa necessari entrar, una a una, a cada fitxa descriptiva. És una tasca que no s'ha fet en aquest projecte, però es considera interessant apuntar-ho, a fi i efecte que es tingui present aquesta casuística, i es puguin començar a plantejar alternatives que siguin més inclusives quant a gènere.

Paral·lelament, en un full de càlcul s'han recollit els codis de referència, juntament amb el nom i el tipus de fons de cada cas seleccionat.

4. 1. 1. Sol·licitud de les dades en formats reutilitzables

Des d'un bon principi s'ha considerat inviable obtenir la informació manualment des del web donat l'esforç que això suposa i que, fer-ho així, a dia d'avui no representa cap proposta de valor. Així doncs, el pas següent ha estat sol·licitar les dades en un format reutilitzable. Tant l'ANC com la Subdirecció General d'Arxius i Museus han mostrat interès pel projecte i han col·laborat facilitant l'exportació de dades des dels seus respectius centres d'arxiu⁵⁶.

Des de l'ANC s'ha rebut l'exportació de dades de 44 guies de fons, unitats documentals de 23 fons textuais i d'1 fons no textual. S'han rebut en un únic fitxer que agrupa, en 4 taules: guies de fons, unitats documentals textuais, unitats documentals no textuais amb descriptors i sense. Les dades s'han rebut tant en format ACCDB com XLSX.

Des de la Subdirecció d'Arxius s'han rebut un total de 66 guies de fons, unitats documentals de 27 fons i dades descriptives de les unitats d'instal·lació de 10 fons. En aquest cas, les dades de cada fons s'han rebut de manera individual, fent un total de 103 fitxers. En format CSV les guies de fons, i XLS les unitats documentals i d'instal·lació.

L'estructura de les dades facilitades per l'ANC i la Subdirecció d'Arxius no és exactament igual, però presenta similituds destacables. En primer lloc, pel que fa a les guies de fons, els elements comuns que s'han obtingut són: nom de l'arxiu, codi de l'arxiu, codi del fons, nom del fons, codi del tipus de fons, nom del tipus del fons, les dates de creació, descripció, sí l'accés és lliure (L) o restringit (R), si s'ha inventariat (S, N o P) i els metres lineals totals de cada fons. L'exportació de dades de fons del XAC conté molts més elements, la majoria referents a diferents suports i formats.

En segon lloc, pel que fa a les descripcions d'unitats documentals, les dades comunes són el codi d'arxiu, codi del fons, codi de la unitat documental, títol i descripció de la unitat, cronologia i codi de nivell. Interessant destacar el camp <desc_nproced>⁵⁷ en el qual s'informa el nom de la persona que estaria directament vinculada a la producció de la unitat documental que es descriu. Només està emplenat en els casos en què el fons l'ha produït més d'un individu.

Finalment, l'estructura i camps de les taules que descriuen els fons del XAC a nivell d'unitat d'instal·lació, és molt similar a la descripció d'unitats documentals però, donat que no és l'objectiu presentar dades a aquest nivell, s'ha descartat el seu tractament.

4. 2. Tractament de les dades

4. 2. 1. Anàlisi de l'estat inicial de les dades

Una vegada obtingudes les dades, aquestes s'han tractat i preparat per a la seva publicació.

Inicialment les tres principals dificultats que s'han trobat són:

- La descripció pròpiament arxivística. Tant a nivell de fons com d'unitat documental, tots els elements descriptius s'unifiquen en una sola cel·la en fer l'exportació⁵⁸. Això implica que necessàriament s'ha de fer un treball previ per separar cada element en una columna diferent.
- La completesa de les descripcions dels exemples analitzats és molt heterogènia. Així, mentre una bona part dels fons donats d'alta a Arxius en Línia únicament tenen consignats el codi de referència (codis d'arxiu i fons), el nom del fons i les dates extremes, altres són molt més completes. Més enllà del nivell de tractament que hagi rebut cada fons, és lògic que la completesa dels elements descriptius sigui heterogènia, ja que la NODAC defineix camps obligatoris (OB), recomanables (RE) i opcionals (op), de tal manera que no necessàriament s'emplenen els mateixos en cada cas.

- Tot i que se segueix la descripció multinivell en tots els casos, les normes de descripció no defineixen els patrons de codificació per informar cada camp. Això implica que, per tal de desagregar els elements i obtenir dades que puguin ser fàcilment explotables i útils en la seva reutilització, cal una tasca prèvia que, ara per ara, esdevé complexa.

4. 2. 2. Primer tractament

El primer tractament s'ha fet per tal de tenir les dades estructurades segons els elements descriptius NODAC, en tant que és el model que se segueix a Arxius en Línia per mostrar les descripcions, però mantenint el contingut exactament com s'ha rebut. Això ha permès tenir una visió global de quins elements s'han informat i com estan emplenats. El segon tractament, descrit a l'apartat 4.3.1. *Segon tractament*, s'ha fet una vegada decidits quins elements incorporar a la prova de concepte de publicació.

Així doncs, en aquest estadi inicial s'han tractat únicament les guies de fons. Donat que l'estructura i elements que s'han rebut són diferents, per un costat s'han treballat les de fons custodiats per l'ANC i, per un altre costat, les de fons que es troben a arxius comarcals. Aquestes últimes s'han unificat prèviament en un sol fitxer, ja que com s'ha apuntat anteriorment, cada guia s'ha rebut individualment en un fitxer CSV.

Fet aquest pas previ, el procediment ha estat el mateix en els dos casos. S'ha aïllat el camp <descripció> (sempre vinculat al codi d'arxiu i el codi de fons per no perdre la relació), el qual conté tota la informació descriptiva de la unitat a la qual es dona accés des d'Arxius en Línia. Aquest camp s'ha separat segons les etiquetes, les quals fan referència als diferents elements NODAC informats en cada cas.

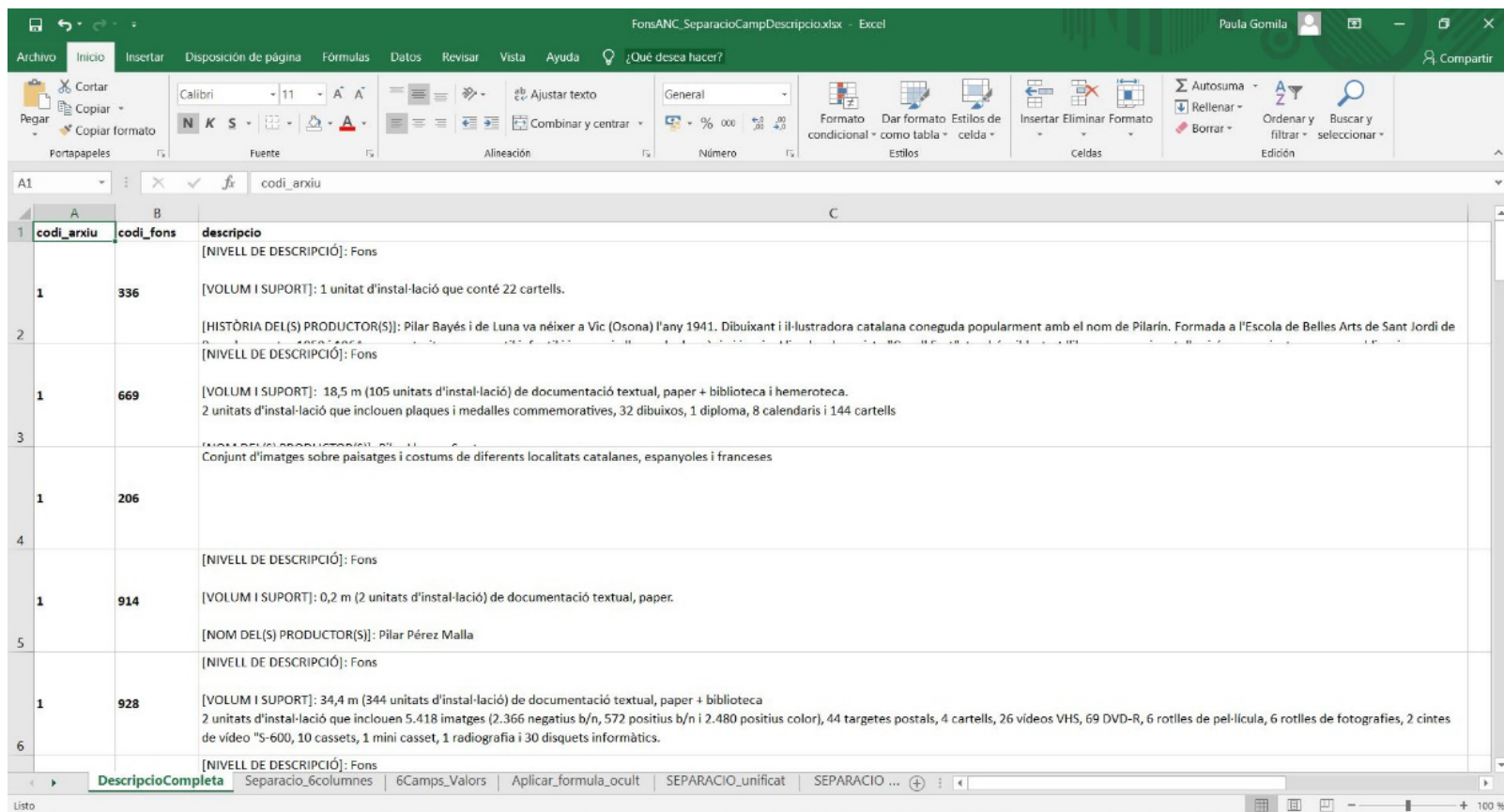
Amb la voluntat d'automatitzar, en la mesura del possible, la separació dels elements, s'ha utilitzat la fórmula següent:

```
=ESPACIOS(EXTRAE(SUSTITUIR($C2;"$";REPETIR  
(" ";LARGO($C2)));1+(COLUMNAS ($D2:D2)*  
LARGO($C2))-LARGO($C2);LARGO($C2)))
```

Per poder aplicar-ho s'han trobat algunes dificultats, especialment a les guies de l'ANC:

- Etiquetes escrites de manera diversa. No només el nom de l'etiqueta, sinó que en alguns casos, aquesta no es troba escrita entre claudàtors sinó parèntesis. Això no només dificulta el procés de separació, sinó que s'ha de tenir present que, en fer la migració a Arxius en Línia, si no s'han escrit les etiquetes correctament, la informació es concatena a l'etiqueta immediatament superior. Així doncs, ja que passa en pocs exemples, tot i que no és la millor opció, la solució adoptada ha estat normalitzar manualment les etiquetes.

El resultat d'aquest primer tractament ofereix una visió global de la informació disponible i dona l'oportunitat d'analitzar amb més facilitat el contingut de cadascun dels elements. A les següents figures es pot veure l'abans i el després d'aquesta primera intervenció:



	A	B	C
	codi_arxiu	codi_fons	descripció
1	1	336	[NIVELL DE DESCRIPCIÓ]: Fons [VOLUM I SUPORT]: 1 unitat d'instal·lació que conté 22 cartells.
2	1	669	[HISTÒRIA DEL(S) PRODUCTOR(S)]: Pilar Bayés i de Luna va néixer a Vic (Osona) l'any 1941. Dibuixant i il·lustradora catalana coneguda popularment amb el nom de Pilarín. Formada a l'Escola de Belles Arts de Sant Jordi de [NIVELL DE DESCRIPCIÓ]: Fons [VOLUM I SUPORT]: 18,5 m (105 unitats d'instal·lació) de documentació textual, paper + biblioteca i hemeroteca. 2 unitats d'instal·lació que inclouen plaques i medalles commemoratives, 32 dibuixos, 1 diploma, 8 calendaris i 144 cartells
3	1	206	[NOM DEL(S) PRODUCTOR(S)]: Conjunt d'imatges sobre paisatges i costums de diferents localitats catalanes, espanyoles i franceses
4	1	914	[NIVELL DE DESCRIPCIÓ]: Fons [VOLUM I SUPORT]: 0,2 m (2 unitats d'instal·lació) de documentació textual, paper.
5	1	928	[NOM DEL(S) PRODUCTOR(S)]: Pilar Pérez Malla [NIVELL DE DESCRIPCIÓ]: Fons [VOLUM I SUPORT]: 34,4 m (344 unitats d'instal·lació) de documentació textual, paper + biblioteca 2 unitats d'instal·lació que inclouen 5.418 imatges (2.366 negatius b/n, 572 positius b/n i 2.480 positius color), 4 targetes postals, 4 cartells, 26 vídeos VHS, 69 DVD-R, 6 rotlles de pel·lícula, 6 rotlles de fotografies, 2 cintes de vídeo "S-600, 10 cassetts, 1 mini casset, 1 radiografia i 30 disquets informàtics.
6			[NIVELL DE DESCRIPCIÓ]: Fons

Figura 5. Visualització del camp <descripció> abans del tractament inicial, amb els elements descriptius unificats en una única cel·la.

FonsANC_SeparacioCampDescripicio.xlsx - Excel Paula Gomila

Archivo Inicio Insertar Disposición de página Fórmulas Datos Revisar Vista Ayuda ¿Qué desea hacer? Compartir

Portapapeles Fuente Alineación Número Estilos de celdas Celdas Edición

	A	B	C	D	E	F	G	H	I	J	K
	codi_arxiu	codi_fons	NIVEL	VOLUM I SUPORT	NOM DEL(S) PRODUCTOR(S)	HISTÒRIA DEL(S) PRODUCTOR(S)	HISTÒRIA ARXIVÍSTICA	DADES SOBRE L'INGR	ABAST I CONTINGUT	SISTEMA D'ORGANITZACIÓ	CONDICIONS D'ACCÉS
1	1	336	Fons	1 unitat d'instal·lació	Pilar Bayés i de Luna va néixer a Vic	Pilar Bayés i de Luna va néixer a Vic	El fons fou conservat pel	La documentació del	Obra impresa entre la qual	Cartells registrats i en procés de	Accés lliure.
2	1	669	Fons	18,5 m (105 unitats)	Pilar Llorens Souto	Ballarina, mestra, traductora, gestora,	El fons fou conservat pel	La documentació del	El fons aplega la	S'han identificat les següents sèries:	Accés lliure, amb excepció
3	1	914	Fons	0,2 m (2 unitats)	Pilar Pérez Malla	Pilar Pérez Malla va néixer a Barcelona	Cal creure que la	El fons va ingressar a	El fons aplega la	S'han identificat les sèries següents:	Accés lliure, amb excepció
4	1	928	Fons	34,4 m (344 unitats)	Lidia Falcón O'Neill	Lidia Falcón O'Neill va néixer a Madrid	El fons fou conservat per	El fons va ingressar a	El fons aplega la	En aplicació del quadre de	Accés lliure, amb excepció
5	1	950	Fons	0,1 m (1 unitat)	Juan Orfila Barrueco / Teresa	Juan Orfila Barrueco va néixer a	Es desconeixen les	El fons va ingressar a	El fons aplega la	La documentació dels dos productors	Lliure.
6	1	901	Fons	0,9 m (3 unitats)	Manuel Miserachs i Rigalt i	Manuel Miserachs i Rigalt va néixer a	La major part del fons va	El fons va ingressar a	El fons aplega la	La documentació dels dos productors	Accés restringit. Adreça
7	1	993	Fons	71.000 Imatges (54.881	Pilar Aymerich i Pulg	Pilar Aymerich i Pulg va néixer a	El fons ha restat a mans	Ingressat a l'Arxiu	Temàticament la producció de	L'arxiu pròpiament dit, està	Accés lliure a còpies (digi
8	1	1007	Fons	0,1 m (1 unitat)	Nan Green	Nan Green, nascuda Nancy Farrow, va	El fons fou conservat per	El fons va ingressar a	El fons és constituït	S'han identificat les sèries següents:	Accés lliure.
9	1	1061	Fons	0,1 m (1 unitat)	Ana Mingorance Pérez	Ana Mingorance Pérez, coneguda també	El fons fou conservat per	El fons va ingressar a	El fons és constituït	D'acord amb el quadre de	Accés lliure.
10	1	1075	Fons	0,1 m (1 unitat)	Alejandra Danilevski i Julia i Elena	Alejandra Danilevski, d'origen rus o	El fons fou conservat per	El fons va ingressar a	El fons és constituït	D'acord amb el quadre estàndard de	Accés lliure, amb excepció
11	1	1080	Fons	0,1 m (1 unitat)	Araceli Ruiz Toribios	Araceli Ruiz Toribios va néixer a	El fons fou conservat per	El fons ingressà a l'ANC,	El fons és constituït	D'acord amb el quadre estàndard de	Accés lliure, amb excepció
12	1	1105	Fons	33,8 m (338 unitats)	Marta Mata i Garriga.	Marta Mata i Garriga va néixer a	El fons fou conservat per	El fons va ingressar a	El fons és constituït	S'han identificat les sèries següents,	Accés lliure, amb excepció
13	1	1106	Fons	1,1 m (11 unitats)	Maria dels Àngels Garriga Martín	Àngels Garriga Martín va néixer el 1898	El fons fou conservat pel	El fons va ingressar a	El fons és constituït	S'han identificat les sèries següents:	Accés lliure, amb excepció
14	1	984	Fons	0,3 m (3 unitats)	Manuel Gil Prieto - Esperanza	Manuel Gil Prieto (Calataiud, 1927 -	El fons fou conservat pel	El fons va ingressar a	El fons és constituït	D'acord amb el quadre estàndard de	Accés lliure, amb excepció
15	1	1131	Fons	0,8 m (1 unitat)	Miquel Victorià Amer i Omar i	Miquel Victorià Amer i Omar va néixer a	La col·lecció fou	La col·lecció va ingressar	La col·lecció és constituïda	La documentació es troba ordenada	Accés lliure.
16	1	761	Fons	0,4 m (4 unitats)	Carmen López Landa	Carmen López Landa va néixer a Madrid	El fons ha estat conservat	El fons va ingressar a	El fons aplega la	S'han identificat les següents sèries:	Accés lliure, amb excepció
17	1	963	Fons	0,1 m (1 unitat)	Francesc Ribó i Rovira i Mariana	Francesc Ribó i Rovira va néixer a	El fons ha estat conservat	El fons va ingressar a	El fons aplega la	La documentació dels dos productors	Accés lliure, amb excepció
18	1	1148	Fons	0,1 m (1 unitat)	Juan Barceló Luqué i Dolores Cabra	Juan Barceló Luqué va néixer a Madrid	El fons ha estat conservat	El fons ingressà a l'ANC,	El fons és constituït	La col·lecció conté la documentació	Accés lliure, amb excepció
19	1	1149	Fons	0,1 m (1 unitat)	Frida Knight	Frida Knight, nascuda Frideswide	El fons fou conservat per	El fons va ingressar a	El fons és constituït	D'acord amb el quadre estàndard de	Accés lliure.
20	1	1170	Fons	1,4 m (14 unitats)	Maria Estrada i Clerch	Destacada assistent social en l'àmbit	El fons ha estat conservat	El fons va ingressar a	El fons conté la documentació	S'han identificat les sèries següents:	Accés lliure, amb excepció
21	1	389	Fons	13,6m (88 unitats)	Montserrat Roig i Fransitorra	Montserrat Roig i Fransitorra, escriptora	El fons ha estat conservat	El fons va ingressar a	El fons conté la documentació	En l'organització del fons s'han	Accés lliure excepte en a
22	1	29	Fons	18 unitats d'instal·lació	De les germanes Rosa, Maria,		El fons ha estat conservat	El fons va ingressar a	El fons conté documentació	La documentació s'ha organitzat en 6	Accés lliure
23	1	102	Fons	9,7 m (97 unitats)	Blanca, Eva i Josep Serra i Pulg	Fills de l'arqueòleg Josep de Calasanz	El fons ha estat conservat	El fons va ingressar a	El fons conté la documentació	En l'organització del fons s'ha	Accés restringit. Adreça
24	1	772	Fons	0,1 m (1 unitat)	Vicenta García Moreno	Vicenta García Moreno va néixer a	El fons fou conservat per	El fons va ingressar a	El fons integra la	S'ha identificat una única sèrie	Accés lliure, amb excepció
25	1	788	Fons	16 unitats d'instal·lació	Isabel Steva Hernández (Colita)	Nascuda a Barcelona el 24 d'agost de	El fons ha estat conservat	Adquirit per	El fons està integrat pels	Fons en tractament.	Accés restringit. Adreça
26	1	190	Fons	0,3 m (3 unitats)	Anna Maria Maluquer i Wahl	Anna Maria Maluquer i Wahl va néixer a	Es desconeixen les	El fons va ingressar a	El fons aplega una part de la	S'han identificat les sèries següents:	Accés lliure.
27	1	428	Fons	26 unitats d'instal·lació	Montserrat Sagarra i Zacarini	Montserrat Sagarra i Zacarini neix al	El fons Montserrat Sagarra	El fons ingressà a l'Arxiu	Material que l'autora va	Classificat per comarques i dins	Accessible la part del for

Separacio_6columnes 6Camps_Valors Aplicar_formula_ocult SEPARACIO_unificat SEPARACIO_unificat_distribuit

Figura 6. Visualització del camp <descripció> després del tractament inicial, amb els elements descriptius separats i distribuïts.

Igualment, ha permès veure ràpidament el tractament a nivell de fons que han rebut els 96 casos treballats, així com quins elements de descripció s'informen més habitualment, quins no i, en conseqüència, quins poden ser més adients per a incorporar en un projecte d'obertura de dades.

Així doncs, després d'aquesta primera revisió veiem que, d'entrada tenim un total de 77 fons personals, 17 col·leccions i 2 fons patrimonials i familiars. Hi ha 23 fons que no disposen de cap mena de contingut al camp <descripció>⁵⁹, és a dir, a Arxius en Línia estan donats d'alta i tenen codi de referència i nom però, en realitat, no han rebut tractament o, si l'han rebut, no s'ha publicat. Uns altres 8 fons tampoc tenen informat aquest camp amb etiquetes seguint l'ús de la NODAC, però sí que presenten una mínima referència descriptiva molt general. En aquests casos, s'ha optat per disposar dita informació a l'element *Abast i contingut*.

Pel que fa als elements que segons la NODAC és obligatori consignar, dels 96 fons que s'han treballat, el 100% tenen el *Codi de referència* i el *Títol* del fons, 82 fons informen sobre la(es) *Data(es) extremes* (85,4%), 67 sobre *Abast i Contingut* (69,8%), 62 sobre el *Volum i suport* (64,6%), 59 sobre la *Història del(s) productor(s)* (61,5%), 58 sobre l'element *Dades sobre l'Ingrés* (60,4%), 56 sobre el *Nom del(s) productor(s)* i sobre les *Condicions d'accés* (58,3%) i 53 la *Història arxivística* (55,2%).

Aquestes dades són sobre el total dels 96 fons, és a dir, tant procedents de l'ANC com dels arxius comarcals. Malgrat tot, 43 dels fons són conservats per l'ANC i, en general, s'ha vist que el percentatge d'elements descrits als seus fons és bastant més elevat que en el conjunt d'arxius del XAC (53 fons treballats en total). S'ha considerat interessant evidenciar-ho, perquè a simple vista es pot veure que el tractament descriptiu que han rebut els fons treballats del XAC és bastant més reduït que en el cas dels fons de l'ANC.

Així, a la **Figura 7** es pot veure el total d'elements NODAC informats per al conjunt de fons de l'ANC (barres de color blau) i el conjunt de fons del XAC (barres de color groc). Amb una línia de color gris que representa el percentatge total d'elements informats, sense fer la distinció segons el custodi dels fons.

A l'**Annex 1** es pot consultar la taula amb les dades, en percentatges i en números absoluts, que s'acaben de comentar, i a partir de les quals s'ha elaborat el gràfic que es presenta a continuació:

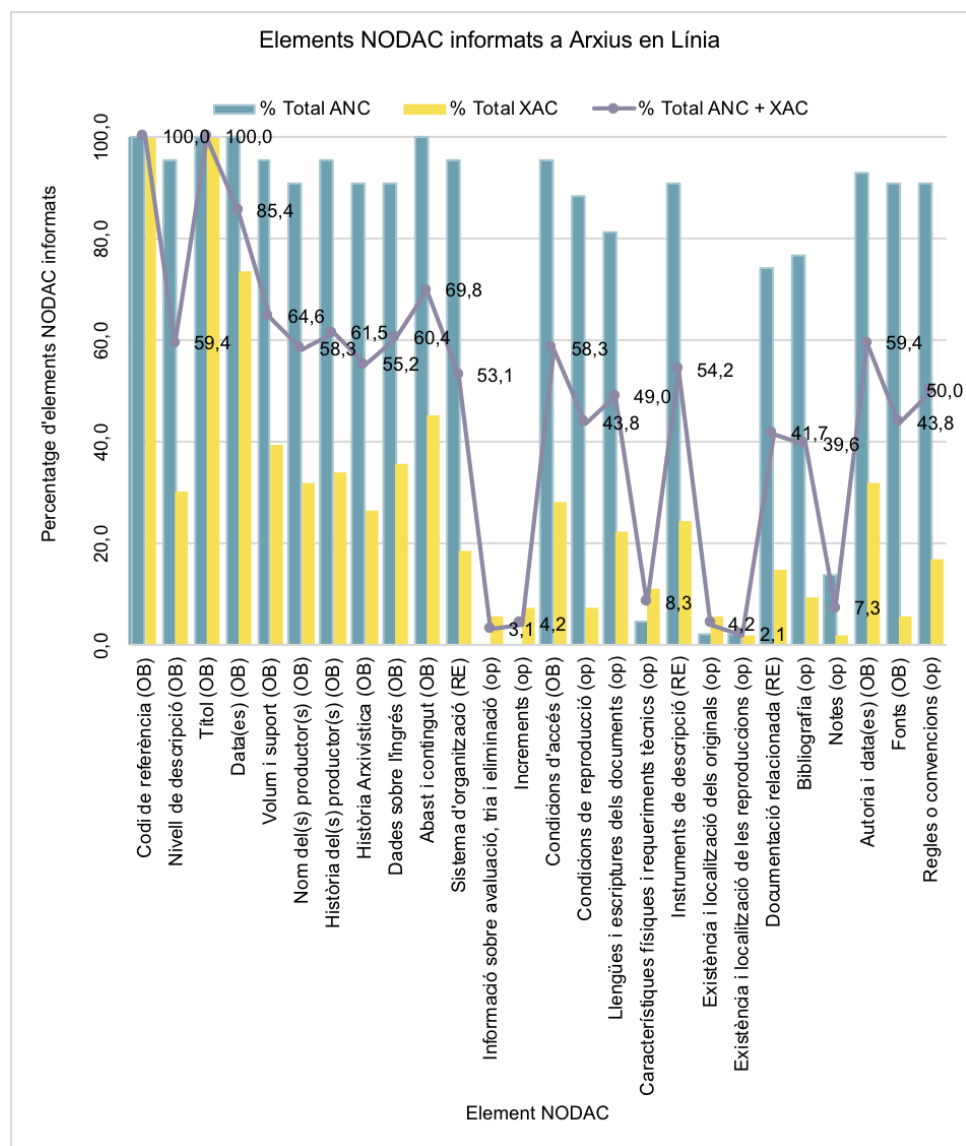


Figura 7. Gràfic que mostra el percentatge d'elements NODAC que s'informen i mostren, del total de 96 fons treballats, a les guies de fons disponibles a Arxius en Línia. Es distingeix entre el total de fons de l'ANC (43) i el total de fons del XAC (53). La línia de color gris mostra el percentatge total dels 96 fons treballats. (Font: Elaboració pròpia).

4. 3. Selecció dels elements de descripció

Una vegada analitzades les dades disponibles i tenint presents els objectius que es volen aconseguir, s'ha procedit a fer la selecció dels elements de descripció. S'ha de buscar l'equilibri entre allò que es vol mostrar i les dades de què realment es disposa, de tal manera que l'obertura sigui realista, i el resultat funcional, lògic i coherent. Així doncs, amb la voluntat de ser pragmàtics, es planteja un model de dades en què una part important dels elements, no tots, es poden obtenir fàcilment.

Com s'ha vist fins al moment, en aquest primer tractament de dades, el resultat obtingut ja es podria utilitzar per fer una proposta d'obertura. Fer-ho directament amb aquestes dades té un avantatge interessant, i és que la informació presentada és fidel al contingut actual dels instruments de descripció disponibles a Arxius en Línia. Tot i això, també presenta certs inconvenients, doncs cal tenir present que la qualitat de les dades que contenen la majoria dels elements ofereix unes possibilitats de reutilització bastant reduïdes. A més, no s'aporta valor més enllà de tenir la informació estructurada i poder presentar-la en taules.

Amb tot, i per tal de possibilitar nous usos, donada l'heterogeneïtat de les dades actuals i l'objectiu de representació en clau de gènere, es proposa la desagregació d'alguns dels elements, fet que requereix esforç i suposa certa complexitat.

Cal remarcar que el model que es proposa no pretén substituir els instruments de descripció actuals, sinó ser-ne un complement i una aposta de valor, tant per donar visibilitat a la dona com a subjecte productor, com per avançar i mostrar un vessant diferent dels arxius i mirar de subvertir la imatge que es té de la professió des de fora⁶⁰.

És necessari fonamentar la proposta en els principis i les normes de descripció arxivística, per tal que el resultat sigui rigorós. Així, seguirà les regles de descripció multinivell que estableixen la ISAD(G) i la NODAC, per garantir que es manté el context i les relacions entre les unitats de descripció, així com la recuperació de la informació i la possibilitat d'unificar i creuar conjunts de dades. D'aquesta manera, s'han seleccionat i analitzat els camps i àrees que la NODAC considera obligatoris (OB) i recomanats (RE), i s'hi han afegit altres elements necessaris tant per incloure la perspectiva de gènere, com per facilitar la geolocalització a diferents nivells. De la mateixa manera, s'han descartat alguns elements descriptius que no aporten informació rellevant per a l'usuari extern.

A continuació es mostra la selecció final, amb la definició de cada element i la seva consignació:

Element	Definició	Consignació
Nom arxiu	Nom complet del servei d'arxiu	OB
Acrònim arxiu	Acrònim del servei d'arxiu	OB
Codi arxiu	Codi que identifica inequívocament el servei d'arxiu	OB
Municipi arxiu	Municipi on es troba físicament el servei d'arxiu	RE
Nom tipus fons	Identifica el tipus de fons segons el Quadre de fons	OB
Codi tipus fons	Codi que identifica el tipus de fons segons el Quadre de fons	OB
Classe fons	Informa la classe genèrica del fons (<i>textual o no textual</i>)	OB
Nivell de descripció	Identifica el nivell d'organització al qual correspon la unitat que es descriu	OB
Codi fons	Codi que identifica unívocament el fons	OB
Codi unitat	Codi que identifica unívocament la unitat documental	OB
Nom fons	Títol que dona nom al fons	OB
Nom unitat	Títol que dona nom a la unitat documental	OB
Data(es)	Dates d'inici i fi de la unitat de descripció. Inclou referència a dates aproximades i altres especificitats	Op
Data inici	Data d'inici de la unitat de descripció	OB
Data fi	Data de fi de la unitat de descripció	OB
Volum (ml)	Metres lineals que ocupa el fons	OB
Nom del/la productor/a	Identifica el/la productor/a del fons ⁶¹	OB
Gènere del/la productor/a	Informa el gènere de la persona física que ha produït, en la seva totalitat o en part, el fons documental	OB
Municipi naixement	Municipi on ha nascut la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE
Municipi mort	Municipi on ha mort la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op
Any naixement	Any de naixement de la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE
Any mort	Any de la mort de la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op
Activitat professional	Enumera les diferents activitats professionals realitzades per la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE

Vincles entitats	Enumera els principals organismes amb els quals s'ha vinculat la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op
Dipositant	Nom de qui ha formalitzat l'ingrés de la documentació al servei l'arxiu	OB
Gènere dipositant	Gènere del dipositant (Dona, Home, Mixt ⁶² o Organisme)	OB
Data ingrés	Data del primer ingrés efectiu	OB
Modalitat ingrés	Modalitat segons la qual s'ha formalitzat l'ingrés de la documentació	OB
Abast i contingut	Descripció general de la unitat de descripció	OB
Condicions d'accés	Condicció d'accés a la documentació del fons (Lliure o Restringit)	OB
Inventariat	Informa sobre el nivell d'inventariat del fons (Sí, No o Parcial)	OB
Instruments de descripció	Enllaç a la fitxa descriptiva d'Arxius en Línia	OB
Documentació relacionada	Fons relacionats o complementaris al qual es descriu	RE
Latitud arxiu	Latitud geogràfica del servei d'arxiu per poder geolocalitzar-lo	RE
Longitud arxiu	Longitud geogràfica del servei d'arxiu per poder geolocalitzar-lo	RE

Taula 2. Selecció dels elements de descripció per a l'obertura de dades de Fons patrimonials i familiars, Fons personals i Col·leccions, amb la corresponent definició i consignació. (Font: Elaboració pròpia)

A. ÀREA IDENTIFICACIÓ

Es mantenen tots els elements NODAC d'aquesta àrea, tant a nivell de fons com d'unitat documental, a excepció del volum que únicament s'informarà en el nivell superior.

L'element *Codi de referència* es treballa desagregat en diferents metadades: *Acrònim de l'arxiu*, *Municipi de l'arxiu*, *Codi de tipus de fons*, *Nom de tipus de fons*, *Codi de l'arxiu*, *Codi de fons* i *Codi d'unitat documental*, a incorporar en els nivells de descripció corresponents. Això permetrà que s'heretin en nivells inferiors, es puguin fer filtres i recuperar informació a diferents nivells. El municipi de l'arxiu es justifica pensant en el futur del model de dades, de tal manera que es

puguin incloure fons que no es trobin en arxius del SAC. Els arxius que no formen part del SAC no tenen un codi numèric assignat i, seguint la NODAC, la manera de referenciar-los és a partir del municipi on es troba. A més, aquesta dada, tant si l'arxiu forma part del SAC com si no, en facilita informació geogràfica útil de cara a la georeferenciació.

L'element *Títol* fa referència al nom de la unitat de descripció. En tant que se seguiran les regles multinivell, les unitats documentals heretaran la relació del títol del nivell superior per tal de no perdre el context, d'aquesta manera s'han assignat dues metadades, *nom fons* i *nom unitat*, doncs són els dos nivells que es treballen en aquesta proposta.

L'element *Data(es)* es manté. A més, a nivell de fons s'han afegit els camps, *data inici* i *data fi*, els quals s'expressarà en anys i informaran les dates extremes de creació. Això té la voluntat de fer que el contingut de l'element sigui més versàtil, manipulable i reutilitzable.

L'element *Volum i suport*, en aquest cas, únicament informarà dels metres lineals al nivell de fons per tal d'oferir una idea general del seu embalum total⁶³. A les exportacions rebudes, tant del XAC com de l'ANC, hi ha un camp en què es facilita aquesta informació.

S'han analitzat models de publicació de dades com el que utilitza l'Arxiu Municipal de Girona per obrir els seus catàlegs⁶⁴, i es considera que seria molt interessant valorar l'aplicació de l'esquema que proposen. Malgrat tot, són dades que suposaria una complexitat enorme d'extrapolar per aquest projecte a partir de la informació que hi ha a l'element *Volum i suport* de les descripcions treballades aquí.

Sí que es tenen dades que identifiquen de manera genèrica si es tracta de fons textuais (T) o no textuais (I), i d'aquesta informació sí que se n'ha contemplat la seva obertura. Al camp que contindrà aquestes dades se li ha donat el nom de *Classe fons*.

B. ÀREA DE CONTEXT

Els elements de l'àrea de context s'informaran a nivell de fons. Únicament en els casos en què s'ha identificat més

d'un individu, amb nom i cognoms, com a productor d'un fons, i s'ha referenciat a nivell d'unitat documental, s'inclourà també al nivell inferior. Es tracta del camp *<codi_nproced>* el qual, ara per ara, només es té disponible per als fons de l'ANC. Seria interessant poder identificar amb un codi numèric tots els productors com a persones físiques, de tal manera que sigui més senzill establir relacions de procedència⁶⁵.

L'element *Nom del(s) productor(s)* es divideix en dos. Per tal d'incloure la perspectiva de gènere es fa necessari incloure una metadada que informi sobre el gènere de qui ha produït la documentació. Així, un camp farà referència al *nom del productor* i l'altre al *gènere del productor*. S'ha vist que, en una part important de les descripcions, no es disposa de l'element *Nom dels(s) productor(s)*, però en canvi, a tots hi ha el nom del fons (element *Títol*). Els casos que es té dades als dos camps, s'ha comprovat que únicament presenta diferències ortotipogràfiques com, per exemple, l'ús de guió alt en un cas i la conjunció 'i' en l'altre. D'aquesta manera, s'ha depurat la informació disponible al camp *nom del fons*, per obtenir tots els noms de les productores dels fons.

Els elements *Història del(s) productor(s)* i *Dades sobre l'ingrés*, es divideixen en diferents camps cadascun, per tal de desagregar el màxim possible aquesta informació i traduir-la a dades que siguin comprensibles i reutilitzables fàcilment. Així, de la *Història del(s) productor(s)* s'extrau-

ran les *dates* (en anys) de *naixement i mort*, el *municipi de naixement i mort*⁶⁶, la principal *activitat professional* portada a terme per la productora i els principals *vincles amb entitats* (de qualsevol mena) amb les quals s'ha relacionat al llarg de la seva trajectòria vital. De les *Dades sobre l'ingrés*, es treballarà el *dipositant*, el *gènere del dipositant*, la *modalitat d'ingrés* i la *data d'ingrés*. Val a dir que, ara per ara, ambdós elements requereixen un esforç considerable per a la seva desagregació, i en molts de casos no es disposa d'informació de tots els camps⁶⁷.

L'element *Història arxivística* no s'inclourà a la proposta, ja que la informació que conté és molt tècnica i no aporta informació substancial per a l'usuari.

C. ÀREA DE CONTINGUT I ESTRUCTURA

Únicament es manté l'element *Abast i contingut*. En aquest cas, el contingut es presentarà exactament com s'ha rebut, tant a nivell de fons com d'unitat documental. Si bé és cert, a les unitats documentals, aquest element acostuma a contenir informació relativa a: topònim, llengua, tècnica i volum. En gairebé tots els casos hi consta, almenys, el volum. És un element útil a nivell de fons, en tant que aporta informació general sobre el seu contingut. No és tan rellevant a nivell d'unitat documental, aportant informació més substancial per a l'usuari l'element *nom unitat*.

Donat el caràcter eminentment tècnic dels elements *Sistema d'organització*, *Informació sobre avaluació*, *tria i eli-*

minació i *Incrementos*, cap d'ells s'incorporarà al model de dades. A més, pel que fa als dos últims, doncs no s'han informat pràcticament en cap cas i, a més, els fons treballats són de conservació permanent.

D. ÀREA DE CONDICIONS D'ACCÉS I ÚS

Els elements d'aquesta àrea únicament s'emplenaran a nivell de fons.

Es manté l'element *Condicions d'accés*. Tant a les dades facilitades per l'ANC com per la Subdirecció General d'Arxius, hi ha una columna on s'informa sobre el règim d'accés amb una L o una R, en funció de si aquest és lliure o restringit. Al model de dades s'utilitzarà la informació d'aquesta columna. Si bé és cert, s'ha pogut comprovar que en un parell de casos, la sigla d'aquesta columna no concorda amb el que hi ha referenciat al camp <descripció> i, per tant, a la fitxa d'Arxius en Línia.

En tant que es vol mantenir la relació dels diferents recursos descriptius, es considera important mantenir el vincle amb la fitxa d'Arxius en Línia, el qual s'incorporarà a l'element *Instruments de descripció*. La voluntat no és tant facilitar la reutilització de dades o crear visualitzacions i mapes, sinó més aviat no perdre el context. Igualment, busca que l'usuari pugui fer les comprovacions i comparacions que cregui pertinents. Com ja s'ha apuntat anteriorment, cal tenir present que és una dada que fàcilment podrà quedar obsoleta si no se'n fa el manteniment adequat.

Relacionat amb aquest element, i donat que a l'exportació facilitada es disposa d'una columna en la qual s'informa de si el fons ha estat inventariat o no, en la seva totalitat o en part, amb les lletres S, N o P, es proposa utilitzar i obrir aquesta informació amb el mateix nom en què s'ha rebut, és a dir, *Inventariat*.

E. ÀREA DE DOCUMENTACIÓ RELACIONADA

Es manté l'element *Documentació relacionada* i únicament s'emplenarà a nivell de fons, ja que no es disposa de dades en nivells inferiors.

No s'inclouran els elements *Existència i localització de reproduccions*, *Existència i localització dels originals* i *Bibliografia*. Es considera que seria interessant incloure un camp en el qual s'informi sobre si la unitat documental disposa d'imatges digitals associades o no. És una dada que molt probablement tenen informada a nivell intern els diferents centres d'arxiu, però no s'ha rebut en l'exportació que han aportat per materialitzar el projecte.

F. ÀREA DE NOTES I ÀREA DE CONTROL DE LA DESCRIPCIÓ

No s'inclouran al conjunt de dades donat que no aporten informació rellevant per a l'usuari.

G. ELEMENTS DE GEOREFERENCIACIÓ

S'incorporen dues metadades més que són necessàries per a la geolocalització dels arxius: la *latitud* i la *longitud*. Donat que no es disposa de l'adreça normalitzada, i que

no es considera que sigui d'imperiosa necessitat disposar de les dades de contacte i localització dels arxius, s'ha procedit a recopilar manualment les coordenades X i Y amb *Google Maps*.

A continuació es presenten dos exemples, un a nivell de fons i l'altre a nivell d'unitat documental, de com es tradueix aquest esquema en un cas concret:

Elements	Exemple: Fons
Nom arxiu	Arxiu Nacional de Catalunya
Acrònim arxiu	ANC
Codi arxiu	1
Municipi arxiu	Sant Cugat del Vallès
Nom tipus fons	Personals
Codi tipus fons	13
Classe fons	Textual
Nivell de descripció	Fons
Codi fons	901
Nom fons	Manuel Miserachs i Rigalt - Montserrat Ribalta i Escoda
Data(es)	Dates de creació: 1891 - 2009
Data inici	1891
Data fi	2009
Volum (ml)	1
Nom del/la productor/a	Montserrat Ribalta i Escoda
Gènere del/la productor/a	Dona
Municipi naixement	Barcelona
Municipi mort	Barcelona
Any naixement	1909
Any mort	1983
Activitat professional	Bibliotecària; Professora
Vincles entitats	Escola de Bibliotecàries de la Generalitat de Catalunya; Escola Normal de Barcelona; Acadèmia Miralles de Sabadell; Comissió de Cultura del Club Femení d'Esports
Dipositant	Antònia Miserachs i Ribalta
Gènere dipositant	Dona
Data ingrés	08/06/2009
Modalitat ingrés	Donació

Abast i contingut	El fons aplega la documentació produïda i rebuda per Manuel Miserachs i Rigalt en el desenvolupament de les seves activitats personals i professionals, així com per la seva esposa Montserrat Ribalta i Escoda, preferentment en l'àmbit perso-nal i creatiu. Pel que fa al primer, el fons inclou principalment documents genea-lògics i identificatius; documentació acadèmica dels estudis de batxillerat i medi-cina; documentació militar i relativa a la seva participació en la Guerra Civil espanyola i depuració política posterior; correspondència personal amb Montserrat Ribalta; historials professionals, activitat al Col·legi de Metges de Barcelona; activitat com a professor; activitat mèdica a l'Hospital de la Santa Creu i Sant Pau i altres centres; separates, guions i esborranys de conferències, intervenci-ons, comunicacions i articles mèdics; activitat al Sindicat de Metges de Catalu-nya i la Reial Acadèmia de Medicina de Barcelona, entre altres entitats; distincions i necrològiques; i, finalment, diverses publicacions alienes relatives als inicis de la transfusió de sang a Catalunya, aspecte que la documentació del fons permet documentar amb riquesa. Pel que fa a Montserrat Ribalta i Escoda, el fons aplega, fonamentalment, certi-ficats i documents identificatius; documentació acadèmica dels estudis de mestra, bibliotecària i infermera; dietari personal, notes per a diverses novel·les i arti-cles publicats a la revista "Cavall Fort". El fons inclou, així mateix, diversos testi-monis de la seva relació amb la poetessa Rosa Leveroni. METGES / TRANSFUSIÓ / BANCS DE SANG / HEMATOLOGIA / HEMOTERÀ-PIA / GUERRA CIVIL ESPANYOLA / EPISTOLARIS / CARTES DES DEL FRONT / DIETARIS / NOVEL·LA / BIBLIOTECÀRIES / MESTRES
Condicions d'accés	Lliure
Inventariat	No
Instruments de descripció	Fitxa Arxius en Línia: MANUEL MISERACHS I RIGALT - MONTSERRAT RIBALTA I ESCODA
Documentació relacionada	Joan Colomines i Puig (CAT ANC 1 - 0720) Documentació de Rosa Leveroni (CAT XXX BARCELONA BIBLIOTECA DE CA-TALUNYA Fons Carles Riba i Clementina Arderiu)
Latitud arxiu	41.465586
Longitud arxiu	2.069320

Taula 3. Exemple de descripció d'un fons personal segons l'esquema proposat en aquest projecte. (Font: Elaboració pròpia a partir de les dades descriptives del fons de l'ANC)

Com es pot veure a l'exemple de la **Taula 3**, es tracta d'un fons personal que presenta algunes de les característiques que s'han anat comentant fins ara. D'una banda, inclou documentació de dues persones, en aquest cas, un matrimoni. D'una altra banda, la descripció del fons, espe-

cialment l'element *Abast i contingut*, és poc respectuosa amb la perspectiva de gènere. Es diu que "*el fons aplega la documentació produïda i rebuda per Manuel Miserachs i Rigalt (...), així com per la seva esposa Montserrat Ribalta i Escoda, (...)*". Així doncs, per un costat, caldria conèixer

els motius pels quals s'ha considerat més adient mantenir unida la documentació per poder determinar amb més fonament si és o no una decisió encertada, si caldria separar-los o, si en lloc d'un fons personal seria més correcte identificar-lo com a fons familiar. Per un altre costat, in-

dependentment de la resposta anterior, si es vol fer referència al vincle entre els dos individus, caldria fer-ho de forma neutra per evitar donar, encara que sigui de manera inconscient, més valor o rellevància a un que a l'altre.

Elements	Exemple: unitat documental
Nom arxiu	Arxiu Comarcal de l'Alt Empordà
Acrònim arxiu	ACAE
Codi arxiu	110
Nivell de descripció	Unitat Documental
Cofi fons	58
Nom fons	Francesca Vidal Comalat
Codi unitat	138
Nom unitat	Esborrany d'un discurs sobre Guillem Comalat, oncle de Francesca Vidal Comalat.
Data(es)	1950
Abast i contingut	Manuscrit de sis fulls. Datació aproximada.

Taula 4. Exemple de descripció d'una unitat documental segons l'esquema proposat en aquest projecte. (Font: Elaboració pròpia a partir de les dades descriptives del fons de l'ACAE)

4. 3. 1. Segon tractament

Així doncs, per tal de poder presentar les dades tal com s'ha vist en els dos exemples anteriors, tot destacant la figura de la dona com a productora de fons, ha estat necessari fer un segon tractament, una vegada decidits els elements de descripció:

En aquest pas s'ha incidit, especialment, en tres punts:

- La desagregació manual de l'element *Dades sobre l'ingrés*
- La desagregació manual de l'element *Història del(s) productor(s)*
- La depuració i normalització de l'element *Data(es)*, amb la finalitat d'obtenir les metadades referents a les dates extremes de creació dels fons.

Així doncs, la desagregació dels dos primers elements s'ha fet manualment. S'han extret les dades del nostre interès i s'han disposat a la columna corresponent. Això ha permès veure que, encara que l'element estigui emplenat, en la majoria de casos no conté informació rellevant. Pel que fa a l'element *Data(es)*, està informat pràcticament a la totalitat dels casos. En els fons dels arxius comarcals disposem de quatre camps: *inici*, *fi*, *continuitat* i *cronologia*. En els fons de l'ANC únicament comptem amb el camp *cronologia*.

És interessant detenir-se mínimament en aquest element per exemplificar la complexitat que suposa l'ús exclusiu de la NODAC per al tractament digital de dades, ja que no

ofereix normalització pel que fa a la codificació del contingut de cadascun dels seus elements.

A nivell de fons trobem que la data s'informa en anys i, si és aproximada, s'expressa entre claudàtors, amb un *ca.* o un interrogant al final. Si parlem d'aquest camp, a nivell d'unitat documental, la diversitat de patrons és encara major. Així, trobem dates expressades en anys, en mesos i anys, o en dies, mesos i anys. En alguns casos, apareixen els mesos expressats amb lletres, no amb números. En el cas d'unitats que fan referència a un període curt en dies, acostumen a referenciar-se els dos dies separats per un guió alt, i tot seguit, se separa amb punts el mes i l'any.

A continuació es presenten exemples representatius de l'heterogeneïtat d'aquest element:

1930-1940	[1930-1940]	març 1940
[1930]-1940	12-14.03.1940	14-16 de febrer de 1982
1993 ca. 1930-1940	14.03.1940	juliol-agost 1993
1930-1940?	03.1940	finals juny 1978

Donada la gran diversitat de possibilitats que trobem, la solució que es proposa és:

- A nivell de fons. Per poder treballar les dades en mapes i gràfics, s'ha optat per eliminar qualsevol element que no sigui numèric. A les dades del XAC es tenen els anys

d'inici i fi per separat, no és així en els fons de l'ANC, pels quals ha estat necessari fer la separació i depurar qualsevol element no numèric.

- A nivell d'unitat documental. Mantenir, tal com s'ha obtingut, el camp *cronologia* per tal de no desvirtuar totes aquestes possibilitats, especialment quan es tracta de dates aproximades. La metadada passarà a dir-se *Data(es)* per tal de normalitzar-ho segons NODAC. Tot i això, de cara a futures intervencions, es proposa expressar la data en DD/MM/AAAA. I, per aquelles unitats que no disposen de tota la informació, utilitzar les franges extremes, per exemple:

	Data inici	Data fi
Juliol-agost 1993	01/07/1993	31/08/1993
1993	01/01/1993	31/12/1993

D'aquesta manera es desvirtua mínimament el contingut actual i es disposa d'unes dates expressades amb criteris unificats i usables.

4. 4. Els estàndards de descripció

Tots els elements del model de dades s'han escollit amb el focus posat en incorporar la perspectiva de gènere i en ampliar les possibilitats de visibilització i difusió dels fons. S'ha fet a partir de les dades disponibles i de les normes i estàndards de descripció, per tal de crear una pro-

posta homogènia i coherent que permeti la reutilització i la interoperabilitat.

Els elements relatius a la descripció de fons i unitats documentals corresponen, en termes generals, als esquemes NODAC i ISAD(G), encara que ha estat necessari desagregar el contingut d'alguns en diferents camps per assolir els objectius de gènere marcats. Aquestes normes, juntament amb la ISAAR (CPF), no defineixen l'estructura per a representar i mostrar les dades de tal manera que, en entorn Web, resulta limitada la seva aplicació, tal com afirma De la Puente (2018, 16) referint-se a la ISAD(G):

“el contenido de los elementos descriptivos que componen dicha estructura no está normalizado y ello limita las posibilidades de su aplicación en un entorno digital para poder garantizar una recuperación de la información archivística más eficiente y precisa, lo que conlleva también problemas en la representación de interrelaciones en la información descriptiva, sobre todo en la propiedad de herencia de propiedades entre nivel jerárquicos superiores e inferiores y en las posibilidades de navegación a lo largo de esta estructura, todo lo cual dificulta la indización y posterior recuperación de elementos particulares”.

Per aquest motiu, és interessant tenir presents els esquemes de codificació normalitzats EAD (*Encoded Archival Description*) i EAC (*Encoded Archival Context*)⁶⁸, en tant

que unifiquen la manera de representar les descripcions arxivístiques a Internet per tal de facilitar la consulta i la interoperabilitat. La idea principal d'EAD i EAC és codificar el llenguatge entès per màquines per interpretar-lo i retornar-lo en una visualització llegible per humans. Un llenguatge utilitzat per codificar els instruments de descripció, no el material d'arxiu pròpiament. Tal com afirma Moyano Collado (2013) es tracta d'estàndards que permeten codificar i normalitzar en metadades l'estructura dels elements descriptius de les entitats.

EAD i EAC són esquemes normalitzats que ofereixen la possibilitat de crear instruments de descripció que siguin independents de qualsevol maquinari i programari comercialitzat. Defineixen una estructura de dades normalitzades que reproduïen en format digital els instruments de descripció arxivística i que determinen, no només els tipus d'elements a utilitzar sinó el possible contingut de cadascun d'ells. Aquests estàndards, a més, utilitzen atributs que permeten vincular els elements EAD i EAC als seus corresponents d'ISAD(G) i ISAAR(CPF) respectivament, per tant, es millora l'accessibilitat i es facilita la creació de registres ISAD(G) i ISAAR(CPF) a partir d'EAD i EAC. En definitiva, estandarditzen les descripcions facilitant el seu ús mitjançant XML, un llenguatge de marques que permet estructurar la informació descriptiva d'un document textual. Tot plegat, facilita la interoperabilitat en tant que les dades estructurades i normalitzades podran ser utilitzades per altres aplicacions i plataformes Web.

Cal esmentar també Dublin Core, un model de metadades que utilitza XML i normalment està basat en RDF. L'esquema està definit a la ISO 15836 i defineix l'estructura i sintaxi, en un llenguatge d'esquema formal, del conjunt d'un total de 15 elements per descriure els objectes digitals. Conté elements ISAD(G). Les principals característiques d'aquest model són la seva simplicitat, la interoperabilitat, l'adopció a nivell mundial -ha estat traduït a vint llengües-, i la seva transversalitat amb altres esquemes de metadades.

4. 5. Estructura del model de dades

4. 5. 1. Selecció del model de dades

Una base de dades és una relació estructurada de dades emmagatzemades en un sistema informàtic. Hi ha diferents tipus de models, poden ser: jeràrquics, en xarxa, relacionals, basats en objectes o multidimensionals.

Per abordar aquest projecte s'ha escollit el model relacional, el qual és utilitzat habitualment per gestionar grans volums d'informació i és prou potent per a la seva recuperació. En aquest tipus de model, les dades s'estructuren a partir de *taules* o *entitats*. Cada taula pot tenir molts *registres* (files) i cada registre pot tenir múltiples *campes* (columnes).

El benefici més important que aporta aquest model al projecte és evitar duplicar dades per a poder recuperar la

informació, de tal manera que és més eficient, redueix la redundància i ocupa menys espai de memòria. En aquest cas, la informació s'ha estructurat en un conjunt de quatre taules principals, la descripció de les quals i la relació entre les mateixes s'explica a continuació. Es reflectirà també com aquestes relacions, en funció de l'anàlisi de les possibles casuístiques, podran variar a mesura que s'incorporin noves dades al *dataset*.

Un altre aspecte interessant d'aquest model de dades és que garanteixen la integritat de la informació, perquè si es vol eliminar un registre, igualment s'eliminen aquells que hi estan relacionats i que en són dependents.

4. 5. 2. Contingut de les taules

Fins ara s'ha fet l'anàlisi de les dades de què es disposa i s'ha definit què es vol obtenir per a obrir. S'ha establert l'esquema de metadades sobre el qual es treballarà i s'ha descrit el model de dades que s'utilitzarà per presentar-les.

El següent pas ha estat estructurar les dades en taules i crear les relacions necessàries per evitar duplicar la informació, facilitar la consulta i la recuperació de la informació de manera fàcil i ràpida. Així doncs, inicialment s'han plantejat les següents quatre taules⁶⁹:

Taula 1. Arxius. Inclou les dades geogràfiques dels arxius del XAC i de l'ANC a partir de les coordenades X i Y. Incorpora

també el nom del municipi on es troba cada arxiu, i els codis que identifiquen unívocament cada centre (nom, acrònim i codi d'arxiu⁷⁰). Seria interessant ampliar aquesta taula i incorporar-hi totes les dades d'adreça i contacte de cada arxiu. Ara bé, actualment no es troben en formats reutilitzables i s'ha considerat que no correspon als objectius d'aquest projecte recollir aquestes dades.

Taula 2. Fons. Aquesta taula incorpora la majoria de metadades referents a la descripció a nivell de fons que s'han considerat com a mínimes per a obrir els instruments de descripció. De la mateixa manera, aquelles que permeten dotar de major context i aporten informació sobre les característiques dels fons, com són el tipus, la classe i els diferents camps desagregats de l'element *Dades sobre l'ingrés*.

Taula 3. Productors(es). S'ha creat una taula per a les dades dels/les productors/es. En aquest projecte el que es vol és evidenciar el paper de la dona productora de fons, per tant, cal establir mecanismes que permetin definir el gènere del productor i consignar informació identificativa d'aquests.

Donat que no es treballen únicament fons personals, sinó que també s'han inclòs els fons familiars i patrimonials i les col·leccions, s'ha treballat el productor a partir de cadascuna de la/les persona/es físiques que han aplegat la documentació en l'exercici de les activitats que li són pròpies⁷¹. Així, per evitar duplicar la informació general

dels fons (la *Taula 2. Fons*) en els casos en què hi hagi més d'un individu productor, s'ha creat una altra taula que incorpora la informació sobre les persones físiques que han produït els fons. Les dades es presenten desagregades per gènere i informen sobre les dates i els municipis de naixement i mort, les principals activitats professionals i els vincles amb entitats i organismes.

Taula 4. Unitats documentals⁷². Inclou les dades principals que descriuen les unitats documentals, simples i compostes, de cadascun dels fons. Inclouen les dates, l'abast i contingut, el codi i el nom de la unitat.

4. 5. 3. Les relacions entre les taules

Per tal de fer efectiu el model relacional de dades cal definir les relacions, les quals poden ser de tres tipus: relacions d'un a un (1,1), relacions d'un a varis (1,n) i relacions de varis a varis (n,n).

- Les relacions (1,1) es donen quan un registre de la Taula 1 només pot tenir un registre coincident a la Taula 2, i viceversa. No és la més habitual.
- Les relacions (1,n) es donen quan un registre de la Taula 1 pot tenir múltiples coincidències a la Taula 2, però un registre de la Taula 2 només pot coincidir amb un registre de la Taula 1.

- Les relacions (n,n) es donen quan diversos registres de la Taula 1 tenen coincidència amb més d'un registre de la Taula 2, i viceversa. Són les més complexes de treballar.

En aquest treball les tres relacions principals que s'estableixen són (1,n). Així i tot, a continuació s'expliquen els canvis que es podrien donar en un futur, en el cas que s'incorporin noves dades al model. Així, les relacions entre les taules són les següents:

Relació Arxius – Fons (1,n). En els casos analitzats i treballats aquí, un arxiu està relacionat amb diferents fons, però aquests estan relacionats únicament amb un arxiu. D'aquesta manera, la relació que ara es proposa és d'un a varis. Tot i així, sabem que és possible que un mateix fons es trobi dispers en diferents arxius i, si això passa, a mesura que es vaguin incorporant noves dades, aquesta relació variarà i es farà necessari establir-la de varis a varis (n,n). Si es dona el cas, caldrà crear una nova taula que relacioni les taules Arxius i Fons, a través de les seves claus principals.

Relació Fons – Productor(es) (1,n). Com ja s'ha dit, en aquest treball el productor s'aborda des del punt de vista de la o les persones físiques que han aplegat la documentació. La relació establerta és d'un a varis, ja que en aquest cas, un productor genera un únic fons. Ara bé, tot i que és poc probable, s'ha de tenir en compte que podria

donar-se el cas, que en futures incorporacions d'informació, aquesta relació esdevingui de varis a varis.

Pensem, per exemple, en un fons familiar en què un dels seus membres tingui clarament definit, identificat i separat el seu fons personal. Si els dos fons s'acaben incorporant en aquesta base de dades, caldrà establir una relació de varis a varis per contemplar aquesta casuística. Aquesta situació es podria donar també si en un futur s'amplia l'abast i s'hi inclouen fons comercials i empresarials o d'associacions i fundacions, doncs serà més fàcil que una mateixa persona física pugui estar vinculada tant amb el seu fons personal com amb el fons de l'empresa que ha creat o l'associació a la qual ha estat vinculada.

En aquest estudi de cas, la majoria de fons treballats són personals i això no passa, de tal manera que la relació establerta és (1,n).

Relació Fons – UD (1,n). Aquesta relació és més senzilla. Un fons té múltiples unitats documentals i aquestes només poden estar relacionades amb un fons.

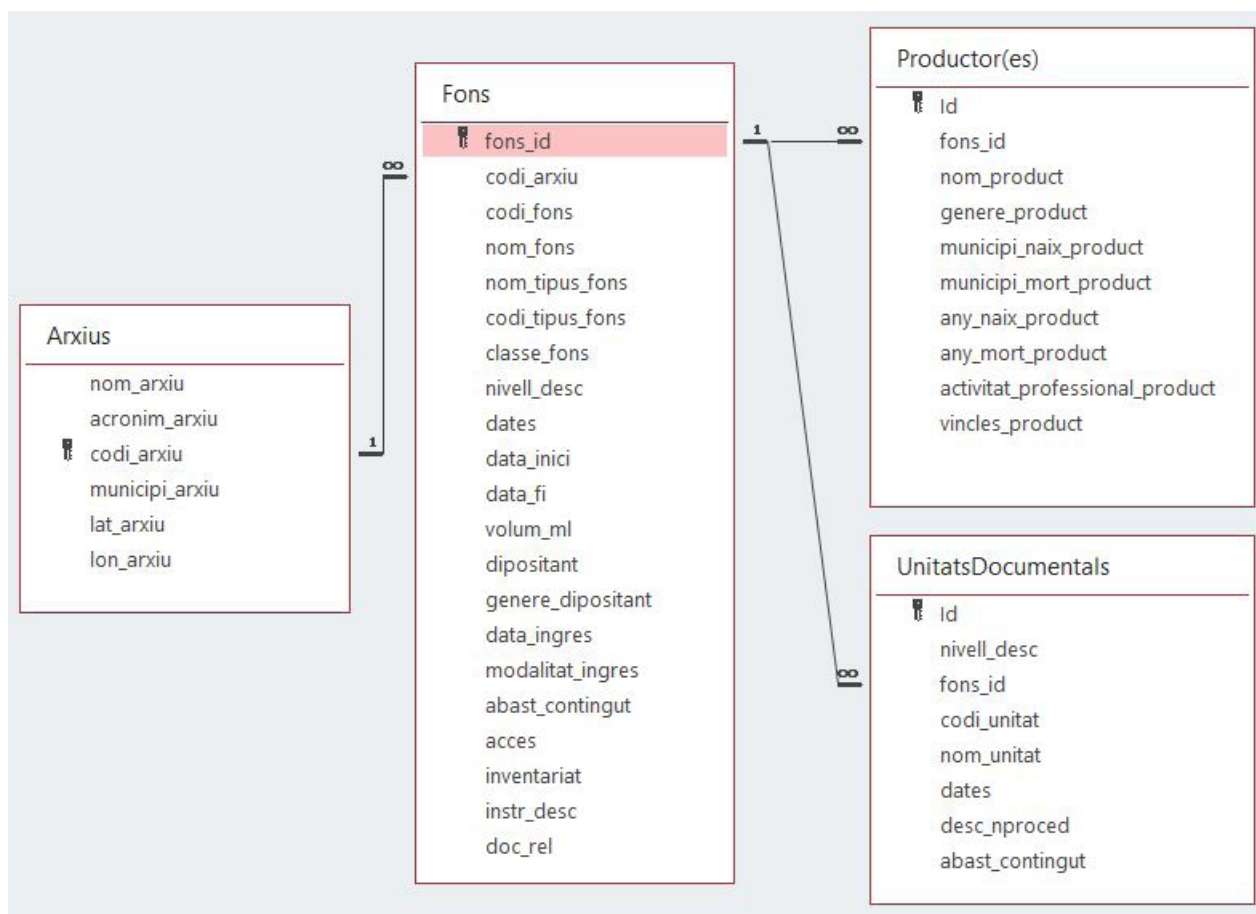


Figura 8. Esquema de les relacions entre les diferents taules del *dataset*. (Font: Elaboració pròpia)

Ara per ara, aquest *dataset* serà estàtic, doncs les dades s'han treballat en local i no hi ha manera de mantenir-lo actualitzat sense la col·laboració i compromís dels diferents arxius.

De tota manera, és poc probable que les dades pateixin grans canvis. Es tracta de fons tancats, i a no ser que se'n faci el tractament o noves intervencions, no s'espera que hi hagi canvis rellevants, almenys a curt termini. És també per aquest motiu que, tot i el risc que suposa tenir una dada que pugui quedar obsoleta, es considera important disposar d'un camp amb l'enllaç a la fitxa d'Arxius en Línia, doncs permet mantenir la relació de context amb l'instrument de descripció.

5. Publicació de les dades

74



5. 1. Definició dels termes de publicació de les dades

Una vegada s'ha disposat de les dades que es volen publicar i s'ha definit l'estructura i el model de dades, s'ha decidit en quins formats oferir les dades, sota quina llicència publicar-les i quina plataforma utilitzar per fer la prova de concepte.

5. 1. 1. Els formats

En aquest projecte, el tractament i la gestió de les dades, així com la prova de les relacions entre les taules del *dataset*, s'ha fet amb Microsoft Office Excel i Access. Malgrat tot, per tal de complir amb la iniciativa de dades obertes, és important que aquestes es publiquin en formats oberts, reutilitzables i no propietaris.

Així doncs, tot i que la gestió s'hagi fet amb formats propietaris, es proposa la prova d'obertura en XLSX, ODS i CSV, de tal manera que almenys dos dels formats sí compleixen amb els principis de dades obertes.

5. 1. 2. Llicències d'ús

Les llicències i termes d'ús són igual d'importants que la reutilització i la interoperabilitat tècnica, i també han de complir amb els principis de les dades obertes. Les llicències utilitzades per publicar dades en obert han de permetre l'accés i la reutilització, i poden atorgar llibertat per compartir, distribuir, adaptar o utilitzar les dades, així com la creació d'obres derivades.

Actualment hi ha dues opcions: *Creative Commons* (CC)⁷³ i *Open Data Commons* (ODC)⁷⁴. Les primeres són aplicables tant a dades com a documents, mentre que les segones s'utilitzen únicament per a dades. Segons Clabo i Ramos-Vielba (2015) la tendència internacional es basa en l'ús de llicències CC i acostumen a tenir especificacions diferents segons el portal on es publiquen.

De les sis llicències que proposa CC, quatre compleixen amb els principis d'obertura:

- *Creative Commons Zero* (CC0): cap tipus de restricció. El creador renuncia a qualsevol dret d'autoria sobre l'obra.

- *Creative Commons Reconeixement* (CC BY): únicament requereix que qui reutilitza les dades faci referència a la font o autor original.
- *Creative Commons Reconeixement-NoDerivades* (CC BY-ND): permet la reutilització i la difusió sempre que faci referència a l'autor de la mateixa, però no es pot alterar per crear obres derivades.
- *Creative Commons Reconeixement-NoComercial* (CC BY-NC): es permet la creació d'obres derivades i la difusió de l'original, sempre que es reconegui l'autoria de la mateixa, però no permet l'ús amb finalitats comercials. Les obres derivades no necessàriament han de mantenir la mateixa llicència que l'original.

Les llicències ODC van ser desenvolupades i són gestionades per l'*Open Knowledge Foundation*, i se centren únicament en dades. Actualment les llicències són tres:

- *Open Data Commons Public Domain Dedication and License* (PDDL). Permet fer difusió, reutilitzar i adaptar les dades sense cap tipus de restricció.
- *Open Data Commons Attribution License* (ODC-By). Exigeix fer referència a l'autoria o font d'origen de les dades per poder reutilitzar-les legalment.

- *Open Data Commons Open Database License* (ODbL). Permet la reutilització sempre que es reconegui l'autoria de la informació original i cal mantenir la mateixa llicència a les obres derivades.

Com ja s'ha vist a l'apartat 2.1.3 *Legislació relativa a dades obertes i reutilització*, aquesta no especifica ni obliga a que els instruments de descripció es publiquin en formats oberts i reutilitzables, però sí que s'exigeix en compliment de la llei de transparència que se'n faci publicitat activa. Així, les dades que es proposa obrir, tractades o no, s'han extret de les exportacions que els propis arxius han facilitat per al projecte, i per tant, s'entén que no presenten restriccions d'accés i poden ser difoses lliurement. D'aquesta manera, els *datasets* que es publicaran no hauran de tenir cap tipus de restricció d'ús, en tant que l'objectiu és que puguin ser utilitzats per qualsevol persona i amb qualsevol finalitat.

Així i tot, el més adient és establir unes condicions mínimes per a la reutilització del conjunt de dades: no alterar ni desnaturalitzar el contingut i citar la font. Així mateix, seria interessant també tenir coneixement de les obres derivades, i dels possibles suggeriments dels usuaris. Això permetria, no només la difusió per arribar a un públic més ampli, sinó engegar actuacions de millora en actualitzacions posteriors.

Les llicències que s'adapten a aquestes necessitats són: CC BY i ODC-By.

5. 1. 3. Plataforma de publicació

La publicació del conjunt de dades es proposa fer amb CKAN (*Comprehensive Knowledge Archive Network*), en la versió Demo⁷⁵. Es tracta d'una plataforma *open source* per a l'emmagatzemament i distribució de catàlegs de dades. Disposa d'eines per a la publicació, l'intercanvi, la recerca i l'ús de conjunts d'informació. És una plataforma gratuïta que, actualment, empreses i organitzacions governamentals d'arreu del món, utilitzen per a desenvolupar els seus portals de dades.

La plataforma proporciona una manera senzilla de fer que les dades siguin visibles i la seva cerca fàcil. Al mateix portal de CKAN es defineixen les seves característiques principals, entre les quals destaquen:

- Publicació i gestió de conjunts de dades.
- Emmagatzemament de dades.
- Personalització. Permet seleccionar i escollir les funcions que cada organització vol implantar al seu portal de dades, amb més de 200 extensions disponibles.

- Visualitzar dades en gràfics i mapes. Tot i que en tractar-se d'una plataforma per a la publicació de dades, aquestes funcionalitats són força limitades i caldrà anar a aplicacions externes si es volen obtenir resultats més complexos i elaborats.

5. 2. Prova de concepte

5. 2. 1. El procés de publicació a CKAN i la creació de visualitzacions

Per poder fer la prova de concepte, prèviament cal registrar-se a la versió Demo de CKAN, per tal de tenir un perfil d'usuari a partir del qual publicar, administrar i gestionar les dades.

Una vegada dins de la plataforma, el primer pas ha estat crear una organització, a la qual se li ha donat el nom de *TFM_FonsDones*. És a aquesta organització fictícia a la qual se li ha atribuït l'autoria dels diferents conjunts de dades publicats. A continuació, s'ha procedit a crear els conjunts de dades i a pujar-hi els diferents fitxers en els tres formats seleccionats.

L'objectiu de publicar dades en obert en portals d'aquestes característiques és que l'usuari pugui utilitzar-les i reutilitzar-les lliurement, de tal manera que no es publiquen les quatre taules que conformen el model de dades que s'ha elaborat prèviament, sinó conjunts de dades extrets

d'aquest model, els quals han estat pensats amb finalitats de visualització concretes. Així, per realitzar aquesta prova de concepte, s'han publicat els *datasets* que es descriuen a continuació⁷⁶:

1. Volums de fons per arxiu

Es publica el contingut de la taula d'arxius amb dues columnes addicionals: a una s'indica si l'arxiu custodia fons de dones (1) o no (0). A l'altra columna, s'indica el número de fons de dones que custodia cadascun dels arxius treballats en aquest projecte. Per obtenir aquestes dues columnes, s'ha procedit de la manera següent:

- S'ha fet una consulta a la base de dades completa i relacionada, on s'ha extret únicament les columnes: codi arxiu i codi fons, de tal manera que s'ha obtingut un relació dels 96 fons amb els seus respectius codis d'arxiu.
- S'ha exportat el resultat a XLSX.
- S'ha afegit una columna amb tots els codis d'arxiu.
- S'ha aplicat la funció *frequència*. És una fórmula matricial, de tal manera que per obtenir tots els resultats caldrà prémer la combinació de tecles CTRL + MAYUS + ENTER.
- S'ha creat un duplicat de la taula Arxius i s'hi ha enganxat el resultat. A la nova columna, se li ha donat el nom de *num_fons*.
- A continuació, s'ha filtrat aquesta nova columna i s'ha assignat un 0 a aquells arxius que no custodien fons, i un 1 als que sí. El nom d'aquesta columna és *fons_si_no*.

A partir d'aquesta taula es poden crear algunes visualitzacions amb les eines que facilita el mateix CKAN. Se'n proposen dues:



Figura 9. Visualització d'un gràfic generat amb CKAN que mostra el volum de fons custodiat per cada arxiu. L'eix X fa referència al codi d'arxiu, mentre que l'eix Y indica el número de fons custodiats en cada cas.

Volums de fons per arxius

🔧 Gestiona

📄 Anar al recurs

🔗 API de dades

URL: <https://demo.ckan.org/dataset/c2493a1d-c6f5-4104-802a-1e08bb88227c/resource/73f731c5-a57b-4739-a668-69b643d5b87e/download...>

Del resum del conjunt de dades

El dataset conté els següents camps: nom de l'arxiu, acrònim de l'arxiu, codi d'arxiu, municipi, coordenades geogràfiques. La columna 'fons_si_no', indica amb un 1 aquells centres d'arxiu...

Font: Volum de fons de dones per arxiu

📊 Data Explorer

📄 Num_Fons_arxiu

📍 Arxius i Fons

Visualització en un mapa dels arxius que custodien fons de dones. Clicant sobre el punt de georeferenciació es pot veure el nom de l'arxiu, l'acrònim, el codi del centre, el municipi i el número de fons de dones que conserva.

🔗 Incrustar

Add Filter

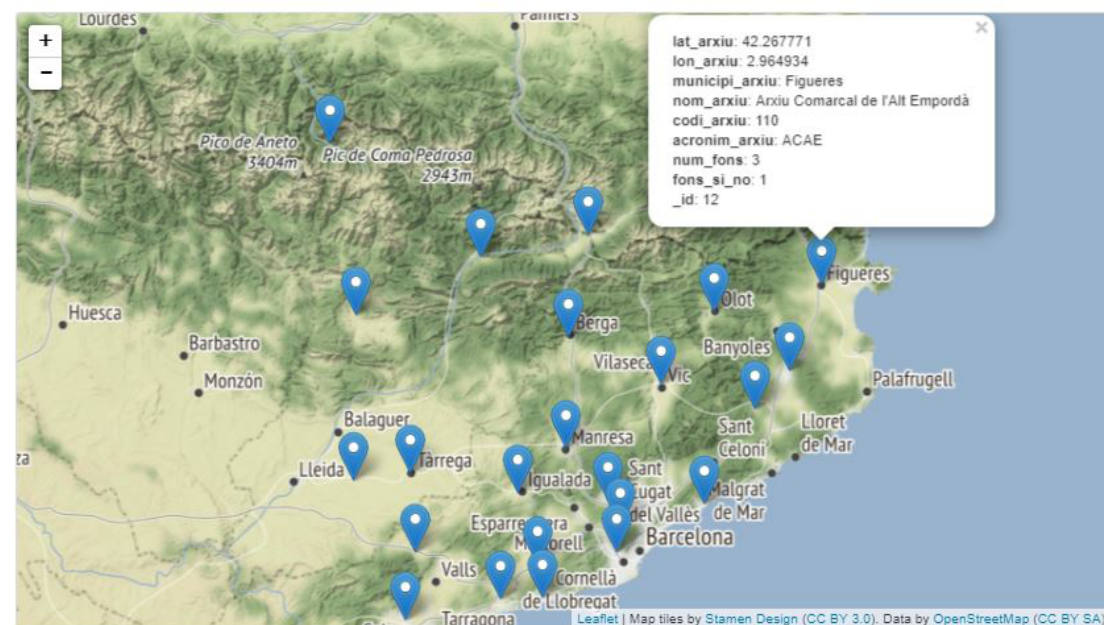


Figura 10. Visualització del mapa generat amb CKAN on es georeferencien únicament els arxius que sí custodien fons de dones. A la miniatura es poden veure les metadades associades a cada punt.

Com es pot veure a les dues visualitzacions anteriors, les opcions que ofereix CKAN són senzilles i no permeten grans personalitzacions ni detalls. Malgrat tot, compleixen amb la funció d'oferir una mostra de les possibilitats que tenen aquestes dades.

2.Descripció a nivell de fons

En aquest cas, es publica un conjunt de dades obtingut a partir de la combinació d'elements de la taula Arxius i la taula Fons. A la publicació a CKAN no s'hi associen visualitzacions, doncs com s'acaba de dir, les opcions que ofereix són limitades. En canvi, s'han treballat les dades en local per tal de crear diferents gràfics⁷⁷. Concretament tres:

- El primer mostra ràpidament el gènere de la persona dipositant dels fons.
- El segon, combina el règim d'accés i el nivell de tractament que han rebut els fons -a partir del camp que indica si el fons ha estat inventariat o no-.
- El tercer, mostra els anys en que s'han efectuat ingressos i les modalitats dels mateixos, juntament amb una línia de tendència i la graella de dades que permeten interpretar més fàcilment el gràfic.

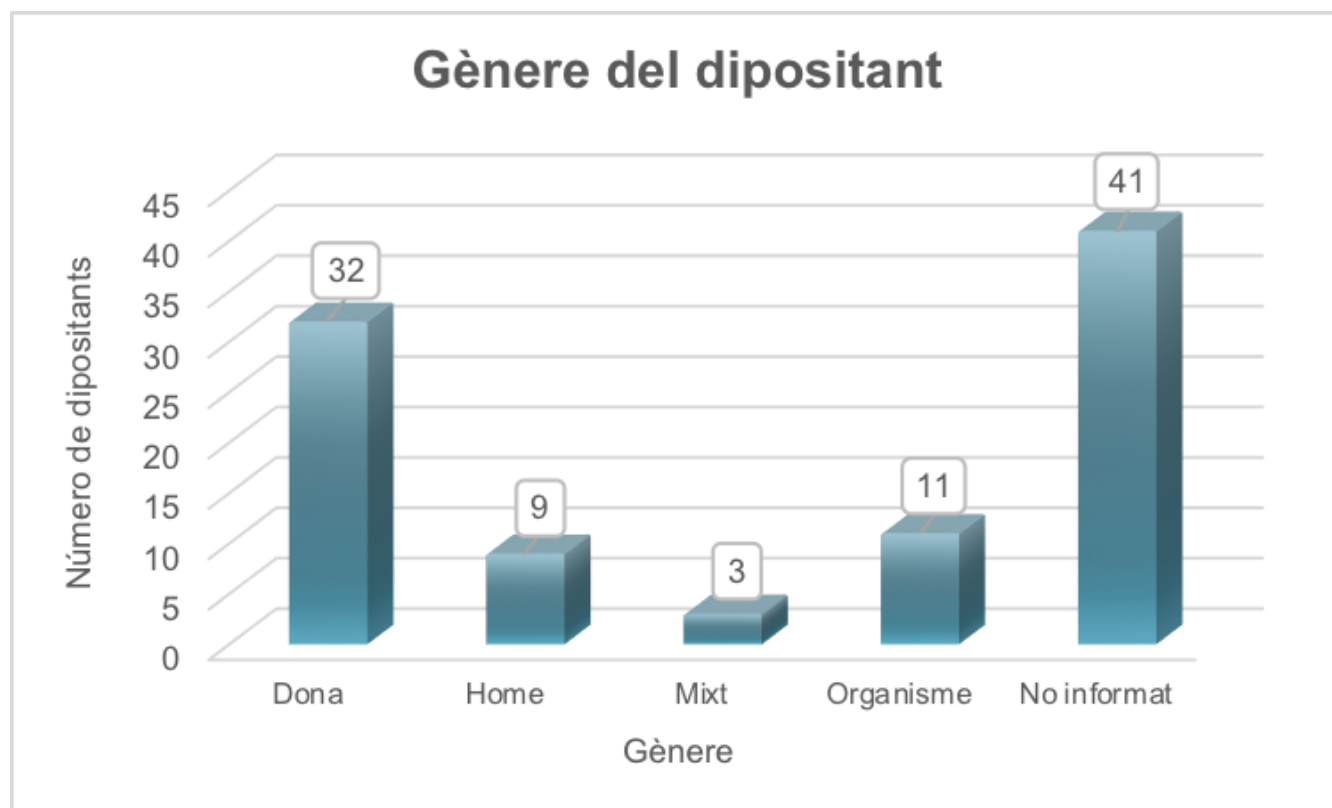


Figura 11. Gràfic amb el número de dipositants segons el seu gènere. (Font: Elaboració pròpia)

A la **Figura 11** es pot veure que, del total de 96 fons treballats, de 41 no es disposa d'informació sobre la persona o persones que han formalitzat l'ingrés i, en conseqüència, se'n desconeix el seu gènere.

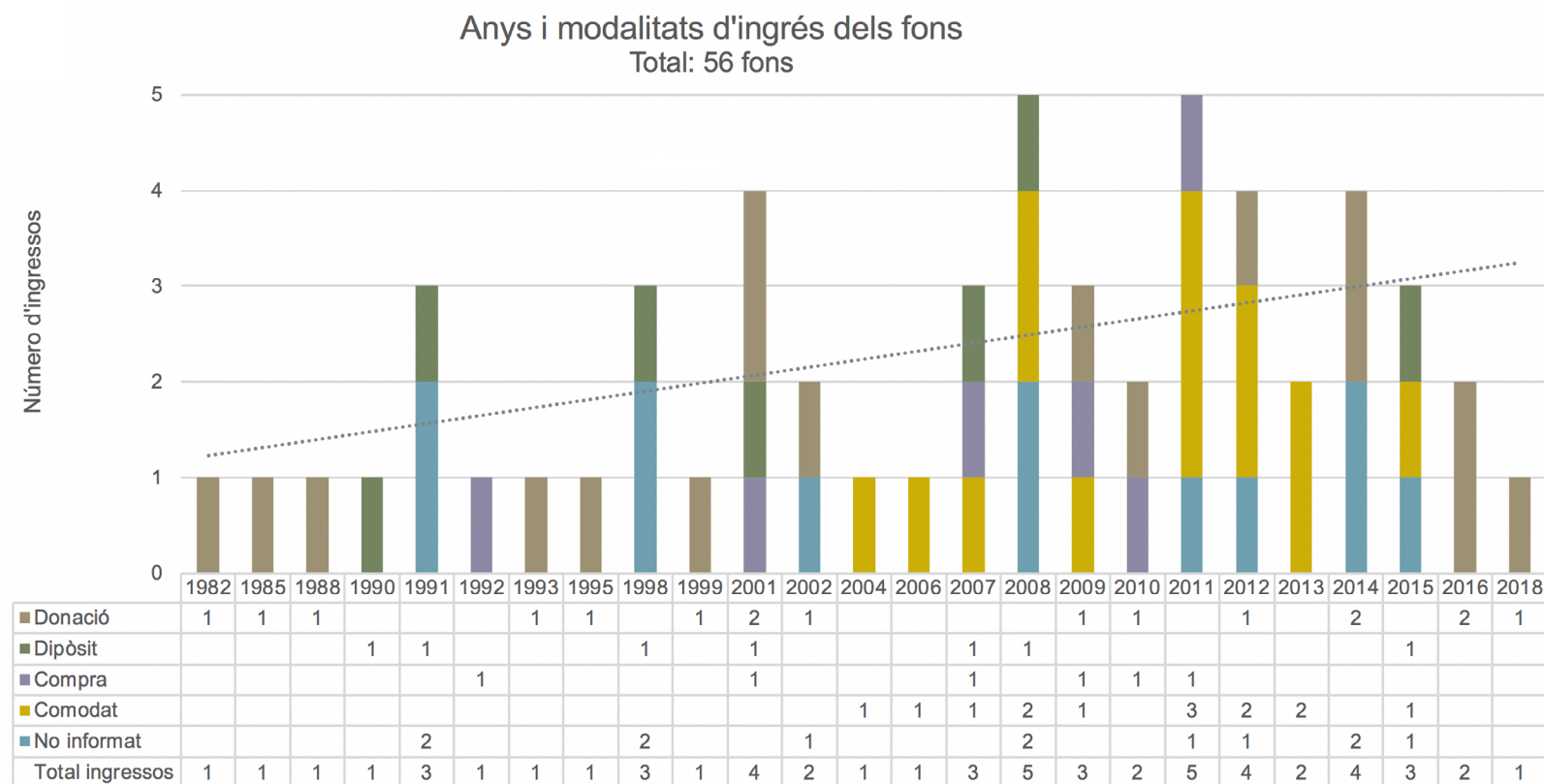


Figura 12. Gràfic que combina dades referents a les modalitats d'ingrés dels fons i els anys en els quals s'han fet efectius. (Font: Elaboració pròpia)

A la **Figura 12** es reflecteixen, tant al gràfic de barres com a la quadrícula inferior, el número d'ingressos per any. El gràfic mostra únicament els anys en que s'han registrat in-

gressos, i només d'aquells fons les descripcions dels quals inclouen la data d'ingrés. La línia discontinua mostra la tendència lineal dels volum d'ingressos.

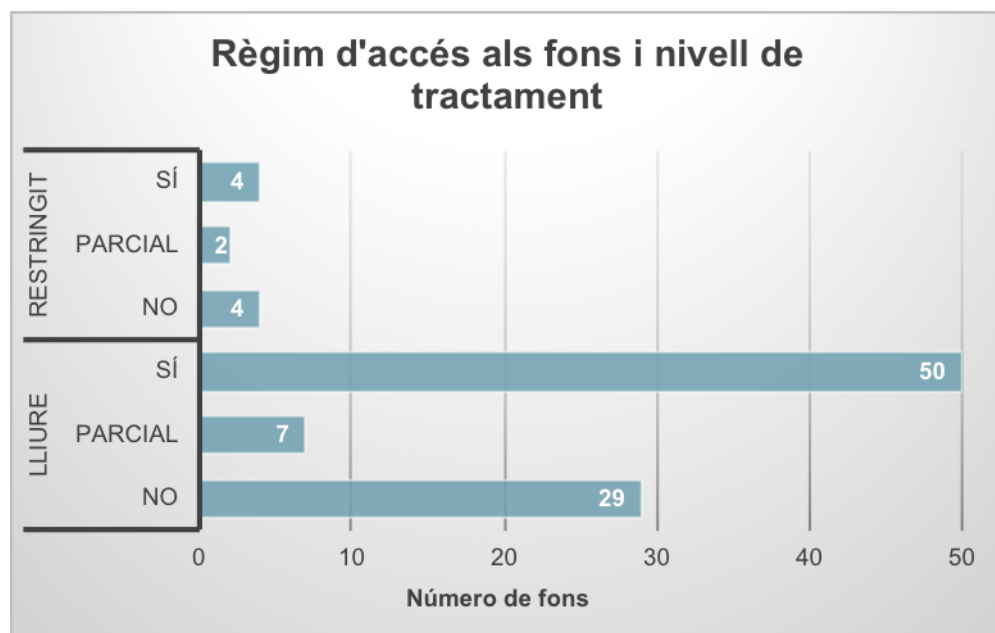


Figura 13. Gràfic que combina el règim d'accés als fons amb el nivell de tractament que han rebut. (Font: Elaboració pròpia)

En aquest cas, s'ha utilitzat el camp *Inventariat* com a dada de referència per al nivell de tractament, el qual està informat a la totalitat dels fons treballats.

3.Descripció dels/les productors/es

Aquest conjunt de dades aporta la informació bàsica sobre la història de les productores. S'hi combinen alguns camps de la taula d'Arxius i de la taula Fons, amb els de la taula Productor(es). Igual que en el cas anterior, el *dataset* s'ha publicat a CKAN sense visualitzacions associades. S'ha treballat amb Microsoft Excel per crear un gràfic en el qual es poden veure les professions i activitats que han

desenvolupat aquestes dones, els fons de les quals es troben als arxius del XAC i de l'ANC.

El procediment seguit per crear el gràfic que es pot veure a la **Figura 14** ha estat el següent:

- A la taula s'ha filtrat per gènere per tal d'obtenir únicament informació sobre dones.
- S'ha seleccionat el camp <professió> i, en tant que pel gràfic que volem fer no ens interessa la seva relació amb els altres camps, sinó únicament el seu contingut, s'ha aïllat en un nou full.

- S'han separat les dades delimitades per punt i coma “;” i s'han disposat totes en una única columna.
- S'ha comptat la freqüència de cada professió i, amb les dades obtingudes, s'ha creat el gràfic al mateix full de càlcul.

A l'hora d'analitzar el resultat obtingut hem de tenir en compte dues qüestions. D'una banda, que de les 104 dones que s'han treballat, només 48 tenen informació relacionada amb l'activitat professional, la qual s'ha extret de l'element *Història del(s) productor(s)*. D'una altra banda, que una mateixa dona pot haver exercit o desenvolupat més d'una professió o activitat. Així doncs, el gràfic mostra 82 activitats professionals portades a terme per un total de 48 dones.

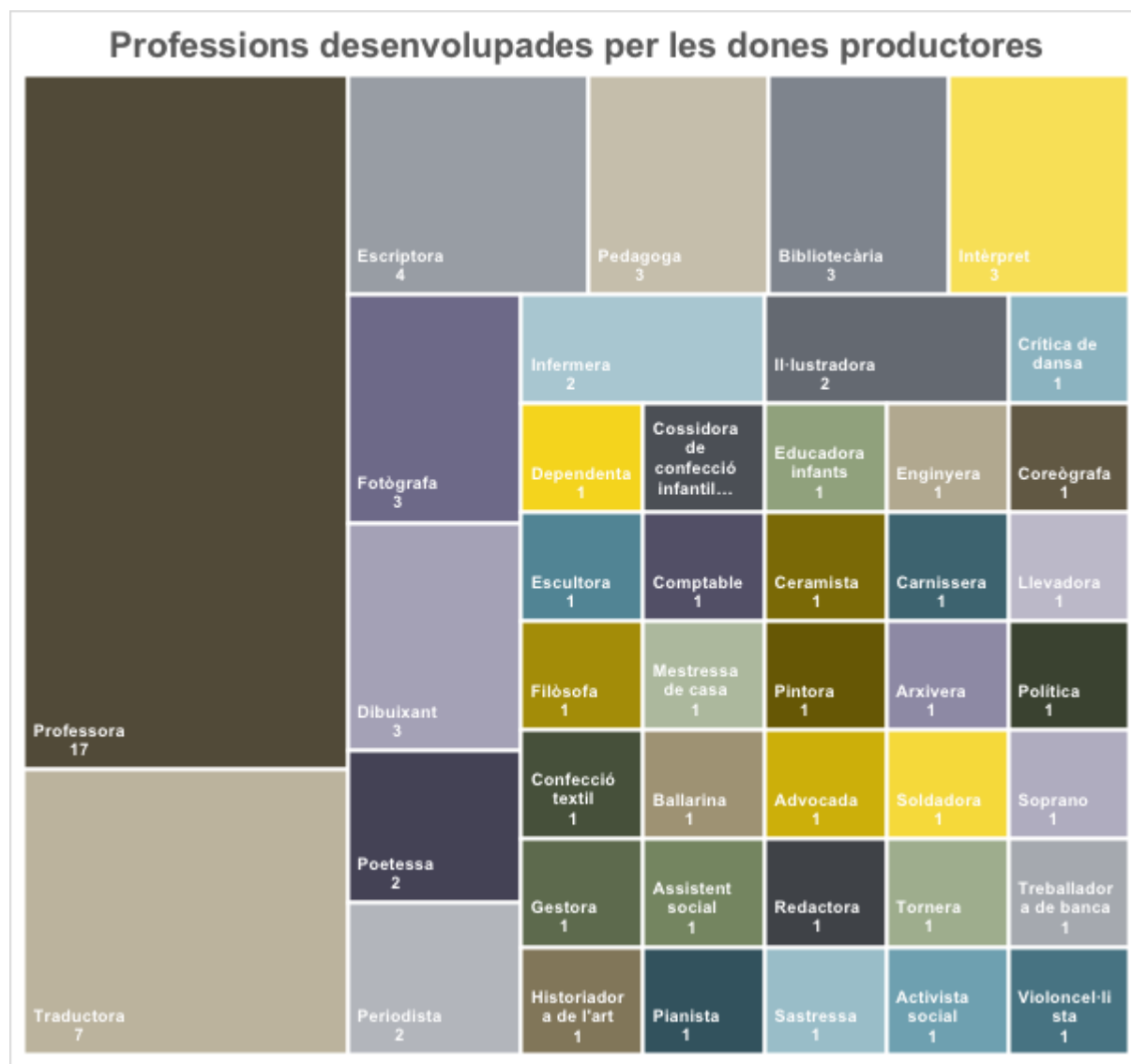


Figura 14. Gràfic on es representen les diferents professions i activitats que han desenvolupat les dones dels fons treballats. (Font: Elaboració pròpia)

Malgrat la incompletesa de les dades, l'elaboració del gràfic ha permès veure com, tot i que les activitats que desenvolupen les dones són ben diverses, destaquen aquelles relacionades amb l'educació, les arts (molt especialment la música) i les lletres.

4. Municipis de naixement de les productores dels fons

En tant que l'objectiu principal és donar visibilitat a la dona productora, s'han explorat altres possibilitats de reutilització de la informació per veure algunes de les oportunitats ofereix la combinació dels instruments de descripció dels arxius amb altres tipus de *datasets*.

En aquest cas, és proposa l'ús d'un conjunt de dades extern que conté informació geogràfica i que ha de permetre geolocalitzar els municipis de naixement i/o mort de la *Taula 3. Productor(es)*.

Per fer-ho, el primer pas ha estat descarregar *dataset BD_Municipios-Entidades*⁷⁸ de l'Institut Geogràfic Nacional⁷⁹. Es tracta d'un *dataset* molt complet i amb molta informació, del qual s'ha importat únicament la taula MUNICIPIOS, la qual s'ha incorporat al conjunt de dades del projecte. A continuació, s'ha relacionat el camp que conté el nom del municipi <NOMBRE_ACTUAL> amb el camp del *Municipi de naixement* de la taula *Productor(es)*. Això permet que, fent una nova consulta, puguem obtenir automàticament les coordenades geogràfiques⁸⁰ d'aquests municipis.

Ara bé, per tal que aquesta combinació sigui efectiva, ha estat necessari normalitzar el vocabulari de noms dels municipis de fora de Catalunya, perquè als instruments de descripció apareixen en la seva traducció al català. Això fa que, en el moment de fer l'encreuament de dades, no es reconeguin els municipis i, en conseqüència, no se'ns retornin a la consulta.

En el procés de normalització s'han pres les següents decisions⁸¹: d'una banda, els llocs de naixement en què coincideixen el nom del municipi i el de província, i no s'especifica a què fa referència exactament, s'ha entès com a municipi (per exemple: Barcelona o Madrid). D'una altra banda, els llocs de naixement que a la *Història del(s) productor(s)* fan referència a barris d'un municipi, s'han modificat i s'ha indicat el nom normalitzat del municipi. Fet aquest pas, s'ha procedit a exportar la taula, publicar-la al perfil de la organització del projecte creada a CKAN i, a continuació, generar el mapa geolocalitzat.

Productor(es) - Municipi de naixement

[Gestiona](#)[Anar al recurs](#)[API de dades](#)

87



URL: <https://demo.ckan.org/ca/dataset/9ea9bec4-10ad-4210-b9a3-b2267494c844/resource/7716d6ec-74a1-481a-b2c2-c25ae6323260/download...>

Del resum del conjunt de dades

Taula que mostra informació bàsica de la història de les dones productores de fons documentals. Inclou el Nom de l'arxiu, l'Acrònim de l'arxiu, el Codi d'arxiu, el Codi fons, el Nom del...

Font: Productor(es) per municipi de naixement

[Explorador de dades](#)[Origen de les dones productores dels fons](#)[Incrustar](#)

Add Filter



Figura 15. Visualització d'un mapa generat a CKAN, en el qual es geolocalitzen els municipis de naixement de les productores, únicament dones nascudes a territori espanyol, dels fons treballats.

En aquest mapa es geolocalitzen 44 de les 104 dones treballades i s'agrupen per proximitat. A mesura que es va fent zoom sobre el mapa es pot veure amb detall el municipi d'origen.

Els resultats obtinguts presenten dues característiques que cal tenir presents a l'hora d'analitzar les dades, i és que només apareixen les dones que tenen informat el municipi de naixement, i d'aquestes només aquelles nascudes a Espanya. Això és perquè el conjunt de dades amb que s'ha creuat, únicament informa dels municipis a l'àmbit estatal. Per tal que apareguessin també les dones nascudes fora d'Espanya caldria, o bé informar les coordenades manualment, o bé buscar altres *datasets* que ofereixin aquesta informació en un àmbit territorial molt més ampli.

5. Descripció a nivell d'unitat documental amb descriptors

La taula amb dades a nivell d'unitat documental de tots els fons no conté descriptors. Per aquest motiu s'utilitzarà, únicament com a exemple, el fons fotogràfic ANC 1-788 COLITA, inclòs en una de les taules facilitades per les exportacions de l'ANC, ja que és l'únic fons de qual es disposa de descriptors a nivell d'unitat documental. Aquests estan separats per columnes i es poden treballar amb certa facilitat. Així doncs, es proposa utilitzar aquesta informació per fer una prova de geolocalització a partir dels topònims, els quals majoritàriament fan referència a municipis.

El procés per obtenir la taula amb les coordenades geogràfiques ha estat en mateix que en el cas anterior, és a dir, que per a l'obtenció de les coordenades dels municipis de naixement de les productores.

Primer s'han normalitzat els topònims⁸² i, a continuació, donat que ja es disposa de la taula MUNICIPIOS de l'IGN al model de dades, s'ha importat directament la taula amb les descripcions del fons ANC 1-788 COLITA amb els descriptors corresponents. S'ha relacionat la columna <geo_des>, és a dir, el topònim, amb la de <NOMBRE_ACTUAL> de la taula MUNICIPIOS⁸³. Un cop exportada la taula, s'ha publicat a CKAN i s'ha creat el recurs de visualització sobre un mapa. Cal tenir present que únicament mostra aquelles unitats de les quals es disposa de descriptor geogràfic referent a municipis espanyols.

Fons ANC 1-788 COLITA

[Gestiona](#)[Afegir al recurs](#)[API de dades](#)

URL: <https://demo.ckan.org/ca/dataset/1a531a07-7a48-4341-940c-48f6a36aeff3/resource/3908cabf-2809-419a-a4a3-b507bb63af68/download>

Del resum del conjunt de dades

Conjunt que inclou les dades descriptives de les unitats documentals del fons. Les coordenades geogràfiques s'han extret de la taula MUNICIPIOS del...

Font: Fons 1-788 COLITA

[Explorador de dades](#)[Geolocalització Unitats documentals](#)

Mapa que geolocalitza els municipis als quals fan referència les unitats documentals del fons ANC 1-788 COLITA. Cal tenir en compte que no reflecteix la totalitat del fons, sinó únicament aquelles unitats vinculades al territori espanyol.

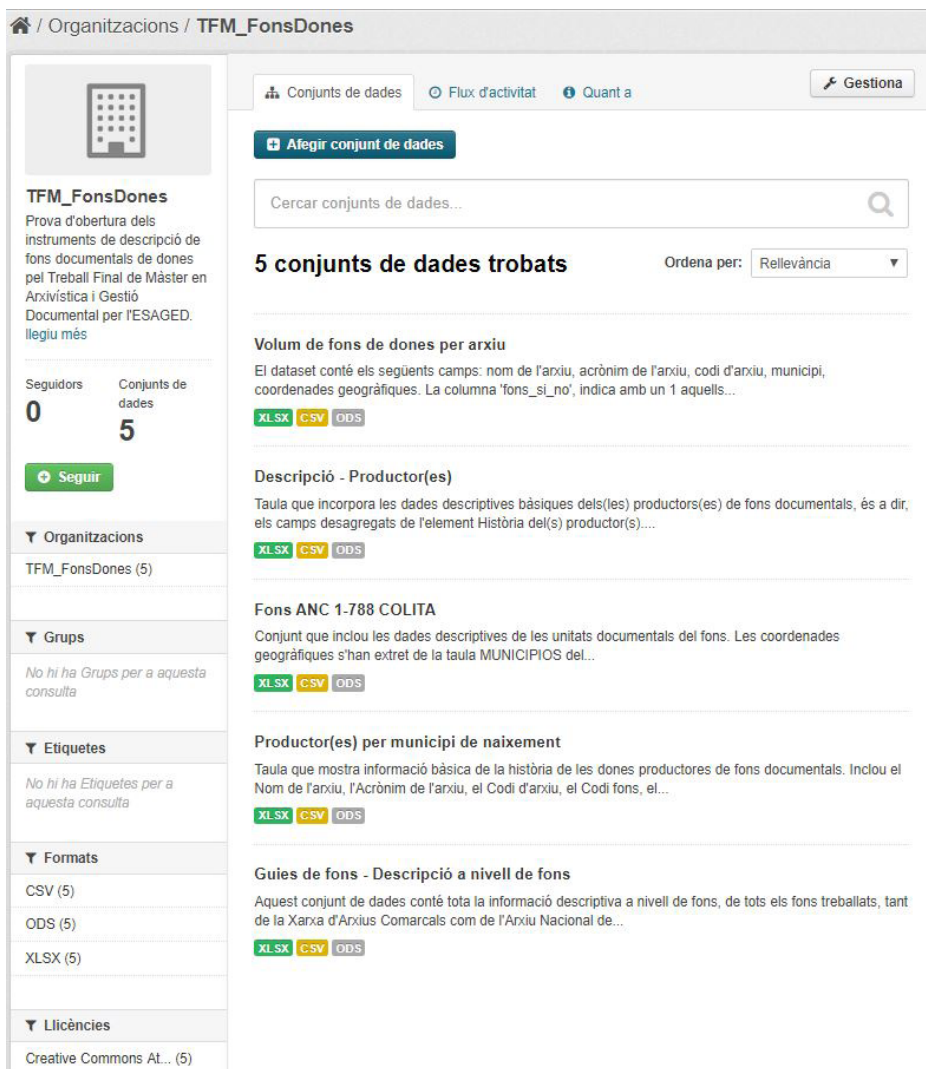
[Incrustar](#)

Add Filter



Figura 16. Visualització d'un mapa generat a CKAN, en el que es geolocalitzen les unitats documentals del fons ANC 1-788 COLITA.

Així doncs, una vegada realitzada la publicació de tots els fitxer de dades a CKAN, el resultat és el següent:



The screenshot displays the CKAN interface for the organization **TFM_FonsDones**. The left sidebar shows the organization's profile with 0 followers and 5 datasets. The main content area lists 5 datasets found, sorted by relevance. The datasets are:

- Volum de fons de dones per arxiu**: El dataset conté els següents camps: nom de l'arxiu, acrònim de l'arxiu, codi d'arxiu, municipi, coordenades geogràfiques. La columna 'fons_si_no', indica amb un 1 aquells... (Formats: XLSX, CSV, ODS)
- Descripció - Productor(es)**: Taula que incorpora les dades descriptives bàsiques dels(les) productors(es) de fons documentals, és a dir, els camps desagregats de l'element Història del(s) productor(s).... (Formats: XLSX, CSV, ODS)
- Fons ANC 1-788 COLITA**: Conjunt que inclou les dades descriptives de les unitats documentals del fons. Les coordenades geogràfiques s'han extret de la taula MUNICIPIOS del... (Formats: XLSX, CSV, ODS)
- Productor(es) per municipi de naixement**: Taula que mostra informació bàsica de la història de les dones productores de fons documentals. Inclou el Nom de l'arxiu, l'Acrònim de l'arxiu, el Codi d'arxiu, el Codi fons, el... (Formats: XLSX, CSV, ODS)
- Guies de fons - Descripció a nivell de fons**: Aquest conjunt de dades conté tota la informació descriptiva a nivell de fons, de tots els fons treballats, tant de la Xarxa d'Arxius Comarcals com de l'Arxiu Nacional de... (Formats: XLSX, CSV, ODS)

The sidebar also shows filters for formats (CSV (5), ODS (5), XLSX (5)) and licenses (Creative Commons At... (5)).

Figura 17. Vista de la pantalla d'inici de l'organització *TFM_FonsDones* creada a CKAN, amb els cinc conjunts de dades en els tres formats en què s'han publicat.

Com es pot veure a la **Figura 17**, s'han publicat un total de cinc *datasets* vinculats a l'organització *TFM_FonsDones*, creada expressament per a la realització del projecte. Els cinc *datasets* estan en formats oberts CSV i ODS, i també en XLSX. Tots amb la mateixa llicència d'ús *Creative Commons Atribució*.

Tal com es pot veure a l'esquerra de la imatge, es podrien assignar etiquetes en el moment de publicar els conjunts de dades, quelcom que permetria aplicar filtres, de la mateixa manera que es pot fer per formats de dades o segons el tipus de llicència d'ús amb la qual s'han publicat.

6. Conclusions

Al llarg d'aquest treball s'han explorat algunes de les oportunitats de difusió que ofereixen l'obertura i la reutilització dels instruments de descripció dels arxius, així com la importància d'incorporar la perspectiva de gènere per facilitar la recerca i millorar l'accessibilitat i la recuperació de fonts originals de dones.

El plantejament inicial ha marcat treballar l'obertura dels inventaris i les guies de fons produïts per dones. Per fer-ho, s'ha creat un model de dades que ha permès estructurar i emmagatzemar la informació que posteriorment s'ha preparat per a la publicació i difusió a la plataforma *open source* CKAN. Aquesta tasca s'ha complementat amb la geolocalització i l'elaboració de recursos visuals en forma de gràfics que han permès, d'una banda, assolir l'objectiu principal del treball i donar visibilitat a la dona productora de fons documentals, i d'una altra banda, exemplificar diferents anàlisis que es poden fer a partir d'aquestes dades.

Després de tot el procés de treball hi ha diverses qüestions que cal posar sobre la taula. D'entrada, és necessari que l'obertura de dades es faci des dels mateixos arxius. Només el mateix arxiu que conserva els fons n'és vertader coneixedor del seu contingut i pot verificar i atorgar autenticitat i fiabilitat a les dades que publica. En cas contrari, es poden generar errors de compressió de la informació que desvirtuïn, o fins i tot transformin, el contingut dels

instruments de descripció i, per tant, s'obtingui un resultat que no s'ajusti a la realitat. A més, l'obertura des dels arxius ofereix l'oportunitat d'automatitzar l'actualització de dades en la creació d'obres derivades per part de tercers, com seria el cas d'aquest projecte. Igualment, es poden treballar conjuntament dades de diferents orígens, independentment de qui custodia la informació, facilitant així apropar-nos als usuaris i captar nous públics.

En segon lloc, la preparació de les dades ha evidenciat la relativa qualitat del contingut descriptiu dels fons analitzats, així com la necessitat de normalitzar la codificació dels camps d'informació per tal que siguin, no només interoperables sinó, útils i usables de cara a la seva reutilització. A més, s'ha comprovat la incompletesa d'una part important de les guies de fons, així com la inexistència de descripcions a nivell d'unitat documental en molts dels exemples, tal com s'ha reflectit a les **Figures 7 i 13**.

En tercer lloc, cal remarcar la importància que l'obertura de dades es faci pensant en els usuaris finals, és a dir, anant molt més enllà de garantir l'accés a la informació, i disposar d'instruments de descripció més complets i amb contingut de major significat i qualitat, de tal manera que si es redueixen els camps de descripció a expressions més petites, aquestes siguin igual o més significatives, i que ofereixin noves possibilitats d'anàlisi i reutilització.

Ara bé, en tant que necessitem enfocar-nos als usuaris potencials dels arxius, es fa evident que calen estudis que defineixin detalladament els perfils d'aquests, especialment en entorns digitals. Els arxius han de tenir clar no només qui són els seus usuaris, sinó sobretot a qui volen adreçar-se. Per això mateix, és necessari que es facin anàlisis que aprofundeixin en els seus interessos i necessitats: què consulten, com ho busquen i quins criteris de cerca fan servir, on fan la recerca, si entenen els instruments de descripció arxivística i, més concretament, si els hi són útils, el seu nivell de satisfacció, així com propostes o suggeriments de millora. Estudis que donin resposta a aquestes preguntes són essencial per tal que els productes dels arxius compleixin amb les seves expectatives i cobreixin les seves necessitats.

En quart lloc, quant a anàlisi de les dades per a la visibilitat de la dona, l'aspecte més destacable del projecte són els recursos visuals i gràfics que es proposen a partir de les dades obertes publicades, els quals donen evidència de la falta de representació de les dones als arxius i permeten demostrar que:

- Hi ha pocs fons de dones. Pel que fa concretament als fons personals, trobem que a data 8 d'agost de 2019 hi ha donats d'alta a Arxius en Línia 1050 fons dels quals 953 són custodiats per arxius que formen part de l'abast del projecte. D'aquests, s'han identificat i analitzat en aquest projecte un total de 77 fons personals de dones,

és a dir, només ho són un 8.07%⁸⁴ dels fons personals conservats en aquests arxius públics catalans.

- 15 dels 38 arxius treballats, no conserven cap fons produït per una dona. És probable que sí que conservin documents produïts per dones inclosos dins fons familiars, però donada la complexitat de localització d'aquest tipus de fons si volem fer cerca per gènere, no s'han pogut identificar.
- Dels fons que són d'accés lliure (86), n'hi ha 29 que resten per inventariar i fins a 7 ho estan parcialment.
- Quant als ingressos, s'ha pogut veure que només 56 fons disposen de la data del primer ingrés efectiu de la documentació a l'arxiu. Així, el 2008 i 2012 són els anys amb major quantitat d'ingressos, un total de cinc en cada cas. Pel que fa a la modalitat d'ingrés, de 12 no se'n té informació, 17 són donacions, 14 comodats, 7 dipòsits i 6 compres. Són dades força parcials i caldria fer una anàlisi més extensa perquè sigui més concluent i fonamentat. Amb tot, a partir de la informació que s'ha obtingut en aquest projecte, ara mateix es pot dir que, tot i que el nombre de donacions és el més elevat, 21 dels ingressos no comporten un canvi de titularitat, i per uns altres 6 ha estat necessari un desemborsament econòmic per adquirir-los. Tot plegat, reforça la idea que cal promoure les donacions de fons de dones als arxius i dissenyar polítiques d'adquisició de fons que siguin més inclusives i que promoguin la donació de documents originals de dones. Igualment, cal incentivar i destinar més recursos al tractament i la difusió

d'aquesta documentació. S'ha d'apostar per estratègies feministes a l'hora d'abordar l'ingrés, la gestió i el tractament de fons als arxius.

- Pel que fa a les activitats professionals desenvolupades per les dones productores dels fons, en general es tendeix cap a la branca de lletres i educació, destacant les professores de música i idiomes, traductores i intèrprets.

En cinquè lloc, l'anàlisi i preparació de les dades a publicar ha mostrat que hi ha certa confusió o incoherència a l'hora d'identificar i distingir els fons personals dels fons familiars i patrimonials, quelcom que caldria revisar doncs pot dificultar la seva recuperació. Dita revisió, passa també per aprofundir en la teoria arxivística per ampliar el concepte de productor i l'aplicació del principi de provenença i de respecte de l'ordre originals dels fons.

En conclusió, amb aquest projecte s'ha demostrat la hipòtesi inicial i és que, efectivament, les dades obertes, la geolocalització, l'anàlisi i la visualització de dades, donen visibilitat als arxius i proporcionen recursos per a la millora de la difusió dels seus fons i documents. I en aquest cas concret, a més, es permet analitzar les dades des d'una perspectiva de gènere, donant visibilitat i protagonisme a les dones productores de documents.

Pel que fa a les línies de continuació d'aquest projecte caldria, d'una banda, plantejar-se allotjar els *datasets* a un servidor propi, o en tot cas, fer una inversió econòmica per tal d'allotjar-los als servidors de CKAN. Això és necessari perquè, per fer la prova de concepte, s'han publicat fent ús de la versió Demo de CKAN, i com ja s'ha comentat, això comporta que cada cert temps s'esborraran les dades.

D'una altra banda, és interessant i necessari ampliar l'abast d'arxius a analitzar, considerant especialment relevant disposar de dades d'àmbit municipal de tot el territori català.

7. Bibliografia

95



BERNAL, À., i DURAN, A. (2001). La identificació dels fons de l'Administració a l'Arxiu Nacional de Catalunya. *LLIGALL Revista Catalana d'arxivística*, 18, 165–216. Recuperat de <https://www.raco.cat/index.php/lligall/article/view/339532/430486>

CARREÑO CORCHETE, E. (2017). Memorias de silencios. Espacio y función de las mujeres en el archivo. *Tabula*, 20, 65–80.

CASWELL, M., i CIFOR, M. (2016). From Human Rights to Feminist Ethics: Radical Empathy in Archives. *Archivaria*, 81, 23–43.

CIFOR, M., i WOOD, S. (2017). Critical Feminism in the Archives. *Critical Archival Studies*, 1(2), 1–27. <https://doi.org/10.24242/jclis.v1i2.27>

CLABO, N., i RAMOS-VIELBA, I. (2015). Reutilización de datos abiertos en la administración pública en España y uso de licencias-tipo. *Revista Española de Documentación Científica*, 38(3), e097. <https://doi.org/10.3989/redc.2015.3.1206>

CODINA-CANET, M. A., i SAN SEGUNDO MANUEL, R. (2016). Propuesta de un centro de archivo del feminismo tras el análisis de los fondos documentales del movimiento feminista. *Revista Española de Documentación Científica*, 39(1), e117. <http://redc.revistas.csic.es/index.php/redc/article/view/920/1328>

COOK, T., i HERNÁNDEZ OLIVERA, L. (2007). Introducción. *Tabula*, 10, 11–17.

DE HAAN, F., i MEVIS, A. (2008). The iav/iiav's Archival Policy and Practice. Seventy Years of Collecting, Receiving, and Refusing Women's Archives (1935-2005) Francisca. In S. E. Weiringa (Ed.), *Travelling Heritage: New Perspectives on Collecting, Preserving and Sharing Women's History* (pp. 23–46). Amsterdam: Aksant Academic Publishers.

DE LA PUENTE, M. (2018). *EAD y EAC: Aplicación de metadatos en la descripción de colecciones archivísticas*. Buenos Aires: Consultora de Ciencias de la Información.

DE LOS RIÓS, N. (2013). Gobierno Abierto. *LLIGALL Revista Catalana d'arxivística*, 35, 139–150. Recuperat de <https://www.raco.cat/index.php/lligall/article/view/340099/431038>

DEPARTAMENT DE CULTURA, GENERALITAT DE CATALUNYA. (1994a). El quadre de fons dels arxius. Norma per a l'elaboració (I). *Arxius. Butlletí Del Servei d'Arxius*, 2.

DEPARTAMENT DE CULTURA, GENERALITAT DE CATALUNYA. (1994b). El quadre de fons dels arxius. Norma per a l'elaboració (II). *Arxius. Butlletí Del Servei d'Arxius*, 3.

DOUGLAS, J. (2018). A call to rethink archival creation: exploring types of creation in personal archives. *Archival Science*, 18(1), 29–49.

DUCHEIN, M. (1983). Theoretical Principles and Practical Problems of Respect des fonds in Archival Science. *Archivaria*, 16(Summer), 64–82. Recuperat de <https://archivaria.ca/index.php/archivaria/article/view/12648/13813>

EICHHORN, K. (2013). *The Archival Turn in Feminism: Outrage in Order*. Philadelphia: Temple University.

HENNINGHAM, N., EVANS, J., i MORGAN, H. (2017). The Australian Women's Archives Project: Creating and Co-curating Community Feminist Archives in a Post-custodial Age. *Australian Feminist Studies*, 32(91–92), 91–107.

HOBBS, C. (2001). The character of personal archives: Reflections on the value of records of individuals. *Archivaria*, 52, 126–135. Recuperat de <https://archivaria.ca/index.php/archivaria/article/view/12817/14027>

JORNET BENITO, N., i TUSET PÁEZ, N. (2016). Construint la memòria dels feminismes: arxius, biblioteques i centres de documentació Una mirada al passat, una reflexió per al futur. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, 36(juny). <https://dx.doi.org/10.1344/BiD2016.36.9>

KETELAAR, E. (2001). Tacit narratives: The meanings of archives. *Archival Science*, 1(2), 131–141.

MASON, K. M., i ZANISH-BELCHER, T. (1999). A Room of One's Own: Women's Archives in the Year 2000. *Archival Issues*, 24(1), 37–54. Recuperat de http://lib.dr.iastate.edu/speccoll_pubs/1

MASON, K. M., i ZANISH-BELCHER, T. (2007). Raising the Archival Consciousness: How Women's Archives Challenge Traditional Approaches to Collecting and Use, Or, What's in a Name? *Library Trends*, 56(2), 344–359. Recuperat de http://lib.dr.iastate.edu/speccoll_pubs/6

MEEHAN, J. (2010). Rethinking original order and personal records. *Archivaria*, 70 (Fall 2010), 27–44. Recuperat de <https://archivaria.ca/index.php/archivaria/article/view/13294/14605>

MEINHARDT LLOPIS, C. (2014). *Arximaps: descripció arxivística i geolocalització. Estudi de cas dels fons notariais catalans*. Universitat Autònoma de Barcelona. Recuperat de <http://hdl.handle.net/2072/261027>

MOSELEY, E. (1973). Women in Archives: Documenting the History of Women in America. *The American Archivist*, 36(2), 215–222. Recuperat de <https://www.americanarchivist.org/doi/pdf/10.17723/aarc.36.2.36744h4q226234j7>

MOYANO COLLADO, J. (2013). La descripció archivistica. De los instrumentos de descripció hacia la web semántica. *Anales de Documentacion*, 16(2). <https://doi.org/10.6018/analesdoc.16.2.171841>

MUNTÉ VAN ENKHUIZEN, K. (2017). *Datos abiertos: geocodificación y representación visual de un instrumento de descripció. El catálogo de licencias de obras particulares del Archivo municipal de Santa Perpètua de Mogoda*. Universitat Autònoma de Barcelona. Recuperat de <http://hdl.handle.net/2072/338660>

PASTRANA ROMAN, V. (2018). *Projecte de geolocalització per a la difusió del fons de la Comissió Provincial de Regions Devastades i Reparacions de Girona*. Universitat Autònoma de Barcelona. Recuperat de <http://hdl.handle.net/2072/336793>

POLLARD, R. A. (2001). The appraisal of personal papers: A Critical literature Review. *Archivaria*, 52(1), 136–150. Recuperat de <https://archivaria.ca/index.php/archivaria/article/view/12818/14029>

RÖMKENS, R., i WIERSMA, A. (2017). Gender and Archiving: Past, Present, Future. *Yearbook of Women's History*, 37.

TERRADELLAS PRAT, E. (2016). *Obrint dades a l'Arxiu Nacional de Catalunya*. Universitat Autònoma de Barcelona. Recuperat de <http://hdl.handle.net/2072/271335>

VOSS-HUBBARD, A. (1995). “No Documents - No History”: Mary Ritter Beard and the Early History of Women's Archives. *The American Archivist*, 58(1), 16–30. <https://doi.org/10.17723/aarc.58.1.hr300127g3142157>

WERNIMONT, J., i FLANDERS, J. (2010). Feminism in the Age of Digital Archives: The Women Writers Project. *Tulsa Studies in Women's Literature*, 29(2), 425–435.

YAKEL, E. (2005). Els arxius a l'era de l'accessibilitat. *LLI-GALL Revista Catalana d'arxivística*, 23, 117–134. Recuperat de <https://www.raco.cat/index.php/lligall/article/view/339611/430565>

8. 1. Annex 1. Quantitat d'elements NODAC informats als fons de dones

Element	ANC	XAC	Total	ANC %	XAC %	Total %
Codi de referència	43	53	96	100,0	100,0	100,0
Nivell de descripció	41	16	57	95,3	30,2	59,4
Títol	43	53	96	100,0	100,0	100,0
Data(es)	43	39	82	100,0	73,6	85,4
Volum i suport	41	21	62	95,3	39,6	64,6
Nom del(s) productor(s)	39	17	56	90,7	32,1	58,3
Història del(s) productor(s)	41	18	59	95,3	34,0	61,5
Història Arxivística	39	14	53	90,7	26,4	55,2
Dades sobre l'ingrés	39	19	58	90,7	35,8	60,4
Abast i contingut	43	24	67	100,0	45,3	69,8
Sistema d'organització	41	10	51	95,3	18,9	53,1
Informació sobre avalua-ció, tria i eliminació	0	3	3	0,0	5,7	3,1
Incrementos	0	4	4	0,0	7,5	4,2
Condicions d'accés	41	15	56	95,3	28,3	58,3
Condicions de reproducció	38	4	42	88,4	7,5	43,8
Llengües i escriptures dels documents	35	12	47	81,4	22,6	49,0
Característiques físiques i requeriments tècnics	2	6	8	4,7	11,3	8,3
Instruments de descripció	39	13	52	90,7	24,5	54,2
Existència i localització dels originals	1	3	4	2,3	5,7	4,2
Existència i localització de les reproduccions	1	1	2	2,3	1,9	2,1
Documentació relacionada	32	8	40	74,4	15,1	41,7
Bibliografia	33	5	38	76,7	9,4	39,6
Notes	6	1	7	14,0	1,9	7,3

Autoria i data(es)	40	17	57	93,0	32,1	59,4
Fonts	39	3	42	90,7	5,7	43,8
Regles o convencions	39	9	48	90,7	17,0	50,0

Taula 5. Taula on es mostren els diferents elements NODAC i el número de fons que en tenen alguna informació publicada a Arxius en Línia.
(Font: Elaboració pròpia)

Les primeres tres columnes mostren les dades amb números sencers. Les altres tres representen la mateixa informació amb percentatges. Total de fons de l'ANC (43); total de fons del XAC (53); i total del conjunt de fons treballats (96).

8. 2. Annex 2. Distribució dels elements descriptius al model de dades

Per a la columna d'exemple s'ha seleccionat un fons amb una descripció molt completa. Els camps que no tenen exemple aquí referenciat és perquè la dada descriptiva

manca i, per tal de no barrejar dades de diferents fons, s'ha deixat incomplet.

Taula 1. Arxius				
Element	Definició	Consignació	Tipus dada	Exemple
Nom arxiu	Nom complet de l'arxiu o centre que custòdia la documentació	OB	Text	Arxiu Comarcal del Vallès Occidental
Acrònim arxiu	Acrònim del servei d'arxiu o centre de custòdia	OB	Text	ACVOC
Codi arxiu	Codi que identifica inequívocament el servei d'arxiu	OB	Text	90
Municipi arxiu	Municipi on es troba físicament el servei d'arxiu	RE	Text	Terrassa
Latitud arxiu	Latitud geogràfica de l'arxiu per poder geolocalitzar-lo	RE	Text	41.553044
Longitud arxiu	Longitud geogràfica de l'arxiu per poder geolocalitzar-lo	RE	Text	2.016302

Taula 2. Fons

Element	Definició	Consignació	Tipus dada	Exemple
Identificador fons	Combina el codi d'arxiu i el codi fons, de tal manera que serveix de clau per relacionar les diferents taules	OB	Text	90-15
Codi arxiu	Codi que identifica inequívocament el servei d'arxiu	OB	Text	90
Codi fons	Codi que identifica inequívocament el fons	OB	Text	15
Nom tipus fons	Identifica el tipus de fons segons el Quadre de fons	OB	Taula validada	Personals
Codi tipus fons	Codi que identifica el tipus de fons segons el Quadre de fons	OB	Taula validada	13
Classe fons	Informa la classe genèrica del fons, segons si és textual o no textual	OB	Taula validada	Textual
Nivell de descripció	Identifica el nivell d'organització al qual correspon la unitat que es descriu	OB	Taula validada	Fons
Nom fons	Títol que dona nom a la unitat documental	OB	Text	Anna Murià i Romaní, escriptora
Data(es)	Dates d'inici i fi de la unitat de descripció. Inclou referència a dates aproximades i altres especificitats	Op	Text	1830-1999
Data inici	Data d'inici de la unitat de descripció	OB	Text	1937
Data fi	Data de fi de la unitat de descripció	OB	Text	2012
Volum (ml)	Metres lineals que ocupa el fons	OB	Text	17
Dipositant	Informa sobre qui ha formalitzat l'ingrés de la documentació al servei d'arxiu	OB	Text	Anna Murià
Gènere dipositant	Informa del gènere del diposita: Dona, Home, Mixt o Organisme	OB	Taula validada	Dona
Data ingrés	Informa de la data del primer ingrés efectiu	OB	Data	09/11/1999
Modalitat d'ingrés	Informa de la modalitat d'ingrés de la documentació	OB	Taula validada	Donació

Abast i contingut	Descripció general del fons	OB	Text	El fons conté documentació personal amb records, quaderns de notes, documents identificatius, certificats, autoritzacions i d'homenatge, mecanoscrits d'obres literàries, correspondència i reculls de premsa. Completen el fons fotografies i documents audiovisuals, alguns objectes artístics, a més d'una interessant bibliografia literària (...)
Condicions d'accés	Informa sobre la condició d'accés a la documentació, pot ser lliure o restringit	OB	Taula validada	Lliure
Inventariat	Informa sobre el nivell d'inventariat del fons. La seva consignació pot ser Sí, No o Parcial	OB	Taula validada	Parcial
Instruments de descripció	Enllaç a la fitxa descriptiva d'Arxius en Línia	RE	Hipervincle	Fitxa Arxius en Línia: Anna Murià i Romaní, escriptora
Documentació relacionada	Informa sobre fons relacionats o complementaris al fons que es descriu	RE	Text	És important consultar el Fons d'Agustí Bartra que també es conserva en aquest arxiu (CAT ACVOC AHT Fons Agustí Bartra, 13/22 i fons Jaume Canyameras i Cortázar 13/56).

Taula 3. Productor(es)

Element	Definició	Consignació	Tipus dada	Exemple
Identificador fons	Combina el codi d'arxiu i el codi fons, de tal manera que serveix de clau per relacionar les diferents taules	OB	Text	90-15
Nom productor/a	Identifica el/la productor/a del fons	OB	Text	Anna Murià i Romaní
Gènere del/la productor/a	Informa el gènere de la persona física que ha produït, en la seva totalitat o en part, el fons documental	OB	Taula validada	Dona

Municipi de naixement	Municipi on ha nascut la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	Barcelona
Municipi de mort	Municipi on ha mort la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	
Any de naixement	Any de naixement de la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	1904
Any de mort	Any de mort de la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	2002
Activitat professional	Enumera les diferents activitats professionals realitzades per la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	Esriptora
Vincles entitats	Enumera els principals organismes amb els quals s'ha vinculat a la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	Institut de les Lletres Catalanes; Diari de Catalunya

Taula 3. Productor(es)

Element	Definició	Consignació	Tipus dada	Exemple
Identificador fons	Combina el codi d'arxiu i el codi fons, de tal manera que serveix de clau per relacionar les diferents taules	OB	Text	90-15
Nom productor/a	Identifica el/la productor/a del fons	OB	Text	Anna Murià i Romaní
Gènere del/la productor/a	Informa el gènere de la persona física que ha produït, en la seva totalitat o en part, el fons documental	OB	Taula validada	Dona
Municipi de naixement	Municipi on ha nascut la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	Barcelona
Municipi de mort	Municipi on ha mort la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	

Any de naixement	Any de naixement de la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	1904
Any de mort	Any de mort de la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	2002
Activitat professional	Enumera les diferents activitats professionals realitzades per la persona física que ha produït, en la seva totalitat o en part, el fons documental	RE	Text	Esriptora
Vincles entitats	Enumera els principals organismes amb els quals s'ha vinculat a la persona física que ha produït, en la seva totalitat o en part, el fons documental	Op	Text	Institut de les Lletres Catalanes; Diari de Catalunya

Taula 4. Unitats documentals

Element	Definició	Consignació	Tipus dada	Exemple
Identificador fons	Combina el codi d'arxiu i el codi fons, de tal manera que serveix de clau per relacionar les diferents taules	OB	Text	90-15
Nivell de descripció	Identifica el nivell d'organització al qual correspon la unitat que es descriu	OB	Taula validada	Unitat documental
Codi unitat	Codi que identifica unívocament la unitat documental	OB	Text	38
Nom unitat	Títol que dona nom a la unitat documental	OB	Text	Notes manuscrites i retalls d'A. Murià. Escrits de Mercè Ibarz.
Data(es)	Dates d'inici i fi de la unitat de descripció. Inclou referència a dates aproximades i altres especificitats	OB	Text	1998
Procedència	Identifica el productor del fons al qual fa referència la unitat, únicament en aquells casos en que hi ha més d'un individu	Op	Text	
Abast i contingut	Descripció general de la unitat de descripció	OB	Text	6 papers manuscrits, un retall de diari i dos impresos amb anotacions d'Anna Murià; tres mecanoscrits de Mercè Ibarz.

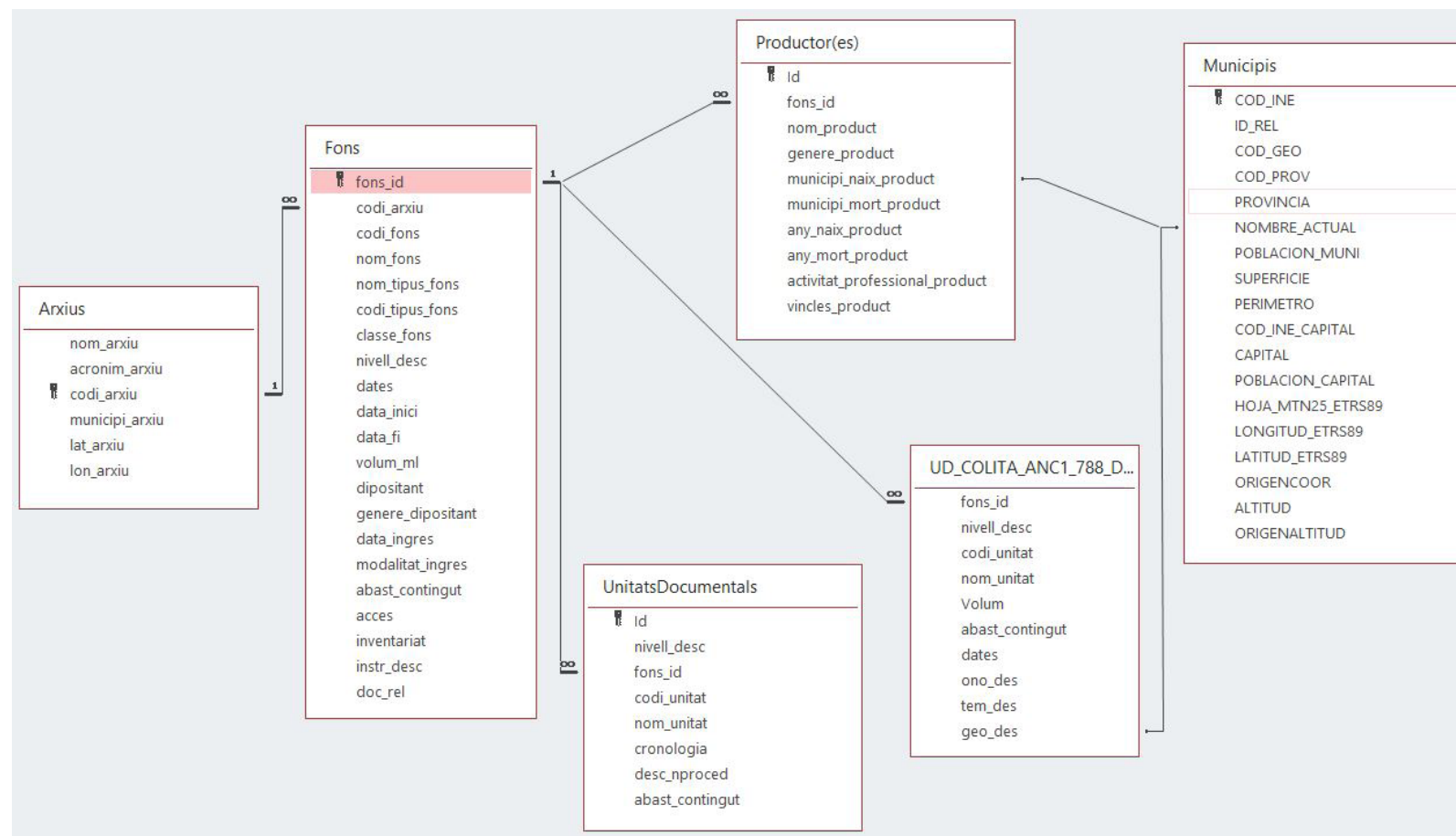


Figura 18. Esquema complet de les relacions definides al model de dades, incorporant un *dataset* extern. (Font: Elaboració pròpia)

8. 4. Annex 4. Vista prèvia de les graelles i metadades de cada *dataset*

105

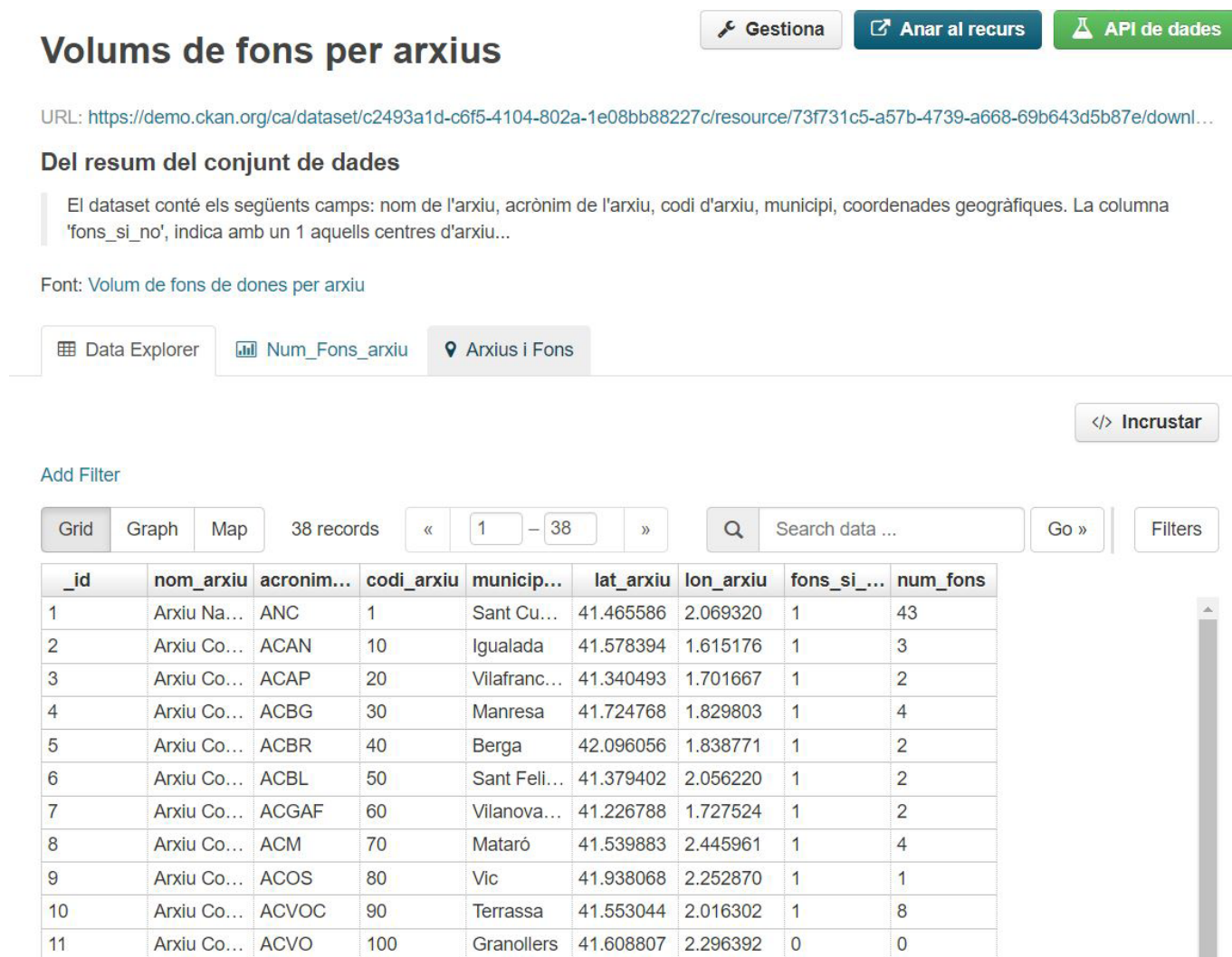


Figura 19. Vista prèvia de la graella generada per CKAN per al *dataset* 'Volums de fons per arxiu'. Les metadades que conté són: *Nom de l'arxiu*, *Acrònim de l'arxiu*, *Codi d'arxiu*, *Municipi de l'arxiu*, *latitud i longitud de l'arxiu*, *Fons sí o no* i *Número de fons*.

Guies de fons

[Gestiona](#)
[Anar al recurs](#)
[API de dades](#)

URL: <https://demo.ckan.org/ca/dataset/f4688e4d-961b-44b7-bd1f-62ea4cd0a691/resource/4bcd83af-33ac-4b45-9454-d6032b2dcf51/download>

Del resum del conjunt de dades

Aquest conjunt de dades conté tota la informació descriptiva a nivell de fons, de tots els fons treballats, tant de la Xarxa d'Arxius Comarcals com de l'Arxiu Nacional de Catalunya. Conté...

Font: [Guies de fons - Descripció a nivell de fons](#)

[Explorador de dades](#)
[Incrustar](#)

Add Filter

Grid	Graph	Map	96 records	«	1	–	96	»	Q	Search data ...	Go »	Filters
nom_arxiu	acronim...	codi_arxiu	lat_arxiu	lon_arxiu	codi_fons	nom_fons	nom_tip...	codi_tip...	classe_f...	nivell_d...		
Arxiu Na...	ANC	1	41.465586	2.069320	336	Col·lecci...	Personals	13	No textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	669	Pilar Llor...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	206	Dolors S...	Personals	13	No textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	914	Pilar Pér...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	928	Lidia Fal...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	950	Juan Orfi...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	901	Manuel ...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	993	Pilar Ay...	Personals	13	Imatge fixa	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	1007	Nan Gre...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	1061	Ana Min...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	1075	Alejandro...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	1080	Araceli R...	Personals	13	Textual	Fons		
Arxiu Na...	ANC	1	41.465586	2.069320	1105	Marta M...	Personals	13	Textual	Fons		

Figura 20. Vista prèvia de la graella generada per CKAN per al dataset 'Guies de fons'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, latitud i longitud de l'arxiu, Codi fons, Nom fons, Nom tipus fons, Codi tipus fons, Classe fons, Nivell descripció, Data(es), Data d'inici, Data de fi, Volum (ml), Dipositant, Gènere dipositant, Modalitat d'Ingrés, Data d'ingrés, Abast i contingut, Accés, Inventariat, Instrument de descripció i Documentació relacionada.

Productor(es)

[Gestiona](#)[Anar al recurs](#)[API de dades](#)

107



URL: <https://demo.ckan.org/ca/dataset/b9905736-be67-44fa-9726-9eb1e34f90fb/resource/3258ee4e-9cf9-44fb-b586-0abdc502a92d/download>

Del resum del conjunt de dades

Taula que incorpora les dades descriptives bàsiques dels(les) productors(es) de fons documentals, és a dir, els camps desagregats de l'element Història del(s) productor(s). També inclou...

Font: [Descripció - Productor\(es\)](#)

[Explorador de dades](#)[Incrustar](#)

Add Filter

Grid	Graph	Map	119 records	«	1	–	100	»	Q	Search data ...	Go »	Filters
codi_arxiu	codi_fons	nom_fons	nom_pr...	genere_...	municip...	municip...	any_nai...	any_mo...	activitat...	vincles_...		
1	336	Col·lecci...	Pilar Bay...	Dona	Vic		1941		Dibuixan...	Escola d...		
1	669	Pilar Llor...	Pilar Llor...	Dona	A Coruña	Sant Cu...	1929	1999	Ballarina...	Jove Ball...		
1	206	Dolors S...	Dolors S...	Dona								
1	914	Pilar Pér...	Pilar Pér...	Dona	Barcelona	Barcelona	1909	2008	Violoncel...	Conserv...		
1	928	Lidia Fal...	Lidia Fal...	Dona	Madrid		1935		Advocada	Gabinete...		
1	950	Juan Orfi...	Juan Orfi...	Home	Barcelona		1918		Professor			
1	950	Juan Orfi...	Teresa P...	Dona	Olot	Barcelona	1918	2006				
1	901	Manuel ...	Manuel ...	Home	Barcelona	Manresa	1906	1980	Metge; H...	Serveis ...		
1	901	Manuel ...	Montserrat...	Dona	Barcelona	Barcelona	1909	1983	Bibliotec...	Escola d...		
1	993	Pilar Ay...	Pilar Ay...	Dona	Barcelona		1943		Fotògrafa	Escola D...		
1	1007	Nan Gre...	Nan Green	Dona	Beeston		1904	1984	Traductora	Partit Co...		
1	1061	Ana Min...	Ana Min...	Dona			1916					
1	1075	Alejandr...	Alejandr...	Dona								
1	1075	Alejandr...	Julia Ro...	Dona			1908		Traducto...			

Figura 21. Vista prèvia de la graella generada per CKAN per al dataset 'Productor(es)'. Les metadades que conté són: *Nom de l'arxiu*, *Acrònim de l'arxiu*, *Codi d'arxiu*, *Codi fons*, *Nom fons*, *Nom productor(a)*, *Gènere productor(a)*, *Municipi de naixement*, *Municipi de mort*, *Any de naixement*, *Any de mort*, *Activitat professional* i *Vincles entitats*.

Productor(es) - Municipi de naixement

[Gestiona](#)
[Anar al recurs](#)
[API de dades](#)

URL: <https://demo.ckan.org/ca/dataset/9ea9bec4-10ad-4210-b9a3-b2267494c844/resource/7716d6ec-74a1-481a-b2c2-c25ae6323260/dow...>

Del resum del conjunt de dades

Taula que mostra informació bàsica de la història de les dones productores de fons documentals. Inclou el Nom de l'arxiu, l'Acrònim de l'arxiu, el Codi d'arxiu, el Codi fons, el Nom del...

Font: Productor(es) per municipi de naixement

[Explorador de dades](#)
[Origen de les dones productores dels fons](#)
[Incrustar](#)

Add Filter

ns	nom_fons	nom_pr...	genere_...	municip...	municip...	any_nai...	any_mo...	activitat...	vincles_...	LONGIT...	LATITU...
	Frances...	Frances...	Dona	Igualada		1900		Pianista;	Palau de...	1.61829...	41.571...
	Maria Ve...	Maria Ve...	Dona	Roda de ...	Figueres	1925	1990	Professora	Patronat ...	2.31045...	41.981...
	Maria Do...	Maria Do...	Dona	Sabadell		1950		Política;	Partit So...	2.10778...	41.541...
	Fons Àn...	Àngels T...	Dona	Sant Per...	Sant Per...	1940	2004	Professora	Escola T...	1.70331...	41.451...
	Rosa Pui...	Beneta T...	Dona	Terrassa	Terrassa	1913	2012	Professora	Escola ...	2.01215...	41.561...
	Rosa Pui...	Rosa Pui...	Dona	Terrassa	Terrassa	1901	1999	Sastress...	Gremi d...	2.01215...	41.561...
	Antònia ...	Antònia ...	Dona	Terrassa		1923	2012	Comptab...	Acadèmi...	2.01215...	41.561...
	Frances...	Frances...	Dona	Terrassa	Terrassa	1906	1982	Infermer...	Conserv...	2.01215...	41.561...
	Col·lecci...	Pilar Bay...	Dona	Vic		1941		Dibuixan...	Escola d...	2.25573...	41.921...
	Pilar Llor...	Pilar Llor...	Dona	A Coruña	Sant Cu...	1929	1999	Ballarina...	Jove Ball...	-8.39550...	43.371...
	Manuel ...	Esperan...	Dona	Arcas		1927				-2.11495...	39.981...

Figura 22. Vista prèvia de la graella generada per CKAN per al dataset 'Productor(es) – Municipi de naixement'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, Codi fons, Nom fons, Nom productor(a), Gènere productor(a), Municipi de naixement, Municipi de mort, Any de naixement, Any de mort, Activitat professional, Vincles entitats, Latitud i Longitud del municipi de naixement.

Fons ANC 1-788 COLITA

URL: <https://demo.ckan.org/ca/dataset/1a531a07-7a48-4341-940c-48f6a36aef3/resource/3908cabf-2809-419a-a4a3-b507bb63af68/download>

Del resum del conjunt de dades

Conjunt que inclou les dades descriptives de les unitats documentals del fons. Les coordenades geogràfiques s'han extret de la taula MUNICIPIOS del...

Font: Fons 1-788 COLITA

 Explorador de dades

 Geolocalització Unitats documentals

[</> Incrustar](#)

Add Filter

GridGraphMap

4615 records

«1–100»

Go »

Filters

ons	nom_fons	nivell_d...	codi_un...	nom_un...	abast_c...	dates	ono_des	tem_des	geo_des	LONGIT...	LATITU
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Ridruejo ...	família	Alella	2.29439...	41.4 ▲
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Ridruejo ...	escriptors	Alella	2.29439...	41.4
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Ridruejo ...	retrat	Alella	2.29439...	41.4
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Ridruejo ...	polítics	Alella	2.29439...	41.4
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Moix i M...	escriptors	Alella	2.29439...	41.4
	Colita	Unitat D...	526	Reportat...	Reportat...	1968	Antonio ...	flamenc	Altea	-0.05144...	38.5
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Moix i M...	família	Alella	2.29439...	41.4
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Moix i M...	retrat	Alella	2.29439...	41.4
	Colita	Unitat D...	148	Reportat...	Reportat...	agost 1973	Moix i M...	polítics	Alella	2.29439...	41.4
	Colita	Unitat D...	153	Reportat...	Reportat...	gener 1974	Vidal, Al...	pintors	Alella	2.29439...	41.4
	Colita	Unitat D...	153	Reportat...	Reportat...	gener 1974	Vidal, Al...	retrat	Alella	2.29439...	41.4
	Colita	Unitat D...	526	Reportat...	Reportat...	1968	Antonio ...	retrat	Altea	-0.05144...	38.5
	Colita	Unitat D...	864	Reportat...	Reportat...	[1979]	Joglars, ...	company...	Barcelona	2.17634...	41.3

Figura 23. Vista prèvia de la graella generada per CKAN per al dataset ‘Fons ANC 1-788 COLITA’. Les metadades que conté són: *Nom de l’arxiu*, *Acronim de l’arxiu*, *Codi d’arxiu*, *Codi fons*, *Nom fons*, *Nivell de descripció*, *Codi unitat*, *Nom unitat*, *Abast i contingut*, *Data(es)*, *Descriptor Onomàstic*, *Descriptor Temàtic*, *Descriptor Geogràfic*, *Latitud i Longitud*.

9. Índex de taules i figures

110



Taula 1. Categorització dels formats de fitxers per a dades obertes. (Font: <https://www.europeandataportal.eu/>)...18

Taula 2. Selecció dels elements de descripció per a l'obertura de dades de Fons patrimonials i familiars, Fons personals i Col·leccions, amb la corresponent definició i consignació. (Font: Elaboració pròpia)59

Taula 3. Exemple de descripció d'un fons personal segons l'esquema proposat en aquest projecte. (Font: Elaboració pròpia a partir de les dades descriptives del fons de l'ANC)64

Taula 4. Exemple de descripció d'una unitat documental segons l'esquema proposat en aquest projecte. (Font: Elaboració pròpia a partir de les dades descriptives del fons de l'ACAE).....65

Taula 5. Taula on es mostren els diferents elements NO-DAC i el número de fons que en tenen alguna informació publicada a Arxius en Línia. (Font: Elaboració pròpia)...99

Figura 1. Esquema de cinc estrelles per a la publicació de dades obertes proposat per Tim Berners-Lee. (Font: <https://5stardata.info/es/>)19

Figura 2. Detall del clúster de Barcelona. Mapa Origen de les persones executades a Catalunya a través d'un sumariísim. (Font: Instamaps. ICGC. Mapa elaborat per l'ANC) 26

Figura 3. Visualització del mapa Origen de les persones executades a Catalunya a través d'un sumariísim. (Font: Instamaps. ICGC. Mapa elaborat per l'ANC) 27

Figura 4. Exemple de visualització de 'Cerca per inventari' a Arxius en Línia49

Figura 5. Visualització del camp <descripció> abans del tractament inicial, amb els elements descriptius unificats en una única cel·la.53

Figura 6. Visualització del camp <descripció> després del tractament inicial, amb els elements descriptius separats i distribuïts.54

Figura 7. Gràfic que mostra el percentatge d'elements NO-DAC que s'informen i mostren, del total de 96 fons treballats, a les guies de fons disponibles a Arxius en Línia. Es distingeix entre el total de fons de l'ANC (43) i el total de

fons del XAC (53). La línia de color gris mostra el percentatge total dels 96 fons treballats. (Font: Elaboració pròpia) .
.....56

Figura 8. Esquema de les relacions entre les diferents taules del dataset. (Font: Elaboració pròpia)72

Figura 9. Visualització d'un gràfic generat amb CKAN que mostra el volum de fons custodiat per cada arxiu. L'eix X fa referència al codi d'arxiu, mentre que l'eix Y indica el número de fons custodiats en cada cas.78

Figura 10. Visualització del mapa generat amb CKAN on es georeferencien únicament els arxius que sí custodien fons de dones. A la miniatura es poden veure les metadades associades a cada punt.79

Figura 11. Gràfic amb el número de dipositants segons el seu gènere. (Font: Elaboració pròpia) 81

Figura 12. Gràfic que combina dades referents a les modalitats d'ingrés dels fons i els anys en els quals s'han fet efectius. (Font: Elaboració pròpia)82

Figura 13. Gràfic que combina el règim d'accés als fons amb el nivell de tractament que han rebut. (Font: Elaboració pròpia)83

Figura 14. Gràfic on es representen les diferents professions i activitats que han desenvolupat les dones dels fons treballats. (Font: Elaboració pròpia)85

Figura 15. Visualització d'un mapa generat a CKAN, en el qual es geolocalitzen els municipis de naixement de les productores, únicament dones nascudes a territori espanyol, dels fons treballats.....87

Figura 16. Visualització d'un mapa generat a CKAN, en el qual es geolocalitzen les unitats documentals del fons ANC 1-788 COLITA.....89

Figura 17. Vista de la pantalla d'inici de l'organització TFM_FonsDones creada a CKAN, amb els cinc conjunts de dades en els tres formats en què s'han publicat.90

Figura 18. Esquema complet de les relacions definides al model de dades, incorporant un dataset extern. (Font: Elaboració pròpia) 104

Figura 19. Vista prèvia de la graella generada per CKAN per al dataset 'Volums de fons per arxiu'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, Municipi de l'arxiu, latitud i longitud de l'arxiu, Fons sí o no i Número de fons. 105

Figura 20. Vista prèvia de la graella generada per CKAN per al dataset 'Guies de fons'. Les metadades que conté són:

Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, latitud i longitud de l'arxiu, Codi fons, Nom fons, Nom tipus fons, Codi tipus fons, Classe fons, Nivell descripció, Data(es), Data d'inici, Data de fi, Volum (ml), Dipositant, Gènere dipositant, Modalitat d'Ingrés, Data d'ingrés, Abast i contingut, Accés, Inventariat, Instrument de descripció i Documentació relacionada..... **106**

màstic, Descriptor Temàtic, Descriptor Geogràfic, Latitud i Longitud..... **109**

Figura 21. Vista prèvia de la graella generada per CKAN per al dataset 'Productor(es)'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, Codi fons, Nom fons, Nom productor(a), Gènere productor(a), Municipi de naixement, Municipi de mort, Any de naixement, Any de mort, Activitat professional i Vincles entitats. ..**107**

Figura 22. Vista prèvia de la graella generada per CKAN per al dataset 'Productor(es) – Municipi de naixement'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, Codi fons, Nom fons, Nom productor(a), Gènere productor(a), Municipi de naixement, Municipi de mort, Any de naixement, Any de mort, Activitat professional, Vincles entitats, Latitud i Longitud del municipi de naixement..... **108**

Figura 23. Vista prèvia de la graella generada per CKAN per al dataset 'Fons ANC 1-788 COLITA'. Les metadades que conté són: Nom de l'arxiu, Acrònim de l'arxiu, Codi d'arxiu, Codi fons, Nom fons, Nivell de descripció, Codi unitat, Nom unitat, Abast i contingut, Data(es), Descriptor Ono-

- 1 [Open Knowledge Foundation](#) [Consulta: 8 de juny de 2019]
- 2 Open Knowledge Foundation. The Open Data Handbook. *¿Qué son los datos abiertos?* Recuperat de: <http://opendatahandbook.org/guide/es/what-is-open-data/> [Consulta: 8 de juny de 2019]
- 3 Sunlight Foundation. *The principles for opening up government information*. Recuperat de: <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/> [Consulta: 9 de juny de 2019]
- 4 Open Data Charter. *Carta Internacional de Datos Abiertos. Principios* Recuperat de: <https://opendatacharter.net/principles-es/> [Consulta: 9 de juny de 2019]
- 5 El document signat el 17 de juliol de 2018 per Ernest Maragall i Mira, Conseller d'Acció Exterior, Relacions Institucionals i Transparència, pel qual el Govern de la Generalitat de Catalunya s'adhereix a la Carta Internacional de Dades Obertes es pot consultar a: <https://drive.google.com/file/d/0B4pnyLgEJbI6OWFrZ0EyMFpmNHJQN3Y2OUZOXzJEdjFJRlUw/view> [Consulta: 9 de juny de 2019]
- 6 [ACORD GOV/154/2018](#), de 20 de desembre, pel qual s'aprova l'estratègia de dades obertes de la Generalitat de Catalunya i l'adhesió a la Carta Internacional de Dades Obertes. [Consulta: 9 de juny de 2019]
- 7 Open Knowledge Foundation. The Open Data Handbook. *¿Cómo abrir datos?* <http://opendatahandbook.org/guide/es/how-to-open-up-data/> [Consulta: 9 de juny de 2019]
- 8 [Datahub](#) [Consulta: 23 de juny de 2019]
- 9 Bennet, D. i Harvey, A. (2009). *Publishing Open Government Data*. World Wide Web Consortium. Recuperat de: <https://www.w3.org/TR/gov-data/> [Consulta: 9 de juny de 2019]
Aquesta guia de bones pràctiques és també interessant per a
- establir els passos a seguir per a obrir dades a les organitzacions. Descriu el procediment bàsic per fer-ho.
- 10 European Data Portal. *Why file formats matter for Open Data* Recuperat de: https://www.europeandataportal.eu/sites/default/files/data_formats_v1.pdf [Consulta: 9 de juny de 2019]
- 11 Tim Berners-Lee (2009). *The next web*. TED Ideas Worth Spreading. Recuperat de: https://www.ted.com/talks/tim_berners_lee_on_the_next_web#t-188536 [Consulta: 9 de juny de 2019]
Es tracta d'un TED Talk on el mateix Berners-Lee expressa què és el *Linked Data* i la importància de posar les dades al Web, i fer-ho en brut per afavorir les relacions entre elles.
- 12 És una idea similar a l'hipervincle utilitzat comunament fins a l'actualitat, però en aquest cas són els recursos els que es relacionen entre sí.
- 13 Cinc Estrelles Dades Obertes. Recuperat de: <https://5stardata.info/es/> [Consulta: 9 de juny de 2019] En aquest enllaç també s'especifiquen alguns dels costos i beneficis de la publicació de dades en obert a cadascun dels cinc nivells que proposa.
- 14 European Data Portal. *¿Cómo de útiles son mis datos?* Recuperat de: <https://www.europeandataportal.eu/elearning/es/module10/#/id/co-01> [Consulta: 23 de juny de 2019]
- 15 Un identificador és una forma abreviada de fer referència a una cosa més complexa. Serien identificadors un Codi Postal o el NIF d'una empresa.
- 16 [Directiva 2003/98/CE](#) del Parlament Europeu i del Consell, de 17 de novembre de 2003 relativa a la reutilització de la informació del sector públic. [Consulta: 25 de juny de 2019]
- 17 [Directiva 2013/37/UE](#) del Parlament Europeu i del Consell, de 16 de juny de 2013, per la qual es modifica la Directiva 2003/98/CE relativa a la reutilització de la informació del sector públic. [Consulta: 25 de juny de 2019]

- 18 [Directiva 2019/1024](#), de 20 de juny de 2019, del Parlament Europeu i del Consell, de dades obertes i reutilització d'informació del sector públic [Consulta: 27 de juny de 2019]
- 19 [Llei 37/2007](#), de 16 de novembre, sobre la reutilització de la informació del sector públic. Text consolidat. [Consulta: 25 de juny de 2019]
- 20 [Llei 18/2015](#), de 9 de juliol, per la qual es modifica la Llei 37/2007, de 16 de novembre, sobre reutilització de la informació del sector públic. [Consulta: 25 de juny de 2019] Aquesta llei incorpora les modificacions que s'introdueixen amb la Directiva 2013/37/UE
- 21 [Llei 19/2013](#), de 9 de desembre, de Transparència, Accés a la Informació Pública i Bon govern [Consulta: 25 de juny de 2019]
- 22 [Llei 19/2014](#), de 29 de desembre, de Transparència, Accés a la Informació Pública i Bon govern [Consulta: 25 de juny de 2019]
- 23 [Reial Decret 4/2010](#), de 8 de gener, pel qual es regula l'Esquema Nacional d'Interoperabilitat en l'àmbit de l'Administració Electrònica [Consulta: 26 de juny de 2019]
- 24 [Resolució de 19 de febrer de 2013](#), de la Secretaria d'Estat d'Administracions Públiques, per la qual s'aprova la Norma Tècnica d'Interoperabilitat de Reutilització de recursos de la informació. [Consulta: 26 de juny de 2019]
- 25 [Llei 20/2015](#), de 29 de juliol, de modificació de la Llei 10/2001, de 13 de juliol, d'arxius i documents. [Consulta: 25 de juny de 2019]. Text consolidat de la [Llei 10/2001](#), de 13 de juliol, d'arxius i documents. [Consulta: 25 de juny de 2019]
- 26 The home of the U.S. Government's open data. *Data Catalog* Recuperat de: <https://catalog.data.gov/dataset> [Consulta: 26 de juny de 2019]
- 27 Open Government at the National Archives. *Archival Descriptions from the National Archives Catalog*. Recuperat de: <https://www.archives.gov/open/dataset-arc.html> [Consulta: 26 de juny de 2019]

En aquest apartat també trobem un document amb l'esquema de metadades utilitzat per fer les descripcions: *Elements Used to Describe Record Groups, Collections, and Archival Materials*. Recuperat de: <https://www.archives.gov/files/open/data/2016/das-xml-mapping.pdf> [Consulta: 26 de juny de 2019]

- 28 The National Archives. *Discovery*. Recuperat de: <https://discovery.nationalarchives.gov.uk/advanced-search> [Consulta: 26 de juny de 2019]
- 29 [Archivo de la Ciudad de Arganda del Rey](#) [Consulta: 26 de juny de 2019]
- 30 [Open Archives Initiative – Protocol for Metadata Harvesting](#) [Consulta: 26 de juny de 2019]
- 31 Dades Obertes Generalitat de Catalunya. *Llista de reparació jurídica de víctimes del franquisme (1938-1978)* Recuperat de: <https://analisi.transparenciacatalunya.cat/Mem-ria/Llista-de-reparaci-jur-dica-de-v-ctimes-del-franqu/3bjt-k7vu> [Consulta: 27 de juny de 2019]
- 32 A les figures 2 i 3 d'aquest projecte es poden veure visualitzacions d'un mapa creat a partir de les dades d'aquest *dataset*.
- 33 Ajuntament de Girona. Servei de Gestió de Documents, Arxiu i Publicacions. *Guies de fons: dades obertes*. Recuperat de: http://www.girona.cat/sgdap/cat/dades_obertes.php [Consulta: 26 de juny de 2019]. Podem trobar l'explicació bàsica de què es pot trobar en aquest apartat, el format en què es presenten les dades, el perquè i com utilitzar-les. En el moment de consultar aquesta informació al web s'indica que l'arxiu custòdia un total de 128 fons, dels quals ofereix en obert dades bàsiques de tots ells, 27 inventaris i 81 catàlegs.

- 34 Ajuntament de Girona. Servei de Gestió de Documents, Arxiu i Publicacions. *Llibre de Sindicat de Remença 1448*. Recuperat de: http://www.girona.cat/sgdap/cat/recurs_remences_dades.php [Consulta: 26 de juny de 2019]. En aquest enllaç es poden consultar els diferents *datasets*, inclòs el referent a 'Dones del Sindicat Remença' amb un total de 113 referències. En tots, clicant sobre 'Més informació+' es pot veure l'estructura del fitxer i una breu descripció del contingut de cada camp.
- 35 Dirección General de Estadística y Censos. Gobierno de la Ciudad de Buenos Aires. *Sistema de Indicadores de Género BA*. Recuperat de: <https://www.estadisticaciudad.gob.ar/si/genero/principal> [Consulta: 26 de juny de 2019]
- 36 [Datos 2x](#) [Consulta: 26 de juny de 2019]
- 37 [Open Heroines](#) [Consulta: 26 de juny de 2019]
- 38 Open Government Partnership. [Gender](#) [Consulta: 26 de juny de 2019]
- 39 European Institute for Gender Equality. *Gender Equality Index*. EU Open Data Portal. Recuperat de: <https://data.europa.eu/euodp/data/dataset/gender-equality-index> [Consulta: 26 de juny de 2019]
- 40 NUNES, J. Institut Cartogràfic i Geològic de Catalunya, 2013. *Diccionaris*. Recuperat de: <http://www.icgc.cat/Ciutada/Informa-t/Diccionaris> [Consulta: 8 de juliol de 2019] En aquest enllaç es pot accedir a les definicions completes dels termes 'geolocalització', 'georeferenciació' i 'geocodificació' aquí referenciats, així com a altres relacionats amb la cartografia, fotogrametria, sistemes d'informació geogràfica o teledetecció.
- 41 Archivo Histórico de Protocolos de Gipuzkoa. *Mapa de fondos notariales de Gipuzkoa*. Recuperat de: <http://oinati.gipuzkoakultura.net/fondos/10/41?phpMyAdmin=7wkhbX-qHew9zde79NPtG7Ylt8ac> [Consulta: 9 de juliol de 2019]
- 42 Archivo Histórico de Protocolos de Gipuzkoa. *Mapa de las contadurías de hipotecas de Gipuzkoa*. Recuperat de: <http://oinati.gipuzkoakultura.net/fondos/27-fondos/161-mapa-de-las-con-tadurias-de-hipotecas-en-gipuzkoa.html> [Consulta: 9 de juliol de 2019]
- 43 Archivo de Arganda del Rey. Mapas. *Planos parcelarios del municipio de Arganda del Rey*. Recuperat de: <http://archivo.ayto-arganda.es/Mapas.aspx> [Consulta: 9 de juliol de 2019]
- 44 Entraríem aquí en un debat al voltant del concepte de productor que s'aborda amb més deteniment a l'apartat 3.2.2 *Els fons* d'aquest treball.
- 45 [Diccionari Biogràfic de Dones](#). [Consulta: 23 de juliol de 2019]
- 46 Perpinyà, R. (2016, 8 de juny). Dones als arxius: una presència invisible. *Núvol, el digital de cultura*. Recuperat de: <https://www.nuvol.com/opinio/dones-als-arxius-una-presencia-invisible/> [Consulta: 01/06/2019]
- 47 Perpinyà, R. (2019, 29 de maig). Els arxius han de ser feministes. *La DaDa. El diari digital de l'Associació d'Arxivers i Gestors de Documents de Catalunya*. Recuperat de: <http://arxivers.com/la-dada/els-arxius-han-de-ser-feministes/> [Consulta: 01/06/2019]
- 48 Rubio, C. (2019, 29 de maig). Què passa amb la memòria de les dones? *Núvol, el digital de cultura*. Recuperat de: <https://www.nuvol.com/noticies/que-passa-amb-la-memoria-de-les-dones/> [Consulta: 1 de juny de 2019]
- 49 Arxiu Històric de Terrassa i Arxiu Comarcal del Vallès Occidental. *Compromís de l'Arxiu amb la perspectiva de gènere*. Recuperat de: <http://xac.gencat.cat/ca/detalls/Noticia/Protocol-de-protocol-en-perspectiva-de-genere> [Consulta: 2 de juny de 2019]
- 50 [#Cuéntalo](#) [Consulta: 2 de juny de 2019]
Alguns articles que parlen de la iniciativa del projecte, el procés de tractament i visualització de les dades: Ruíz, V. i Maria, A. (2018, 25 de juliol). Reportatge: #Cuéntalo: preservació, tractament i pluralització d'un fons documental social. *La DaDa. El diari digital de l'Associació d'Arxivers i Gestors de Documents de Catalunya*. Recuperat de: <https://arxivers.com/la-dada/reportatge-cuentalo-preservacio-tractament-i-plura->

- [litzacio-dun-fons-documental-social/](#) [Consulta: 2 de juny de 2019]
The BSC Viz Corner – A visualization Blog (2018, 12 de desembre). #Cuéntalo – La Visualitzación. Recuperat de: <http://www.bsc.es/viz/corner/?p=210&lang=es> [Consulta: 2 de juny de 2019]
- 51 The National Foundation for Australian Women (NFAW); The University of Melbourne. Australian Women's Archives Project. *The Australian Women's Register*. Recuperat de: <https://www.womenaustralia.info/> [Consulta: 25 de juny de 2019]
 - 52 Principalment l'Arxiu Municipal de Barcelona, l'Arxiu Municipal de Girona, i aquells que centralitzen les seves descripcions a Arxius en línia.
 - 53 Generalitat de Catalunya. Departament de Cultura. *Arxius en línia* Recuperat de: <http://arxiusenlinia.cultura.gencat.cat/ArxiusEnLinia/> [Consulta: 28 de juny de 2019].
 - 54 Es tracta de la traducció a l'anglès promoguda per *Archivaria* del text originalment publicat a *La Gazette des Archives* 97 (1977), i sobre el qual es basa també en gran mesura el mètode adoptat per l'ANC per a establir el mètode d'identificació dels seus fons de l'administració.
 - 55 Arxius en Línia. Fitxa del fons [ACBL50-19 Consultorio para la mujer Elena Francis](#) [Consulta: 03 de juliol de 2019]
 - 56 A la sol·licitud de dades en formats reutilitzables també s'han inclòs alguns fons comercials i d'empreses i alguns d'associacions i fundacions, per tal d'analitzar el seu contingut i valorar la idoneïtat i coherència amb els objectius del projecte. Finalment, no s'han utilitzat. És per aquest motiu que s'han rebut dades de més fons dels que s'han acabat treballant a la proposta.
 - 57 Aquest camp <desc_nproced> figura a les unitats documentals facilitades per l'ANC i està degudament complimentada en el casos que procedeix. Per exemple, al Fons 1-26 Carles Riba – Clementina Arderiu, en aquest camp s'identifica si la unitat documental que es descriu està directament relacionada amb Carles Riba o amb Clementina Arderiu.
 - 58 A l'hora de fer les descripcions s'utilitzen etiquetes que identifiquen cada [Element Nodac], per tal que en el moment de fer la migració a Arxius en Línia aquestes siguin detectades. En fer l'exportació, aquestes elements s'han concatenat i unificat en una sola cel·la, al camp <descripció>.
 - 59 A les exportacions que s'han rebut, el codi de referència (codi d'arxiu i codi de fons), el títol del fons i les dates, disposen d'un camp per a cada element, independent del camp <descripció>. Per aquest motiu, en aquest primer recompte veiem que hi ha fons que no disposen de contingut descriptiu però, tot i això, tenen la majoria de dades identificatives bàsiques informades.
 - 60 Vilanova Godoy, G. (2018) *#Comsomcomensveuen* Treball de recerca del Màster d'Arxivística i Gestió de Documents de l'Escola Superior d'Arxivística i Gestió de Documents. (Treballs fi de màster i de postgrau) <https://recercat.cat/bitstream/handle/2072/338658/Treball%20de%20recerca.pdf?sequence=1> [Consulta: 5 de juliol de 2019] En aquest treball, l'autor ofereix una panoràmica completa de la situació actual de la professió.
 - 61 El *Nom o títol del fons* acostuma a ser idèntic o, si més no, fa referència directa al *Nom del(s) productor(s)*. En tant que això és així, i per tal de donar major visibilitat a la dona com a subjecte productor, s'ha decidit treballar aquest element no tant en el sentit estricte del què la teoria arxivística defineix com a productor, sinó per identificar i treballar individualment les persones físiques que han produït un fons. Per aquest motiu, en aquest projecte, un mateix fons podrà tenir més d'un productor.
 - 62 S'indica que el gènere de qui fa l'ingrés és mixt en aquells casos en que l'ha formalitzat més d'una persona i hi ha tant dones com homes.
 - 63 Utilitzar els metres lineals com a mesura pot no ser el més comprensible per a l'usuari, però donada la informació disponible actualment, s'ha considerat que és l'única manera possible per donar una idea del volum total del fons, i que a més pugui tenir una utilitat en la creació de visualitzacions. Presentar les dades tal qual es consignen ara per ara en aquest element no ho permet.

64 L'Arxiu Municipal de Girona, en els catàlegs que ofereix en obert, inclou metadades com:

- *Classe de document* (imatge fixa, textual, audiovisual...)
- *Subclasse* (fotografia, manuscrit, pel·lícula, vídeo, sonor...)
- *Tipus de suport* (plàstic, vidre, pergamí, paper, òptic, magnètic...)
- *Format* (6x9, 13x18, casset...)
- *Subtipus de suport* (nitrat, acetat...)
- *Procediment* (gelatinobromur, gelatina DOP...)
- *Polaritat* (negatiu, positiu)
- *Color* (b/n, color)

Totes elles consignades a nivell d'unitat documental. Una proposta molt interessant, ja que no només dona informació detallada de cada unitat descrita, que pot ser reutilitzada i fàcilment analitzada, sinó que a partir d'aquestes dades és molt més senzill extrapolar el volum total d'unitats del fons. Permet saber de quina tipologia de documents es tracta i les seves característiques. Igualment, es facilita l'actualització i reescriptura de l'element *Volum i suport* a nivell de fons un cop aquest ha estat tractat en la seva totalitat.

Donat que a la sol·licitud de dades que es va fer expressament per aquest projecte es va demanar únicament aquella informació descriptiva difosa a Arxius en Línia, cal contemplar la possibilitat que a nivell intern els arxius aquí treballats puguin disposar també de metadades molt similars a les que ofereix l'Arxiu Municipal de Girona.

65 Assignar un codi a cada persona física facilitaria establir les relacions de varis a varis entre fons, unitats documentals i productors. Aquestes relacions s'expliquen amb més detall a l'apartat 4.5.3 *Les relacions entre les taules* d'aquest treball.

66 Disposar dels municipis de naixement i mort de les productores de fons permetrà geolocalitzar-les, per exemple, a Instamaps. S'ha de tenir en compte però, que no tots els programes de creació de mapes permeten la geolocalització per municipi, i es farà necessari obtenir les coordenades XY o fer ús d'altres tècniques de geocodificació. Igualment, haurem de tenir present que si ho fem així, instamaps només reconeix els municipis de Catalunya, de tal manera que per als casos de dones nascudes o mortes fora de Catalunya, aquesta estratègia no servirà i caldrà buscar una alternativa. És una casuística que en

aquest projecte es dona en molt pocs casos. Com es pot veure a l'apartat 5.2.1. *El procés de publicació a CKAN*, disposar dels municipis utilitzant vocabulari normalitzat, permet l'obtenció automàtica de coordenades mitjançant l'encreuament de *datasets* que sí que contenen aquesta informació geogràfica.

67 Desagregar aquestes dades, encara que hagi requerit un esforç important d'anàlisi i tractament, ha permès veure clarament com, encara que el camp estigui informat, en la majoria de casos falta molta informació. I no s'acostuma a especificar el desconeixement de les dades que manquen.

Pel que fa al camp *dipositant*, com s'ha comentat, les dades s'han extret de l'element *Dades sobre l'ingrés*. Ara bé, en el cas dels arxius comarcals en la majoria de fons no hi ha informació sobre el nom del dipositant. En canvi, a l'exportació facilitada per la Subdirecció hi ha un camp en el qual sí que hi ha aquesta informació. Així, per als fons d'arxius comarcals, s'han treballat els dos camps conjuntament per obtenir el màxim d'informació possible.

68 EAD és l'equivalent a ISAD(G), mentre que EAC és la codificació de la norma ISAAR(CPF).

69 A l'**Annex 2** es pot consultar la distribució dels elements de descripció a les diferents taules, la definició de cadascun d'ells, la consignació, el tipus de dades i un exemple.

70 Aquesta taula incorpora dades de tots els arxius treballats (XAC i ANC), independentment de si custodien o no fons de dones. Per elaborar aquesta taula, s'ha utilitzat el codi d'arxiu que s'ha facilitat a les exportacions rebudes. Ara bé, no tots els arxius custodien fons de dones, i en aquests casos s'ha hagut de buscar a Arxius en Línia i anotar manualment l'acrònim i el codi d'arxiu. En aquest sentit, cal destacar que:

- En el moment de preparar el model de dades, l'Arxiu Comarcal de la Terra Alta no està donat d'alta a Arxius en Línia, de tal manera que no se n'han pogut consultar cap fons ni tampoc conèixer el seu codi d'arxiu. No s'ha inclòs a la taula.
- L'Arxiu Comarcal del Baix Ebre, figura amb acrònims d'arxiu diferents: a Arxius en Línia (ACTE) i a la seva pàgina web (ACBEB). S'ha optat per identificar-lo amb el codi que apareix al web.

- 71 Hem de tenir present que, inclús si es fes el prototip únicament per a fons personals, la casuística seria la mateixa. Com hem pogut veure a l'anàlisi de les dades, alguns fons identificats com a personals inclouen documentació de dues o més persones, generalment matrimonis o germans. Així doncs, si volem distingir la documentació que han produït homes de la que han produït dones, cal donar la mateixa importància a un i altre, individualitzant-los com a punts d'accés al fons. Analitzant l'element *Història del(s) productor(s)* s'ha pogut comprovar que, en els fons personals que es dona aquesta condició, s'acostuma a redactar la història de cadascuna de les persones per separat.
- 72 Ara per ara, a la *Taula 4. Unitats documentals*, no s'hi han inclòs els camps per a descriptors, doncs aquests només s'han obtingut en aquest nivell per al fons ANC 1-788 COLITA. Per tal de veure les possibilitats que ofereixen els topònims, a l'apartat 5.2.1 *El procés de publicació a CKAN* s'ha treballat una taula addicional que incorpora únicament les descripcions d'unitats documentals del fons COLITA amb els seus descriptors. El resultat es pot veure a la **Figura 16**.
- 73 Creative Commons. *Sobre las licencias*. Recuperat de: https://creativecommons.org/licenses/?lang=es_ES [Consulta: 15 de juliol de 2019]
- 74 Open Data Commons. *Licenses*. Recuperat de: <https://opendatacommons.org/licenses/index.html> [Consulta: 15 de juliol de 2019]
- 75 **CKAN Demo** [Consulta: 15 de juliol de 2019]
CKAN és un *software open source*. Hi ha disponible una versió Demo que et permet fer-ne ús sense necessitat de disposar d'un servidor, doncs s'utilitza el seu. El funcionament és igual en qualsevol cas, però s'ha de tenir present que a la versió Demo cada cert temps s'esborren les dades. Si es vol evitar perdre la informació hi ha dues opcions: o bé pagar per disposar d'un servidor a les seves infraestructures, o bé instal·lar el *software* a un servidor propi.
- 76 A l'**Annex 4** es poden consultar la vistes prèvies de les graelles dels diferents *datasets* publicats.
- 77 Els gràfics es poden veure a les **Figures 11 a 13**. Els tres s'han creat a partir del mateix *dataset* i les dades s'han treballat mitjançant taules dinàmiques a Microsoft Excel.
- 78 Instituto Geográfico Nacional. Centro de Descargas. *Nomenclátor Geográfico de Municipios y Entidades de Población* (NGMEP). Recuperat de: <http://centrodedescargas.cnig.es/CentroDescargas/equipamiento.do?method=mostrarEquipamiento> [Consulta: 17 de juliol de 2019]
Es tracta d'un fitxer en format ZIP que conté la base de dades en formats MDB i ODB. A més, inclou un PDF amb la descripció del contingut de les taules i les seves relacions.
- 79 Es podrien utilitzar altres conjunts de dades com, per exemple, els que hi ha disponibles a Muncat i que es poden descarregar al següent enllaç: http://municat.gencat.cat/ca/Temes/els-ens-locales-de-catalunya/descarrega_fitxers/ [Consulta: 18 de juliol de 2019]
Malgrat tot, s'ha considerat més oportú el que ofereix Institut Geogràfic Nacional, doncs encara que l'abast del projecte sigui a l'àmbit autonòmic català, s'ha vist que no totes les dones que s'han treballat són nascudes aquí i, per tant, en permet una combinació més àmplia.
- 80 Aquesta relació es podria fer també amb el municipi de mort però, en aquest cas, per fer la prova s'ha fet únicament amb el de naixement. Hem de tenir present que quan es publiqui aquesta informació a CKAN s'haurà de fer referència a l'ús de la base de dades de l'IGN, doncs la reutilització de la mateixa es permet sota llicència oberta que requereix el reconeixement de la seva autoria.
- 81 A continuació s'especifiquen els canvis concrets que s'han fet: Calataiud (Calatayud), Canillas (Madrid), Ciutat de Mallorca (Palma), Sant Vicenç de Calders (El Vendrell), Atalaya de Villar del Saz de Arcas (Arcas), Corunya (A Coruña).
- 82 En aquest cas, els canvis que s'han fet per normalitzar els nom són els següents: Càdis (Cádiz), Castell i Platja d'Aro (Platja d'Aro), Escala, l' (L'Escala), Hospitalet de Llobregat, l' (L'Hospitalet de Llobregat), Màlaga (Málaga), Pamplona(Iruña) (Pamplona/Iruña), Segòvia (Segovia).

En els casos en què el topònim fa referència a la província, la unitat documental no quedarà reflectida ni al mapa ni al *dataset* publicat. També hi ha unitats en què el topònim informa d'altres països (França, Grècia, Egipte o Itàlia), els quals tampoc apareixeran en el mapa, doncs a banda de ser fora del territori espanyol, no es disposa d'informació a nivell de municipi.

- 83 A l'**Annex 3** es pot consultar l'esquema que mostra com s'ha incorporat i relacionat la taula MUNICIPIOS amb les taules *Fons ANC 1-788 COLITA* i *Productor(es)*.
- 84 953 fons personals entre l'ANC i els arxius del XAC, sense comptar els seus fons d'Arxiu de Complement. Cal tenir present que el recompte total de fons a Arxius en Línia s'ha fet una vegada acabat el projecte, de tal manera que és molt probable que en el temps transcorregut entre la sol·licitud de dades als arxius i el recompte que es fa aquí, s'hagin donat d'alta nous fons al cercador, quelcom que faria variar una mica el percentatge total.

Publicació conjunta



FUABformació
Escola d'Arxivística
i Gestió de Documents



Aquesta obra està subjecta a llicència Creative Commons Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya (<http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>). Es permet la reproducció total o parcial i la comunicació pública de l'obra, sempre que no sigui amb finalitats comercials, i sempre que es reconegui l'autoria de l'obra original. No es permet la creació d'obres derivades.