

Gestió de les dades obertes: models i protocols per al govern obert i reutilització

Montserrat Garcia Alsina

Aquest document és fruit d'una recerca promoguda a partir de la subvenció a treballs de recerca sobre l'Administració pública de l'Escola d'Administració Pública de Catalunya (Resolució GAH/590/2018, de 20 de març). Es presenta tal com va ser lliurat pels autors en la seva versió definitiva.

Data de lliurament de la versió definitiva: setembre de 2020

Memòria justificativa de recerca per a la convocatòria de concessió de subvencions a treballs de recerca sobre l'Administració pública (EAPC 2018)

La memòria justificativa consta de tres parts:

- 1.- Dades bàsiques i resums
- 2.- Document de síntesi
- 3.- Memòria del treball (informe científic)

1.- Dades bàsiques i resums

Nom de la convocatòria
EAPC 2018

Títol del projecte

Gestió de es dades obertes: models i protocols per al govern obert i reutilització

Dades de l'investigador/a responsable

| | |
|--------------------|---------------|
| Nom | Cognoms |
| Montserrat | Garcia Alsina |
| Adreça electrònica | |

Dades de la universitat / centre al qual s'està vinculat

Universitat Oberta de Catalunya

Número d'expedient

2018 EAPC 000017

Paraules clau: cal que esmenteu cinc conceptes que defineixin el contingut de la vostra memòria.
dades obertes, reutilització de la informació, sector infomediari, intel·ligència territorial,

Data de presentació de la justificació

02/12/2019

Nom, cognoms i signatura
de l'/la investigador/a responsable

Altres signatures, si escau



Montserrat Garcia Alsina

Resum en català (de 300 a 400 paraules)

El projecte té com a focus central les pràctiques que estan efectuant les administracions públiques per a gestionar dades obertes en el marc de la transparència i govern obert, però sobretot, en el marc de la reutilització de la informació com a generador de valor i motor de desenvolupament econòmic en els territoris. La reutilització de la informació promou la innovació i la presa de decisions internes. La recerca feta fins ara apunta que les pràctiques de gestió de dades obertes compta amb una dècada de desenvolupament. Tanmateix, la recerca feta fins ara identifica la necessitat de treballar més àmbits que formen part del projecte proposat: a) major sistemàtica de pràctiques, b) definició de models de gestió, c) identificació de les barreres que impedeixen la implementació i la reutilització de la informació per contribuir al desenvolupament econòmic, i d) models per avaluar la implementació de pràctiques de gestió de dades obertes. El projecte està basat en metodologies mixtes amb predomini de qualitatives, i ha tingut com objecte d'estudi les administracions públiques catalanes, concretament els ajuntaments, la Generalitat de Catalunya i els centres de recerca i universitats. La recollida de dades ha estat feta mitjançant entrevistes en profunditat per a tota la mostra, i enquestes addicionals adreçades als ajuntaments. Els resultats senyalen que encara queda per endavant camí per recórrer en molts dels ajuntaments, sobre tot aquells amb menys recursos econòmics. Reptes per endavant són molts, dels que destaquen els següents: 1) desenvolupar protocols que recullin de manera clara els processos i procediments que pautin la sistemàtica de les pràctiques i publicació de les dades obertes, i 2) impulsar el recolzament a nivell polític i directiu materialitzat en forma de recursos, formació i estructures organitzatives dedicades; 3) major vincle amb la indústria de reutilització per fomentar el desenvolupament econòmic; i 4) unir esforços per desenvolupar metadades i vocabularis comuns per afavorir la federació de dades.

Resum en anglès (màxim 400 paraules) **optatiu**

2.- Document de síntesi. (Ha de tenir 4.500 paraules, és a dir, unes 15 pàgines)
Cal fer constar l'objecte de la recerca i els principals resultats obtinguts.

DOCUMENT DE SÍNTESI

1. Objecte de la recerca

Aquesta recerca té com a objectiu principal conèixer el context i les pràctiques que s'estan duent a terme en les administracions públiques, inclosos els centres de recerca, per a gestionar les dades obertes. En concret aquest objecte s'ha concretat en els següents objectius de recerca.

- Conèixer les pràctiques de la publicació de dades obertes a Catalunya en els portals, tant de les administracions públiques com dels centres de recerca universitària, i com contribueixen a la transparència i la governança de la informació a nivell intern.
- Identificar els factors que faciliten i inhibeixen la publicació de dades en obert.
- Compilar indicadors per avaluar la qualitat de les dades obertes.
- Localitzar estàndards i models de gestió de dades obertes.
- Documentar bones pràctiques de gestió de dades obertes i dades de recerca a Catalunya i en països del nostre entorn.
- Establir aspectes que haurien de contenir els protocols de gestió de les dades obertes.

2. Principals resultats obtinguts

Aquest projecte de recerca ha estat una primera immersió en el back-office de les administracions públiques catalanes. Els resultats són una base per seguir fent recerca i conèixer en profunditat les pràctiques i els reptes que les administracions tenen per al ple desplegament de l'obertura de les dades, que redundi en l'impuls de la indústria de reutilització.

Aquest coneixement ahora contribueix a la reflexió per pensar en quines accions són necessàries per superar els reptes identificats per tal de seguir impulsant l'obertura de dades.

Els reptes identificats pel que fa a ajuntaments i Generalitat de Catalunya, també extrapolables a centres de recerca són:

- Necessitat de **definir processos i procediments** que abastin tot el **cicle de gestió de la dada**, per tal de garantir la seva integritat. Aquests procediments es podrien recollir en ordenances o ordres per tal d'impulsar la gestió de dades, de manera similar a com la llei de transparència ha impulsat la publicació d'informació.
- **Unificar esforços** en el desenvolupament de metadades homogènies de grups de datasets.

- Establir **vincles amb la indústria de reutilització** més forts que activin mecanismes de publicació proactiva de datasets, tenint en compte els requisits que precisa aquesta indústria per a l'elaboració eficient de productes i serveis. Aquest punt és rellevant atès que és un sector econòmic que contribueix al desenvolupament econòmic d'un territori. Aquest podria ser un àmbit especialment d'interès per aquells municipis amb disminució de població per manca d'ocupació. La identificació de fórmules d'atracció, podria ser una línia de treball vinculada amb les dades obertes.
- Treballar **quadres d'indicadors d'avaluació** per garantir la gestió i publicació de dades en termes de qualitat. Hi ha iniciatives a nivell nacional i internacional que han de ser explorats, perquè poden ser aprofitables. Una autoavaluació dels datasets oferts i com s'estan utilitzant facilita seguir avançant en la direcció adequada.
- Recollir les **necessitats formatives** en l'àmbit de les dades obertes i trobar la forma de planificar la formació i compartir el coneixement entre les diferents administracions públiques.
- Definir accions per impulsar un **canvi de cultura organitzativa** que fomenti l'obertura de dades, i deixi de banda el concepte de "propietat" de la dada a nivell departamental.

Pel que fa als centres de recerca i universitats senyalem que la manca d'estudis sobre aquest àmbit ha constituït una limitació i un repte per a identificar les àrees que poden ser les interlocutores per extreure informació de les pràctiques. Tot i així hem identificat interlocutors i hem identificat que als reptes senyalats abans podem afegir altres reptes.

- **La força motriu de la compartició de les dades en obert està constituïda pels finançadors i les polítiques públiques europees. Per tant, cal impulsar pràctiques i eines que ajudin a compartir les dades de recerca, com un tret propi de la cultura científica. En definitiva, cal promoure accions per generar una cultura de compartició de dades de recerca.**

En el cas de l'accés obert, es va iniciar una força motriu en les declaracions per l'accés obert a les publicacions científiques de Budapest (2001, Bethesda (2003) i Berlin (2003) i que va ser impresa pels investigadors. Tanmateix, en el cas de la compartició de dades de recerca l'evolució ha estat més de dalt a baix, començant, parlant de l'ecosistema europeu, de les institucions europees a través de les seves polítiques en els darrers 6-8 anys, però també en l'estratègia d'adaptació a la ciència oberta en la qual s'està treballant. Això té uns determinats efectes en la cultura científica i en l'adaptabilitat, visualitzant-se una visió burocratitzadora de la compartició de coneixements científics.

- **Existeix un marc conceptual diferent, on la transparència, la reutilització queden enrere de l'obligatorietat, el FAIR play (*Findable, Accessible, Interoperable and Reusable*) i les dinàmiques disciplinars. Per tant, caldria impulsar pràctiques de transparència i reutilització.**

Tan en les converses com en la literatura científica (informes i articles científics) es veu clarament que les aproximacions al fet OBERT té arrels diferents. En el cas de les universitats i centres de recerca, l'obertura de les dades està lligada a la ciència oberta i la ciència ciutadana. En el cas de les administracions públiques l'obertura sembla emergir com a evolució lògica de la transparència, les polítiques d'accés a la informació i la participació ciutadana i la reutilització de la informació. Creiem oportú una reflexió conjunta entre el marc general de l'administració pública, i la comunitat acadèmica per a compartir experiències, i també, sinergies, atès que algunes de les

dades que es poden emprar poden estar o bé en els portals de dades obertes o bé en els repositoris institucionals o temàtics, pensant en una lògica d'interoperabilitat que actualment no es planteja a les universitats. En aquest sentit, tot i que la ciència ciutadana esdevé un dels aspectes de la ciència oberta, no es considera en la part de compartició de dades de recerca, sovint molt especialitzada.

- **El més similar a un portal de dades obertes de les administracions públiques serien el conjunt de datasets en els repositoris digitals. Per tant, cal potenciar una política pública de repositoris de dades obertes, temàtiques o institucionals, fàcilment accessibles al món acadèmic i a la ciutadania en general.**

Els portals de transparència i els portals de recerca, entesos com a metadades del funcionament de les institucions, es treballen des d'unitats i equips diferents, de forma que no comparteixen dinàmiques ni visions relacionades amb la interoperabilitat comuna o bé la reutilització. Així, els repositoris tenen la lògica FAIR i s'estructuren a nivell de directoris internacionals com ara DOAR.

Els repositoris digitals en els centres de recerca estaven pensats inicialment per a documents tradicionals de la comunicació científica (tesis doctorals, articles, capítols de llibre, memòries). S'ha hagut d'adaptar-los com repositori per emmagatzemar datasets, i per a això ha calgut una feina de repensar metadades i millorar aspectes com la capacitat d'emmagatzematge, la validació tècnica, ètica i legal dels conjunts de dades i els permisos. La forma de resoldre no ha seguit la lògica d'un portal de dades obertes com els de les administracions públiques, sinó que s'ha aprofundir en una dinàmica ja existent relacionada amb l'accés obert als documents científics, on ha calgut considerar una nova escala. Per tant, caldrà revisar la plataforma d'emmagatzematge i accés a la informació en termes de cerca, accessibilitat, interoperabilitat i reutilització per a garantir que les dades siguin realment FAIR.

- **Els investigadors continuen tenint necessitats relacionades amb la gestió de les dades obertes.** Per tal, cal promoure la formació en temes com són:
 - Necessitats de formació
 - Necessitats d'incentius pel fet de posar actius en forma de dataset a disposició de la comunitat científica
 - Necessitat de polítiques clares referents a les dades obertes, no només a requeriment extern
- **A diferència de les dades obertes de l'administració, les comunitats acadèmiques disposen també dels seus propis repositoris digitals com a alternativa. Tot i així, el repte és crear un repositori comú pels centres e recerca.**

Per a les biblioteques universitàries – les encarregades de gestionar les dades obertes – ha estat més oportú plantejar-se i debatre si els repositoris eren l'eina adequada per als objectius de la difusió de dades de recerca que no pas crear-ne de nous. Tot i això, i tenint en compte la gran quantitats de dades que es poden necessitar, es planteja la creació d'un repositori consorciat de dades de recerca promogut pel CSUC. És un debat encara obert. Actualment, ens trobem amb un informe que en parla, i descriu molt adequadament els reptes eminentment tècnics que planteja. Són les institucions universitàries participants del CSUC les que han d'acabar veient la viabilitat i el full de ruta per a la creació del repositori.

Dins les universitats, són els serveis bibliotecaris els qui lideren de forma tradicional els aspectes de la ciència oberta, primer en l'accés obert a articles científics, i ara en l'accés a les dades de recerca. Així, els repositoris neixen amb una lògica de catàleg bibliotecari amb accés final al text complert.

- **Existeix un ecosistema de repositoris de dades de recerca ampli, incloent-ne de comercials, entesos com a model de negoci ampli, i repositoris internacionals com ara Zenodo. A més, hi ha asimetria entre disciplines. Per tant, el repte és la construcció d'eines que siguin útils i compatibles amb les dinàmiques que cada comunitat disciplinària està proveïnt.**

A diferència dels portals de dades obertes de les institucions, el ventall d'opcions de difusió de dades de recerca té tres àmbits. En primer lloc l'àmbit **institucional**, a nivell d'universitat, pensant en les dades generades pels investigadors; en segon lloc l'àmbit **europèu**, com ara Zenodo, per a donar servei sobretot a tots els projectes finançats per les institucions europees (ERC, Horizon 2020). Finalment, el tercer àmbit és el comercial, on s'estan desenvolupant nous productes en el marc de la nova estratègia de les grans empreses relacionades amb la comunicació científica (Elsevier, Springer Nature, American Chemical Society). Aquestes empreses volen oferir serveis globals d'acompanyament als investigadors en tots els estadis de la recerca. Un exemple és Figshare que ofereix algunes avantatges competitives, com ara el fet d'oferir DOIs als datasets dipositats en el seu servei. En aquest cas, no cal passar per processos de validació, cosa que pot permetre competitivitats asimètriques com succeeix per exemple en el cas de xarxes socials acadèmiques tipus ResearchGate en comparació amb els repositoris institucionals, que a més a més també perden visibilitat a efectes de Search Engine Optimization (SEO).

D'altra banda, l'asimetria entre disciplines fa difícil una visió homogènia dins d'una mateixa universitat, emergint una necessària sinergia entre professionals de disciplines independents del centre universitari

- **Les universitats i centres de recerca es troben construint les transicions a marxes forçades entre la necessitat d'infraestructures, recursos humans i formacions i un corpus de bones i males pràctiques. Per tant, cal construir un marc de treball que contempli de manera global necessitats dels usuaris investigadors, infraestructura, persones dedicades, usuaris, processos i procediments.**

Així hem pogut comprovar que mentre es fan estratègies a mig i llarg termini relacionades amb la creació d'infraestructures i la redacció de polítiques de dades obertes de tipus institucional, els investigadors tenen ja necessitats concretes i tècniques per a complir els requeriments dels finançadors. D'alguna forma, com ja hem dit, cal considerar l'efecte extern a l'hora de generar el canvi.

- **La reutilització no és el principal motor de la compartició, tant si es visualitza o no un possible valor econòmic de les dades. Per tant, cal homogeneïtzar objectius i sensibilitzar sobre la diversitat del valor de les dades obertes originada per la reutilització.**

Tot i que la finalitat darrera per part dels decisors de les polítiques europees és la reutilització, no és així per part d'aquells que les comparteixen. Aquests ho fan sovint per la necessitat de complir els requeriments dels finançadors. Això genera algunes disfuncions, i costa trobar exemples de datasets que no hagin estat requerits a ser compartits. La filosofia i estratègia FAIR ha de permetre en darrera instància la reutilització, però aquesta filosofia encara no està del tot establerta. Un dels aspectes que mostra que no es tenen del tot clars els aspectes relacionats amb la reutilització

és la manca d'un debat clar sobre si les llicències d'accés incorporen o no l'ús comercial. En la lògica activista de l'accés obert s'entenia que en quedava exclòs. Tanmateix aquest supòsit queda qüestionat per les noves accions lligades al Pla S (iniciativa per a l'accés obert de publicacions científiques iniciat per Science Europe). El Pla S senyala que també els articles científics han de permetre una utilització posterior comercial. Per tant, la reutilització planteja nous reptes a les dades de recerca. La xarxa temàtica Maredata (a nivell estatal) va redactar unes recomanacions dirigides als investigadors sobre com fer que les dades fossin FAIR.

- **Atesa la gran experiència en el treball en l'àmbit de l'accés obert, existeix una dinàmica de col·laboració en l'àmbit de les bibliotecaris universitaris. Per tant, cal mantenir i incrementar les facilitats per mantenir la col·laboració.**

Les biblioteques són les unitats dins les universitats que tenen principalment com a repte la gestió de les dades obertes. Experiències com els grups específics de REBIUN o el grup de treball del CSUC estan permetent treballar de forma eficient en l'aprenentatge comú. Aquest fet cal relacionar-ho amb la necessitat de formació que han experimentat els professionals dedicats a la gestió de dades. Aquests han hagut d'assolir competències i habilitats relacionades amb el que es coneix com a *data librarian* o bé *data curator*, per tal de donar la qualitat necessària als paquets de dades i a més assolir els requeriments FAIR. En el cas dels centres de recerca, són altres les unitats encarregades de la gestió de dades.

Compartir dades de recerca és un fet irreversible, i és d'esperar que l'obligatorietat lligada a projectes avanci en les actuacions de tots els finançadors de recerca, tan públics com privats. Així mateix, quan es pensa en els possibles usuaris de les dades posades a disposició, es pensa en un ecosistema tancat, compost per la resta de la comunitat científica i investigadora.

- **Cal associar un valor afegit al qui comparteix i també anar afavorint noves dinàmiques i polítiques d'avaluació de la recerca i els recercaires on donin valor a les pràctiques de compartició de dades**

La comunicació científica com a fet i l'avaluació de la qualitat i l'impacte de la recerca estan destinades a ser diferenciades. Per inèrcia, criteris tecnològics i facilitat en les mètriques ofertes per productes comercial com ara Web of Science, l'avaluació de projectes, grups de recerca, investigadors i institucions estan encara massa relacionats amb cada cop més un sol tipus de document, l'article científic, i un indicador preferent, el factor d'impacte. Les crítiques tradicionals (de fet, des de la pròpia creació) ha dut al món acadèmic a plantejar-se i replantejar-se com s'avalua la qualitat d'una recerca, i també a generar nous indicadors per a poder avaluar la ciència en un nou paradigma obert, i les dades de recerca haurien de ser uns elements de qualitat i distinció de la bona recerca.

3.- Memòria del treball (informe científic). Ha d'estar escrit en català. Ha de tenir una extensió mínima de 18.000 paraules (o 108.000 caràcters), aproximadament unes 60 pàgines. Pot incloure altres fitxers de qualsevol mena, no més grans de 10 MB cadascun d'ells.

MEMÒRIA JUSTIFICATIVA PER A PROJECTES DE RECERCA

Orientacions generals

TAULA DE CONTINGUTS

Contenido

| | | |
|---------|--|----|
| 1 | Introducció..... | 11 |
| 1.1 | Justificació de l'estudi..... | 13 |
| 1.2 | Objectius de recerca..... | 14 |
| 1.3 | Antecedents | 14 |
| 1.3.1 | <i>Pràctiques de gestió de dades</i> | 14 |
| 1.3.2 | <i>Factors que faciliten o inhibeixen les dades.....</i> | 16 |
| 1.3.3 | <i>Indicadors de qualitat.....</i> | 16 |
| 1.3.4 | <i>Estàndards.....</i> | 20 |
| 1.3.5 | <i>Dades de recerca en obert, un ecosistema d'informació diferenciat.....</i> | 21 |
| 2 | Metodologia..... | 25 |
| 2.1 | Selecció de la mostra | 26 |
| 2.1.1 | <i>Ajuntaments.....</i> | 26 |
| 2.1.2 | <i>Generalitat de Catalunya</i> | 26 |
| 2.1.3 | <i>Centres de Recerca i Universitats</i> | 31 |
| 2.1.4 | <i>Diputacions</i> | 32 |
| 2.2 | Recollida d'informació | 32 |
| 2.2.1 | <i>Experiències en gestió de dades obertes i identificació de portals de dades obertes</i> | 32 |
| 2.2.2 | <i>Focus group.....</i> | 33 |
| 2.2.3 | <i>Entrevistes.....</i> | 33 |
| 2.2.4 | <i>Enquestes.....</i> | 34 |
| 2.3 | Anàlisi de dades..... | 35 |
| 3 | Seguiment i avaluació..... | 36 |
| 4 | Resultats..... | 37 |
| 4.1 | Pràctiques per gestionar dades obertes | 37 |
| 4.1.1 | Estructures organitzatives..... | 37 |
| 4.1.1.1 | <i>Ajuntaments.....</i> | 37 |
| 4.1.1.2 | <i>Generalitat</i> | 38 |
| 4.1.1.3 | <i>Centres de recerca i universitats</i> | 39 |
| 4.1.2 | Recursos | 39 |
| 4.1.2.1 | <i>Ajuntaments.....</i> | 39 |
| 4.1.2.2 | <i>Generalitat</i> | 40 |
| 4.1.2.3 | <i>Centres de recerca i universitats</i> | 40 |
| 4.1.3 | Procediments | 42 |
| 4.1.3.1 | <i>Ajuntaments.....</i> | 42 |
| 4.1.3.2 | <i>Generalitat</i> | 43 |
| 4.1.3.3 | <i>Centres de recerca i universitats</i> | 43 |
| 4.1.4 | Reutilització..... | 44 |

| | | |
|---------|--|----|
| 4.1.4.1 | Ajuntaments..... | 44 |
| 4.1.4.2 | Generalitat | 45 |
| 4.1.4.3 | Centres de recerca i universitats | 46 |
| 4.2 | Factors que faciliten o inhibeixen les pràctiques | 47 |
| 4.3 | Indicadors de qualitat i reutilització | 48 |
| 4.4 | Estàndards emprats | 49 |
| 5 | Conclusions | 50 |
| 6 | Impacte previst..... | 55 |
| 6.1 | Impacte científic i socioeconòmic | 55 |
| 6.2 | Publicacions..... | 56 |
| 6.3 | Propostes d'actuacions | 56 |
| 7 | Referències | 58 |
| 8 | Annexos | 65 |
| 8.1 | Annex 1: Municipis de Catalunya i capitals de comarca | 65 |
| 8.2 | Annex 2: Llistat de portals de transparència de les universitats catalanes i dels instituts CERCA..... | 67 |
| 8.2.1 | Universitats..... | 67 |
| 8.2.2 | Centres de Recerca | 67 |
| 8.3 | Annex 3: Guió de les entrevistes | 70 |
| 8.3.1 | Ajuntaments, Generalitat de Catalunya i Diputacions | 70 |
| 8.3.2 | Universitats i Centres de recerca | 72 |
| 8.4 | Annex 4: Carta de consentiment..... | 74 |
| 8.5 | Annex 5: Enquestes | 75 |

1 Introducció

Les dades obertes són un dels pilars del govern obert, la transparència i les ciutats intel·ligents (Abella et al., 2014) atès que faciliten l'accés a la informació pública. El lliurament actiu o sota demanda d'informació i dades públiques en formats obertes està en la base de la transparència i la rendició de comptes de les administracions públiques (Catalunya, 2014; Janssen, 2012). Per tant, l'accés a la informació i el govern obert estan vinculats (Yannoukakou i Araka, 2014). L'administració pública a Estats Units, en el període legislatiu d'Obama, va focalitzar la gestió de les dades obertes en la transparència, col·laboració i participació ciutadana (Zuiderwijk i Janssen, 2014; Veljkovic, et al., 2014).

Tanmateix, el govern obert no és l'únic àmbit on s'apliquen les dades obertes. Aquestes, a més d'enfortir la democràcia i l'ètica, tenen potencial per generar desenvolupament econòmic (Ferro i Osella, 2013; Pabon et al., 2013; Gurstein, 2011) i innovar (Rohunen et al., 2014; Zuiderwijk et al., 2014a). Part important de l'impuls de l'economia és la gestió de les dades i la tecnologia que suporta la seva gestió i explotació (Pabon et al., 2013; Gurstein, 2011). Per tant, les dades obertes són la matèria primera de l'economia digital, que és considerada per part de la Unió Europea com un dels motors de generació de valor per crear ocupació i desenvolupament sostenible (Comissió Europea, 2011; Comissió Europea, 2014). De fet, hi ha estudis que identifiquen diversos models de negoci basats en la reutilització de dades (Ferro i Osella, 2013).

Aquesta matèria primera es posa a disposició de la ciutadania tant per part d'empreses com per part de les administracions i empreses públiques, hospitals, centres de recerca, i universitats per a què siguin reutilitzades. D'aquesta manera, la indústria de reutilització de la informació (administrativa i científica) impacta sobre l'economia digital al crear nous productes i serveis a partir del tractament de les dades obertes (ASEDIE, 2018; Comissió Europea, 2014; Abella et al., 2014; Weiskopft i Weng, 2013; Comissió Europea, 2011; Marcos-Martin i Soriano-Maldonado, 2011).

Hi ha un altre aspecte de la reutilització que és l'aplicació d'aquesta en el si de les mateixes organitzacions i administracions públiques. Les dades generades en el si de les mateixes administracions públiques poden ser reutilitzades per a la pròpia gestió i per generar coneixement sobre el territori de cara a la presa de decisions o el disseny d'estratègies o plans de desenvolupament territorial (Garcia et al., 2018).

Menció especial tenen les dades obertes generades en projectes de recerca, pel fet que quan es parla de transparència i govern obert, són les que queden més oblidades en tractar-se d'un ecosistema d'informació diferent. Tot i això, cal no oblidar que no per això deixen de tenir el seu espai en la indústria de reutilització. És un fet que la societat també pot disposar de les dades obertes resultants dels projectes de recerca realitzats pels centres de recerca i universitats. Efectivament, la gestió de dades, siguin o no massives (*big data*) afecta totes les disciplines científiques, amb un paper evident en l'àmbit de la salut (Groves et al., 2013), però també en ciències socials.

Podem recordar, per exemple, les implicacions de gestionar les dades procedents de les xarxes socials (Panger 2016). Metzler et al. (2016) en el seu estudi sobre les ciències socials computacionals conclou que aquestes dades ofereixen un gran potencial per als científics socials. No obstant això, la col·lecció reeixida i l'anàlisi rigorosa d'aquestes dades requereix noves habilitats, noves col·laboracions, nous mètodes d'investigació i noves eines computacionals. Les dades obertes en recerca poden ser interpretades, també, com un moviment que fa que les dades recopilades siguin també reutilitzables per a múltiples usos, inclosa la creació de valor econòmic (per exemple, a les ciutats intel·ligents), o per al control

dels pressupostos públics emprats per a finançar la recerca. Al destacar dades grans obertes podem parlar, per exemple, del moviment *open* en el món acadèmic.

Com més oberta sigui la informació, millor i més àmplia és la recerca. Com a exemple, destaquem la gran transcendència de compartir dades en el cas de la crisi de l'Ébola (Chretien et al., 2016). En aquest sentit, l'acadèmia està impulsant la promoció de la ciència oberta, i les dades d'investigació obertes formen part d'aquest procés. Així, hi ha un alt interès en les polítiques científiques per considerar la ciència oberta i la seva promoció. Per exemple, destaquem l'obligatori dipòsit de dades brutes de projectes de recerca en el programa Horizon 2020 (López-Borrull 2014). A més a més, la gestió de dades presenta importants reptes quant a la difusió dels conjunts de dades. Així, s'introdueix un nou actor en la comunicació científica format pels conjunts de dades que es presenten, ja sigui en forma de material complementari als articles o fins i tot en forma de *data papers*, articles publicats en revistes de dades (García-García et al., 2015).

Des del nostre punt de vista, creiem més oportú que mai una recerca que miri d'entendre i posar en relació dues visions de dades obertes que deslliguin el món acadèmic i el món de l'administració pública. En ambdós casos, parlem de dades generades per la pròpia administració (i la major part del món universitari en forma part), finançades la major part amb pressupost públic (regional, estatal o europeu) i amb consum d'aquestes dades per part d'altres actors del propi entorn (Canals, 2017).

En conseqüència, aquesta temàtica es pot estudiar des de diferents àmbits. Vinculat amb aquest projecte, destaquem tres. El primer, i el predominant o més conegut, és l'àmbit social i polític (democràcia, transparència, rendiment de comptes). En segon lloc, l'altre àmbit des del qual s'estudien les dades obertes és el de l'economia digital, basada en la reutilització de la informació pública. Per últim, destaquem el tercer àmbit que és el de la ciència oberta, basat en la publicació en obert de les dades resultants de la recerca.

A més, els dos primers àmbits estan relacionats. De fet, hi ha autors que consideren requisit imprescindible per a que hi hagi transparència que les dades es publiquin en formats que permetin la seva reutilització. Per tant, la reutilització acaba redundant en el govern obert i la participació ciutadana. En conseqüència, per a que les dades obertes compleixin la funció que s'espera d'elles, aquestes requereixen tenir una sèrie de característiques. Aquest és un dels temes que s'han de tenir en compte en tot estudi que es faci sobre les dades obertes.

La publicació de dades obertes sobre tot procedeix d'organismes de l'àmbit de l'administració pública. Per tant, en aquest marc és rellevant conèixer les polítiques de gestió de dades obertes per part d'aquestes administracions públiques, i en concret quins processos i procediments tenen establerts per tal de publicar en obert, no només per facilitar la transparència, sinó també la reutilització de les dades.

Tenint en compte aquest marc, en aquest apartat presentem la justificació de l'estudi, els objectius de recerca i els antecedents de la literatura.

1.1 Justificació de l'estudi.

Les pràctiques de gestió de dades per ser efectives haurien d'estar regides per directrius plasmades en processos i procediments específics per tal d'assegurar que les dades compleixin les característiques requerides i sistematitzar la seva gestió. Aquests processos haurien de definir polítiques d'ús i privacitat (Abella et al., 2014). Tanmateix, són pocs els indicis de com les administracions públiques estan organitzant-se per gestionar les dades obertes, i garantir publicar-les regides per paràmetres de qualitat.

Tenint en compte això, l'estudi ve justificat pels següents quatre aspectes:

En primer lloc, la implantació de la llei de transparència a Catalunya per part de les administracions públiques ha estat centrada en la publicitat activa influïda per la necessitat de combatre la corrupció. El resultat ha estat publicar en formats diversos, que predominantment no són formats oberts i, per tant, difícilment reutilitzables (Garcia et al., 2018). A més, la publicació de dades en obert ha de protegir les dades personals de manera segura i fiable, per tal de complir la llei de protecció de dades (Espanya, 2018). En conseqüència, cal desenvolupar protocols i models que fomentin les pràctiques en el si de les administracions públiques per augmentar la publicació de dades, informació i documents en formats oberts d'acord a la llei de transparència, i la llei de reutilització de la informació pública. Aquest aspecte està poc treballat, més enllà de la diversitat de metodologies i processos definits per diferents organismes implicats en la publicació de dades en obert (Kučera, 2014).

En segon lloc, les carències en la publicació de formats oberts repercuteixen en el desenvolupament de la indústria infomediària dedicada a reutilitzar la informació a Catalunya. D'acord a l'Informe de l'Asociación Multisectorial de la Información (ASEDIE), Catalunya és la segona comunitat autònoma a Espanya amb empreses infomediàries (un total de 127, amb 5 més de les identificades en el 2018), per darrera de la Comunitat de Madrid (amb 256, amb 4 més que el 2018) (ASEDIE, 2019). Per tant, sembla que el foment de la publicitat de dades obertes, ja sigui en el portal de la transparència o en el portal de dades obertes de la administració pública, està potenciant la indústria de la reutilització a Catalunya. Tot i així, ens falta conèixer si el potencial de les dades obertes a Catalunya encara es pot desenvolupar més.

En tercer lloc, les agències finançadores de projectes de recerca amb diners públics requereixen als centres de recerca que publiquin en formats oberts les dades generades en el treball de camp, per tal que puguin ser també reutilitzades. A més, la nova directiva de dades obertes i reutilització de la informació impulsa molt més l'ús de les dades resultants de projectes de recerca (Unión Europea, 2019). Per tant, aquestes dades també han de ser gestionades per garantir la seva fiabilitat, autenticitat, usabilitat i la seva accessibilitat. Sovint els portals de dades obertes de l'administració es separen de la gestió de dades obertes fruit de la recerca universitària, quan totes requereixen els mateixos tractaments i formats. A més, són escassos els estudis sobre la gestió de dades resultants de la gestió a les universitats. És a dir, haurien de ser objecte de gestió i de recerca no només les dades resultants dels projectes de recerca, sinó també les resultants de la gestió administrativa. Per aquest motiu, les incloem en aquest projecte de recerca.

Per últim, una bona gestió de les dades en el si de les administracions públiques pot repercutir en l'ús de les pròpies dades per a la pròpia gestió de l'administració, tant per a gestionar coneixement, innovar i prendre decisions. L'ús d'aquestes dades de manera interna, aplicant tecnologia de grans volums de dades i la metodologia d'intel·ligència territorial, també faciliten conèixer en profunditat el territori i prendre decisions sobre aquests (Garcia et al., 2018).

1.2 Objectius de recerca.

Per orientar la nostra recerca, plantegem els següents objectius:

- Conèixer les pràctiques de la publicació de dades obertes a Catalunya en els portals, tant de les administracions públiques com dels centres de recerca universitària, i com contribueixen a la transparència i la governança de la informació a nivell intern.
- Identificar els factors que faciliten i inhibeixen la publicació de dades en obert.
- Compilar indicadors per avaluar la qualitat de les dades obertes.
- Localitzar estàndards i models de gestió de dades obertes.
- Documentar bones pràctiques de gestió de dades obertes i dades de recerca a Catalunya i en països del nostre entorn.
- Establir aspectes que haurien de contenir els protocols de gestió de les dades obertes.

1.3 Antecedents

Presentem els antecedents dels temes que estudiem en aquest projecte, en quatre apartats. En primer lloc exposem les pràctiques identificades en la gestió de dades. En segon lloc, recollim els factors facilitadors i inhibidors que s'han identificat en la gestió de les dades. En tercer lloc recollim els indicadors per avaluar la qualitat de les dades. En quart lloc, trobem els models i estàndards que serveixen per sistematitzar la gestió de les dades. I per últim, fem un apartat concret dedicat a les dades de recerca en obert, per ser un ecosistema d'informació diferenciat.

1.3.1 Pràctiques de gestió de dades

En aquests moments la gestió de dades obertes està en el punt de mira de les polítiques locals, regionals, nacionals i internacionals (Nugroho et al., 2015) com un motor de desenvolupament econòmic i com un fonament de la transparència en l'acompliment del govern (Zuiderwijk i Janssen, 2014; Ferro i Osella, 2013). Tot i que hi ha innumbrables metodologies per publicar dades obertes, encara manca un model estandarditzat en el si de les organitzacions que acompanyi l'adopció d'aquestes pràctiques i eviti desconeixement de com procedir (Kučera, 2014; Hossein et al., 2016).

La gestió de les dades implica tenir en compte un seguit d'accions que abasten des de la seva producció fins a la publicació d'aquestes, tenint en compte d'altres aspectes com: a) protecció de dades personals i sensibles; b) respecte vers la propietat intel·lectual; c) garantia de la qualitat de les dades, per a preservar la qualitat dels productes i serveis que es deriven de la reutilització de les dades, i d) quins indicadors poden reflectir el benefici que comporta la publicació de dades.

Per a això, a més d'emprar una tecnologia adequada, cal establir un marc sistemàtic de gestió regulat, que quedi recollit i documentat en processos i procediments (Hossain et al., 2016; Lemieux et al., 2014; Whitmore, 2014; Eppler, 2006; Gasser, 2004; Piattini, 2002). Aquests, també han d'incloure els documents generats a les organitzacions, i integrar la gestió de dades als sistemes de gestió documental que cada organització tingui definits (Lemieux et al., 2014; McDonald i Leveille, 2014; Thurston et al., 2014; McLeod, 2012). A més és precís tenir guies formals que indiquin com avaluar l'èxit de les iniciatives de dades

obertes (Abella, et al., 2014; Whitmore, 2014). D'aquesta manera les organitzacions asseguren una correcta entrada i tractament de dades que garanteixi la seva qualitat.

Aquests processos i procediments han de definir què es fa, com i qui és responsable de fer les accions. Els temes que poden englobar els processos són: 1) condicions de producció de les dades incloent els formats de producció i publicació, recollida, classificació i emmagatzematge, i condicions de posada a disposició de la ciutadania, tenint en compte sempre aspectes de qualitat de les dades (Kenett i Shmueli, 2014; Chung et al, 2002; Wang et al., 2002); 2) protecció de dades personals i sensibles (Espanya, 2018); 3) definició de llicències i condicions d'us, que també venen regulades per la legislació sobre reutilització de la informació (Espanya 2007; Espanya, 2011; Espanya 2015; Unión Europea, 2019) 4) gestió de sol·licituds i llicències en els termes indicats per la legislació acabada d'esmentar; 5) quadre d'indicadors per valorar les accions i l'impacte de les dades publicades en la indústria de reutilització (Abella, et al., 2014; Whitmore, 2014).

Són escassos els treballs existents sobre processos de gestió de les dades i governança de les dades, per combatre les illes informacionals existents a les organitzacions. Per això hem de destacar la publicació de treballs sobre gestió i publicació de dades obertes, com el de la metodologia de Kučera (2014), que recull els problemes que implica la seva gestió i publicació, o el treball de Currie (2016) que explora els processos duts a terme a Los Angeles per publicar dades obertes o el treball d'AlAnazi i Chafield (2012) que analitzen les pràctiques a Orient Mig (Aràbia Saudí, Kuwait, Bahrain, Qatar, Emirats Àrabs, Oman, Iemen, Jordània, Israel, Iran, Síria, Líban i Iraq). Són rellevants també els esforços fets per algunes administracions públiques i organitzacions internacionals per redactar guies o metodologies per implantar la gestió de dades obertes (Kučera, 2014), de les que destaquem les publicades per Sunlight Foundation (2014) o la de les Nacions Unides (2013).

Exemples posteriors als esmentats per Kučera (2014) són guies com la del Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia (Gobierno de Colombia, 2016, 2019) o la Guí a de la Federación Española de Municipios y Provincias (FEMP, 2017 i 2019), o més recentment la guia elaborada per la Consejería de Fomento y Medio Ambiente de la Junta de Castilla y León en el marc del projecte Comunidad Rural Digital, que detalla les accions i fases a tenir en compte en la implantació de la gestió de dades obertes en un administració (Junta de Castilla y León, 2019). Aquesta última guia ens subministra un gràfic de la seqüencialitat d'accions per implantar la gestió de dades obertes en una administració pública.

El tema sobre la qualitat de les dades, i els indicadors per avaluar-les, està desenvolupat en el subapartat 1.3.3.

Aquestes iniciatives es poden completar amb d'altres procedents de la disciplina arxivística i gestió de documents. Aquesta disciplina ofereix una metodologia que aporta orientacions per definir processos de gestió de dades al llarg de tot el seu cicle de vida, que garanteixin la fiabilitat, actualització i rellevància de les dades (Lemieux et al., 2014; McDonald i Leveille, 2014; Thurston et al., 2014; McLeod, 2012). A més, facilita integrar dades obertes i els documents i expedients electrònics que formen part del corpus de les dades publicades en els portals, en el marc del dret d'accés a la informació pública.

En definitiva, la pràctica d'aplicar polítiques de gestió de dades obertes compta ja amb una dècada de desenvolupament. Tanmateix la recerca en aquest àmbit és més recent, i s'ha identificat la necessitat de realitzar més projectes de recerca en diferents àmbits, dels que senyalem els que formen part del projecte proposat: a) major sistemàtica de pràctiques, b) definició de models de gestió, c) identificació de les barreres que impedeixen la implementació i la reutilització de la informació per contribuir al desenvolupament econòmic, i d) models per avaluar la implementació de pràctiques de gestió de dades obertes (Hossain et al., 2014)

1.3.2 Factors que faciliten o inhibeixen les dades

Malgrat les iniciatives existents en la gestió de dades obertes, trobem estudis que senyalen un seguit de factors que actuen com a barreres de la seva implantació (Piovesan, 2015; Kučera, 2014; Conradie i Choeni, 2014; Bichard i Knight, 2012). Treballar sobre aquestes barreres facilitarà la bona implantació i l'ús de les dades obertes.

Kučera (2014) presenta un resum dels problemes detectats per ell i autors com Ubaldi (2013) i Both i Schieferdecker (2012). Aquests problemes els categoritza en els següents grups:

- Polítics y socials per referir-se als problemes relacionats amb el suport polític, presa de decisions i problemes socials.
- Econòmics, que engloben tres tipus de problemes: els relacionats amb els costos i beneficis de les dades obertes; com identificar i mesurar-los, i els problemes vinculats al finançament de les iniciatives de dades obertes.
- Procedimentals i organitzatius, que fan referència a l'existència i com són els procediments vinculats a la gestió de les dades, les estructures organitzatives existents, i els processos definits mitjançant els quals es lliuren les dades obertes.
- Legals: són els problemes relacionats amb els processos de l'obertura de les dades, per tal de protegir les dades personals i sensibles.
- Tècnics: són els problemes relacionats amb la tecnologia, formats de dades i infraestructura necessària per publicar dades en obert.

1.3.3 Indicadors de qualitat

Tenint en compte que la pròpia gestió de les dades es converteix en una activitat clau de l'economia digital, el tractament d'aquestes dades ha de vetllar per la seva qualitat, de manera que les dades que es generin i que s'utilitzin siguin precises, fiables, íntegres, rellevants, actuals, completes, interpretables i accessibles (Chung et al., 2002; Wang et al., 2002). D'aquesta manera, vetllant per la qualitat de les dades, es garanteix la qualitat de la informació generada (Kenett i Shmueli, 2014). Però tant important és validar la qualitat de les dades, i com poder accedir a elles amb facilitat. Tots dos aspectes són rellevants per copsar els diferents tipus de valor que les dades poden generar (Máchová i Lněnička, 2017).

Els indicadors de qualitat de les dades obertes s'han estudiat des de dos punts de vista. Un punt de vista és el de les dades en si, i sobretot els grups de datasets. L'altre punt de vista és el dels portals de dades obertes, que en molts casos estan vinculats als portals de transparència, atès que transparència s'uneix en molt casos a la publicació d'informació en formats oberts (Klein et al., 2018; Máchová i Lněnička, 2017; Bello, 2016; Lourenço, 2015). Tot i així, els indicadors que estan construïts pensant amb els portals, acaben incloent indicadors de data sets. Per tant, s'hauria de fer un quadre d'indicadors que englobi tant el portal en si en quant a usabilitat, navegació i estructura de la informació recollida en el mateix portal, mecanismes de cerca de les dades i de la informació, i característiques pròpies dels datasets, de les dades i facilitat de descàrrega.

Vetro et al. (2016) i Daves et al (2016) apunten els costos de la baixa qualitat de les dades com un impediment per aconseguir que les dades siguin reutilitzades i aprofitades. Per tant, l'aprofitament i el benefici que es fa de les dades és escàs, comparat amb el gran nombre d'iniciatives de publicar dades obertes. Això coincideix amb altres investigacions com la feta per COTEC (Abella, et al., 2019; Zuideerwijk, 2015). Entre els problemes a tractar, destaquen: la incompletitud de les dades, la manca de traçabilitat, incongruència de les dades, us de formats no estàndards, dades no actualitzades, manca de metadades, errors o duplicats o necessitat de temps per entendre les dades (Vetro et al., 2016). Això es pot

traduir en una sèrie d'indicadors per avaluar les dades publicades en els portals. Hi ha consens en que els indicadors poden ser: precisió (*accuracy*), completitud (*completeness*), rellevància, compliment, traçabilitat, consistència, provenença, autenticitat, veracitat, reputació, data d'actualització o de caducitat, incomprensió, us d'estàndards oberts, dades processables per màquina, sense restriccions d'us, entre d'altres (Vetro, et al., 2016; ISO 8000; Open Government Data; McDonald i Léveillé, 2014; Gironés, 2013; Thurston, 2012; Zhao i Hartig, 2012).

Hi ha una àmplia varietat de marcs i mètriques per avaluar la qualitat de les dades. Vetro et al. (2016) aporten una revisió dels enfocaments existents per mesurar la qualitat de les dades obertes. Hi ha des de mètriques que recullen un o pocs aspectes, fins mètriques que engloben un gran ventall d'aspectes. El model més restrictiu d'avaluació de la qualitat és el nomenat de 5 estrelles creat per Berners Lee (2006) consistent en cinc nivells de requeriments. És útil només per valorar aspectes referents al format en que la dada es publica en el web, ja que els cinc nivells apunten a com està disponible la dada, des d'una llicència propietària a la capacitat d'enllaçar amb altres dades. Aquests cinc nivells són: 1) disponible en el web però sense llicència oberta; 2) disponible com dada estructurada llegible per màquina (per exemple, un Excel en lloc d'una imatge escanejada); 3) Disponible en format no propietari com per exemple, format CSV en lloc de format Excel; 4) us d'estàndards oberts procedents del W3C (RDF i SPARQL) i URLs per a identificar les coses; i 5) dades enllaçables a altres proveïdors de dades per a subministrar context. Tenint en compte aquest aspecte, tot i assolir la publicació 5 estrelles, la qualitat pot ser pobre, ja que no te en compte aspectes com la precisió de les dades exemptes d'errors ortogràfics o tipogràfics, o la completitud, consistència i actualització (*timeliness*).

D'entre els models més detallats per avaluar les dades obertes, Vetro et al. (2016) apunten dos: Ubaldi (2013) i Maurino et al. (2014). El marc d'anàlisi desenvolupat per Ubaldi (2013) inclou mètriques que cobreixen diferents punts de vista com el polític, l'organitzatiu i el tècnic, per mesurar si la dada està fàcilment disponible (per exemple, nombre de datasets i metadades disponibles en un portal específic), demanda (vegades vista per dia). Un altre punt de vista que també incorporen en les mètriques és la reutilització com, per exemple, nombre d'apps desenvolupades amb les dades. El marc de Maurino et al. (2014) avalua el nivell del dataset i està centrat en tres aspectes: a) completitud en el sentit de disponibilitat i accés al document mitjançant enllaços interns i externs; b) precisió en termes del format, i referència temporal en quant actualització de les dades.

Vetro et al, (2016) es basen en Ubaldi (2013), Maurino et al. (2014), Behkamal et al. (2014), Open Linked Data Quality i la ISO/IEC 25012 per desenvolupar un altre marc d'avaluació que inclou model de qualitat de les dades i dades enllaçables. Aquest model inclou un ampli ventall de mètriques vinculades a datasets o cel·les: traçabilitat, vigència, caducitat, compliment d'estàndards, completitud, intel·ligibilitat i precisió.

A part d'aquests indicadors de qualitat més centrats en les dades i els datasets, a la literatura també trobem indicadors per avaluar els portals de transparència, alguns dels quals també inclouen aspectes relacionats amb els datasets. Són diversos els estudis que existeixen, d'entre els quals destaquem els aportats per Máchová i Lněnička, (2017); Bello et al. (2016) Oliveira (2016) i Lourenço (2015) perquè son compatibles i sumen.

Lourenço (2015) proposa un seguit de requeriments per avaluar els portals de transparència dels governs, alguns dels quals coincideixen amb els indicadors de marcs anteriors. Aquests són:

- a) Qualitat. Aquest indicador hauria d'estar articulat mitjançant mecanismes externs i independents a qui genera les dades.

- b) Completitud de les dades. Aquest indicador hauria d'incloure datasets rellevants, que cobreixin períodes de temps determinats, etc. Per fer-ho cal que els portals incloguin metadades per a cada dataset que identifiquin aspectes com l'autoria, entre d'altres.
- c) Accés i visibilitat. Aquest indicador proporciona diferents opcions de cerca per tal de facilitar a l'usuari la consulta de les dades i que siguin fàcils d'entendre.
- d) Usabilitat i comprensió. Aquest indicador té en compte que els portals haurien d'incloure descripcions clares del contingut dels datasets i com categoritzar-los.
- e) Actualització. L'indicador avalua que les dades han d'estar actualitzades i en el portal s'hauria d'especificar el període d'actualització.
- f) Valor i utilitat. Els usuaris haurien de poder suggerir millores o la introducció de dades rellevants per a ells.
- g) Granularitat. Caldrà indicar el nivell/unitat d'anàlisi de les dades que ofereix el portal.

Podem completar aquests indicadors amb els proposats per Bello et al. (2016) per a avaluar com estan implementats els portals de dades obertes:

- a) Tipus de tecnologia: on s'avalua si el portal inclou aplicacions web, i una combinació de plataformes en obert i propietàries.
- b) Ús de l'esquema *Open-data 5-stars evaluation*, ideat per Berners-Lee (2006), que com hem apuntat abans avalua la presència de diferents atributs en el portal: accessibilitat de les dades, estructuració de les dades, ús de formats no propietaris, ús d'URLs significatives, i enllaç de les dades a d'altres dades o a llocs que donin context a les mateixes.
- c) Format: tipus de formats i adherència a formats no-propietaris.
- d) Tipus de datasets publicats (àrees de contingut que ofereix el portal, de manera que els datasets puguin cobrir diferents interessos i preferències).
- e) Acord del tipus de llicència de dades obertes que s'utilitza.
- f) Funcionalitat: característiques del portal que permet als usuaris realitzar diferents tasques, com són: accedir a les dades sense restriccions, descarregar fàcilment els documents o les dades, exportar les dades en altres formats de forma fàcil, treballar els elements dels datasets de forma desagregada per facilitar l'anàlisi, visualitzar el significat dels diferents datasets, i cercar i navegar de forma fàcil.

Oliveira (2016) proposa avaluar com es publiquen els datasets en els portals de govern obert de Brasil en base als següents criteris:

- a) Accés a les dades tenint en compte les següents característiques: quantitat i mida, formats, àrees temàtiques que abasta, evolució del dataset des de la seva creació, i històric d'actualització.
- b) Metadades: estàndards, llicències i procedència de la informació.
- c) Esquema del dataset: sintaxi, noms dels fitxers, mida del dataset, i dispersió de les dades.

Per últim, Máchová i Lněnička (2017) centren els seus indicadors d'avaluació en la descripció i la classificació vinculada a com es troben les dades, i per tant, en com és l'accessibilitat a les dades obertes. La idea central és la necessitat de que les dades estiguin ben descrites en termes de qualitat per a que hi hagi persones que les puguin transformar en coneixement i fer-les útils. Els portals estan basats en un sistema web que recull els datasets en diferents formats procedents de diferents fonts, i des d'on es poden descarregar mitjançant API. Per

trobar els datasets, aquests han d'estar etiquetats i organitzats en un catàleg, i han d'estar disponibles mitjanant un cercador. Per tant, un aspecte rellevant que senyalen aquests autors és la qualitat de les metadades. A més inclouen una sèrie d'elements apuntats per altres autors, abans esmentats, a partir dels quals creen un marc per a comparar portals de dades obertes. Aquest marc es basa en dos blocs: a) característiques del portal de dades obertes (aspectes tècnics, disponibilitat i accessibilitat dels datasets, i canals de participació i comunicació com fòrums, formularis de petició, FAQ i social media); i b) característiques dels datasets com títol i descripció, qui publica, actualització, cobertura geogràfica, llicència, URL, mida, nombre de visites i descàrregues, formats llegibles per màquina, visualitzacions i ocasió de que l'usuari pugui qualificar i comentar.

Per últim, hem de senyalar un àrea de recerca que treballa la qualitat de les dades obertes, des del punt de vista de la capacitat de ser reutilitzades, i sobretot de l'impacte que tenen en la reutilització. Encara hi ha pocs estudis sobre mètriques per avaluar els beneficis de l'obertura de dades públiques i l'impacte en el sector de la reutilització. És difícil avaluar l'impacte de les polítiques públiques de dades en la indústria de la reutilització, per diferents raons, entre elles la falta de mètriques adequades (Abella et al., 2014; Fournier-Tombs, 2011). A més, encara són pocs els governs o entitats públiques que estimen quantitativament els beneficis de les seves polítiques d'obertura de dades públiques. Això, en part, pot ser degut a que les seves estratègies de reutilització de dades es basen més en el recolzament de les llibertats ciutadanes, el compliment de la legislació, o declaracions públiques de transparència (Abella et al., 2014; Huijboom; Van-den-Broek, 2011).

Fournier-Tombs (2011) revisa els treballs que han estudiat la influència dels portals sobre la comunitat d'usuaris i observa que els mètodes basats en la medició de l'ús, com per exemple enquestes sobre reutilitzadors i usuaris, són d'utilitat limitada perquè no se poden obtenir mostres estadísticament significatives (Abella et al., 2014). D'altra banda, com apunta Abella et al. (2014), els informes que publica l'entitat pública Red.es analitzen l'impacte del sector econòmic que reutilitza la informació. Tanmateix, manca una relació casual de l'impacte econòmic amb les seves polítiques públiques de dades (Abella et al., 2014). En aquest sentit és rellevant la proposta d'indicadors elaborada per Abella et al. (2014) pensades per avaluar el grau de reutilització de dades. Els seus autors han nomenat a aquesta mètrica Meloda (MEtric for reLeasing Open DATa). Aquesta mètrica està basada en els principis de l'open government data (Eaves, 2009) i considera quatre dimensions (estàndards tècnics, accés a la informació llegible per màquina de forma automàtica, marc legal, i models de dades) a les que associa una sèrie de nivells per a cadascuna (Figura 1). Aplicacions d'aquesta mètrica es poden observar en la pàgina web del projecte MELODA (<https://www.meloda.org/what-is-meloda-for/>)

| Estàndares tècnics | Acceso | Legal | Modelo de datos |
|--|--|---|---|
| 1. Estándar privativo Ej.: .xls, .shp, .doc | 1. Sin acceso Ej.: mail no automático o acceso en persona | 1. Copyright Ej.: copyright | 1. Sin modelo publicado Ej.: tabla de datos sin descripción de los campos |
| 2. Estándar abierto Ej.: .csv, .ods, wms | 2. Acceso vía web con registro Ej.: formulario manual | 2. Uso privado Ej.: copyright permitiendo uso personal | 2. Modelo con campos de datos Ej.: tabla de datos con descripción de los campos |
| 3. Estándar abierto con metadatos Ej.: rdf, rss, json | 3. Acceso directo vía web Ej.: url único | 3. Uso no comercial Ej.: CC BY-NC 4.0 | 3. Modelo con especificaciones de campos Ej.: vocabularios disponibles |
| | 4. Acceso vía web con parámetros Ej.: url con parámetros | 4. Uso comercial Ej.: CC BY-SA 4.0 | 4. Modelo externo normalizado Ej.: vocabularios disponibles aceptados por organización de normalización |
| | 5. Acceso completo (API) Ej.: punto de acceso <i>Sparql</i> | 5. Uso no limitado con autoría Ej.: CC BY 4.0 | 5. Modelo externo y generalizado Ej.: vocabularios disponibles aceptados por organización de normalización y reconocidos |

Figura 1: Dimensiones i nivells de la mètrica Meloda (Font: Abella, et al., 2014)

1.3.4 Estàndards

Els processos per publicar dades obertes estan mancats d'una estandardització (Kucera, 2014). Tot i així, aquesta afirmació la podem matisar, perquè disposem de diferents estàndards que poden ajudar en la definició de processos i procediments, cadascun des de la seva àrea de coneixement. Aquests els presentem en tres blocs: 1) forma d'avaluar la publicació de les dades d'una ciutat; 2) com avaluar la qualitat de les dades; i 3) com sistematitzar la gestió de documents. Aquests tres blocs el descrivim tot seguit.

En primer lloc, tenim la norma UNE 178301 de Ciutats Intel·ligents i Dades Obertes. Aquesta norma té com a objecte establir la **forma d'avaluar la publicació de dades obertes d'una ciutat**, mitjançant un conjunt de mètriques i indicadors per mesurar el grau de maduresa de l'obertura de dades elaborades i custodiades pel sector públic, de manera que es faciliti la reutilització, en l'àmbit de les ciutats intel·ligents (AENOR, 2015).

En segon lloc disposem de dues normes que apunten a les característiques de la **qualitat de les dades**. La primera es una sèrie de normes (ISO 8000) que ofereix un marc d'indicadors per avaluar la qualitat de les dades i els datasets (Figura 2). La segona norma vinculada en la gestió de dades és la ISO 25012, que té com a objectiu definir els indicadors de la qualitat de les dades pensant en els requeriments de qualitat dels productes de software (ISO, 2008).

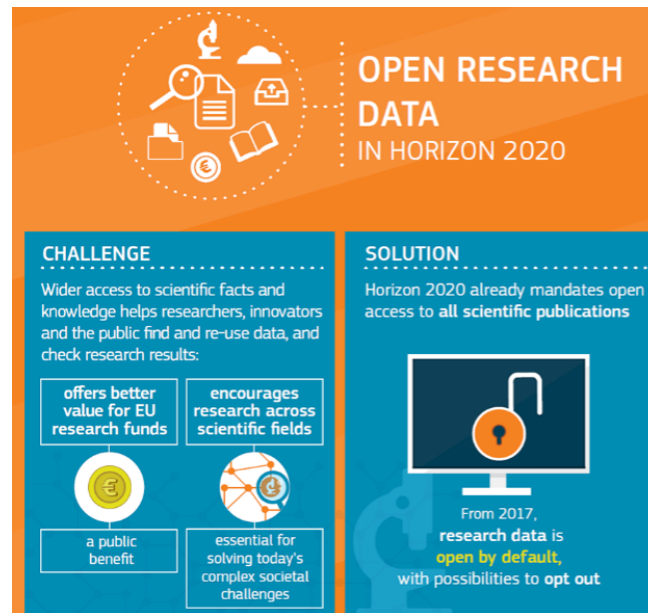
| |
|---|
| ISO/TS 8000-1:2011 Data quality -- Part 1: Overview. |
| ISO/TS 8000-100:2009 Data quality -- Part 100: Master data: Overview. |
| ISO 8000-2:2012 Data quality -- Part 2: Vocabulary. |
| ISO/TS 8000-120:2009 Data quality -- Part 120: Master data: Exchange of characteristic data: Provenance. |
| ISO/TS 8000-130:2009 Data quality -- Part 130: Master data: Exchange of characteristic data: Accuracy. |
| ISO 8000-110:2009 Data quality -- Part 110: Master data: Exchange of characteristic data: Syntax, semantic encoding, and conformance to data specification. |
| ISO/TS 8000-140:2009 Data quality -- Part 140: Master data: Exchange of characteristic data: Completeness. |
| ISO/TS 8000-150:2011 Data quality -- Part 150: Master data: Quality management framework. |
| ISO/TS 8000-311:2012 Data quality -- Part 311: Guidance for the application of product data quality for shape (PDQ-S). |
| ISO 8000-8:2015 Data quality -- Part 8: Information and data quality: Concepts and measuring. |

Figura 2: Normes ISO per gestionar la qualitat de les dades (Font: ISO)

Aquesta sèrie de normes 8000 té en compte un seguit d'aspectes i tres rols genèrics. Els aspectes que es tenen en compte són: a) les persones que han de participar en la gestió de les dades; b) els processos que es responsabilitzen de la gestió efectiva de les dades; i c) la millora continua dels processos dedicats a garantir la qualitat de les dades. Els rols que s'identifiquen són: a) gestor de dades, responsable dels factors organitzatius per gestionar la qualitat de les dades; b) administrador de dades, responsable de coordinar i supervisar el

II-lustració 1 Principis FAIR per a gestionar dades obertes
(Font: https://www.openaire.eu/images/Guides/FAIRdataprinciples_foster.png)

Tal i com veurem més endavant, podem parlar també d'un marc conceptual i una força motriu que neix amb una dinàmica pròpia, en general com a conseqüència de la força motriu dels finançadors europeus, bàsicament els projectes Horizon 2020 i els European Research Council (ERC).



<https://ec.europa.eu/research/openscience/index.cfm?pg=openaccess>

En el cas dels projectes Horizon 2020 foren els primer en obrir la ruta cap a la compartició de les dades de recerca. Així en els plans pilot del període 2014-2016, determinades disciplines foren requerides a compartir les dades de recerca en brut (raw data) generades durant la realització del projecte. Així, podem parlar que les dades de recerca ja s'introdueixen en el bell mig del projecte, a les quals cal tenir present ja des del disseny del propi projecte.

En aquest moment apareix un document clau per a la gestió de dades de recerca en el disseny d'un projecte, que és el pla de gestió de dades conegut amb les sigles PGD o DMP en anglès (ambdós termes s'utilitzen indistintament).

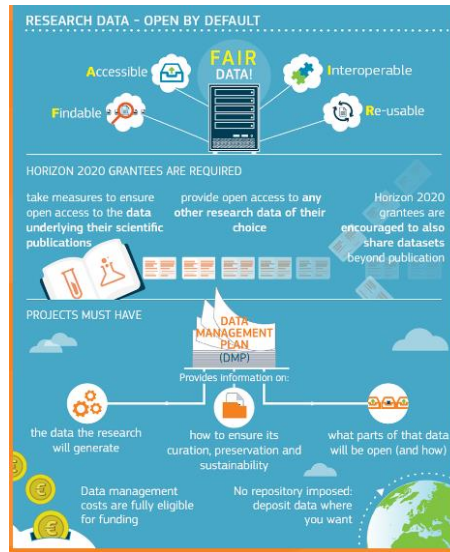


Figura 4: Esquema per a la gestió de dades amb PGD o DMP
(Font: <https://ec.europa.eu/research/openscience/index.cfm?pg=openaccess>)

Aquest document permet fer una distinció entre les dades que es generen, les responsabilitats, les qüestions tècniques, ètiques i legals de les dades així com altres aspectes que s'incorporen ja des de l'inici per a poder avançar en la lògica "open by default", que és l'escenari en el qual treballa la Comissió Europea, consistent en fer que l'escenari sigui tan obert com sigui possible, tan tancat com sigui necessari (as open as possible, as closed as necessary)

Encara ara els resultats d'enquestes globals com per exemple la realitzada per Digital Science o la realitzada pel CSUC, mostren que hi ha un percentatge de científics familiaritzats amb les dades FAIR, essent aquest coneixement creixent a cada onada anual de l'enquesta. Quan es demana específicament sobre com pensen gestionar les dades de recerca, els investigadors enquestats pel CSUC mostren el comportament següent:

Difusió de les dades

Un 60% dels enquestats han indicat que tenen previst dipositar i difondre les dades (totes o algunes) en obert. 184 (50%) volen fer-ho publicant les dades associades a les publicacions, 58 (15%) un cop s'hagin tractat les dades, 57 (15%) a mida que es vagin obtenint i 49 (13%) un cop finalitzi el projecte. Les altres opcions són minoritàries (vegeu gràfic 1).



Gràfic 1. Previsió sobre quan es publicaran les dades

Figura 5: Com es gestionen les dades.

(Font:

https://recercat.cat/bitstream/handle/2072/351567/1810%20InformeNovaEdicioEnquesta_def.pdf?sequence=5)

Finalment, cal recordar la importància de les dades de recerca en el Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020, que relaciona aquest aspecte amb la llei de la Ciència del 2011. Aquesta llei ja esmentava l'accés obert a la ciència, tot i que encara no estava desenvolupat el Plan. També cal fer esment que els projectes a nivell estatal tenen uns requeriments a hores d'ara no tan estrictes com els europeus.

2 Metodologia

L'univers d'estudi és l'administració pública de Catalunya, i més concretament els departaments o àrees responsables de gestionar les dades obertes. També hem tingut en compte les àrees responsables de facilitar a les persones l'accés a la informació pública en format reutilitzable y garantir la reutilització de les dades en el termes indicats en la llei de transparència, concretament articles 16 y 6e) (Catalunya, 2014).

Atès que aquest univers d'estudi és ampli, vàrem seleccionar quatre tipologies d'administracions públiques:

- 1) **Ajuntaments** per ser una tipologia d'administració pública amb una ampla representativitat per nombre en tot el territori;
- 2) La **Generalitat de Catalunya**, com a autoritat representativa en matèria legislativa i alhora – i sobre tot - com a impulsora de les polítiques de dades obertes;
- 3) Els **centres de recerca** incloses les **universitats**, per ser una administració pública, que des de fa uns anys precisa gestionar dades obertes resultants de la recerca, d'acord als requeriments establerts en les convocatòries públiques per finançar projectes de recerca (tant a nivell nacional com europeu). La disponibilitat de dades de recerca a més, des del juny de 2020 està contemplada en la recent directiva europea relativa a les dades obertes y la reutilització de la informació del sector públic (Unió Europea, 2019) on s'impulsa a que els estats membres recolzin la disponibilitat de les dades de recerca.
- 4) Les quatre **Diputacions** com a impulsores / facilitadores d'estructures per facilitar la gestió diària als ajuntament. Aquest grup, es va integrar com administració objecte de gestió de dades obertes, un cop vàrem veure la dificultat per contactar amb ajuntaments, i a la llum de les pràctiques que observàvem en els ajuntaments, després de realitzar les primeres entrevistes.

Un cop definit l'univers d'estudi vàrem iniciar el treball empíric, basant-nos en metodologies mixtes, amb predomini de les qualitatives. El treball empíric l'hem dividit en quatre fases, sent les tres primeres qualitatives i la quarta quantitativa descriptiva. L'elecció d'aquesta metodologia ha estat donada pel caràcter incipient de la temàtica de recerca d'aquest projecte, i molt especialment en l'àmbit geogràfic on s'aplica. Basant-nos amb aquesta metodologia hem definit criteris de selecció de la mostra, instruments i procediments de recollida de dades i procediments d'anàlisi de les dades empíriques de caire qualitatiu. La metodologia quantitativa ha estat per acotar els resultats de recerca mitjançant enquesta amb respostes més de caire quantitatiu. A continuació presentem criteris de selecció, instruments i procediments metodològics emprats per a recollir i analitzar les dades.

2.1 Selecció de la mostra

Per a cada tipologia d'administració pública hem establert com a criteris per definir la mostra, els que descrivim a continuació:

2.1.1 Ajuntaments

En el cas dels ajuntaments hem seleccionat inicialment tots els que son capital de comarca (43) atès que entre aquests ajuntaments ja hi ha varietat de context, pel que fa a extensió i nombre de població, havent-hi municipis petits, mitjans i grans, i municipis per sota dels 10.000 habitants (Annex 1). La selecció dels departaments responsables de la gestió de dades obertes ha estat feta a partir de l'organigrama i responsabilitats descrites per cada ajuntament en les seu portal i seu electrònica.

2.1.2 Generalitat de Catalunya

Pel que fa la Generalitat de Catalunya hem identificat aquells departaments o organismes que tenen competència amb la gestió de dades obertes. Ho hem fet per tres vies, en primer lloc mitjançant el cercador d'organismes i responsables de la Generalitat de Catalunya (Figura 6).

Cerca d'organismes i responsables de la Generalitat de Catalunya

Trobeu informació bàsica i de contacte dels organismes que componen la Generalitat de Catalunya i dels seus responsables.

Introduiu les paraules a cercar

 Cerca automàtica [Ajuda](#)

Cerca per: Nom Responsable Població Codi postal Comarca

- + **Comissió de Dades Obertes del Departament de Salut i del seu sector públic**
 Depèn de: Departament de Salut
 08028 | Barcelona | Barcelonès
- + **Direcció General de Transparència i Dades Obertes**
 Núria Espuny i Salvadó
 Depèn de: Departament d'Acció Exterior, Relacions Institucionals i Transparència
 08002 | Barcelona | Barcelonès
- + **Sub-direcció General de Transparència i Dades Obertes**
 Joaquim Dorda i Ventura
 Depèn de: lacions Institucionals i Transparència
 08002 | Barcelona | Barcelonès
- + **Àrea de Dades Obertes**
 Depèn de: Departament d'Acció Exterior, Relacions Institucionals i Transparència
 08002 | Barcelona | Barcelonès

Avis legal | Accessibilitat | Sobre gencat | © Generalitat de Catalunya

Figura 6: Cercador d'organismes i responsables de la Generalitat de Catalunya

En segon lloc, hem considerat dos àmbits més que des del punt de vista de la lògica influeixen la gestió de dades, independentment de les polítiques i la planificació definides en cada administració. Un àmbit està seleccionat des del punt de vista tecnològic, per identificar les infraestructures tecnològiques (telecomunicacions i sistemes informàtics) que hi ha al darrera de la gestió de dades obertes. L'altre àmbit l'hem inclòs per dues raons: a) la necessitat de classificar i recuperar informació, per facilitar als ciutadans l'accés a la informació pública, i b) l'impuls de polítiques en matèria d'arxius, gestió de documents i actius digitals, en tant que contenidors de dades i informació pública, susceptible de ser

reutilitzades, gestionades, mantingues i preservades al llarg del temps. Aquest punt s'ha de tenir molt en compte atès que s'està tractant de definir el model de governança de les dades corporatives de la Generalitat de Catalunya, des de l'Oficina d'Innovació i Administració Digital (OIAD).

Per últim, ho hem fet mitjançant el mètode de la bola de neu, emprant terminologia de la teoria fonamentada (grounded theory). Aquest mètode ens ha portat a incloure el Consorci Administració Oberta de Catalunya, perquè ha estat identificat posteriorment com impulsor de la gestió de dades obertes als ens locals.

Per tant, tenint en compte aquests criteris, en l'àmbit de la Generalitat de Catalunya ens hem focalitzat en cinc àrees, sent les dues primeres àrees les que hem identificat mitjançant el cercador d'organismes de la Generalitat. Aquestes quatre àrees són

- Departament d'Acció Exterior, Relacions Institucionals i Transparència.
- Departament de Salut.
- Centre de Telecomunicacions i Tecnologies de la Informació de la Generalitat de Catalunya.
- Servei de Suport al Sistema d'Arxius de Catalunya.
- Consorci de l'Administració Oberta de Catalunya

En relació a la primera àrea – el **Departament d'Acció Exterior, Relacions Institucionals i Transparència** – hem analitzat l'estructura departamental de la institució, fent atenció especial als àmbits on es treballaven processos i especialment gestió de les dades obertes. A tal fi hem analitzat les funcions indicades en els diferents decrets. De tots ells, concretament hem identificat la Secretaria de Transparència i Govern Obert, on hi ha la Direcció General de Transparència i Dades Obertes (Figura 7).



Figura 7: Departament Direcció General de Transparència i Dades Obertes (Font: <http://sac.gencat.cat/sacgencat/AppJava/organigrama.jsp?codi=19816&jq=200001>)

Aquesta secretaria té entre les seves funcions, la de dirigir el procés d'obertura de dades públiques de la Generalitat i facilitar la reutilització de la informació del sector públic (art. 36 a Catalunya, 2019a) (Figura 8).

Secretaria de Transparència i govern obert

FUNCIONS

- a) Impulsar i promoure el govern obert.
- b) Dirigir i impulsar les polítiques de transparència dins l'Administració de la Generalitat i el seu sector públic instrumental.
- c) Dirigir les estratègies per a garantir el compliment de la normativa de transparència i accés a la informació pública per part dels departaments i entitats del sector públic de l'Administració de la Generalitat i organismes que en depenen.
- d) **Dirigir el procés d'obertura de dades públiques de la Generalitat i facilitar la reutilització de la informació del sector públic.**
- e) Impulsar i promoure polítiques d'integritat en l'administració pública i el sector públic, així com la seva avaluació, coordinadament amb la Secretaria d'Administració i Funció Pública i sens perjudici de les competències que li corresponen.
- f) Impulsar **el Pla de Govern Obert de la Generalitat.**
- g) Donar **suport al món local en la implementació** i el desplegament de les polítiques de govern obert.
- h) Establir els criteris de qualitat democràtica que han de complir els canals, processos i espais de participació ciutadana de la Generalitat.
- i) Qualsevol altra funció de naturalesa anàloga que li sigui encomanada.

Figura 8: Funcions de la Secretaria de Transparència i Govern Obert (Font: elaboració pròpia a partir de Catalunya 2019a)

Dins de la Secretaria de Transparència i govern obert, la Subdirecció General de Transparència i Dades Obertes s'encarrega de la part operativa de la gestió. Concretament, d'entre les seves funcions destaquem les funcions d'assessorar els departaments en matèria de dades obertes, supervisar i coordinar el Portal de dades obertes de la Generalitat de Catalunya, i participar en iniciatives interinstitucionals d'obertura de dades (art. 38, Catalunya, 2019a) (Figura 9)

Sub-direcció de transparència i Dades Obertes

FUNCIONS

- a) Coordinar l'aplicació de les directrius i els criteris corporatius necessaris per a garantir el compliment de la normativa de transparència i accés a la informació pública per part dels diferents departaments i entitats del sector públic de l'Administració de la Generalitat.
- b) **Assessorar els departaments** de la Generalitat en matèria de transparència i **dades obertes.**
- c) Supervisar i desenvolupar el Portal de la transparència i coordinar el procés d'integració i garantia d'interoperabilitat dels portals, les aplicacions i la informació que s'hi ha de vincular, coordinadament amb la direcció general competent en matèria de modernització i innovació de l'Administració i sens perjudici de les competències que li corresponen.
- d) **Supervisar i coordinar el Portal de dades obertes** de la Generalitat de Catalunya.
- e) Participar en **iniciatives interinstitucionals d'obertura de dades.**
- f) Donar suport al món local i a la resta de subjectes obligats per al compliment de la normativa de transparència i accés a la informació.
- g) Qualsevol altra funció de naturalesa anàloga que li sigui encomanada.

Figura 9: Funcions de la subdirecció de transparència i dades obertes el Decret 83/2019 (Font: elaboració pròpia a partir de Catalunya, 2019a)

Pel que fa a la segona àrea seleccionada, en el **Departament de Salut**, hem localitzat la Comissió de Dades Obertes creada l'abril del 2019, dependent de la Secretaria General de Salut (Catalunya, 2019b) (Figura 10).



Figura 10: Gestió de dades obertes al Departament de Salut (Font <http://sac.gencat.cat/sacgencat/AppJava/organigrama.jsp?codi=21043&jq=93>)

Concretament aquesta Comissió té les següents funcions (Catalunya, 2019b):

- Formular els criteris de planificació de l'oferta conjunta de dades obertes a publicar, i elevar-la al Consell de Direcció del Departament de Salut per aprovar-la.
- Crear i gestionar el Catàleg de dades obertes de Salut, com a instrument de planificació, amb el contingut següent:
 - La relació dels conjunts de dades publicats.
 - La relació dels conjunts de dades aprovats pel Consell de Direcció del Departament de Salut i pendents de publicació.
 - La relació dels conjunts de dades que estiguin en fase d'elaboració per part de la Comissió.
- Supervisar els processos organitzatius i tècnics per publicar i actualitzar l'oferta conjunta de dades obertes aprovada pel Consell de Direcció del Departament de Salut.

La tercera àrea seleccionada ha estat la de gestió de les telecomunicacions i sistemes informàtics, que està inclosa en les funcions del **Centre de Telecomunicacions i Tecnologies de la Informació de la Generalitat de Catalunya** (CTTI) (Figura 11). Aquest organisme té entre les seves funcions la planificació, adquisició i instal·lació d'equipaments dels sistemes de telecomunicacions i informàtics, i l'establiment de directius, coordinació, i control de l'execució (Figura 11).

CTTI

Funcions

- Realització de la **planificació tècnica i l'establiment de les directrius de la gestió i l'explotació dels serveis i dels sistemes de telecomunicacions i informàtics** que siguin necessaris per al funcionament de la Generalitat d'acord amb les atribucions que atorga l'Estatut d'autonomia i de conformitat amb el que estableix la legislació de l'Estat sobre ordenació de les telecomunicacions i amb el que estableixen les legislacions catalana i europea en l'àmbit de les competències respectives.
- **Coordinació, la supervisió i el control de l'execució dels sistemes i dels serveis de telecomunicacions i d'informàtica** aptes per a satisfer les necessitats de l'Administració de la Generalitat en aquesta matèria.
- **Assessorament tècnic** en relació amb les telecomunicacions i la informàtica corporatives i la formació del personal tècnic, mitjançant l'organització i la promoció de cursos i altres possibilitats de formació per a l'elaboració de sistemes i la prestació de serveis.
- Realització de les propostes dels **plans d'actuació** en matèria de telecomunicacions i informàtica.
- Realització de la **gestió** de forma directa o indirecta **de l'execució de les infraestructures** necessàries per implantar les xarxes i **els serveis de telecomunicacions i informàtics** requerits i la seva explotació i manteniment posteriors.
- Establiment i la fixació de les **especificacions tècniques** dels terminals d'usuaris en l'àmbit de les telecomunicacions i la informàtica per garantir-ne la correcta connectivitat a les xarxes.
- **Adquisició i la instal·lació dels diferents equipaments** que constitueixen les xarxes de telecomunicacions o telemàtiques necessàries per al desenvolupament de la seva activitat.
- Gestió de les freqüències assignades a la Generalitat de Catalunya i a les entitats, organismes i empreses que en depenen per garantir-ne l'ús més correcte.
- Establiment de la **política d'explotació de la xarxes** de telecomunicacions i telemàtiques.

Figura 11: Funcions del CTTI (Font: elaboració pròpia a partir de Catalunya, 2019c)

La quarta àrea tinguda en compte ha estat l'**Àrea d'arxius i gestió documental** on hem identificat la Secció de Gestió Documental i Estandardització i la Coordinació de Sistemes de Gestió Documental Corporativa dependents de la Direcció General del Patrimoni Cultural del Departament de Cultura (Figura 12).



Figura 12: Organismes vinculats a la gestió de documents (Font: Cercador d'organismes)

Entre les funcions de la Secció de Gestió Documental i Estandardització destaca la planificació, implementació i actualització el Sistema General de la Gestió Documental corporativa inclòs l'arxiu electrònic únic de l'Administració de la Generalitat (Figura 13).

Secció de Gestió Documental i Estandardització

Funcions

- Participar en la **planificació, implementació i actualització del Sistema General de Gestió Documental** en l'àmbit de la Generalitat de Catalunya.
- **Proposar mesures de foment de l'ús de les tecnologies de la informació i la comunicació en l'àmbit de la gestió documental corporativa i vetllar per la seva implementació**, en coordinació amb les polítiques corporatives d'administració digital i ús de les tecnologies de la informació a l'Administració de la Generalitat i a la resta d'administracions.
- Implementar polítiques d'innovació i d'estàndards de qualitat en el seu àmbit d'actuació.
- Donar suport a l'impuls de l'acció de la Comissió Nacional d'Accés, Avaluació i Tria Documental.
- **Administrar l'arxiu electrònic únic de l'Administració de la Generalitat de Catalunya.**
- **Planificar i supervisar l'execució del programa d'actuació en gestió documental per als arxius centrals** administratius dels departaments i entitats adscrites.
- Qualsevol altra funció de naturalesa anàloga que se li encarregui.

Figura 13: Funcions de la Secció de Gestió Documental i Estandardització
(Font: http://sac.gencat.cat/sacgencat/AppJava/organisme_fitxa.jsp?codi=14316)

Per últim, hem analitzat el Consorci d'Administració Oberta de Catalunya (AOC) per la seva implicació amb el sector públic de Catalunya. L'AOC es constitueix per a l'impuls, la implantació i la utilització de les tecnologies de la informació i de les comunicacions en el sector públic de Catalunya per a la millora de la gestió pública (http://sac.gencat.cat/sacgencat/AppJava/organisme_fitxa.jsp?codi=13865).

En un projecte futur s'ha d'incloure un altre departament d'estudi que està en procés de definir la governança de la dada i la transformació digital en el si de la Generalitat: Departament de Polítiques Digitals i Administració Pública (Catalunya, 2019d). Just a principis d'octubre de 2019 la Generalitat ha presentat el nou model d'Administració Digital. És rellevant sobre tot per com s'organitza la gestió de les dades, i sobretot, com s'utilitza aquesta a nivell intern.

2.1.3 Centres de Recerca i Universitats

En el cas de les universitats i centres de recerca es va triar una mostra que incorporés les diferents visions complementàries, ja fos des de l'àmbit universitari o des de l'àmbit d'un centre de recerca. Aquestes visions són la vessant tècnica, o la de gestió o la direcció o estratègica. Així, els perfils triats són:

- Direcció de l'Àrea Oberta del CSUC, com a visió general i estratègica de país al voltant d'aquests aspectes. Ells coordinen grups de treball entre universitats i com veurem tenen un paper proactiu en la gestió de les dades de recerca.
- Un director de biblioteca universitària amb l'encàrrec de gestionar, dipositar i preservar dades de recerca. A més a més, és el coordinador del grup de repositoris a nivell de l'Estat (REBIUN) i pot oferir també una visió de la col·laboració que es duu a terme en aquests àmbits.

- Un responsable d'un dels repositoris digitals universitaris, on s'ha decidit acollir els datasets de la recerca duta a terme per part dels investigadors.
- Un responsable de la gestió de les dades de recerca d'un institut CERCA, per a conèixer la visió específica d'aquest tipus d'institucions, amb una tradició bibliotecària diferent i unes necessitats d'alta exigència per part de finançadors de la recerca en temes de compartició de dades.

Creiem que cada un dels entrevistats pot aportar una visió complementària

2.1.4 Diputacions

Atès que només hi ha quatre diputacions, no hem establert cap criteri per seleccionar la mostra. Tan sols hem identificat les àrees objecte d'estudi, que corresponen a aquells departaments que gestionen dades i recolzen el món local pel que fa a la gestió de dades.

2.2 Recollida d'informació

La recollida d'informació ha estat feta en quatre fases, per a les quals hem emprat diferents fonts i instruments. A continuació detallarem les fases seqüencials i els instruments emprats per recollir la informació.

2.2.1 Experiències en gestió de dades obertes i identificació de portals de dades obertes

Per tal de conèixer experiències i bones pràctiques d'altres administracions públiques tant nacionals com internacionals – a part de les que són objecte d'estudi – hem fet un buidatge de literatura científica i sectorial.

Després, hem localitzat i buidat literatura gris i portals d'organismes i administracions públiques, tant genèrics como de dades obertes. L'objectiu era conèixer les experiències en gestió de dades obertes a Catalunya i a d'altres àmbits geogràfics. A més, de manera molt específica, hem explorat el portals de les administracions objecte d'estudi, per observar i analitzar quina informació donen les administracions públiques de si mateixes. Hem posat atenció a si tenen portal de dades obertes, quines dades s'ofereixen i com els ciutadans poden cercar informació.

Aquestes fonts també faciliten identificar quines són les àrees que gestionen les dades obertes i quines són les accions dutes a terme o les polítiques d'aplicació. En el grup de fonts de literatura gris hem analitzat documents com, per exemple, plans estratègics, informes de seguiment o d'avaluació, documents públics de grups de treball, guies o legislació. Aquests documents, a més de donar informació sobre la gestió interna de les dades, també donen informació sobre les accions que es realitzen per impulsar la gestió de dades obertes. Els portals de les administracions públiques ofereixen informació sobre les estructures organitzatives, que informen del context en el que es lidera la gestió de dades obertes i la reutilització.

De cara a preparar les fases següents de recerca, prenem en consideració que els portals de les administracions públiques objecte d'estudi donen informació sobre els canals per comunicar amb elles (telèfons, adreces electròniques o plataformes d'administració electrònica). Per tant, aquest tipus d'informació també l'hem recollida en un formulari preparat a tal efecte.

2.2.2 Focus group

Aquest instrument l'hem seleccionat amb un objectiu doble. Per una banda, amb els focus group volíem unir peces d'informació, sistematitzar la informació recollida en la fase prèvia, i confirmar o identificar els interlocutors per fer entrevistes. D'altra banda, l'objectiu ha estat obtenir mitjançant el seu anàlisi una primera aproximació a la temàtica, amb contextos amplis i rics en matisos per preparar la fase següent.

Es va fer un focus group amb representants de dos àrees que, com hem esmentat abans, podrien contribuir o estar contribuint a la gestió de dades en el si de la Generalitat de Catalunya, tot i no haver estat identificades en aquestes àrees, d'entre les funcions que tenen encomanades, funcions explícites en la gestió de dades obertes. Aquestes àrees cobreixen dos àmbits, d'una banda, el tecnològic, i l'altre, des del punt de vista d'impuls de polítiques en matèria d'arxius i gestió de documents/objectes digitals. Concretament, varen participar un representant de CTTI i dos de l'àmbit de gestió documental i arxius.

Posteriorment estava previst realitzar més focus group amb ajuntaments, centrats per províncies, un cop finalitzades les entrevistes. Tanmateix, a la llum dels resultats de la primera fase, i del primer focus group, es va posposar per a la segona fase. Finalment, atenent als resultats d'aquesta segona fase, amb baixa participació, pel context d'eleccions i constitució de consistoris, hem descartat realitzar focus group.

2.2.3 Entrevistes.

Hem fet entrevistes en profunditat a interlocutors d'Ajuntaments, Diputacions, Generalitat de Catalunya, Universitats i centres de recerca. Per a identificar i contactar amb els interlocutors primer de tot hem cercat adreces electròniques de contacte, a la web de l'organisme o en el directori de persones, accessible des de la seu electrònica. En el cas de no identificar a la persona i/o el correu electrònic, hem contactat per telèfon. En dos casos se'ns ha remés a la seu electrònica per fer la sol·licitud. Aquesta via ha resultat un atzucac, atès que no hem obtingut resposta.

En els cas dels ajuntaments hem localitzat 43 interlocutors dels 42 contactats. Pel que fa a la Generalitat els interlocutors han estat identificats a partir del directori disponible a la web. En el cas de les Diputacions la localització dels interlocutors ha estat més difícil, atès que no sempre els directoris són oberts. I per últim, les universitats i centres de recerca han estat localitzats mitjançant l'exploració del web i trucades a les unitats.

Un cop identificada la persona interlocutora, hem fet un primer contacte per correu electrònic, i en cas de no obtenir resposta, hem trucat per telèfon. Hem concertat entrevista sempre per correu electrònic. En alguns casos, un cop concertada l'entrevista, aquesta ha estat anul·lada, per diversos motius. En alguns casos s'ha pogut reprogramar, i en d'altres ha estat posposada sine die.

Per a cada organisme hem definit un guió similar, però adaptat a cada tipus d'institució, segons funcions (annex 3). Totes les entrevistes han estat enregistrades, prèvia autorització signada de la persona entrevistada i compromís de confidencialitat per part de l'entrevistador, en representació del grup de recerca (annex 4).

Concretament, dels 43 municipis seleccionats, hem fet entrevistes als 13 Ajuntaments que van acceptar mantenir una entrevista. Sobte el baix índex de resposta dels ajuntaments. És de remarcar que la fase del treball de camp es va iniciar en el moment previ a les eleccions municipals, i va continuar durant el període de constitució dels consistoris. Tot i que el

interlocutors seleccionats eren de l'àmbit tècnic i no del polític, la majoria dels tècnics varen posposar les entrevistes a la constitució del nou equip de govern. En alguns casos es va perllongar en el temps, pel que alguns municipis van quedar exclosos de la fase del treball de camp. La segona raó de la baixa participació, la podem trobar basant-nos en les respostes d'alguns municipis apuntant a poca activitat en la gestió de dades obertes. Aquesta explicació podria ser aplicable a la falta de resposta d'altres municipis, a la nostra sol·licitud a participar en el treball de camp, contrastat amb la informació obtinguda de les seues electròniques.

Pel que fa a la Generalitat de Catalunya hem aconseguit realitzar un total de quatre entrevistes. En relació a les Diputacions, són dos les que han participat. I per últim, un total de quatre universitats o centres de recerca han participat en les entrevistes. La taula següent resumeix les entrevistes realitzades.

| ORGANISME | NOMBRE ENTREVISTES |
|--|--------------------|
| Ajuntaments | 13 |
| Generalitat de Catalunya | 4 |
| Diputacions | 2 |
| Universitats i Centres de Recerca | 3 |
| Consorti serveis universitaris de Catalunya (CSUC) | 1 |

2.2.4 Enquestes

Per últim, a partir dels resultats de les entrevistes realitzades hem elaborat un formulari per recollir les dades mitjançant una enquesta en línia als ajuntaments (annex 5). Totes les preguntes tenen respostes tancades. Tot i així, en totes les preguntes s'ha deixat l'opció "altres", per tal de recollir possibles respostes no contemplades en l'anàlisi resultant de les entrevistes.

D'acord als criteris de selecció de la mostra abans apuntat, hem enviat el qüestionari de l'enquesta a tots els ajuntament capitals de comarca, que són 43 dels 949 existents a Catalunya. En alguns casos (35) s'han enviat directament a la bústia de la persona identificada com responsable d'alguna manera de les dades obertes, i en la resta (8) a la bústia genèrica de l'ajuntament o de l'àrea de la que se'ns ha informat per telèfon. Aquesta mostra representa l'administració catalana amb un marge d'error del 14,16% amb un nivell de confiança del 95%. Dels 10 que han respost, 6 són els que ja havien respost a les entrevistes.

L'enquesta ha estat oberta un mes, i s'han enviat dos recordatoris. Han respost 10 ajuntaments dels 43, que equival al 23.26% de resposta, amb un nivell de confiança del 95%, hi ha un marge d'error del 27.47%. El baix percentatge de resposta coincideix amb el baix percentatge de resposta dels ajuntaments a participar en les entrevistes. Per tant, la baixa participació de les unitat objecte d'anàlisi en el treball de camp és significativa de cara a las conclusions de l'estudi. Aquestes podran ser contrastades i/o corroborades per un futur estudi, que tingui en compte els avenços que es facin en la gestió de dades obertes en les administracions públiques catalanes.

Queda pendent de fer enquestes als responsables de gestió de dades obertes, per les mateixes raons que no hem efectuat entrevistes.

Finalment, pel que fa les universitats i centres de recerca, a la llum de les entrevistes efectuades, vàrem optar per excloure-les de les enquestes. Durant les quatre entrevistes en profunditat els entrevistats donaven una visió encara primerenca de la gestió de les dades de recerca per part de les universitats. De fet, encara s'estan debatent internament algunes de les qüestions que un qüestionari podia fer emergir, com ara la utilització o no del mateix

repositori institucional per a dipositar les dades obertes o la necessitat de crear-ne un de nou. Aquest aspecte serà descrit amb més amplitud en els següents apartats.

Així mateix per a conèixer la visió d'investigadors, el CSUC ja va donar a conèixer el 2019 la segona enquesta que havien dut a terme entre els investigadors, la finalitat de la qual era "conèixer les necessitats dels investigadors en matèria de gestió de dades de recerca així com tenir una visió general del grau de coneixement del servei de suport a la gestió de dades de recerca que s'ofereix des de les universitats".

És per això que es va creure oportú parlar a les entrevistes en profunditat de tots aquests aspectes i no replicar l'enquesta, sinó aprofitar els resultats i l'anàlisi de l'enquesta ja duta a terme. Al cap i a la fi, aquests és un dels objectius de la pròpia ciència oberta, optimitzar recursos materials i de recerca per a millorar l'eficiència en la captura i compilació de dades en el marc de la recerca. Els resultats de l'enquesta s'esmentaran i comentaran en l'apartat de resultats.

2.3 Anàlisi de dades

Les dades recollides durant el buidatge de les fonts documentals i pàgines web, les hem traslladat a un document, estructurat d'acord a les preguntes de recerca i els objectius del projecte.

Per analitzar el focus group, les entrevistes i les respostes obertes de les enquestes, prèviament hem transcrit els enregistraments, i hem construït uns codis d'anàlisi d'acord als objectius de recerca definits. La gestió de l'anàlisi ha comptat amb el suport del programari d'anàlisi qualitatiu ATLAS.TI. Posteriorment hem analitzat les respostes agrupades pels codis resultants.

Les enquestes han donat unes dades ja estructurades, d'acord a les preguntes tancades. Aquestes han estat tractades amb estadística descriptiva, i posteriorment les hem analitzat. Hem posat en diàleg l'anàlisi d'aquestes amb el resultat previ de l'anàlisi de les entrevistes, per tal de complementar, confirmar o apuntar dissonàncies entre entrevistes i enquestes. Per últim hem integrat l'anàlisi resultants amb els resultats preliminars de les entrevistes.

3 Seguiment i avaluació

El projecte des del seu inici ha comptat amb una cronograma que recollia el detall i la periodificació de les tasques identificades en cadascuna de les cinc fases en que s'ha dividit el projecte. En el cronograma s'especificava les dates d'inici i finalització de cada fase i de les activitats a executar en cadascuna d'elles.. Aquestes fases han estat les següents: 1) Revisió de la literatura; 2) construcció del marc d'anàlisi, 3) treball de camp i 4) anàlisi de les dades i 5) redacció de l'informe final (Figura 14).

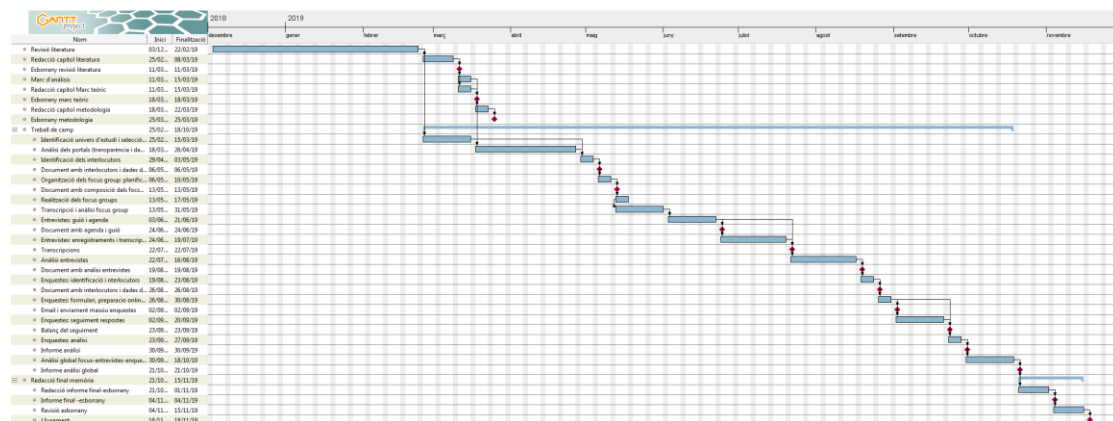


Figura 14: Planificació del projecte

Durant les dues primeres fases hem establert reunions quinzenals per repartir la feina i comentar els avenços assolits. Al finalitzar cada fase, dins de l'equip analitzàvem les dades obtingudes, a partir de les quals concretàvem les següents accions a realitzar. En cas de retard respecte a la planificació, definíem els ajustos i correccions a realitzar.

El major retard ha estat en el treball de camp, atès que és una fase que depèn de l'agenda dels actors involucrats en la gestió de les dades obertes, i que treballen en les unitats d'anàlisi seleccionats. Aquesta fase ha estat especialment difícil per tres causes. En primer lloc, ens hem trobat amb la dificultat d'identificar i contactar amb les persones que podrien ser les interlocutores en els ajuntaments per fer les entrevistes. En segon lloc, això ha estat influït per les eleccions municipals, com hem esmentat abans, tant pel que fa el període previ com el posterior. Els dos fets han retardat la recollida de dades i, en conseqüència, la resta de fases. En tercer lloc, pocs municipis estaven treballant intensament en la gestió de dades obertes. Aquests fets han obligat a prendre dues decisions. La primera ha estat la d'eliminar la recollida de dades mitjançant l'instrument dels focus group, que ha quedat només acotat a un acció dins la Generalitat. La segona decisió ha estat la de reduir la previsió d'entrevistes fetes, donada la baixa resposta per participar en les entrevistes en profunditat. Per la resta, s'ha complert el pla de treball establert en el projecte presentat a la sol·licitud, tot i que amb menys temps del previst.

En el cas del món acadèmic, les quatre entrevistes han permès, juntament amb l'estudi de la literatura i enquestes publicades en l'àmbit a nivell català, espanyol i internacional obtenir les dades suficients per a l'extracció de conclusions òptimes i les respostes als objectius marcats.

4 Resultats

En aquest apartat presentem les respostes plantejades en el projecte de recerca, que es desprenen dels objectius. En primer lloc descrivim quines són les pràctiques de les administracions públiques estudiades, per a gestionar les dades obertes. En segon lloc presentem quins han estat els factors que han facilitat o inhibit aquestes pràctiques. En tercer lloc recollim quins indicadors de qualitat i reutilització d'informació s'estan fent servir a les administracions públiques catalanes. Per últim exposem quins estàndards de gestió de dades obertes o en general s'han implantat a les administracions públiques.

4.1 Pràctiques per gestionar dades obertes

En aquest apartat presentem les pràctiques de les administracions públiques per gestionar les seves dades, obrir-les i posar-les a disposició de la ciutadania. Ho fem tenint en compte els següents aspectes: 1) Quines estructures administratives y orgàniques han creat les administracions públiques. 2) Quins recursos han assignat a aquestes estructures per poder definir estratègies i plasmar-les en una operativa; 3) quins processos y procediments s'han definit i com s'han plasmat; i 4) quin és el vincle amb la reutilització.

4.1.1 Estructures organitzatives

En aquest apartat exposem quines estructures han definit les administracions públiques per a gestionar les dades obertes. Entenem per estructura com s'ha organitzat la funció per definir com serà la gestió de les dades obertes .

4.1.1.1 Ajuntaments

Són pocs els ajuntaments que tenen definida un estructura dedicada a la gestió de les dades obertes. En uns casos s'ha reestructurat algun departament i l'han dotat de la funció de gestió de les dades obertes. En les entrevistes veiem que només en un cas han creat un departament específic. Amb l'enquesta s'ha confirmat aquesta tendència, un 50% dels ajuntaments que han respost tenen definida la funció de gestió de dades obertes dins d'un departament que te assignades d'altres funcions. Tant en les entrevistes com en les enquestes constatem que hi ha un predomini de l'àrea d'informàtica com la responsable de la gestió de dades obertes, seguit pel Departament encarregat de gestionar la transparència (Figura 15).

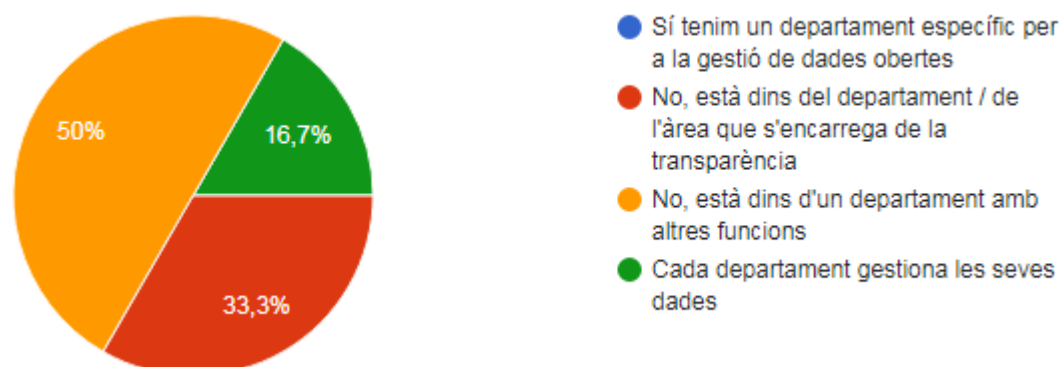


Figura 15: Estructura organitzativa per a gestionar les dades (Font: elaboració pròpia)

En el cas de les entrevistes ha emergit encara una altra figura que te assignada la responsabilitat de gestionar les dades obertes recollida pel departament d'informàtica. Aquesta figura és la persona que ocupa la Secretaria de l'Ajuntament.

És de remarcar que segons les enquestes, el 33,3% d'ajuntaments tenen unida la funció de transparència a la de gestió de les dades (Figura 15). Això pot tenir implicacions pel que fa a la supeditació de les dades obertes al que marca la llei de transparència. Una d'aquestes implicacions és la de limitar el ventall de temes sobre el que preparar dades, com sembla apuntar el 50% dels enquestats que declara publicar les dades obertes referent als temes que senyala la llei de transparència (Figura 16). En percentatges més baixos es publiquen dades referents a altres departaments. Això pot estar influït a la part de la indústria de reutilització interessada en consumir les dades obertes que publiquin els ajuntaments.

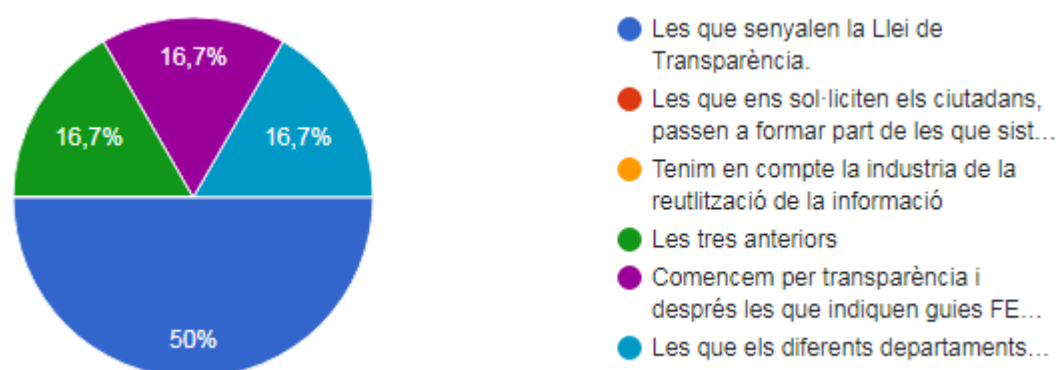


Figura 16: Dades obertes que es publiquen (Font: elaboració pròpia)

És baix el nombre de departaments que gestionen les seves dades. És un indicatiu de manca d'una política clara de gestió de dades, i de que pot haver-hi descoordinació entre les àrees. Això podria tenir com a conseqüència l'existència d'illes informacionals (Figura 15). Aquest fet està corroborat a les entrevistes, i a les enquestes ja que el 66,7% dels enquestats afirmen no tenir cap procediment per vehicular dades obertes generades als diferents departaments.

4.1.1.2 Generalitat

En els darrers anys s'ha fet una estructura organitzativa que en certa manera està dissenyada en xarxa y coordinada per la Subdirecció de Transparència i Dades Obertes. Aquesta àrea dissenya de manera estratègica les accions que tots els Departaments de la Generalitat i empreses públiques han de desenvolupar en matèria de dades obertes. La Subdirecció es coordina amb totes les conselleries mitjançant les Unitats d'Informació, i amb el CTTI. Les Unitats d'Informació, creades amb l'objectiu de desplegar la llei de transparència, en aquests moments també desenvolupen tasques de gestió de dades obertes. Per la seva banda, el CTTI també té un representant als diferents Departaments de la Generalitat, per temes de caire tecnològic, però alhora també es coordina per temes de dades obertes. Aquest representants mantenen reunions periòdiques per tractar temes comuns.

Recentment, el Departament de polítiques digitals ha rebut l'encàrrec de treballar sobre la governança de la dada, per al seu ús a nivell intern. D'alguna manera s'hauran d'articular mecanismes de coordinació entre departaments.

4.1.1.3 Centres de recerca i universitats

La pràctica de compartir dades obertes es troba encara en un estadi inicial a nivell d'universitats i centres de recerca. El principal finançador que les demana com a requeriment en les seves convocatòries és la Unió Europea per als projectes Horizon 2020. A partir de l'any 2017 va deixar de ser un pla pilot en algunes disciplines i es va començar a demanar en tots els projectes. Per tant, encara està previst que han d'arribar els datasets associats a aquests projectes, que normalment són de 3 anys. A partir d'aquell moment es podrà avaluar amb encert i eficiència quines han estat les principals pràctiques.

De moment, s'ha pogut comprovar a les entrevistes que les pràctiques habituals passen per la proactivitat de certs investigadors que estan més compromesos amb el moviment per l'accés obert, i en conseqüència han començat a desar les dades obertes en el repositori institucional. Algun d'aquests investigadors són de ciències socials, que generen dades de mida i memòria més petita que d'altres disciplines, més intensives en gran consum i generació de dades. També s'ha detectat que hi ha disciplines que treballen en repositoris temàtics especialitzats, on el que marca més és la dinàmica de la comunitat de compartir.

L'aprenentatge que estan duent a terme les institucions acadèmiques ha fet que sovint alguns d'aquests projectes a més a més de pioners hagin servit per a veure les implicacions i els fluxos de treball associats, com en el cas del projecte TESLA a la Universitat Oberta de Catalunya. La creació de grups de treball a nivell del CSUC ha permès poder compartir adequadament dubtes i la creació de plantilles d'un document molt rellevant en els grans projectes finançats, el Data Management Plan. Aquest document és un pla de gestió de dades que els grups de recerca han de presentar juntament amb la sol·licitud del projecte.

4.1.2 Recursos

Un altre aspecte vinculat a les pràctiques de gestió de les dades obertes està format pels recursos dedicats a la gestió. Els recursos estan constituïts per elements com persones, xarxes de relacions i maquinari molt especialment el portal de dades obertes. En aquest apartat també hem inclòs el vincle amb d'altres xarxes com a part del recursos que utilitzen les administracions per desplegar la gestió de les dades obertes, perquè són un suport per a gestionar coneixement, compartir experiències i optimitzar esforços. A continuació presentem aquests aspectes.

4.1.2.1 Ajuntaments

El nombre de persones dedicades a la gestió de les dades obertes, en general, és baix. Les entrevistes majoritàriament senyalen la manca de recursos. El resultat de les enquestes confirma aquesta tendència, ja que la mitjana de persones treballant es de 2,5, oscil·lant entre 0 persones en alguns ajuntaments a 6 en d'altres.

Pel que fa la infraestructura tecnològica, pràcticament tots els ajuntament disposen de portal de dades obertes, tot i que en molts casos correspon al que el Consorci AOC posa a disposició dels ens locals. Tots afirmen tenir un cercador, seguit per una majoria que disposa de cercador i categories, i per últim un menor nombre de plataformes disposa d'un quadre de classificació a més del cercador i categorització (Figura 17). Això no coincideix amb un aspecte tractat en el focus group, i corroborat a les entrevistes de la Generalitat, quan es va afirmar que uns dels datasets més descarregat era el del quadre de classificació dels ens locals, amb els codis. En futurs estudis haurem de veure les raons de descàrrega del quadre de classificació i l'ús que es fa d'aquest.

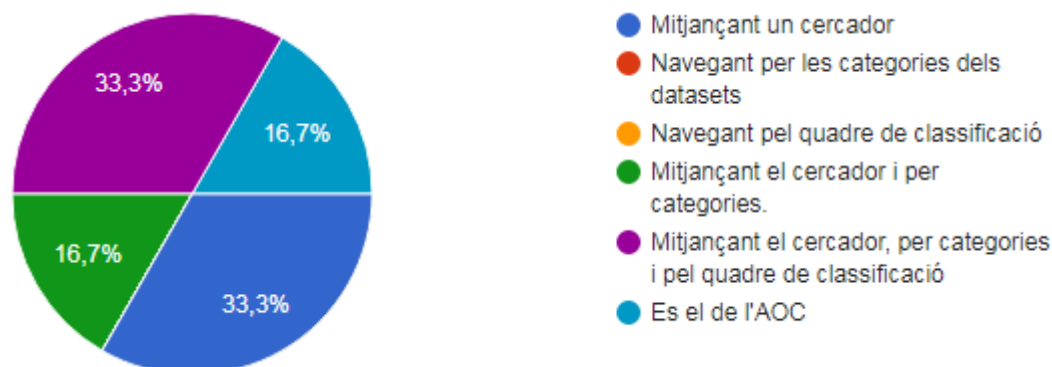


Figura 17: Instruments de cerca en el portal de dades obertes (Font: elaboració pròpia)

L'observació d'experiències en altres administracions ha estat una pràctica majoritària (60%), tenint com a referents tant ajuntaments catalans, com internacionals.

Per últim, senyalar que hi ha escassa participació en grups de treball sobre dades obertes en concret, i en xarxes de municipis en general.

4.1.2.2 Generalitat

A les entrevistes s'observa l'optimització dels recursos existents com una clau rellevant per a desplegar l'estratègia i l'operativa de la gestió de les dades obertes. Tot i així, a les entrevistes es manifesta escassetat de recursos i la necessitat de convocar places per tal de contractar persones més especialitzades en el àmbit de les dades obertes. Aquesta escassetat de recursos es materialitza en l'endarreriment de la posta en marxa d'uns projectes, o bé en la durada del desplegament d'altres projectes.

En relació a la plataforma de dades obertes, la Generalitat compta amb una basada en un programari propietari, Socrates. És una plataforma que facilita efectuar les consultes que precisen els ciutadans, i molt especialment els periodistes de dades o altres empreses de reutilització d'informació. Aquest maquinari també ha estat implantat per algun dels ajuntaments entrevistats.

4.1.2.3 Centres de recerca i universitats

Des del punt de vista del món acadèmic i investigador, hem observat que els recursos emprats a l'hora de treballar per a posar les dades de recerca en obert són principalment de dos tipus: humans i materials.

Pel que fa als recursos humans, han estat els serveis bibliotecaris en les universitats els que majoritàriament han assumit l'encàrrec de la gestió per compartir les dades obertes. Cal distingir en aquest punt dos aspectes. Per una banda, els grups de recerca poden tenir la necessitat d'espais i entorns per a la gestió continuada de les dades generades en la seva recerca. D'altra banda hi ha la tasca de preservació i difusió. És en aquest segon aspecte on el rol de les biblioteques universitàries apareix. Així, hi ha institucions on un dels seus treballadors a la biblioteca ha passat a encarregar-se d'aquests aspectes (incloent formació adequada i continuada). Aquesta persona encarregada de les dades obertes és el que es coneix com a data librarian o data curator. Les tasques associades són tant l'acompanyament en el disseny dels Data Management Plans com vetllar per implantar una infraestructura adequada per a poder garantir la preservació mínima de les dades durant 10

anys, tal com demanen els principals finançadors. En el cas dels instituts de recerca es valora que hi ha una barreja de perfils assignats a la tasca de la gestió de dades, amb un perfil més tecnològic/informàtic.

Quant als recursos materials, a les universitats el camí ha estat majoritàriament ajudar en dos aspectes: a) s'ha recomanat un repositori extern, per a la que s'ha treballat una guia conjunta a nivell del grup de treball del CSUC (Figura 18) o b) s'ha apostat pel repositori institucional ja existent.

Infraestructures de gestió de dades de recerca

Recull de portals i repositoris de dades de recerca.

És recomanable utilitzar recursos europeus per a assegurar-se que compleixen la llei de protecció de dades.

| | |
|---|--|
|  Zenodo Repositori multidisciplinar d'accés obert desenvolupat pel CERN. |  EUDAT Infraestructura col·laborativa que proporciona serveis per a la gestió de dades de recerca, formació i consultoria. |
|  Dariah Infraestructura per a investigadors d'arts i humanitats que treballen amb mètodes de recerca digital. |  CLARIN - European Research Infrastructure for Language Resources and Technology Infraestructura europea per a gestionar i difondre recursos lingüístics seguint els principis FAIR. |
|  Figshare Repositori multidisciplinar d'accés obert desenvolupat per Digital Science. |  re3data.org Directori de repositoris de dades de recerca. |

Figura 18: Portals i repositoris de dades obertes
(Font: <http://biblioteca.uoc.edu/ca/recerca/eines>)

Pel que fa al repositori institucional, en tots els casos ha calgut fer una avaluació sobre si conceptualment i tecnològicament eren els instruments més adients per al dipòsit de les dades de recerca. A nivell català, es treballa amb 2-3 programaris diferents per a suportar aquests repositoris. Això comporta haver d'adaptar camps i metadades, pensats inicialment per a la compartició d'articles científics i altres tipus de treballs, a les característiques dels datasets. Els camps i metadades d'articles i treballs són més textuals que no pas els data sets. Aquests plantegen reptes tècnics que ha calgut anar treballant, com ara la tria del millor format, l'existència o no de requeriments legals i ètics associats a les dades (privacitat, anonimització, propietat intel·lectual, etc.), entre d'altres.

4.1.3 Procediments

En aquest subapartat presentem com s'estan organitzant les administracions públiques per gestionar les dades obertes. Ho fem des del punt de vista de la sistematització de les tasques, i si aquestes s'han materialitzat en directrius, legislació, protocols, mapa de processos y procediments.

4.1.3.1 Ajuntaments

Referent a la definició de processos i sistemàtica de gestió, atenent a les entrevistes són pocs els ajuntaments que treballen el cicle de la gestió de les dades amb processos definits i recollits en protocols. A les enquestes el 66,7% no disposa de procediments per publicar dades al portal, cap te uns procediments per identificar les dades que s'han de posar al servei de la ciutadania, i el 83,3% no disposa d'un procediment específic per atendre les sol·licituds de dades en obert (Figura 19).



Figura 19: Procediment per atendre sol·licituds de dades obertes (Font: elaboració pròpia)

La poca tendència a tenir procediments establerts pot guardar paral·lelisme amb el baix percentatge en que els ajuntaments disposen d'un sistema de gestió (Figura 20), sent el més freqüent un sistema de gestió de la qualitat (ISO 9000). A les entrevistes hem constatat que son pocs els que tenen implantat la UNE 178301 de Ciutats Intel·ligents. Recordem que aquest estàndard estableix la forma d'avaluar la publicació de dades obertes de una ciutat.

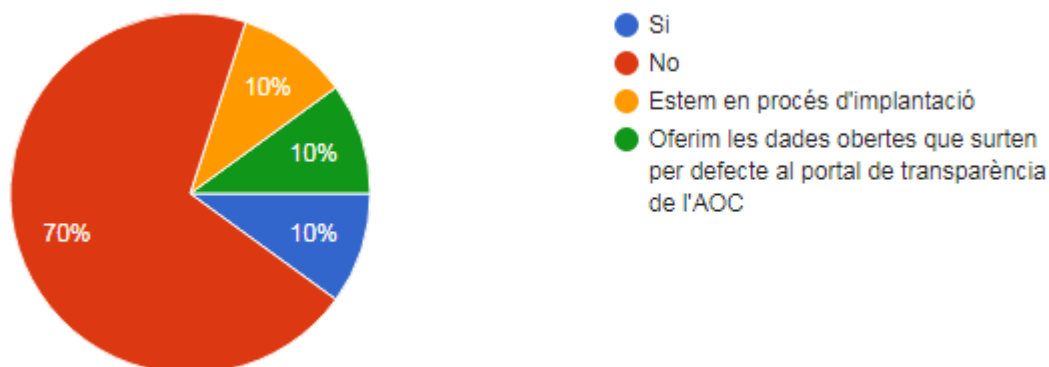


Figura 20: Sistemes de gestió implantats als ajuntaments (Font: elaboració pròpia)

Tot i així a les entrevistes constatem que això pot canviar, atès que hi ha ajuntaments que estan definint processos (Figura 21), i el 37,5% li agradaria tenir una ordenança sobre dades obertes (Figura 22)

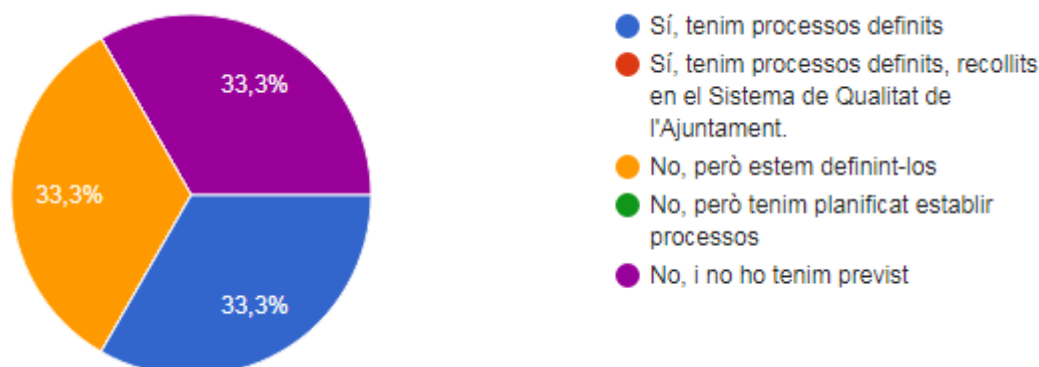


Figura 21: Directrius per gestionar dades obertes (Font: elaboració pròpia)

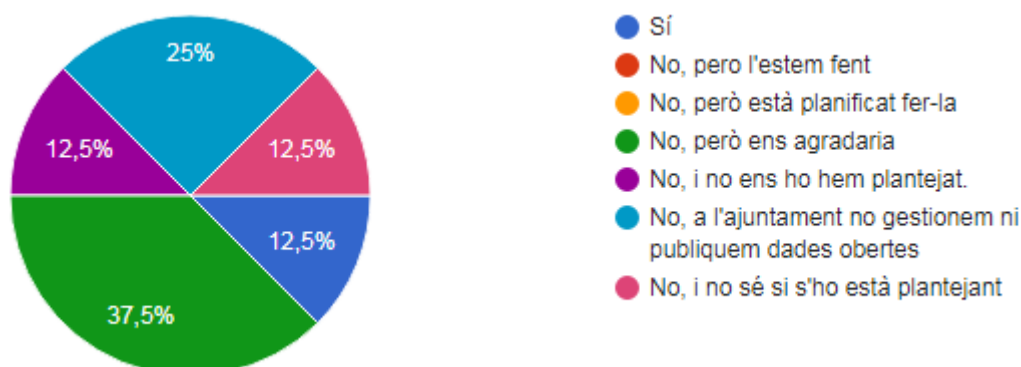


Figura 22: Existència d'ordenança sobre dades obertes als ajuntaments (Font: elaboració pròpia)

4.1.3.2 Generalitat

Des de la Subdirecció General de Transparència i Dades Obertes s'estan dissenyant els processos i els procediments de les tasques que s'han de dur a terme. L'objectiu és harmonitzar les maneres de tractar les dades en totes les oficines de la Generalitat. Això passa també per definir formats, codificació per tal de facilitar la publicació i l'accés als datasets.

De la mateixa manera s'estan definint responsabilitats i identificant repositoris. El CTTI juga un paper important en aquest àmbit i l'AOC també pel que fa sobre tot al recolzament als ens locals. En un futur quan el Departament d'Administració Digital desplegui les seves funcions, s'anirà cap a una harmonització dels sistemes d'informació on s'allotgen les dades, actuant així, potencialment, contra les illes informacionals.

4.1.3.3 Centres de recerca i universitats

No existeixen a hores d'ara unes pràctiques totalment estandarditzades a l'hora de treballar amb la compartició de dades en obert, però sí en cada institució determinats projectes i recerques que donen suport a la creació de dinàmiques i entendre de forma adequada les

implicacions. Així, en alguna universitat s'ha creat un grup de treball entre biblioteca, recercaires, àrea jurídica i tecnologia per a poder donar la millor resposta.

S'ha pogut veure que a l'hora de treballar la gestió de dades, s'ha treballat i es treballa de forma paral·lela en la pràctica i la necessitat més o menys urgents dels investigadors juntament amb la necessitat de crear i aprovar una política de dades. De nou, el treball col·laboratiu del grup de treball del CSUC ha ajudat a crear una certa estructura comuna del que implica una política de dades obertes i què implica.

Aquest mateix repte no només el tenen les universitats sinó que també el tenen les revistes científiques. Per exemple, una de les passes més rellevants ha estat que revistes com PLOS One passessin també a demanar que els articles enviats a la seva revista havien de contenir un enllaç a un repositori on les dades estiguessin a disposició dels investigadors.

Així mateix, finançadors privats com Recercaixa estan seguint les passes de fundacions com Wellcome o Melinda and Bill Gates a l'hora de demanar en els seus plans de finançament la posada a disposició de les dades.

4.1.4 Reutilització

Una de les finalitats de gestionar les dades obertes és la de reutilitzar-les ja sigui al servei del govern obert i el rendiment de comptes, com al servei de la indústria de reutilització. Però també, les dades poden ser reutilitzades a nivell intern per a la presa de decisions. En aquest apartat presentem quin és el vincle de les administracions amb la indústria de reutilització de la informació, què estan fent al respecte i si reutilitzen la informació que generen per a consum intern.

4.1.4.1 Ajuntaments

Pel que fa als vincles que els ajuntaments estableixen amb la indústria de la reutilització, observem, tant en les entrevistes com en les enquestes, que el vincle és baix i esporàdic. El 87,5% dels enquestats afirma no tenir identificades empreses del sector de la reutilització en el municipi, o que utilitzin les seves dades. Pocs ajuntaments han informat tenir identificades empreses usuàries de les dades obertes que l'ajuntament publica. Tot i així, de les entrevistes es desprèn interès en que les dades publicades siguin emprades per la ciutadania, y per empreses per al seu ús en la conversió de les dades en productes i serveis d'informació.

Aquesta preocupació fa que els ajuntaments sensibilitzats amb la gestió de dades obertes treballin per publicar en formats oberts, i ampliïn el catàleg de datasets per tal de satisfer la demanda que els va arribant dels ciutadans. Malgrat aquesta sensibilitat i voluntat de servir totes les demandes, no sempre les sol·licituds poden ser ateses. La limitació ve donada per la limitació de recursos dedicats, com són la dificultat de preparar els datasets requerits, degut a la necessitat d'invertir temps i recursos dels que no es disposa.

Es detecta més interès en subministrar dades al ciutadà que a les empreses. Tot i així emergeix certa sensibilització cap a les empreses del sector infomediari. En cap cas s'ha detectat iniciativa per part d'àrees municipals amb la responsabilitat de desenvolupament del territori.

Un aspecte que també ha emergit en les entrevistes, és el valor de federar les dades amb altres repositoris per facilitar la feina als reutilitzadors. En aquest aspecte juga un paper important la homogeneïtzació de les metadades i les normes tècniques d'interoperabilitat. En

aquest sentit, els ajuntaments hi estan treballant, el 37,5% les te implantades, i el 25% hi està treballant per implantar-les. Per tant, es va per bon camí, però encara queda feina per fer.

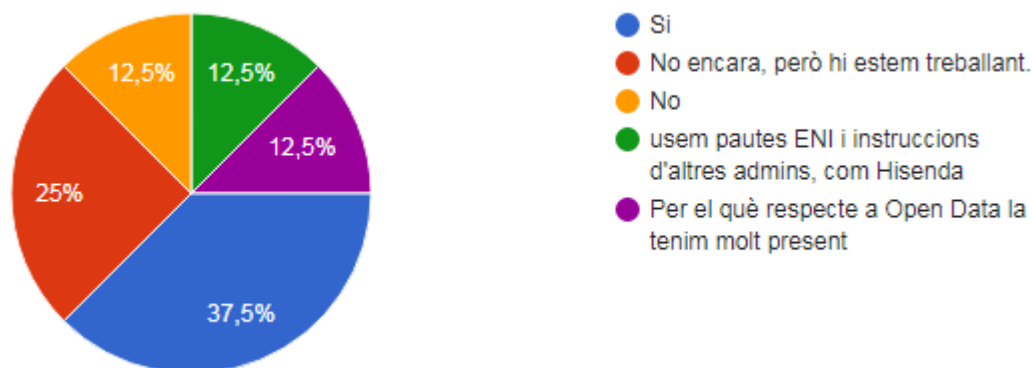


Figura 23: Implantació de Normes Tècniques d'Interoperabilitat (Font: elaboració pròpia)

Per últim, hem d'indicar que en els darrers anys hi ha ajuntaments que han començat a veure els avantatges de disposar d'un portal de dades obertes, per extreure dades per a la pròpia gestió. És un punt útil per localitzar la informació que tenen bolcada. Aquesta informació l'utilitzen per a consum intern amb la finalitat de fer informes o fins i tot per prendre decisions. Tot i així, pel que veiem en les enquestes, encara és majoritari (77,8%) el nombre d'ajuntaments que no utilitzen les dades obertes per a consum intern (Figura 24).

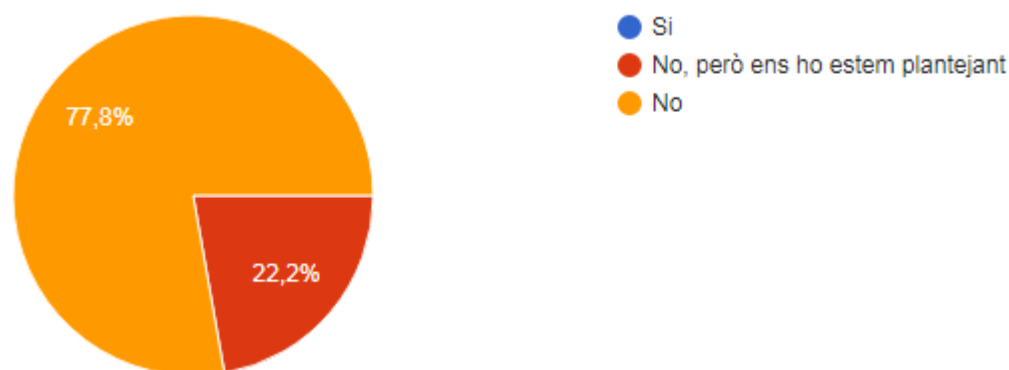


Figura 24: Us intern de les dades (Font: elaboració pròpia)

4.1.4.2 Generalitat

De les entrevistes fetes amb representants de diferents àrees de la Generalitat, es desprèn especial sensibilització per publicar dades obertes i la màxima quantitat possible. Es té voluntat de satisfer a un ampli ventall de ciutadania. La idea és publicar dades tant amb finalitat interna com per respondre la demanda externa. Un exemple de consum de les dades a nivell intern és poder emprar el quadre de classificació documental, o el llistat de centres educatius. Un clar exemple de subministrament de dades a la ciutadania és la demanda de dades per part de periodistes i premsa en general.

Un tret comú de la Generalitat amb els Ajuntaments, és que de moment, les dades obertes també es publica predominantment sota demanda. Per tant, s'obren de manera reactiva. Tot i així, està previst establir mecanisme per ser proactius. Per això es vol detectar necessitats de la societat, i avançar-se a aquestes. Per tant, la tendència és ser més actiu en la publicació de les dades, tot i que continuï el servei de publicar sota demanda, sempre i quan aquesta no

suposi una càrrega de treball desorbitada de temps, pel tractament previ que les dades necessiten per respondre a la petició.

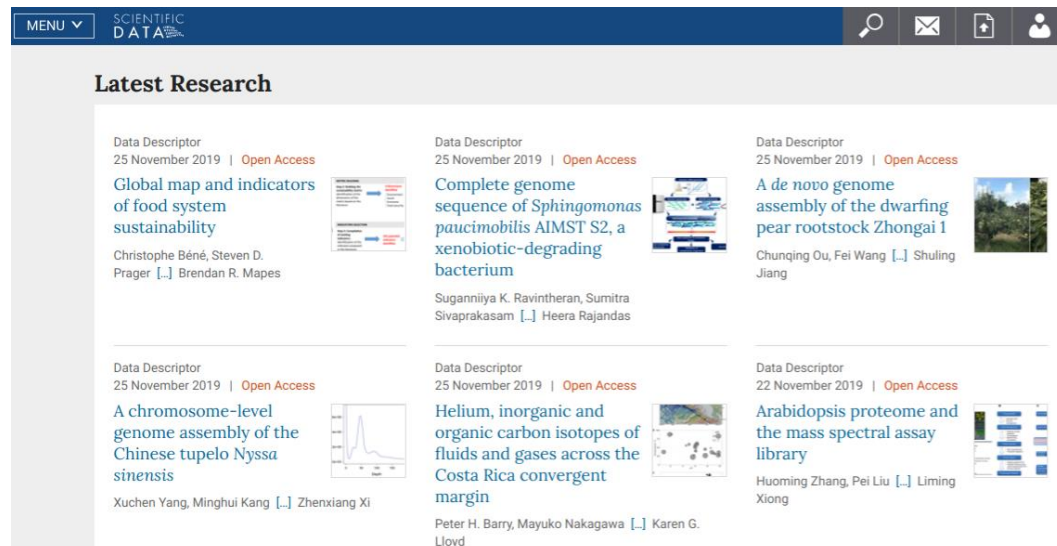
Per impulsar la reutilització de les dades hem de senyalar el desenvolupament de quadres de metadades i la homogeneïtzació de terminologia per contribuir a la qualitat dels datasets i facilitar la reutilització.

4.1.4.3 Centres de recerca i universitats

La reutilització és un dels components del conegut acrònim per a les dades obertes, FAIR que promou que les dades de recerca siguin trobables, accessibles i interoperables. Així, s'entén que aquests principis només tenen sentit perquè al final han de poder ser reutilitzables.

Tot i que la reutilització es el sentit darrer que les institucions europees li han volgut donar també a la compartició de dades, cal tenir en compte que a hores d'ara no sembla que es tingui prou clar quin és l'avantatge directe per als investigadors. Així, no existeixen encara pràctiques clares sobre quina és la millor forma d'afavorir la reutilització. ni sobretot com poder cercar i detectar datasets d'alt valor afegit.

Existeixen, tal com s'ha comentat en els antecedents, revistes de dades dedicades a publicar datasets. Algunes d'aquestes revistes fins i tot estan indexades amb factor d'impacte. Aquestes revistes són una forma de donar també un reconeixement d'output de producció científica. Un exemple és Scientific Data (Figura 25), que compila i recull datasets lliurats pels investigadors de la mateixa manera que es lliuren els article científics. Aquestes revistes disposen també d'un peer review, on s'avalua tant la qualitat com els aspectes formals i tècnics que facilitin la reutilització.



The screenshot shows the 'Latest Research' section of the Scientific Data journal website. It features a grid of six article cards, each with a title, authors, and a small thumbnail image. The articles are:

- Global map and indicators of food system sustainability** by Christophe Béné, Steven D. Prager, and Brendan R. Mapes.
- Complete genome sequence of *Spingomonas paucimobilis* AIMST S2, a xenobiotic-degrading bacterium** by Suganniya K. Ravintheran, Sumitra Sivaprakasam, and Heera Rajandas.
- A de novo genome assembly of the dwarfing pear rootstock Zhongai 1** by Chungqing Ou, Fei Wang, Shuling Jiang, and others.
- A chromosome-level genome assembly of the Chinese tupelo *Nyssa sinensis*** by Xuchen Yang, Minghui Kang, and Zhenxiang Xi.
- Helium, inorganic and organic carbon isotopes of fluids and gases across the Costa Rica convergent margin** by Peter H. Barry, Mayuko Nakagawa, Karen G. Lloyd, and others.
- Arabidopsis proteome and the mass spectral assay library** by Huoming Zhang, Pei Liu, Liming Xiong, and others.

Figura 25: Exemple de revista de publicació de dades obertes (Font: elaboració pròpia, captura de pantalla)

4.2 Factors que faciliten o inhibeixen les pràctiques

En aquest apartat senyalem els aspectes que estan facilitant o retardant la gestió de dades obertes. El que plasmem aquí és un compendi del que ha anat sortint a les entrevistes i en el cas dels Ajuntaments, també el que han recollit les enquestes. Factors que han sortit són: recolzament, recursos, formació, canvi cultural, reconeixement de la funció de gestió de dades, influència de la legislació, i tecnologia.

El comú denominador en totes les administracions públiques (ajuntaments i Generalitat) és la necessitat d'un major recolzament polític per impulsar projectes de gestió de dades obertes. Aquest recolzament s'ha de concretar en el dia a dia, no només amb dotació de recursos i estructures adequades, sinó amb la implicació i resolució dels problemes.

La manca de recursos per recollir i convertir els documents i les dades en formats oberts és un altre dels aspectes que està frenant el ple desplegament de la gestió de dades obertes a les administracions públiques. En la Generalitat el desplegament està en un estadi més avançat, però en els ajuntaments hi ha varies velocitats. Els Ajuntaments grans varen ser els pioners del procés de gestió i publicació de dades obertes, en front dels més petits. Tot i així encara constatem municipis amb un estadi baix, on es va a remolc de la publicació procedent de l'obligatorietat de la llei de transparència. El baix índex de les respostes per part dels ajuntaments, en si ja és una dada representativa, i podria ser l'índex del baix recolzament per impulsar projectes de dades obertes.

La formació és un altre dels aspectes que ha sortit sobretot en les converses amb els ajuntaments. La Generalitat sí que potencia la formació en la gestió de les dades obertes, especialment entre les persones que fan d'enllaç en els diferents departaments, concretament les Unitats d'Informació. L'Escola d'Administració Pública de Catalunya és un referent per impartir formació, amb materials que la Subdirecció elabora. S'està treballant també en la formació a d'altres nivells. Pel que fa als ajuntaments de manera generalitzada es demanda més formació i recolzament per impulsar iniciatives eficients de gestió de dades.

Unit al tema de la formació, des del punt de vista organitzatiu es detecta la necessitat d'un canvi cultural per obrir les dades. Sovint a les entrevistes ha emergit el temor que tenen alguns departaments en alliberar la "seva" dada. En els casos dels ajuntaments que estan més avançats en la gestió de les dades i que ja han vist els avantatges, sobretot a nivell intern, el canvi cultural ja s'està produint. En l'altre extrem es constata que els ajuntaments amb estadis més incipients d'obertura de les dades, el canvi cultural organitzatiu es percep com un factor que està frenant la gestió de les dades.

Unit al canvi cultural, en alguns ajuntaments es percep que l'obertura de les dades va a remolc de la transparència, i per tant aquesta d'alguna manera està actuant com un fre en ampliar les temàtiques sobre les que crear datasets. Això, al mateix temps, està actuant com un fre a la reutilització. Aquest resultat està associat a l'existència d'un departament específic per a gestionar les dades, i per tant al reconeixement de la funció de dades obertes en el si de les organitzacions.

Des del punt de vista legal, l'aplicació de la llei de protecció de dades i la necessitat d'anonimitzar les dades actua com un fre en l'obertura de les dades. En molts casos se senyala la manca de tecnologia que afavoreix i facilita el tractament de dades personals i sensibles.

Pel que fa als centres de recerca i universitats, a través de les entrevistes, hi ha factors coincidents amb els que han sorgit en les administracions públiques. Cal ressenyar que en

les dades de recerca sorgeixen dos principals factors que inhibeixen les pràctiques, però també iniciatives que senyalen com revertir la inhibició:

- En primer lloc, hi ha un factor cultural, més marcat en unes disciplines que en unes altres, de costums i inèrcies. La compartició de dades dins la ciència oberta no troba encara instruments clars, excepte en aquelles disciplines on la pràctica ja era usual abans de les obligatorietats de les polítiques europees i estatals associades al finançament.
- En segon lloc, hi ha tot un seguit de reptes tecnològics, ètics i legals que només és possible superar a través de l'acompanyament i la formació. És a dir, el factor que emergeix és la facilitar al màxim que es pugui dur a terme la pràctica de gestionar i obrir les dades, per compartir-les. Destaquen iniciatives com l'Open Access Week, on el microMOOC en ciència oberta oferta per diferents universitats liderades per la UOC, ajuda a promoure les dades obertes. Per tant, un factor que emergeix és el de seguir fent iniciatives de promoció i formació en la gestió de dades obertes.

A tots aquests factors cal associar els més generals lligats a la ciència oberta, que són els relacionats amb les recompenses i incentius. S'ha detectat com necessari anar afavorint noves dinàmiques i polítiques d'avaluació de la recerca i dels recercaires on donin valor a les pràctiques de compartició de dades. En aquests factors ja es treballa a nivell de polítiques europees, perquè s'entén que cal associar un valor afegit al qui comparteix.

4.3 Indicadors de qualitat i reutilització

La qualitat de les dades és un tema que ha sorgit a les entrevistes i preocupa en totes les administracions. Tanmateix, podríem dir que no és un tema prioritari. Tampoc hem detectat que s'hagin dissenyat i implementat a les administracions públiques indicadors de reutilització per avaluar o al menys per a conèixer com les dades obertes publicades en els seus portals propis estan influïent en la indústria de reutilització.

En el cas dels ajuntaments, el 83% ha manifestat no avaluar la publicació de dades obertes al portal. Tot i així, alguns ajuntaments sí estan desenvolupant un quadre d'indicadors.

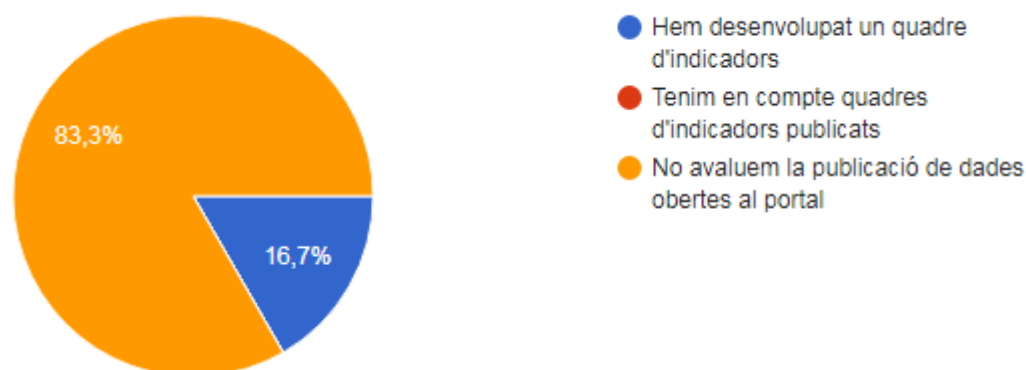


Figura 26: Avaluació de publicació de dades obertes en el portal (Font: elaboració pròpia)

La Generalitat també està treballant en aquest tema però encara que no te implantats indicadors de qualitat, ni te una manera exhaustiva de conèixer com impacta la publicació de les seves dades en la indústria de la reutilització de dades.

Pel que fa als centres de recerca i universitats, no existeixen a nivell de les institucions acadèmiques uns indicadors de qualitat clars a emprar, tant per a detectar els datasets més

rellevants i amb possible valor afegit, econòmic o social, ni per a fer un seguiment oportú de qui i per a què està reutilitzant els datasets. Així, el principal criteri és el fet que ha de ser accessible per a tothom, sense necessitat de crear usuaris que permetrien saber qui es descarrega els paquets de dades.

Es mostra a continuació un exemple de mètriques d'ús que fa el repositori digital de la Universitat Autònoma de Barcelona, el DDD, que funciona igual per datasets com per articles científics.

Estadístiques globals del DDD: <http://www.uab.cat/web/els-nostres-fons/estadistiques-1345756787773.html>

| Accessos per anys i mesos | | | Accessos per àmbit geogràfic | | |
|---------------------------|-----------|--------------|------------------------------|-----------|--------------|
| Anys | Consultes | Descàrregues | Geogràfic | Consultes | Descàrregues |
| 2019 | 44 | 8 | Catalunya | 29 | 5 |
| | | | Estats Units | 5 | 3 |
| | | | Espanya | 5 | 0 |
| | | | Corea del Sud | 2 | 0 |
| | | | Finlàndia | 1 | 0 |
| | | | Alemanya | 1 | 0 |
| | | | Xina | 1 | 0 |

Download: 214778.csv

- Entenem per consulta l'accés a un registre del DDD, o la descàrrega si aquesta no es fa des del registre complet.
- Entenem per descàrrega l'accés al fitxer que conté el text complet, imatge, àudio o vídeo. En els registres que contenen més d'un fitxer se sumen conjuntament totes les descàrregues.
- Eliminem de les estadístiques els accessos dels robots i del personal de gestió. Definim com a robots aquells ordinadors que consulten almenys un cop l'any el fitxer de robots o el servidor oai, o d'altres detectats per nosaltres. Els ordinadors de gestió són aquells que han accedit a pàgines de manteniment.

4.4 Estàndards emprats

L'ús d'estàndards pel que fa a formats oberts és un tema que està en expansió en les administracions públiques. Tot i així, encara no s'estan utilitzant de manera generalitzada als ajuntaments. Coincideix aquest resultat amb la encara incipient gestió de dades obertes.

Altres estàndards com l'impulsat per la Directiva INSPIRE (Infrastructure for Spatial Information in Europe) per a dades geogràfiques o la norma UNE 178301 referent a dades obertes per a Ciutats intel·ligents estan també en la ment d'alguns ajuntaments. És aquesta última norma de ciutats intel·ligents la que està menys introduïda.

Pel que fa la Generalitat, requereix menció especial l'esforç en definir estàndards en el si de la Generalitat pel que fa a terminologia de noms, prenent com a referents altres iniciatives europees.

En el cas dels centres de recerca i universitats, cada comunitat científica treballa per a la creació de bancs de dades específics per a compartir material científic. Un exemple és el Protein Data Bank en el cas de les estructures descrites de proteïnes, on cal dipositar les dades abans de publicar un article. Tot i així, a nivell institucional no hi ha un estàndard clar. Institucions més petites i menys interdisciplinars troben una forma més senzilla de treballar. Pel que fa a les universitat, constatem que les costa trobar un estàndard clar per a donar sortida tant a la física d'altres energies com a les humanitats digitals. Així, els estàndards emprats serien els ja existents per a compartir documents en els repositoris, que fonamentalment és el Dublin Core.

5 Conclusions

Aquest projecte de recerca ha estat una primera immersió en el back office de les dades obertes en les administracions públiques catalanes. Els resultats són una base per seguir fent recerca i conèixer en profunditat les pràctiques i els reptes que les administracions tenen per assolir el ple desplegament de l'obertura de les dades, i que a més redundi en l'impuls de la indústria de reutilització.

El coneixement resultant del projecte alhora contribueix a la reflexió sobre quines accions són necessàries per superar els reptes identificats, i per tant, seguir impulsant l'obertura de dades.

Els reptes identificats pel que fa a ajuntaments i Generalitat de Catalunya són:

- Necessitat de **definir processos i procediments** que abastin tot el **cicle de gestió de la dada**, per tal de garantir la seva integritat. Aquests procediments es podrien recollir en ordenances o ordres per tal d'impulsar la gestió de dades, de manera similar a com la llei de transparència ha impulsat la publicació d'informació.
- **Unificar esforços** en el desenvolupament de metadades homogènies de grups de datasets.
- Establir **vincles amb la indústria de reutilització** més forts que activin mecanismes de publicació proactiva de datasets, tenint en compte els requisits que precisa aquesta indústria per a l'elaboració eficient de productes i serveis. Aquest punt és rellevant atès que és un sector econòmic que contribueix al desenvolupament econòmic d'un territori. Aquest podria ser un àmbit especialment d'interès per aquells municipis amb disminució de població per manca d'ocupació. La identificació de fórmules d'atracció, podria ser una línia de treball vinculada amb les dades obertes.
- Treballar **quadres d'indicadors d'avaluació** per garantir la gestió i publicació de dades en termes de qualitat. Hi ha iniciatives a nivell nacional i internacional que han de ser explorats, perquè poden ser aprofitables. Una autoavaluació dels datasets oferts i com s'estan utilitzant facilita seguir avançant en la direcció adequada.
- Recollir les **necessitats formatives** en l'àmbit de les dades obertes i trobar la forma de planificar la formació i compartir el coneixement entre les diferents administracions públiques.
- Definir accions per impulsar un **canvi de cultura organitzativa** que fomenti l'obertura de dades, i deixi de banda el concepte de "propietat" de la dada a nivell departamental.

Pel que fa als centres de recerca i universitats podem afegir altres reptes.

- **La força motriu de la compartició de les dades en obert està constituïda pels finançadors i les polítiques públiques europees. Per tant, cal impulsar pràctiques i eines que ajudin a compartir les dades de recerca, com un tret propi de la cultura científica. En definitiva, cal promoure accions per generar una cultura de compartició de dades de recerca.**

En el cas de l'accés obert, es va iniciar una força motriu en les declaracions per l'accés obert a les publicacions científiques de Budapest (2001, Bethesda (2003) i

Berlin (2003) i que va ser empresa pels investigadors. Tanmateix, en el cas de la compartició de dades de recerca l'evolució ha estat més de dalt a baix, començant, parlant de l'ecosistema europeu, de les institucions europees a través de les seves polítiques en els darrers 6-8 anys, però també en l'estratègia d'adaptació a la ciència oberta en la qual s'està treballant. Això té uns determinats efectes en la cultura científica i en l'adaptabilitat, visualitzant-se una visió burocratitzadora de la compartició de coneixements científics.

- **Existeix un marc conceptual diferent, on la transparència, la reutilització queden enrere de l'obligatorietat, el FAIR play (*Findable, Accessible, Interoperable and Reusable*) i les dinàmiques disciplinars. Per tant, caldria impulsar pràctiques de transparència i reutilització.**

Tan en les converses com en la literatura científica (informes i articles científics) es veu clarament que les aproximacions al fet OBERT té arrels diferents. En el cas de les universitats i centres de recerca, l'obertura de les dades està lligada a la ciència oberta i la ciència ciutadana. En el cas de les administracions públiques l'obertura sembla emergir com a evolució lògica de la transparència, les polítiques d'accés a la informació i la participació ciutadana i la reutilització de la informació. Creiem oportú una reflexió conjunta entre el marc general de l'administració pública, i la comunitat acadèmica per a compartir experiències, i també, sinergies, atès que algunes de les dades que es poden emprar poden estar o bé en els portals de dades obertes o bé en els repositoris institucionals o temàtics, pensant en una lògica d'interoperabilitat que actualment no es planteja a les universitats. En aquest sentit, tot i que la ciència ciutadana esdevé un dels aspectes de la ciència oberta, no es considera en la part de compartició de dades de recerca, sovint molt especialitzada.

- **El més similar a un portal de dades obertes de les administracions públiques serien el conjunt de datasets en els repositoris digitals. Per tant, cal potenciar una política pública de repositoris de dades obertes, temàtiques o institucionals, fàcilment accessibles al món acadèmic i a la ciutadania en general.**

Els portals de transparència i els portals de recerca, entesos com a metadades del funcionament de les institucions, es treballen des d'unitats i equips diferents, de forma que no comparteixen dinàmiques ni visions relacionades amb la interoperabilitat comuna o bé la reutilització. Així, els repositoris tenen la lògica FAIR i s'estructuren a nivell de directoris internacionals com ara DOAR.

Els repositoris digitals en els centres de recerca estaven pensats inicialment per a documents tradicionals de la comunicació científica (tesis doctorals, articles, capítols de llibre, memòries). S'ha hagut d'adaptar-los com repositori per emmagatzemar datasets, i per a això ha calgut una feina de repensar metadades i millorar aspectes com la capacitat d'emmagatzematge, la validació tècnica, ètica i legal dels conjunts de dades i els permisos. La forma de resoldre no ha seguit la lògica d'un portal de dades obertes com els de les administracions públiques, sinó que s'ha aprofundir en una dinàmica ja existent relacionada amb l'accés obert als documents científics, on ha calgut considerar una nova escala. Per tant, caldrà revisar la plataforma d'emmagatzematge i accés a la informació en termes de cerca, accessibilitat, interoperabilitat i reutilització per a garantir que les dades siguin realment FAIR.

- **Els investigadors continuen tenint necessitats relacionades amb la gestió de les dades obertes.** Per tal, cal promoure la formació en temes com són:
 - Necessitats de formació
 - Necessitats d'incentius pel fet de posar actius en forma de dataset a disposició de la comunitat científica
 - Necessitat de polítiques clares referents a les dades obertes, no només a requeriment extern
- **A diferència de les dades obertes de l'administració, les comunitats acadèmiques disposen també dels seus propis repositoris digitals com a alternativa. Tot i així, el repte és crear un repositori comú pels centres e recerca.**

Per a les biblioteques universitàries – les encarregades de gestionar les dades obertes – ha estat més oportú plantejar-se i debatre si els repositoris eren l'eina adequada per als objectius de la difusió de dades de recerca que no pas crear-ne de nous. Tot i això, i tenint en compte la gran quantitat de dades que es poden necessitar, es planteja la creació d'un repositori consorciat de dades de recerca promogut pel CSUC. És un debat encara obert. Actualment, ens trobem amb un informe que en parla, i descriu molt adequadament els reptes eminentment tècnics que planteja. Són les institucions universitàries participants del CSUC les que han d'acabar veient la viabilitat i el full de ruta per a la creació del repositori.

Dins les universitats, són els serveis bibliotecaris els qui lideren de forma tradicional els aspectes de la ciència oberta, primer en l'accés obert a articles científics, i ara en l'accés a les dades de recerca. Així, els repositoris neixen amb una lògica de catàleg bibliotecari amb accés final al text complet.

- **Existeix un ecosistema de repositoris de dades de recerca ampli, incloent-ne de comercials, entesos com a model de negoci ampli, i repositoris internacionals com ara Zenodo. A més, hi ha asimetria entre disciplines. Per tant, el repte és la construcció d'eines que siguin útils i compatibles amb les dinàmiques que cada comunitat disciplinària està proveint.**

A diferència dels portals de dades obertes de les institucions, el ventall d'opcions de difusió de dades de recerca té tres àmbits. En primer lloc l'àmbit **institucional**, a nivell d'universitat, pensant en les dades generades pels investigadors; en segon lloc l'àmbit **europeu**, com ara Zenodo, per a donar servei sobretot a tots els projectes finançats per les institucions europees (ERC, Horizon 2020). Finalment, el tercer àmbit és el comercial, on s'estan desenvolupant nous productes en el marc de la nova estratègia de les grans empreses relacionades amb la comunicació científica (Elsevier, Springer Nature, American Chemical Society). Aquestes empreses volen oferir serveis globals d'acompanyament als investigadors en tots els estadis de la recerca. Un exemple és Figshare que ofereix algunes avantatges competitives, com ara el fet d'oferir DOIs als datasets dipositats en el seu servei. En aquest cas, no cal passar per processos de validació, cosa que pot permetre competitivitats asimètriques com succeeix per exemple en el cas de xarxes socials acadèmiques tipus ResearchGate en comparació amb els repositoris institucionals, que a més a més també perden visibilitat a efectes de Search Engine Optimization (SEO).

D'altra banda, l'asimetria entre disciplines fa difícil una visió homogènia dins d'una mateixa universitat, emergint una necessària sinergia entre professionals de disciplines independents del centre universitari

- **Les universitats i centres de recerca es troben construint les transicions a marxes forçades entre la necessitat d'infraestructures, recursos humans i formacions i un corpus de bones i males pràctiques. Per tant, cal construir un marc de treball que contempli de manera global necessitats dels usuaris investigadors, infraestructura, persones dedicades, usuaris, processos i procediments.**

Així hem pogut comprovar que mentre es fan estratègies a mig i llarg termini relacionades amb la creació d'infraestructures i la redacció de polítiques de dades obertes de tipus institucional, els investigadors tenen ja necessitats concretes i tècniques per a complir els requeriments dels finançadors. D'alguna forma, com ja hem dit, cal considerar l'efecte extern a l'hora de generar el canvi.

- **La reutilització no és el principal motor de la compartició, tant si es visualitza o no un possible valor econòmic de les dades. Per tant, cal homogeneïtzar objectius i sensibilitzar sobre la diversitat del valor de les dades obertes originada per la reutilització.**

Tot i que la finalitat darrera per part dels decisors de les polítiques europees és la reutilització, no és així per part d'aquells que les comparteixen. Aquests ho fan sovint per la necessitat de complir els requeriments dels finançadors. Això genera algunes disfuncions, i costa trobar exemples de datasets que no hagin estat requerits a ser compartits. La filosofia i estratègia FAIR ha de permetre en darrera instància la reutilització, però aquesta filosofia encara no està del tot establerta. Un dels aspectes que mostra que no es tenen del tot clars els aspectes relacionats amb la reutilització és la manca d'un debat clar sobre si les llicències d'accés incorporen o no l'ús comercial. En la lògica activista de l'accés obert s'entenia que en quedava exclòs. Tanmateix aquest supòsit queda qüestionat per les noves accions lligades al Pla S (iniciativa per a l'accés obert de publicacions científiques iniciat per Science Europe). El Pla S senyala que també els articles científics han de permetre una utilització posterior comercial. Per tant, la reutilització planteja nous reptes a les dades de recerca. La xarxa temàtica Maredata (a nivell estatal) va redactar unes recomanacions dirigides als investigadors sobre com fer que les dades fossin FAIR.

- **Atesa la gran experiència en el treball en l'àmbit de l'accés obert, existeix una dinàmica de col·laboració en l'àmbit de les bibliotecaris universitaris. Per tant, cal mantenir i incrementar les facilitats per mantenir la col·laboració.**

Les biblioteques són les unitats dins les universitats que tenen principalment com a repte la gestió de les dades obertes. Experiències com els grups específics de REBIUN o el grup de treball del CSUC estan permetent treballar de forma eficient en l'aprenentatge comú. Aquest fet cal relacionar-ho amb la necessitat de formació que han experimentat els professionals dedicats a la gestió de dades. Aquests han hagut d'assolir competències i habilitats relacionades amb el que es coneix com a *data librarian* o bé *data curator*, per tal de donar la qualitat necessària als paquets de dades i a més assolir els requeriments FAIR. En el cas dels centres de recerca, són altres les unitats encarregades de la gestió de dades.

Compartir dades de recerca és un fet irreversible, i és d'esperar que l'obligatorietat lligada a projectes avanci en les actuacions de tots els finançadors de recerca, tan públics com privats. Així mateix, quan es pensa en els possibles usuaris de les dades posades a disposició, es pensa en un ecosistema tancat, compost per la resta de la comunitat científica i investigadora.

- **Cal associar un valor afegit al qui comparteix i també anar afavorint noves dinàmiques i polítiques d'avaluació de la recerca i els recercaires on donin valor a les pràctiques de compartició de dades**

La comunicació científica com a fet i l'avaluació de la qualitat i l'impacte de la recerca estan destinades a ser diferenciades. Per inèrcia, criteris tecnològics i facilitat en les mètriques ofertes per productes comercials com ara Web of Science, l'avaluació de projectes, grups de recerca, investigadors i institucions estan encara massa relacionats amb cada cop més un sol tipus de document, l'article científic, i un indicador preferent, el factor d'impacte. Les crítiques tradicionals (de fet, des de la pròpia creació) ha dut al món acadèmic a plantejar-se i replantejar-se com s'avalua la qualitat d'una recerca, i també a generar nous indicadors per a poder avaluar la ciència en un nou paradigma obert, i les dades de recerca haurien de ser uns elements de qualitat i distinció de la bona recerca.

6 Impacte previst

En aquest capítol de la memòria exposem tres aspectes: a) l'impacte científic i socioeconòmic dels resultats del treball de recerca, b) les activitats de divulgació de resultats previstes mitjançant publicacions i presentacions en reunions científiques i de divulgació; i c) propostes d'actuacions.

6.1 Impacte científic i socioeconòmic

Aquest projecte ha iniciat una línia de recerca consistent en estudiar les accions internes de les administracions públiques, universitat i centres de recerca per implantar la funció de gestió de dades obertes en la seva activitat diària. És a dir, hem treballat el *back-office* de les unitats objecte d'anàlisi. Més concretament, l'**impacte científic** es refereix a identificar alguns dels aspectes que s'han de tenir en compte per conèixer quines pràctiques s'estan duent a terme en l'administració pública i centres públics de recerca de Catalunya per tal de facilitar l'accés dels ciutadans (particulars i empreses) a la informació pública, i més específicament a les dades obertes, per tal de contribuir a reutilitzar la informació. Un aspecte rellevant és conèixer com s'estan definint els processos per a garantir l'accés dels ciutadans a les dades obertes i garantir la seva qualitat. Ha servit també per identificar aspectes que s'han de sistematitzar, accions a dur a terme, i quins punts de millora es poden introduir. Aquest coneixement ajuda a establir les següents línies d'actuació en el si de les administracions:

- 1) Aspectes a tenir en compte per establir un **protocol tipus** per **gestionar dades tipus** que reculli accions preventives i correctives, en cas de que no disposin d'un. En aquest protocol s'ha d'incloure accions per reutilitzar les dades a nivell intern, a més de les accions més interioritzades del protocol per publicar dades per a la ciutadania i molt especialment per a les empreses del sector de reutilització.
- 2) Definir processos de **mineria de dades** per tal d'explotar a nivell intern el gran volum de dades que genera el sector públic, tenint en compte la metodologia de la intel·ligència territorial.
- 3) Dissenyar **indicadors** en dos àmbits: a) indicadors de qualitat de les dades obertes que es publiquen en el portal; i b) indicadors de l'ús de les dades obertes per part de la ciutadania, i molt concretament, per part de les empreses del sector de reutilització de la informació.
- 4) Organitzar sessions formatives a partir de les necessitats detectades, per tal d'introduir **bones pràctiques** obtingudes d'**experiències** de **països** del nostre entorn, segons cada context.

Des del **punt de vista socio-econòmic**, l'impacte pot ser a mig i llarg termini, i està lligat a l'impuls del sector de reutilització de la informació en les àrees d'influència de les administracions públiques. Els resultats han posat en evidència el camí que encara han de recórrer les administracions per publicar dades obertes i impulsar la reutilització de la informació. Aquest punt és important, ja que recordem la indústria infomediària o de reutilització de la informació és causa de motor de desenvolupament econòmic, tant si les dades es reutilitzen per consum intern de les organitzacions com si són emprades com matèria primera de la indústria de la reutilització (Tinholt i Carrara, 2017; Espanya, 2015; Alvarez Espinar, 2014).

6.2 Publicacions

La divulgació dels resultats es farà amb posterioritat a la finalització del projecte. Està previst fer-la mitjançant diferents canals:

- Article per a la Revista Española de Transparencia a presentar el pròxim 10 de desembre, per a ser publicat durant el primer semestre de 2020. L'article durà per títol "*El back-office de la gestión de datos abiertos para su reutilización en Cataluña*".
- Article de caire més científic, sobre la gestió de les dades obertes als ajuntaments per presentar en una revista indexada (WoS o Scopus).
- Article de caire científic centrat en l'àmbit estudiat de les dades de recerca en obert, per presentar en una revista indexada (WoS o Scopus) durant el 2020
- *Presentació de comunicació al V Congreso Internacional de la Transparencia* edició 2020. Estem a l'espera que s'obri el call for papers.
- Organització d'un acte a la UOC sobre dades obertes i reutilització de la informació.

6.3 Propostes d'actuacions

A la llum dels resultats, proposem un seguit d'actuacions que poden contribuir a impulsar o optimitzar els esforços que les administracions públiques estan portant a terme per a gestionar dades obertes i posar-les a disposició de la ciutadania. En tots els casos cal definir què s'ha de fer, qui i com.

En primer lloc cal **definir de manera clara la funció de gestió de dades obertes** mitjançant tres accions:

- a) **Establir de manera clara i pública quin àrea gestiona i coordina** la funció de dades obertes, i com es coordina amb la funció de transparència. Aquesta actuació, a més, implica dos accions:
 - a. Definir un àrea o subàrea responsable que treballi amb coordinació amb altres àrees organitzatives.
 - b. Implantar i executar les accions necessàries per fomentar la reutilització de la informació, tant internament com externament. En aquest darrer cas, cal també establir el vincle amb el sector infomediari establert en el territori on governa l'administració, o fomentar-lo, atès que aquest és una de les fonts per al desenvolupament del territori. Aquestes accions cal fer-les amb coordinació amb l'àrea de promoció econòmica del territori, en cas de que existeixi a l'ajuntament o en la Generalitat de Catalunya, i en el cas de les universitat i centres de recerca, possiblement, amb el departament de transferència.
- b) **Dibuixar els processos** necessaris per identificar i publicar les dades i la informació que precisa la ciutadania, dins dels paràmetres de qualitat i federació de les dades. En els processos s'ha de tenir en compte com facilitar l'accés de la ciutadania a aquestes dades, a la informació pública i a la documentació. Aquests processos cal plasmar-los en un mapa, que estigui relacionat amb la resta de processos existents en l'organització i amb les aplicacions informàtiques i repositori que contenen dades i documents, generats en cadascun d'aquests processos.

- c) **Definir els procediments i recursos necessaris, humans i materials, les responsabilitats associades, i les accions d'acompanyament.** És necessari dedicar millors recursos i infraestructures, incloent recursos humans (equips dedicats específicament a tots els aspectes entorn a dades obertes, i també ciència oberta en el cas de les universitats), i formació. Dins dels recursos materials s'han d'incloure els repositori, les aplicacions informàtiques o els sistemes d'informació necessaris per automatitzar la publicació de la informació, i dels datasets. Aquesta acció és especialment rellevant, i a tal fi cal realitzar una auditoria de sistemes d'informació i de les necessitats existents, des del punt de vista de gestió de la informació, per tal d'alinear els sistemes informàtics amb els processos. En cas d'identificar-se requeriments que no queden coberts pels sistemes d'informació o aplicacions ja existents en el si de les administracions, s'ha de planificar la implantació d'aquests nous recursos, ja sigui mitjançant un desenvolupament propi o en consorci amb altres administracions, o adquirint solucions comercials ja existents.

En segon lloc, cal crear un **quadre d'indicadors de qualitat i requeriments**, per avaluar les tasques de gestió de dades obertes que s'estan duent a terme, i els portals on es publiquen les dades. És especialment rellevant que es tinguin en compte les necessitats que la indústria de reutilització de la informació ha manifestat. Vinculat amb això, cal incorporar **indicadors d'ús de les dades** per part de la indústria de reutilització. Es pot partir dels quadres d'indicadors identificats en estudis previs, i detallats en un altre apartat d'aquest document.

En tercer lloc, aquelles organitzacions que no tenen implantada la seva **política de gestió de documents** ni un sistema de gestió de documents, cal que els defineixin, i també que desenvolupin els instruments associats a cada fase de la gestió de documents, molt especialment el quadre de classificació, taules d'avaluació i accés, i l'esquema de metadades. Aquest sistema de gestió i els instruments de classificació i avaluació poden ajudar en la gestió de les dades obertes. Per a definir el sistema i la política és recomanable comptar amb les normes de gestió de documents, especialment la ISO 15489 i la sèrie ISO 30300. Aquesta última norma permet definir la funció de gestió de documents, i vincular-la amb altres sistemes de gestió, com la norma ISO de qualitat (sèrie 9000) o la mediambiental (sèrie 14000). En aquesta línia de treball cal potenciar i actualitzar, si cal, l'àrea de gestió de documents i arxiu per tal de donar servei a la gestió dels actius digitals en els termes que s'estan definint en el futur Decret d'Administració Digital, que ara està en fase de consulta pública (Catalunya, 2019d)

Per últim, cal que les administracions fomentin les **xarxes amb d'altres administracions** públiques per tal d'intercanviar coneixement, i resoldre problemàtiques comunes, ja sigui vinculades a la funció de gestió de dades obertes o a aspectes més concrets de l'administració electrònica, entre d'altres. A més, en aquestes xarxes també poden sorgir sinèrgies i vies per compartir recursos com, per exemple, el desenvolupament o l'adquisició d'aplicacions informàtiques consorciades, o el suport de professionals com enginyers informàtics o graduats en informació i documentació. Pel que fa a les dades de recerca en obert, cal major col·laboració entre universitats i centres de recerca

7 Referències

Abadal, Ernest., Aleixandre, Rafael, Canals, Agustí, Ferrer, Antonia, Garrigós, Irene, Hernández, Tony, López-Borrull, Alexandre, Mazón, José N., Melero, Remedios, Méndez, Eva, Ollé, Candela, Peset, Fernanda (2018). "Maredata: red temática sobre datos de investigación en abierto". Dins ECA 2017 *Congreso Ecosistemas del Conocimiento Abierto* (pp. 105-114). Salamanca: Ediciones Universidad de Salamanca.

Abella, Alberto. (Coord.); Ortiz-de-Urbina Criado, Marta; de Pablos Heredero, Carmen; Vidal Cabo, Christian; Ferrer Sapena, Antonia (2019). [La reutilización de datos abiertos en España II](#). DesideDatum, Universidad Rey Juan Carlos y Cátedra Transparencia i Participació. [Consulta: Novembre, 2019]

Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen (2014). MELODA, métrica ara evaluar la reutilización de datos abiertos. *El Profesional de la Información*, 23(6), 582-588.

AENOR (2011). *UNE-ISO:30301: 2011. Información y Documentación. Sistemas de gestión para los documentos. Fundamentos y vocabulario*. Madrid: AENOR.

AENOR (2015). *UNE 178301:2015. Ciudades Inteligentes. Datos Abiertos (Open Data)*. Madrid: AENOR.

AENOR (2016). *UNE-ISO 15489-1:2016. Información y Documentación. Gestión de documentos. Parte 1. Conceptos y principios*. Madrid: AENOR

AlAnazi, Jazem Mutared; Chatfield, Akemi T. (2012). [Sharing government-owned data with the public: a cross-country analysis of open data practice in the Middle East](#). *18th Americas Conference on Information Systems (AMCIS) 2012 Proceedings* (pp. 1-10). AIS eLibrary. [Consulta: Febrer, 2019]

Aleixandre-Benavent, Rafael; Vidal-Infer, Antonio Alonso-Arroyo, Adolfo; Ferrer-Sapena, Antonia; Peset, Fernando; García García, Alicia (2015). [Gestión de los datos brutos de investigación en los investigadores españoles en ciencias de la salud](#). *Trauma*, 26(1), 73-81.

Álvarez Espinar, Martín (2014). Apertura y reutilización de datos públicos. *Govern Obert*, núm 2. Barcelona: Generalitat de Catalunya.

ASEDIE (2019). Sector Infomediario. Madrid: ASEDIE.

Behkamal, Behshid; Kahani, Mohsen; Bagheri, Ebrahim; Jeremic, Zoran (2014). A metrics-driven approach for quality assessment of linked open data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(2), 64–79.

Bello, Olayiwola.; Akinwande, Victor; Jolayemi, Olayiwola; Ibrahim, Ahmed (2015). Open Data Portals in Africa: An Analysis of Open Government Data Initiatives. *Africa Journal of Library Archives and Information Science*, 26, agost, pàg. 97-106.

Berners-Lee, Tim (2006). [Linked data-design issues](#). Technical Report, W3C.

Bichard, Jo-Anne; Knight, Gail (2012). Improving public services through open data: public toilets. *Municipal Engineer*, 165(3), 157-165.

Both, Wolfgang; Schieferdecker, Ina (2012). Berliner Open Data-Strategie. *Organisatorische, rechtliche undtechnische Aspekte offener Daten in Berlin. Fraunhofer FOKUS im Auftrag der Senatsverwaltung für Wirtschaft, Technologie und Frauen in Berlin*, pp. 172. Berlin: Fraunhofer Verlag.

Canals, Agustí; López-Borrull, Alexandre (2017). "Big Data for Scientific Knowledge". Dins: *18th European Conference on Knowledge Management*, Barcelona 7-8 Septembre.

Catalunya (2014). Llei 19/2014, del 29 de desembre, de transparència, accés a la informació pública i bon govern. DOGC, núm. 6780. 31/12/2014.

Catalunya (2019a) DECRET 83/2019, de 24 d'abril, de reestructuració del Departament d'Acció Exterior, Relacions Institucionals i Transparència

Catalunya (2019b) ORDRE SLT/74/2019, de 15 d'abril, per la qual es crea la Comissió de Dades Obertes del Departament de Salut i del seu sector públic

Catalunya (2019c) DECRET 124/2019, de 4 de juny, de reestructuració del Departament de Polítiques Digitals i Administració Pública.

Catalunya (2019d) EDICTE de 13 de setembre de 2019, pel qual se sotmet a informació pública el Projecte de decret d'Administració Digital.

Catalunya (2104) Llei 19/2014, del 29 de desembre, de transparència, accés a la informació pública i bon govern.

Cerrillo-i-Martinez, Agusti. 2012. Fundamental interests and open data for re-use. International. *Journal of Law and Information Technology*, 20(3), 203-222.

Chretien, Jean-Paul; Rivers, Caitlin M.; Johansson, Michael A. (2016). Make Data Sharing Routine to Prepare for Public Health Emergencies. *PLoS Med*, 13(8) <https://doi.org/10.1371/journal.pmed.1002109>.

Chung, Woo Young; Fisher, Craig; Wang, Richard (2002). What skills matter in data quality? *The 7th International Conference on Information Quality (ICIQ-'02)*, Boston, MA

Comisión Europea (2011). *Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Datos abiertos. Un motor para la innovación, el crecimiento y la gobernanza transparente*. COM (2011) 882 final. Bruselas.

Comisión Europea (2014). [Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Hacia una economía de los datos próspera](#). COM (2014) 442 final. Bruselas.

Comisión Europea (2016). [Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. European cloud initiative - Building a competitive data and knowledge economy in Europe](#). COM (2016). 178 final.

Comisión Europea (2018). [COMMISSION RECOMMENDATION of 25.4.2018 on access to and preservation of scientific information](#).

Conradie, Peter; Choenni, Sunil (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, 31(suplement 1) 10-17 <https://doi.org/10.1016/j.giq.2014.01.003>

CSUC (2017). [Recomanacions per seleccionar un repositori per al dipòsit de dades de recerca: Versió 3.](#)

CSUC (2019). [Data management plans: Version 3](#), June 2019.

CSUC (2019). [FAIR x FAIR. Requisitos factibles, alcanzables e implementables para un repositorio de datos de investigación FAIR.](#)

Currie, Morgen Elisabeth (2016). [The Data-fication of Openness. The Practices and Policies f Open Government Data in Los Angeles](#) (Tesi Doctoral). University of California, Estats Units

Eaves, David (2009) ["The three laws of open government data"](#). *Conference for Parliamentarians: balancing openness and the public interest in protecting information.* [Consulta: Febrer, 2019].

Eppler, Martin J. (2006). *Managing information quality. Increasing the value of information knowledge-intensive Products and Processes.* Berlin: Springer.

Espanya (2007). Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.

Espanya (2011). Reial decret 1495/2011, de 24 d'octubre, pel qual es desplega la Llei 37/2007, de 16 de novembre, sobre reutilització de la informació del sector públic, per a l'àmbit el sector públic estatal.

Espanya (2015). Ley 18/2015, de 9 de julio, por la que se modifica la ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público.

Espanya (2018). Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

Federación Española de Municipios y Provincias (2017). [Datos abiertos: guía estratégica para su puesta en marcha. Conjunto de datos mínimos a publicar.](#) Wolters Kluwer, S.A. [Consulta: Febrer, 2019]

Federación Española de Municipios y Provincias (2019). [Datos abiertos FEMP 2019. 40 conjuntos de datos a publicar por las Entidades Locales.](#) [Consulta: Octubre, 2019]

Ferro, Enrico; Osella, Michelle (2012). Business models for PSI re-use: a multidimensional framework. Comunicació a: *The European Commission Workshop "Using Open Data: Policy Modeling, Citizen Empowerment, Data Journalism"*, Brusel·les.

Fournier-Tombs, Eleonore (2011). Evaluating the impact of open data websites. Working papers series-SSRN. <http://dx.doi.org/10.2139/ssrn.1926201> [Febrer, 2019]

García García, Alicia; López Borrull, Alexandre; Peset, Fernanda (2015). Data journals: eclosión de nuevas revistas especializadas en datos. *El profesional de la información*, 24(6), 845-854.

Garcia-Alsina, Montserrat; Canals, Agustí; Cobarsi, Josep; Ortoll, Eva; Martinez, José Antonio; Cermeno, Lluís; Sardà, Jaume; Santanach, Daniel. (2018). El back-office de la transparencia: su puesta en marcha en Cataluña. *Actas del II Congreso Internacional de la Transparencia*, Madrid, septiembre 2017

García-García, Alicia; López-Borrull, Alexandre; Peset-Mancebo, Maria Fernanda (2015). Data journals: eclosión de nuevas revistas especializadas en datos. *El Profesional de la Información*, 24(6), 845-854.

Gasser, Urs (ed.) (2004). *Information Quality regulation: foundations, perspectives, and applications*. Baden: Nomos : Schultess.

Gironès Roig, J. (2013). *Data mining*. Barcelona: FUOC.

Gobierno de Colombia - Ministerio de Tecnologías de la Información y las Comunicaciones (2016). *Guía de datos abiertos en Colombia Dirigida a las entidades sujeto de aplicación de la Ley 1712 de 2014 de Transparencia y Acceso a la Información Pública, para la aplicación de orientaciones y buenas prácticas en el desarrollo de estrategias de apertura y reúso de datos abiertos*

Gobierno de Colombia - Ministerio de Tecnologías de la Información y las Comunicaciones (2019). [Guía para el uso y aprovechamiento de Datos Abiertos en Colombia - modificaciones de diagramación](#). [Consulta: juny, 2019]

Gómez, Nancy-Diana; Méndez, Eva; Hernández-Pérez, Tony (2016). Social sciences and humanities research data and metadata: A perspective from thematic data repositories. *El profesional de la información*, 25(4), 545-555. <http://dx.doi.org/10.3145/epi.2016.jul.04>

Groves, Peter; Kayyali, Basel; Knott, David; Kuiken, Steve Van (2013). *The 'big data' revolution in healthcare: Accelerating value and innovation*. Mckinsey&company (ed.). <http://hdl.handle.net/11146/465>

Grupo de Trabajo de "Depósito y Gestión de datos en Acceso abierto" del proyecto Recolecta (2012). La conservación y reutilización de los datos científicos en España. Informe del grupo de trabajo de buenas prácticas. Madrid: FECYT. <http://eprints.rclis.org/21007/>

Gurstein, M. B. (2011). Open data: Empowering the empowered or effective data use for everyone? *First Monday*, vol. 16, núm. 2.

Hernández-Pérez, Tony (2016). En la era de la web de los datos: primero datos abiertos, después datos masivos. *El profesional de la información*, 25(4), 517-525. <http://dx.doi.org/10.3145/epi.2016.jul.01>

Hossain, Mohammad Alamgir; Dwivedi, Yogesh K.; Rana, Nripendra P. (2016). State-of-art-in open data research: insights from existing literature and a research agenda. *Journal of Organizational Computing and electronic commerce*, 26(1-2), 14-40.

ISO (2008). ISO/IEC 25012:2008. *Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model*. Ginebra: ISO

Janssen, Katleen (2012). [Open government data and the right to information: Opportunities and obstacles](#). *The Journal of Community Informatics*, 8(2), 2.

Junta de Castilla y León – Consejería de Fomento y Medio Ambiente (2019). *Open Data. Publicación y reutilización de Datos Abiertos como iniciativa de Gobierno Abierto en la Administración*.

Kenett, Ron S.; Shmueli, Galit (2014). On information quality, *Journal of the Royal Statistical Society*, 177(1), 3-38.

Kučera, Jan. (2014): *Open Government Data*. University of Economics, (Tesis Doctoral) Prague, Faculty of Informatics and Statistics, Xequia.

LEARN (2017). [Toolkit of best practice for research data management](#).

Lemieux Victoria Louise; Gormly, Brianna; Rowledge, Lyse (2014). Meeting Big Data challenges with visual analytics. The role of records management. *Records Management Journal*, 24(2), 122-141.

LERU (2013). [LERU roadmap for research data](#). Advice paper n. 14.

López-Borrull, Alexandre. (2014). Prom-Open: Promoting open science in an on-line e-learning environment. *Edulearn 2014 6th International Conference on Education and New Learning Technologies*. Barcelona 7-9 July 2014.

Lourenço, Rui Pedro (2015). An analysis of open government portals: A perspective of transparency for accountability. *Government Information Quarterly*, 32(3), 323-332.

Máchová, Renate; Lněnička, Martin (2017). Evaluating the Quality of Open Data Portals on the National Level. *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1), 21-41.

Maurino, Andrea, Spahiu, Blerina, Batini, Carlo, Viscusi, Gianluigi (2014). Compliance with Open Government Data Policies: an empirical evaluation of Italian local public administrations. *Twenty Second European Conference on Information Systems*, Tel Aviv.

McDonald, John; Leveille, Valerie. 2014. Whither the retention schedule in the era of big data and open data? *Records Management Journal*, 24(2), 99-121.

McLeod, Julie (2012). Thoughts on the opportunities for records professionals of the open access, open data agenda. *Records Management Journal*, 22(2), 92-97.

Melero, Remedios (2018). "Recomanacions per a la gestió de dades d'investigació dirigides a investigadors" <http://hdl.handle.net/10609/98347>

Melero, Remedios; Hernández-San-Miguel, Javier (2014). Acceso abierto a los datos de investigación, una vía hacia la colaboración científica. *Revista Española de Documentación Científica*, 37(4) <http://dx.doi.org/10.3989/redc.2014.4.1154>

Metzler, Katie; Kim, David A.; Allum, Nick; Denman, Angella (2016). Who Is Doing Computational Social Science? Trends in Big Data Research. *White paper*. London, UK: SAGE Publishing.

Nacions Unides (2013). [Open Government Data and Services](#). [Consulta: febrer: 2019]

Nugroho, Rinita Putri; Zuiderwijk, Anneke; Janssen, Merijn; de Jong, Martín (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy*, 9(3), 286-308.

Moedas, Carlos; Oettinger, Günther H. (2015). [Open science for a knowledge and data-driven economy](#). [Consulta: marzo 2020].

Pabon, Gustaco.; Gutierrez, Claudio.; Fernandez, Javier F.; Martinez-Prieto, Miguell A. (2013). Linked Open Data technologies for publication of census microdata. *Journal of the American Society for Information Science and Technology*, 64(9), 1802-1814.

Panger, Galen (2016). Reassessing the Facebook experiment: critical thinking about the validity of Big Data research. *Information, Communication & Society*, 19(8), 1108-1126.

Peset, Fernanda; González, Luis-Millán (2017). *Ciencia abierta y gestión de datos de investigación*. Gijón: TREA.

Piattini, Mario G.; Calero, Coral; Genero, Marcela F.(ed.) (2002). *Information and database quality*. Boston (etc.): Kluwer Academic.

Piovesan, Federico (2015). *Beyond standards and regulations: obstacles to local open government data initiatives in Italy and France*. (Tesis Doctoral). Politecnico di Torino, Italia.

Rohunen, Anna; Markkula, Jouni; Heikkila, Marikka; Heikkila, Jukka (2014). Open traffic data for future service innovation: addressing the privacy challenges of driving data. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 71-89.

Sunlight Foundation (2014). [Open Data Policy Guidelines](#). [Consulta: febrer, 2019]

Tenopir, Carol; Dalton, Elizabeth D.; Allard, Suzie; Frame, Mike; Pjesivac, Ivanka; Pollock, Danielle; Dorsett, Kristina (2015). Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS ONE*, 10(8) <http://dx.doi.org/10.1371/journal.pone.0134826>.

Tenopir, Carol; Allard, Suzie; Douglass, Kimberly; Aydinoglu, Arsev Umur; Lu, Lei; Read, Eleanor; Manoff, Maribeth; Frame, Mike (2011). Data sharing by scientists: practices and perceptions. *PLoS ONE*, 6(6) doi:10.1371/journal.pone.0021101

Thurston, Anne Catherine (2012). [Trustworthy records and open data](#). *The Journal of Community Informatics*, 8(2). [Consulta: Maig 2016].

Thurston, Anne; Lemieux, Victoria. Louise; Gormly, B.; Rowledge, L. (2014). Meeting Big Data challenges with visual analytics: The role of records Management. *Records Management Journal*, 24(2), 122-141.

Tinholt, Dinand.; Carrara, Wendy (2017). [Re-using Open Data. A study on companies transforming Open Data into economic & societal value](#). European Portal. [Consulta: Juny de 2017].

Ubaldi, Barbara. (2013). [Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives](#). *OECD Working Papers on Public Governance*. OECD Publishing, 22. [Consulta: febrer, 2019]

Unión Europea (2019). Directiva (UE) 2019/2014 del Parlamento Europeo y del Consejo de 20 de junio de 2019 relativa a los datos abiertos y la reutilización de la información del sector público.

Veljković, Nataša.; Bogdanović-Dinić, Sanja; Stoimenov, Leonid (2014). Benchmarking open government: An open data perspective. *Government Information Quarterly*, 31(2), 278-290.

Vetro, Antonio; Canova, Lorenzo; Torchiano, Marco; Orozco Minotas, Camilo; Iemma, Raimondo; Morano, Federico. (2016). Open Data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, 33(2), 325-337.

Wang, Richard Y.; Ziad, Mostapha; Lee, Yang W. (2002). *Data Quality*. New York et al.: Kluwer Academic Publishers.

Whitmore, Andrew. (2014). Using open government data to predict war: A case study of data and systems challenges. *Government Information Quarterly*, 31(4), 622-630.

Yannoukakou, Aikaterini; Araka, Iliana (2014). Access to Government Information: Right to Information and Open Government Data Synergy. *Procedia – Social and Behavioral Sciences*, 147, 332-340.

Zhao, Jun; Hartig, Olaf (2012). Towards interoperability provenance publication on the linked data web. LDOW2012, Lyon.

Zuiderwijk, Anneke; Janssen, Marijn (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17-29.

Zuiderwijk, Anneke, Janssen, Marijn, Davis, Chris. (2014a). Innovation with open data: Essential elements of open data ecosystems. *Information Polity*, 19(1, 2), 17-33.

Zuiderwijk, Anneke, Janssen, Marijn (2015). Participation and data Quality in open data use: Open Data infrastructures evaluated. Dins Adams, Carl (ed.) *Proceedings of the European Conference on e-Government*, ECEG, 351-358. Portsmouth: University of Portsmouth

8 Annexos

8.1 Annex 1: Municipis de Catalunya i capitals de comarca

| Comarca | Hab. (2016) | Sup. (km ²) | Dens. (hab./km ²) | Capital | Nombre municipis |
|------------------|----------------|----------------------------|----------------------------------|-------------------------|---------------------|
| Alt Camp | 44.107 | 538 | 82,00 | Valls | 23 |
| Alt Empordà | 140.118 | 1.357,50 | 103,20 | Figueres | 68 |
| Alt Penedès | 106.275 | 592,7 | 179,30 | Vilafranca del Penedès | 27 |
| Alt Urgell | 20.548 | 1.447,50 | 14,20 | la Seu d'Urgell | 19 |
| Alta Ribagorça | 3.867 | 426,9 | 9,10 | el Pont de Suert | 3 |
| Anoia | 117.504 | 866,3 | 135,60 | Igualada | 33 |
| Aran | 9.930 | 633,6 | 15,70 | Vielha e Mijaran | 9 |
| Bages | 174.381 | 1.092,30 | 159,70 | Manresa | 30 |
| Baix Camp | 188.358 | 697,1 | 270,20 | Reus | 28 |
| Baix Ebre | 78.977 | 1.002,70 | 78,80 | Tortosa | 14 |
| Baix Empordà | 132.598 | 701,7 | 189,00 | la Bisbal d'Empordà | 36 |
| Baix Llobregat | 809.883 | 486 | 1666,50 | Sant Feliu de Llobregat | 30 |
| Baix Penedès | 99.813 | 296,4 | 336,70 | el Vendrell | 14 |
| Barcelonès | 2.232.833 | 145,8 | 15319,60 | Barcelona | 5 |
| Berguedà | 39.178 | 1.185,30 | 33,00 | Berga | 31 |
| Cerdanya | 17.704 | 546,7 | 32,40 | Puigcerdà | 17 |
| Conca de Barberà | 20.220 | 650,2 | 31,10 | Montblanc | 22 |
| Garraf | 146.876 | 185,1 | 793,50 | Vilanova i la Geltrú | 6 |
| Garrigues | 19.090 | 797,7 | 23,90 | les Borges Blanques | 24 |
| Garrotxa | 55.999 | 734,6 | 76,20 | Olot | 21 |
| Gironès | 186.178 | 575,6 | 323,40 | Girona | 28 |
| Maresme | 441.505 | 398,5 | 1107,80 | Mataró | 30 |
| Moianès | 13.209 | 337,9 | 39,10 | Moià | 10 |
| Montsià | 67.646 | 735,4 | 92,00 | Amposta | 12 |
| Noguera | 38.955 | 1.784,10 | 21,80 | Balaguer | 30 |
| Osona | 155.641 | 1.245,20 | 125,00 | Vic | 51 |
| Pallars Jussà | 13.453 | 1.343,10 | 10,00 | Tremp | 14 |
| Pallars Sobirà | 6.986 | 1.377,90 | 5,10 | Sort | 15 |
| Pla d'Urgell | 36.711 | 305,1 | 120,30 | Mollerussa | 11 |
| Pla de l'Estany | 31.561 | 262,8 | 120,10 | Banyoles | 16 |
| Priorat | 9.381 | 498,6 | 18,80 | Falset | 23 |
| Ribera d'Ebre | 22.471 | 827,3 | 27,20 | Móra d'Ebre | 14 |
| Ripollès | 25.101 | 956,6 | 26,20 | Ripoll | 19 |

| Comarca | Hab. (2016) | Sup. (km ²) | Dens. (hab./km ²) | Capital | Nombre municipis |
|-------------------|------------------|----------------------------|----------------------------------|-------------------------|---------------------|
| Segarra | 22.688 | 722,7 | 31,40 | Cervera | 21 |
| Segrià | 208.881 | 1.396,70 | 149,60 | Lleida | 38 |
| Selva | 167.694 | 995 | 168,50 | Santa Coloma de Farners | 26 |
| Solsonès | 13.313 | 1.001,20 | 13,30 | Solsona | 15 |
| Tarragonès | 249.565 | 319,4 | 781,40 | Tarragona | 22 |
| Terra Alta | 11.761 | 743 | 15,80 | Gandesa | 12 |
| Urgell | 36.059 | 579,7 | 62,20 | Tàrraga | 20 |
| Vallès Occidental | 904240 | 583,1 | 1550,7 | Sabadell i Terrassa | 23 |
| Vallès Oriental | 401.338 | 735 | 546,00 | Granollers | 39 |
| Catalunya | 7.522.596 | 32108 | 234,30 | | 949 |

Taula 1: Nombre de municipis distribuïts per comarques.
(Font: Idescat <https://www.idescat.cat/pub/?id=aec&n=249>)

8.2 Annex 2: Llistat de portals de transparència de les universitats catalanes i dels instituts CERCA

8.2.1 Universitats

| Universitat | Portal Transparència |
|----------------------------|---|
| UOC | https://www.uoc.edu/portal/ca/transparencia/index.html |
| UB | https://www.ub.edu/web/ub/ca/universitat/transparencia/index.html |
| UPC | https://www.upc.edu/transparencia/ca |
| UPF | https://www.upf.edu/web/transparencia |
| UAB | https://www.uab.cat/transparencia/ |
| UdG | https://www.udg.edu/ca/transparencia |
| UdL | https://seuelectronica.udl.cat/transparencia.php? |
| URV | http://transparencia.urv.cat/ca/recerca-transferencia/ |
| UVic | https://transparencia.uvic.cat/ |
| URL | https://www.url.edu/la-url/transparencia-a-la-URL |
| UIC | https://www.uic.es/ca/portal-de-transparencia-uic |
| Universitat Abat Oliva CEU | https://www.uaoceu.es/ca/portal-transparencia |

8.2.2 Centres de Recerca

| Centre de Recerca | Existeix portal (o espai) de transparència en el seu lloc web? | Adreça |
|--|--|---|
| Agrotecnio – Centre de Recerca en Agrotecnologia | No | |
| CED – Centre d'Estudis Demogràfics | Sí | https://ced.uab.cat/sobre-el-ced/portal-de-transparencia/ |
| CIMNE – Centre Internacional de Mètodes Numèrics en Enginyeria | Sí | http://www.cimne.com/vpage/2/0/About/Transparency-Portal |
| CMR[B] – Centre de Medicina Regenerativa de Barcelona | Sí | https://transparencia.cmrb.eu/ |

| | | |
|---|----|---|
| CRAG – Centre de Recerca en Agrigenòmica | Sí | https://www.cragenomica.es/ca/transparencia/portal-de-la-transparencia |
| CREAF – Centre de Recerca Ecològica i Aplicacions Forestals | Sí | http://www.creaf.cat/transparencia |
| CREI – Centre de Recerca en Economia Internacional | Sí | http://www.crei.cat/transparency-portal/ |
| CRG – Centre de Regulació Genòmica | Sí | https://www.crg.eu/en/content/transparency-portal |
| CRM – Centre de Recerca Matemàtica | Sí | http://www.crm.cat/en/About/General/Documentation/Pages/default.aspx |
| CTFC – Centre de Ciència i Tecnologia Forestal de Catalunya | Sí | http://www.ctfc.cat/transparencia.php |
| CTTC – Centre Tecnològic de Telecomunicacions de Catalunya | Sí | http://www.cttc.es/transparency/ |
| CVC – Centre de Visió per Computador | Sí | http://www.cvc.uab.es/?page_id=3164 |
| i2CAT – Internet i Innovació Digital a Catalunya | Sí | https://www.i2cat.net/about-us/transparency-portal/ |
| IBEC – Institut de Bioenginyeria de Catalunya | Sí | http://www.ibecbarcelona.eu/about-us/transparency/ |
| ICAC – Institut Català d'Arqueologia Clàssica | Sí | http://www.icac.cat/qui-som/portal-de-la-transparencia/ |
| ICFO – Institut de Ciències Fotòniques | Sí | http://www.icfo.eu/lang/about-icfo/general-information/#GenInfo-5 |
| ICIQ – Institut Català d'Investigació Química | Sí | http://www.iciq.org/about-us/portal-de-la-transparencia/ |
| ICN2 – Institut Català de Nanociència i Nanotecnologia | Sí | https://icn2.cat/en/about-icn2/transparency |
| ICP – Institut Català de Paleontologia Miquel Crusafont | Sí | http://www.icp.cat/index.php/ca/icp-6/transparencia |
| ICRA – Institut Català de Recerca de l'Aigua | Sí | http://www.icra.cat/mes-informacio |
| ICRPC – Institut Català de Recerca en Patrimoni Cultural | No | |
| IDIBAPS – Institut d'Investigacions Biomèdiques August Pi i Sunyer | Sí | http://transparencia.idibaps.org/ |
| IDIBELL – Institut d'Investigació Biomèdica de Bellvitge | Sí | http://www.idibell.cat/ca/institut/portal-de-transparencia/transparencia |
| IDIBGI – Institut d'Investigació Biomèdica de Girona Dr. Josep Trueta | Sí | http://www.idibgi.org/ca/content/portal-de-transpar%C3%A8ncia |
| IEEC – Institut d'Estudis | Sí | http://www.ieec.cat/en/content/8/transparency |

| | | |
|--|----|---|
| Espacials de Catalunya | | |
| IFAE – Institut de Física d'Altes Energies | Sí | http://www.ifaes.es/eng/about-ifaes/transparency-accountability.html |
| IGTP – Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol | Sí | http://www.germanstrias.org/transparency/ca-index/ |
| IISPV – Institut d'Investigació Sanitària Pere Virgili | Sí | http://transparencia.iispv.cat/ca/ |
| IJC – Institut de Recerca Contra la Leucèmia Josep Carreras | Sí | http://www.carrerasresearch.org/ca/portal-de-la-transpar%C3%A8ncia-ijc_42322 |
| IMIM – Institut Hospital del Mar d'Investigacions Mèdiques | Sí | https://www.imim.cat/infocorporativa/portalttransparencia/ |
| IPHES – Institut Català de Paleoecologia Humana i Evolució Social | Sí | http://www.iphes.cat/transpar%C3%A8ncia |
| IR-Sant Pau – Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau | No | |
| IRB Barcelona – Institut de Recerca Biomèdica | Sí | https://www.irbbarcelona.org/ca/portal-de-la-transparencia |
| IRB Lleida – Institut de Recerca Biomèdica de Lleida | Sí | http://www.irbllleida.org/en/about-us/transparency/ |
| IREC – Institut de Recerca en Energia de Catalunya | Sí | http://www.irec.cat/transparencia/la-fundacio-irec.html |
| IrsiCaixa – Institut de Recerca de la Sida | Sí | http://www.irsicaixa.es/ca/transparencia |
| IRTA – Institut de Recerca i Tecnologia Agroalimentàries | Sí | http://www.irta.cat/ca/irta/transparencia/ |
| ISGLOBAL – Institut de Salut Global de Barcelona | Sí | https://www.isglobal.org/ca/transparencia |
| VHIO – Vall d'Hebron Institut d'Oncologia | Sí | http://www.vhio.net/en/transparency-portal/ |
| VHIR – Vall d'Hebron Institut de Recerca | Sí | http://www.vhir.org/portal1/subboxes_transparencia.asp?s=institut&contentid=91878&t=org-institucional-i-estructura-administrativa |

8.3 Annex 3: Guió de les entrevistes

- Agraïment per acceptar
- Breu introducció del projecte,

Com us vàrem dir per correu, estem conduint un projecte de recerca finançat per l'Escola d'Administració Pública de Catalunya, guanyat en convocatòria competitiva.

L'objectiu del projecte és estudiar les pràctiques de les administracions públiques per gestionar les dades obertes orientades a la reutilització, d'acord a la legislació vigent i els factors que les influeixen.

8.3.1 Ajuntaments, Generalitat de Catalunya i Diputacions

I. BLOC CONTEXT GENERAL

1. Descripció breu de les funcions i quins vincles te amb altres àrees o departaments
2. Existència d'ordenances, legislació, directrius, protocols per accedir i gestionar dades en obert
3. estructura organitzativa i mitjans per gestionar dades obertes? // [donar suport a ajuntaments]
4. Lloc per **publicar** les dades en obert
5. Aplicació d'arxiu electrònic

II. BLOC PROCESSOS I PROCEDIMENTS DOCUMENTS, INFORMACIÓ I DADES

1. Quins processos i procediments interns hi ha establerts per gestionar els documents, la informació i les dades, tenint en compte la necessitat de generar i publicar dades en obert / per publicar en el portal
2. Com identifiqueu les dades que es poden posar a disposició de la ciutadania per a la seva reutilització?
3. Quina informació (documents o dades) es publica en obert?
4. Quina informació / dades / documents estan ja incorporades en els fluxos i procediments ja establerts?
5. Quines peticions teniu d'informació i de dades en obert per part de la ciutadania?
6. Com es garanteix el flux de documents i informació d'altres àrees de l'ajuntament cap al servei de les dades obertes?

III. **BLOC NORMATIVA (ACCÉS, AVALUACIÓ I DISPOSICIÓ) I PRESERVACIÓ**

1. Quins criteris / normativa apliqueu pel establir el règim d'accés, avaluació i eliminació de dades/informació i documents
2. Referent a política de preservació/conservació d'aquests actius digitals. Com es preserva la informació, els documents i les dades a la vostra institució?
3. Com es preserven les dades incorporades en l'aplicació d'arxiu que teniu al vostre organisme?

IV. **BLOC PORTAL DE DADES OBERTES I REUTILITZACIÓ**

1. Com avalueu les accions que feu, la publicació en el portal i el vincle amb la reutilització?
2. Disposeu d'indicadors per avaluar la publicació en obert en el portal [si en tenen]i el compliment amb la llei?
3. Vincles teniu amb la indústria de reutilització?
4. Parlem del portal, com es pot accedir a les dades? hi ha cercador? estan classificades les dades?
5. En quin nivell d'implantació esteu pel que fa les normes tècniques d'interoperabilitat?

V. **BLOC Intel·ligència competitiva i Intel·ligència Territorial**

1. Quins us intern feu de les dades obertes que genereu?
2. Heu mirat algunes experiències en d'altres administracions públiques,
3. Les heu pres com a referent?
4. Quina col·laboració hi ha hagut amb altres administracions (Diputació / Generalitat / Ajuntaments]
5. Quins recursos posen a la seva disposició (tecnològics, formularis, plantilles...)
6. Heu fet formació?
7. Formen part d'alguna xarxa de l'administració?

VI. **FINALITZACIÓ**

Vol afegir algun aspecte més que no hagi sortit durant l'entrevista?

8.3.2 Universitats i Centres de recerca

I. BLOC CONTEXT GENERAL

1. Descriu de manera breu quines són les seves funcions dins de la Universitat, quins vincles te amb altres àrees o departaments,
2. Disposa la Universitat d'una política de dades obertes (o de recerca), o bé de transparència
3. Quina estructura organitzativa i mitjans hi ha a la universitat/centre de recerca vinculats amb les dades obertes?
4. Quina solució de repositori digital teniu a la universitat?

II. BLOC PROCESSOS I PROCEDIMENTS DOCUMENTS, INFORMACIÓ I DADES

1. Quins processos i procediments interns hi ha establerts dins de la Universitat per gestionar els documents, la informació i les dades, tenint en compte la necessitat de generar i publicar dades en obert?
2. Com identifiqueu les dades que es poden posar a disposició per a la seva reutilització?
3. Quina informació (documents o dades) es publica en obert?
4. Quina informació / dades / documents estan ja incorporades en els fluxos i procediments ja establerts?
5. Quines peticions teniu d'informació i de dades en obert per part de la comunitat acadèmica o aliena a ella?
6. Com es garanteix el flux de documents i informació d'altres facultats i departaments de la universitat cap al servei de les dades obertes?

III. BLOC NORMATIVA (ACCÉS, AVALUACIÓ I DISPOSICIÓ) I PRESERVACIÓ

1. Quins criteris / normativa apliqueu pel establir el règim d'accés, avaluació i eliminació de dades/informació i documents
2. PARLEM de POLÍTICA DE PRESERVACIÓ/CONSERVACIÓ d'aquests actius digitals. Com es preserva la informació, els documents i les dades a la vostra institució?
3. Com es preserven les dades incorporades en l'aplicació d'arxiu que teniu al vostre organisme?

IV. BLOC PORTAL DE DADES OBERTES I REUTILITZACIÓ

1. Com avalueu les accions que feu, la publicació en el portal/repositori i el vincle amb la reutilització?
2. Disposeu d'indicadors per avaluar la publicació en obert en el portal /repositori[si en tenen] i el compliment amb la llei?
3. Quins vincles teniu amb la indústria de reutilització?
4. Parlem del portal, com es pot accedir a les dades? hi ha cercador? estan classificades les dades?
5. En quin nivell d'implantació esteu pel que fa les normes tècniques d'interoperabilitat? Com garantiu que les dades puguin ser FAIR?

V. BLOC IC - IT

1. Quins us intern feu de les dades obertes que genereu dins de la Universitat?
2. Quins són els vostres referents i aliats en l'àmbit dels repositoris?
3. Quina col·laboració hi ha hagut per part de la Generalitat/CSUC?
4. Formen part d'alguna xarxa de municipis/universitats?

VI. FINALITZACIÓ

1. Vol afegir algun aspecte més que no hagi sortit durant l'entrevista?

8.4 Annex 4: Carta de consentiment



L'entrevista que us farem forma part d'un projecte de recerca que té com a títol "*Gestió de les dades obertes: models i protocols per al govern obert i la reutilització*", dut a terme pel grup de recerca KIMO de la Universitat Oberta de Catalunya.

L'objectiu del projecte és estudiar les pràctiques de les administracions públiques per gestionar les dades obertes per a ser reutilitzades, d'acord a la legislació vigent i els factors que les influeixen.

La vostra participació en aquest estudi és valuosa, però podeu declinar respondre qualsevol pregunta o demanar aclariments sobre qualsevol qüestió.

El contingut de l'entrevista és confidencial. La informació que ens proporcioneu serà enregistrada per la seva posterior anàlisi. Aquesta anàlisi també serà confidencial i la informació obtinguda en cap cas s'emprarà per altres fins que la recerca.

La informació que ens proporcioneu no s'associarà en cap cas a les vostres dades d'identificació. En aquest sentit es realitzarà una codificació de cada institució per a tipificar els resultats.

En qualsevol cas, les dades personals i institucionals seran registrades per separat de la informació que ens proporcioneu i s'empraran exclusivament amb finalitat de control i seguiment del procés de realització de les entrevistes.

Ateses aquestes premisses, autoritzeu la enregistrament de l'entrevista.

Nom del/ de la participant i signatura

Nom de l'investigador/-a i signatura

Localitat:, adede 2019

8.5 Annex 5: Enquestes

Gestió de dades obertes als ajuntaments

Benvolguts,

El grup de recerca KIMO de la Universitat Oberta de Catalunya (UOC) està conduint un projecte de recerca que du per títol "Gestió de les dades obertes: models i protocols per al govern obert i la reutilització". Aquest projecte està finançat per l'Escola d'Administració Pública de Catalunya.

L'objectiu del projecte és estudiar què estan fent les administracions públiques per tal de gestionar les dades obertes per a la seva reutilització, d'acord a la legislació vigent.

Necessitem la vostra col·laboració per recollir dades en la fase de treball de camp. Per això us demanem respondre a aquestes preguntes. Les respostes seran tractades de manera confidencial, i només serem emprades a efectes d'aquest projecte de recerca.

Per avançar, agraïm la vostra col·laboració.

* Obligatòria

1. Adreça electrònica *

Context general

2. Quin és el vostre càrrec i quines són les vostres funcions? *

3. ¿Teniu a l'ajuntament algun sistema de gestió o algun estàndard implantat? *

Maqueu només un oval.

- Sí
- No
- Estem en procés d'implantació
- Altres: _____

Us intern de les dades obertes

4. Utilitzeu les dades obertes que es generen dins de l'ajuntament, per a la presa de decisions? *

Maqueu només un oval.

- Sí *Passeu a la pregunta 4.*
- No, però ens ho estem plantejant *Passeu a la pregunta 36.*
- No *Passeu a la pregunta 36.*

Us intern de les dades obertes

5. En quins àmbits empreu les dades obertes? *

Sistemes de gestió i estàndards

6. Quines normes o sistemes de gestió teniu implantats? *

Seleccioneu totes les opcions que corresponguin.

- ISO 9001 Sistema de Gestión de la Calidad
- ISO 30301 Sistema de Gestión para los documentos
- UNE 178301 Ciudades inteligentes. Datos abiertos
- ISO 27001 Sistema de Gestión de Seguridad de la Información
- Altres: _____

7. Des de quan teniu aquest o aquests sistemes de gestió i estàndards implantats? *

Passeu a la pregunta 7.

Ordenança de dades obertes

8. L'ajuntament té una ordenança per a les dades obertes? *

Maqueu només un oval.

- Sí *Passeu a la pregunta 8.*
- No, però l'estem fent *Passeu a la pregunta 8.*
- No, però està planificat fer-la *Passeu a la pregunta 8.*
- No, però ens agradaria *Passeu a la pregunta 8.*
- No, i no ens ho hem plantejat. *Passeu a la pregunta 8.*
- No, a l'ajuntament no gestionem ni publiquem dades obertes *Passeu a la pregunta 9.*
- Altres: _____ *Passeu a la pregunta 8.*

Publicació de dades obertes

9. On publiqueu les dades obertes? *

Maqueu només un oval.

- Portal de dades obertes facilitat per l'AOC *Passeu a la pregunta 11.*
- Portal de dades obertes dissenyat per l'Ajuntament *Passeu a la pregunta 11.*
- Portal de Transparència *Passeu a la pregunta 11.*
- No publiquem dades obertes *Passeu a la pregunta 9.*
- Altres: _____ *Passeu a la pregunta 11.*

Passeu a la pregunta 11.

Raons per no gestionar ni publicar dades obertes

10. Explica perquè no gestioneu dades obertes a l'Ajuntament *

11. Explica perquè no publiqueu dades obertes *

Passeu a la pregunta 24.

Directrius per gestionar dades obertes

12. L'ajuntament te directrius sobre com gestionar les dades obertes? *

Maqueu només un oval.

- Sí, tenim processos definits *Passeu a la pregunta 12.*
- Sí, tenim processos definits, recollits en el Sistema de Qualitat de l'Ajuntament.
Passeu a la pregunta 12.
- No, però estem definint-los *Passeu a la pregunta 12.*
- No, però tenim planificat establir processos *Passeu a la pregunta 12.*
- No, i no ho tenim previst *Passeu a la pregunta 12.*
- Altres: _____ *Passeu a la pregunta 12.*

Estructura organitzativa per a la gestió de dades obertes

13. Per a la gestió de dades obertes disposeu d'una estructura organitzativa específica *

Maqueu només un oval.

- Sí tenim un departament específic per a la gestió de dades obertes
- No, està dins del departament / de l'àrea que s'encarrega de la transparència
- No, està dins d'un departament amb altres funcions
- Altres: _____

14. Com es diu el departament o l'àrea des de la que gestioneu les dades obertes? *

15. Quantes persones treballen en la gestió de dades obertes? *

16. Obteniu benefici econòmic de la gestió amb dades obertes? *

Maqueu només un oval.

- Sí
- No
- Estem mirant com obtenir-ne
- Altres: _____

Procediments interns

Dins de l'ajuntament hi ha diferents àrees que publiquen dades en obert.

17. Hi ha establert un procediment per vehicular les dades obertes generades en els diferents departaments per a ser tractades? *

Maqueu només un oval.

- Sí
- No
- Altres: _____

18. Hi ha un procediment per publicar les dades en el portal? *

Maqueu només un oval.

- Sí
- No
- Altres: _____

19. Hi ha un procediment per atendre les sol·licituds de dades en obert? *

Maqueu només un oval.

- Sí, el mateix que les sol·licitud on aplica la llei de transparència
- Sí, un específic per a les dades obertes, diferent al canal de la llei de transparència.
- No, entren per la bustia genèrica
- No ho hem planificat
- Altres: _____

20. Teniu definit un procediment per identificar les dades que s'han de posar al servei de la ciutadania? *

Maqueu només un oval.

- Sí
- No
- Altres: _____

21. Quines dades obertes (datasets) es publiquen? *

Maqueu només un oval.

- Les que senyalen la Llei de Transparència.
- Les que ens sol·liciten els ciutadans, passen a formar part de les que sistemàticament es publiquen
- Tenim en compte la indústria de la reutilització de la informació
- Altres: _____

Passeu a la pregunta 21.

Portal de dades obertes

22. Com es cerca en el vostre portal de dades obertes? *

Maqueu només un oval.

- Mitjançant un cercador
- Navegant per les categories dels datasets
- Navegant pel quadre de classificació
- Mitjançant el cercador i per categories.
- Mitjançant el cercador, per categories i pel quadre de classificació
- Altres: _____

23. Com avalueu la publicació de dades obertes en el portal? *

Maqueu només un oval.

- Hem desenvolupat un quadre d'indicadors
- Tenim en compte quadres d'indicadors publicats
- No avaluem la publicació de dades obertes al portal
- Altres: _____

24. Diga's quin quadre d'indicadors utilitzeu? En què consisteix? *

Passeu a la pregunta 24.

Normativa per gestionar documents (o dades i informació)

Per tractar les dades des de la seva generació fins que s'eliminen o es conserven, quins criteris seguïu

25. Teniu en compte els criteris del departament d'arxius per definir el regim d'accés, avaluació i eliminació de dades informació i documents? *

Maqueu només un oval.

- Sí
- No
- Altres: _____

26. La normativa d'accés i avaluació de la informació és coneguda i s'aplica al vostre organisme? *

Maqueu només un oval.

- Sí, la coneixem i l'apliquem
- Sí, la coneixem però no tenim procediments per aplicar-la
- Sí, la coneixem, i tenim previst aplicar-la més endavant.
- No la coneixem, i per tant, no l'apliquem
- Altres: _____

27. Com es preserva la informació, els documents i les dades a la vostra institució? *

28. S'identifiquen les dades essencials susceptibles de ser reutilitzades? *

Maqueu només un oval.

- Sí
- No, però ho tenim previst
- No
- Altres: _____

29. Existeix i/o és coneguda a la vostra institució una política de preservació d'aquests actius digitals? *

Maqueu només un oval.

Sí *Passeu a la pregunta 29.*

No *Passeu a la pregunta 30.*

Altres: _____ *Passeu a la pregunta 29.*

Política de preservació

30. Quina política de preservació d'aquests actius digitals existeix i/o és coneguda a la vostra institució? *

Emmagatzematge

31. Quina solució d'arxiu electrònic teniu en el vostre Ajuntament? *

32. Com es preserven aquestes documents electrònics? *

Normes Tècniques d'Interoperabilitat

33. Teniu implantades les normes tècniques d'interoperabilitat? *

Maqueu només un oval.

Sí

No encara, però hi estem treballant.

No

Altres: _____

Vincles amb la indústria de reutilització

34. En el vostre municipi, teniu identificades empreses del sector de reutilització de la informació? *

Maqueu només un oval.

- Sí
- No
- Altres: _____

Vincles amb la indústria de reutilització

35. Teniu relació amb les empreses del sector? *

Maqueu només un oval.

- Sí
- No, per tenir previst establir algun tipus de relació.
- No
- Altres: _____

36. Teniu tarifes establertes per subministrar informació pública amb fins comercials?

Maqueu només un oval.

- Sí
- No
- Altres: _____

Referents

37. Heu mirat experiències de gestió de dades obertes d'altres administracions? *

Maqueu només un oval.

- Sí *Passeu a la pregunta 37.*
- No *Passeu a la pregunta 38.*
- Altres: _____

Referents

38. Quines experiències heu mirat? *

Suport institucional

39. Teniu suport de la Diputació per implantar la gestió de dades obertes? *

Maqueu només un oval.

- Sí *Passeu a la pregunta 39.*
- No *Passeu a la pregunta 40.*
- Altres: _____

Suport Diputació

40. Quin tipus de suport? *

Seleccioneu totes les opcions que corresponguin.

- Finançament
- Formació
- Solucions tecnològiques per a la gestió de dades obertes
- Altres: _____

Suport institucional

41. Teniu suport de la AOC? *

Maqueu només un oval.

- Sí *Passeu a la pregunta 41.*
- No *Passeu a la pregunta 43.*
- Altres: _____

Suport AOC

42. Quin tipus de suport? *

Seleccioneu totes les opcions que corresponguin.

- Finançament
- Formació
- Solucions tecnològiques
- Altres: _____

Suport AOC

43. Quin altra tipus de suport teniu de l'AOC? *

Relacions

44. Participeu d'algun grup de treball vinculat amb gestió de dades obertes? *

Seleccioneu totes les opcions que corresponguin.

- Sí, el de la Federació de Municipis de Catalunya
- Sí, el de l'Associació Catalana de Municipis i Comarques
- No
- Altres: _____

Relacions

45. Amb quina xarxa de municipis teniu vincles, relacionats amb la gestió de dades obertes? *

Seleccioneu totes les opcions que corresponguin.

- Associació Catalana de Municipis i Comarques
- Federació de Municipis de Catalunya
- Cap
- Altres: _____

46. Amb això finalitzem l'enquesta. Gràcies per la vostra col·laboració!

Deixeu d'emplenar el formulari.

Finalitza el qüestionari

Gràcies per la vostra participació

- Envia'm una còpia de les meves respostes.