



PROTOCOL DE GESTIÓ DE DADES

Índex

Índex	2
Introducció i fonament	3
Visió i propòsit	4
Objectius	4
Àmbit d'aplicació i principis rectors	4
Rols i responsabilitats	5
El cicle de vida de les dades: una visió operativa	6
Pla de Gestió de Dades (PGD / DMP)	7
La Unitat de Banc de Dades del CED	10
Accés, llicències i reconeixement	13
Suport, formació i recursos	14
Alineament institucional, referències normatives i obligacions operatives derivades	14
ANNEXOS	17

Introducció i fonament

Les ciències socials, i molt especialment la Demografia, fonamenten la seva aportació científica en l'anàlisi sistemàtica d'informació sobre poblacions, famílies, trajectòries vitals i dinàmiques socials. Al Centre d'Estudis Demogràfics (CED) es treballa amb conjunts de dades procedents de fonts estadístiques, registres administratius, enquestes i estudis qualitius que, un cop tractats i documentats amb diligència, permeten explicar fenòmens col·lectius i interpretar processos individuals. Per la seva naturalesa i pel seu potencial d'impacte, aquestes dades constitueixen un actiu científic i social que cal preservar, contextualitzar i facilitar per al seu ús responsable i per a la reproducció del coneixement.

La gestió de dades és, per tant, una dimensió essencial del rigor metodològic. Una pràctica de gestió ben planificada assegura la qualitat, la integritat, la traçabilitat i la reproductibilitat dels resultats, i incrementa la utilitat a llarg termini dels actius científics generats pel CED. A més, una bona gestió afavoreix la col·laboració interdisciplinària i interinstitucional, facilita la transferència de coneixement i incrementa l'impacte social de la recerca. En l'actual context d'impuls a la Ciència Oberta i l'Open Data, disposar d'un marc estructurat de gestió de dades multiplica el valor científic i social dels conjunts d'informació generats pel centre, sempre que aquesta obertura es compatibilitzi amb les exigències legals, ètiques i de contractació.

Aquesta política defineix el marc institucional que garanteix una gestió responsable, ètica, segura i alineada amb els principis FAIR (Findable, Accessible, Interoperable, Reusable). Promou l'accés obert quan sigui compatible amb les obligacions derivades de la normativa i del consentiment dels participants, i estableix mesures concretes per protegir la privacitat i els drets de les persones i dels proveïdors de dades. El text incorpora tant criteris generals - principis, responsabilitats i criteris d'accés- com orientacions operatives per a la creació, el tractament, la documentació, la preservació i la difusió de dades.

Així mateix, la política reconeix i articula dues exigències complementàries: d'una banda, la vocació d'apertura i transparència pròpia de la Ciència Oberta, i, de l'altra, la necessitat d'aplicar mesures de seguretat, anonimització i control d'accés quan les dades siguin sensibles o estiguin subjectes a restriccions legals o contractuals. En aquest sentit, el marc normatiu contempla el compliment del Reglament General de Protecció de Dades (RGPD), la normativa sobre secret estadístic i altres disposicions estatals i catalanes aplicables, i s'alinea amb les directrius de la Institució CERCA i amb els requisits dels finançadors nacionals i internacionals.

Finalment, la política expressa el compromís institucional del CED: la gestió de dades és una responsabilitat científica i pública que protegeix els drets dels participants i garanteix la qualitat i la sostenibilitat del coneixement generat. Amb aquest objectiu, el document estableix principis, compromisos i procediments operatius perquè les dades gestionades pel centre siguin útils avui i preservades per al demà, en un marc de transparència, responsabilitat i servei a la societat.

Visió i propòsit

Visió. El CED aspira a gestionar les dades com a actius científics de qualitat: accessibles quan l'apertura és ètica i legalment admissible, documentats perquè siguin útils i preservats per alimentar recerques futures.

Propòsit. La política té per objecte traduir aquesta visió en principis, compromisos i pràctiques operatives. Defineix els requisits mínims (per exemple, la disposició d'un Pla de Gestió de Dades dels projectes), les mesures de seguretat necessàries, les normes de documentació i qualitat, i els mecanismes de suport, formació i recursos perquè tot el personal implicat apliqui bones pràctiques de manera coherent i sostinguda.

Objectius

La política persegueix, principalment:

- Assegurar una gestió planificada, transparent i reproducible de les dades al llarg del seu cicle de vida.
- Promoure l'accés responsable i la reutilització de dades de recerca d'acord amb els principis FAIR i les pràctiques de Ciència Oberta.
- Garantir el compliment de la normativa aplicable (protecció de dades, secret estadístic, clàusules de proveïdors).
- Establir rols, responsabilitats i procediments clars per a la preservació, la difusió i la protecció de dades.
- Facilitar recursos, formació i suport per a l'elaboració i aplicació efectiva dels Plans de Gestió de Dades (PGD/DMP).

Àmbit d'aplicació i principis rectors

Àmbit. Aquesta política s'aplica a totes les dades primàries i secundàries, metadades, codis, documents i materials derivats que es generin o gestionin sota la responsabilitat del CED, i a totes les persones i unitats que hi participin -investigadors, tècnics i personal administratiu- així com a contractes i col·laboracions exteriors.

Principis rectors. L'enfocament del CED se sustenta en sis principis fonamentals:

1. **Ciència Oberta i Accés Obert.** Es prioritza la difusió oberta de dades i productes científics finançats amb fons públics, sempre que això sigui èticament i legalment viable; es garanteix, a més, el reconeixement adequat dels autors i proveïdors.
2. **Principis FAIR.** Les dades han de ser localitzables, accessibles segons permisos, interoperables mitjançant estàndards reconeguts i reutilitzables amb metadades completes que facilitin la interpretació.

3. **Protecció i ètica.** La protecció de les persones i el compliment del RGPD i del secret estadístic són prioritaris; s'aplicaran mesures d'anonimització, pseudonimització i controls d'accés quan s'hi justifiqui.
4. **Qualitat i documentació.** Cada dataset ha d'anar acompanyat de documentació essencial (README, codebook, metadades) que assegurï la comprensió i la reutilització.
5. **Transparència i reproductibilitat.** Es documentaran processos analítics, versions de programari i transformacions per facilitar la reproducció i la validació científica.
6. **Compliment normatiu i institucional.** La política es desplega en coherència amb la normativa europea, estatal i catalana i amb les directrius institucionals i dels finançadors.

Rols i responsabilitats

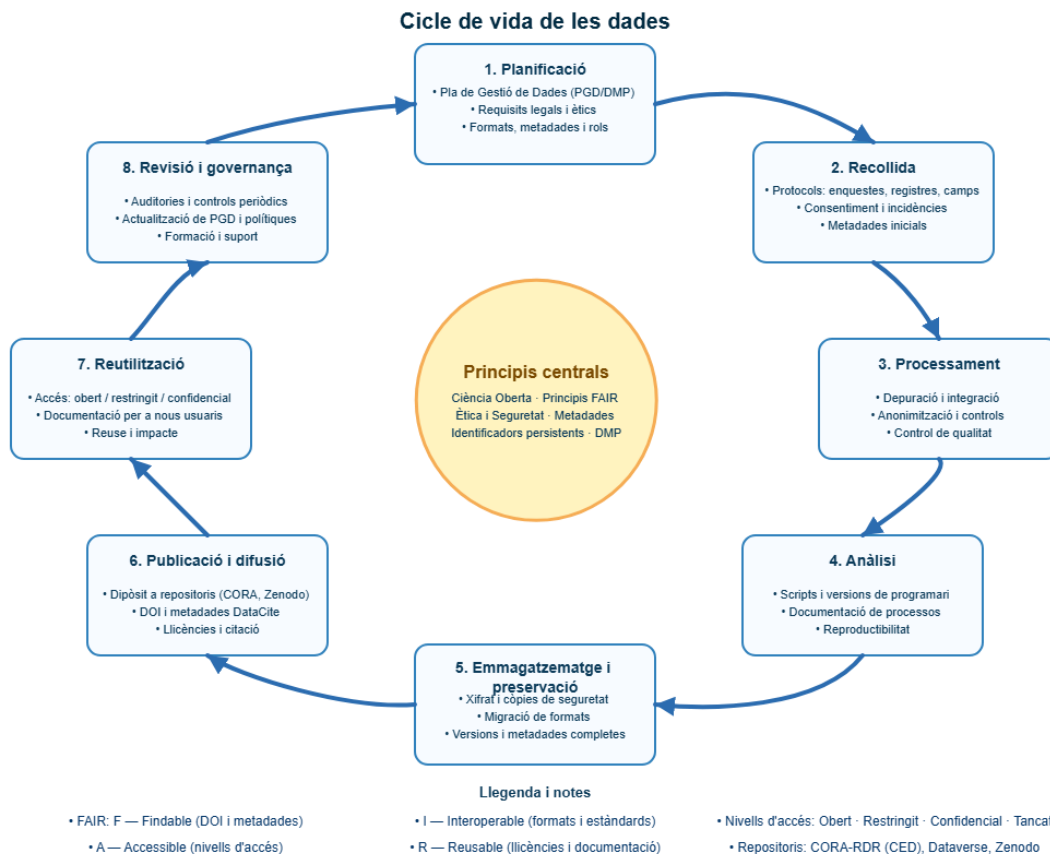
Per evitar ambigüitats, la política defineix responsabilitats concretes:

- **Direcció del CED:** dona suport estratègic i facilita els recursos necessaris perquè la política s'implanti de manera efectiva; també n'aprova les línies generals.
- **Responsable de Gestió de Dades (RGD):** coordina l'aplicació de la política, assessora els equips, manté el registre institucional de datasets i gestiona la relació amb repositoris institucionals (per exemple, l'instància CORA del CED).
- **Delegat/da de Protecció de Dades (DPO):** assessora sobre el compliment del RGPD i participa en la valoració de riscos i incidents relacionats amb dades personals.
- **Unitat de Banc de Dades:** gestiona operativament l'emmagatzematge, els accessos controlats i les garanties de confidencialitat.
- **Unitat d'Informàtica:** proporciona i manté la infraestructura tècnica (servidors, backups, controls d'accés i registres).
- **Investigador/a Principal (IP):** és responsable de garantir que el projecte disposa d'un PGD/DMP, que s'apliquen les mesures de seguretat pactades i que la recerca compleix la política institucional.
- **Personal investigador i tècnic:** executa i documenta les pràctiques establertes (recollida, processament, assegurament de la qualitat).
- **Comitè d'ètica i unitats de suport (Comissió d'Ètica de la UAB, Unitat de Gestió de Projectes):** intervenen per assessorar en decisions ètiques, consentiment informatiu i gestió de riscos o restriccions vinculades a tercers proveïdors.

Aquestes responsabilitats s'executen amb cooperació i amb l'objectiu de facilitar l'activitat investigadora i no de restringir-ne la iniciativa.

El cicle de vida de les dades: una visió operativa

La gestió de dades s'entén com un **cicle de vida** on cada etapa demana decisions documentades i responsables designats:



1. **Planificació.** Abans d'iniciar la recerca, s'elabora el Pla de Gestió de Dades (PGD/DMP), document viu que descriu tipus i formats de dades, metadades, procediments d'emmagatzematge, criteris d'accés, responsabilitats, pla de preservació i mesures de seguretat, així com el compliment de requisits legals i ètics (RGPD, secret estadístic, condicions de proveïdors).
2. **Recollida i generació.** Durant la recollida, es registren sol·licituds de dades a organismes estadístics, protocols de mostreig, instruments (qüestionaris, guions) i metadades inicials (origen, cobertura temporal i geogràfica). Cal documentar incidències, indicadors de qualitat i, quan correspongui, el consentiment informat i les condicions d'ús acordades amb participants o proveïdors.
3. **Processament i control de qualitat.** Es preserva la traçabilitat: es registren totes les transformacions, scripts i versions de programari, així com controls de qualitat (validacions, detecció d'errors). Els scripts i entorns d'execució s'han d'acompanyar de documentació perquè altres puguin reproduir els resultats.
4. **Anàlisi i reproductibilitat.** Els resultats analítics es fonamenten en processos documentats; compartir codi, entorns (contenidors) i instruccions millora la

- transparència i la confiança científica. Quan es treballi amb dades sensibles, s'apliquen tècniques d'anonimització i controls previs a qualsevol difusió.
5. **Emmagatzematge i preservació.** El CED prioritza solucions segures i escalables i defineix ubicacions institucionals segons la sensibilitat del dataset (servidors institucionals, instància CORA RDR, Zenodo, etc.), amb polítiques de còpies redundants, xifrat per dades sensibles i plans de migració de formats per preservar la interpretabilitat. Es recomana l'ús de formats oberts per a la preservació.
 6. **Publicació i difusió.** Sempre que sigui possible i compatible amb la normativa i l'ètica, els datasets es dipositen en repositoris reconeguts que assignin DOI i ofereixin gestió de nivells d'accés (CORA RDR, Dataverse, Zenodo). Cada dipòsit anirà acompanyat de README, codebook, metadades DataCite/Dublin Core i d'una llicència d'ús adequada.
 7. **Reutilització i accés.** Per a dades restringides o confidencials, s'estableix un procediment de sol·licitud formal, avaluació per un comitè i, si s'aprova, accés sota Data Use Agreement i, si cal, en entorns segurs.
 8. **Retenció i eliminació.** Les polítiques de retenció s'ajusten a la categoria de dades, a la normativa i als acords contractats; l'eliminació -quan procedeixi- s'executa amb procediments segurs i registrats.

Pla de Gestió de Dades (PGD / DMP)

Definició i objectiu

El Pla de Gestió de Dades (PGD o DMP — Data Management Plan) és un document formal i operatiu que descriu de manera clara i completa com es gestionarà, durant tot el seu cicle de vida, la informació generada o utilitzada en un projecte de recerca. El PGD explica com es recopilaran o generaran les dades, com s'organitzaran i documentaran, quines mesures de seguretat s'aplicaran, on s'emmagatzemaran, com es preservaran i en quines condicions es posaran a disposició d'altres investigadors o del públic. L'objectiu del PGD és garantir la qualitat, la traçabilitat, la seguretat, la citabilitat i la reutilitzabilitat de les dades, així com demostrar davant dels finançadors i altres agents que s'ha abordat de manera raonada la gestió de la informació produïda pel projecte.

El PGD és l'eina operacional que materialitza la política del CED: no només documenta decisions sinó que guia l'actuació de l'equip de recerca i de les unitats de suport. Un PGD ben elaborat i actualitzat és, en definitiva, la garantia que les dades de recerca del CED seran gestionades amb professionalitat, seguretat i responsabilitat, i que seran tan obertes com sigui possible i tan tancades com sigui necessari.

Caràcter de document viu

El PGD és un document viu: s'elabora a la fase de planificació i s'actualitza al llarg del projecte sempre que hi hagi modificacions rellevants o es produeixin canvis significatius en la gestió de dades (incorporació nous conjunts de dades, canvi de categoria de sensibilitat, noves restriccions contractuals, canvis de política d'accés, resultats imprevistos, canvis normatius o

transformacions metodològiques, etc.), pel que cal registrar una versió nova amb la data i l'autor de l'actualització.

Contingut dels PGD

El PGD ha d'incloure els apartats següents, si bé l'exigència dels requeriments poden variar entre ens de finançament:

- **Identificació i responsabilitats:** Títol del projecte, codi, investigador/a responsable (IP), responsables de gestió de dades (RGD) i responsables tècnics.
- **Resum executiu:** Breu descripció de les dades que es generaran o reutilitzaran i la seva relació amb els objectius del projecte.
- **Tipus, origen i volum estimat de les dades:** Tipologia (enquestes, registres, sensors, imatges, codi...), procedència i volum estimat (GB/TB), freqüència de generació.
- **Formats i estàndards:** Formats d'arxiu previstos (preferència per formats oberts i estandarditzats), codificacions i estàndards de metadades que s'aplicaran.
- **Metadades i identificadors persistents:** Esquema de metadades (p. ex. DataCite) i compromís d'assignar identificadors persistents (DOI per a datasets, ORCID per a autors).
- **Organització i documentació (README / codebook):** Estructura dels fitxers, definició de variables, diccionari de dades, logs i documentació de decisions de processament i neteja.
- **Qualitat i control:** Protocols i processos de control de qualitat (validacions, detecció d'errors, tractament de missings) i gestió de versions del dataset i del codi.
- **Seguretat, emmagatzematge i còpies de seguretat:** Localitzacions d'emmagatzematge (servidors institucionals, repositoris), polítiques de còpia i restauració, xifrat, controls d'accés i mesures específiques per a dades sensibles.
- **Aspectes ètics i legals:** Base legal del tractament (consentiment, interès públic, etc.), clausules de consentiment informat, aplicació del secret estadístic si escau i mesures d'anonimització / DPIA quan sigui necessari.
- **Accés, llicència i difusió:** Nivell d'accés proposat (obert, restringit, confidencial, tancat), embargaments, repositori elegit, llicència d'ús i condicions per a sol·licituds d'accés.
- **Preservació i retenció:** Polítiques de preservació a mig i llarg termini (repositori de preservació, migració de formats, durada de retenció).
- **Recursos i pressupost:** Recursos humans i econòmics previstos per a la gestió, emmagatzematge i preservació de les dades.
- **Riscos, mal ús i mesures de mitigació:** Anàlisi de riscos ètics i jurídiques (possible estigmatització, re-identificació), i protocols per minimitzar-los.
- **Pla d'implantació i calendaris:** Dates clau per a l'activitat de gestió de dades i responsables de cada tasca.
- **Signatures i versions:** Signatura del IP i del RGD, llista de versions i historial d'actualitzacions.

Implementació dels principis FAIR en el PGD

El PGD ha de reflectir explícitament com s'apliquen els principis **FAIR** a cada conjunt de dades:

- **Findable (Localitzable):** assignació d'identificadors persistents (DOI, Handle), metadades enriquides i indexació en recursos de cerca (repositoris institucionals o públics).
- **Accessible (Accessible):** indicació del repositori triat, dels protocols d'accés (obert, per autorització o via entorn segur) i de les condicions per sol·licitar accés. Les metadades han de romandre accessibles encara quan les dades deixin d'estar disponibles.
- **Interoperable (Interoperable):** adopció de formats i estàndards de metadades, ús de vocabularis controlats i descripcions que permetin la combinació amb altres conjunts (per exemple codis ISO, terminologies standard).
- **Reusable (Reusable):** documentació completa (origen, instruccions, scripts, entorn), llicències clares que permetin la reutilització i codis que facilitin la reproducció dels resultats.

Governança, aprovació i responsabilitats

El PGD l'elabora l'Investigador/a Principal en col·laboració amb el Responsable de Gestió de Dades i amb l'assessorament del DPO quan sigui necessari. Procediment d'aprovació: el RGD farà una revisió formal del PGD dins dels terminis establerts i, si cal, sol·licitarà esmenes. La Direcció o la unitat competent validarà la implantació de polítiques generals (recursos, infraestructures). Els canvis substancials al PGD requereixen l'actualització formal i l'enregistrament d'una nova versió amb data i responsable.

Versions, registre i difusió del PGD

Cada PGD ha d'indicar una versió i mantenir un registre d'historial amb els canvis realitzats. Les versions aprovades es custodien al registre institucional de projectes i, quan correspongui, s'acompanyen del dataset publicat (o de la metadada) en el repositori triat. El PGD, en la seva versió pública, ha d'incloure l'abstrac i les condicions d'accés, amb enllaç a la versió completa per a les parts autoritzades.

Suport i eines

La Unitat de Banc de Dades, la Unitat de Gestió de Projectes i la Unitat d'Informàtica ofereixen suport tècnic i plantilles estandarditzades per a l'elaboració del PGD (annexos). També s'ofereix assistència per a la realització de DPIA, per a la implementació de mesures d'anonimització i per a la selecció de repositoris adequats, així com materials formatius destinats als equips de recerca.

La Unitat de Banc de Dades del CED

És la Unitat tècnica del CED que gestiona, guarda i estructura la informació estadística que es fa servir al CED. Dona suport als investigadors del CED en les sol·licituds de dades a proveïdors estadístics i també assessora en tractaments i metodologia relacionada amb diferents tipus d'informació estadística.

Sobre les sol·licituds de dades

El protocol a seguir a les sol·licituds de dades estadístiques, per part del personal investigador del CED.

(1) Si a la sol·licitud de dades es relaciona alguna o diverses d'aquestes opcions:

- que qui ho sol·licita manifesta la pertinença del personal investigador al CED
- que s'aplicarà la normativa de protecció de dades segons els protocols del CED
- que es disposaran o es tractaran les dades en servidors o dispositius ubicats al CED
- que es faran còpies de seguretat segons els protocols del CED
- que es faran les corresponents limitacions d'accés segons els protocols del CED
- que es procedirà, quan calgui -si cal-, a la seva destrucció segons els protocols del CED

aleshores a les sol·licituds s'ha d'incloure el nom i signatura de la persona responsable de la unitat.

Els avantatges, a més que la unitat pugui portar la gestió de la sol·licitud, són:

- disposar del segell del CED quan es requereixi
- disposar de l'autorització del director quan sigui necessari
- aplicar-ho als fons de la recerca del CED, quan s'hagi de pagar per obtenir la informació
- per a les dades d'Eurostat, el CED és un centre reconegut i tenim un registre mitjançant el qual derivem i avancem aquestes sol·licituds.
- per a altres organismes estadístics o institucions nacionals i internacionals, disposem de certs canals a més de la nostra experiència i reconeixement, que poden facilitar els tràmits

(2) Si, per contra, la sol·licitud es fa sense relacionar cap de les opcions de l'apartat (1), aleshores tota la gestió i la responsabilitat recaurà en la persona sol·licitant/s de la informació.

Sobre l'ús i la cessió de fitxers de microdades estadístiques disponibles

Els fitxers de microdades disponibles al CED només es poden utilitzar en projectes de recerca i/o publicacions en què participi almenys una persona contractada, vinculada o adscrita al CED.

Les condicions d'ús i cessió d'aquests fitxers permeten diferenciar-ne dos tipus:

1. Fitxers d'ús general. Els estudiants de doctorat i les persones que realitzen una estada al CED poden accedir a aquests fitxers mentre duri la seva relació amb la institució i prèvia conformitat per part dels investigadors/es principals (IPs) del projecte al qual estan inscrits o dels seus tutors/es. A més, han d'emplenar un document, existent al Banc de Dades, on s'indica per a quin projecte o publicació s'utilitzaran i les persones que hi participaran. Aquest document ha de ser signat per cada persona a qui se cedeixin les microdades, especificant el termini temporal d'ús, i el signant s'ha de comprometre a l'esborrament del fitxer una vegada transcorregut aquest termini.
2. Fitxers amb restricció d'ús. L'accés a aquests fitxers està limitat exclusivament a les persones que van signar el document de confidencialitat corresponent quan es van sol·licitar als organismes productors, com INE, IDESCAT, EUROSTAT, entre d'altres. Per tant, aquests fitxers NO ES PODEN CEDIR a altres persones, fins i tot si ho sol·licités l'investigador/a que va signar el document de confidencialitat. L'accés i ús per part d'altres persones requeriria realitzar una nova sol·licitud al productor de l'estadística, especificant els nous usuaris/es, el projecte o la publicació, i les signatures corresponents del document de confidencialitat.

Aquestes directrius són fonamentals per salvaguardar la custòdia dels fitxers de microdades disponibles al CED i garantir el compliment de les normatives vigents.

Sobre la plantilla DMP

Les accions previstes dins del Grup de treball HRS4R van recomanar de fer una plantilla de Pla de Gestió de Dades del CED. El CED disposa de un document en els tres idiomes, en el qual hi ha indicacions sobre què hauria de contenir cadascun dels apartats que solen requerir-se. Els documents són al Google Drive del CED a la carpeta compartida:

Intercanvis: 4. Protocols CED (4.6. Pla de gestió de dades).

Sobre les mesures de seguretat i la protecció de dades

La seguretat de la informació i la protecció de les dades personals són elements essencials de la responsabilitat científica i institucional del Centre d'Estudis Demogràfics (CED). Per garantir la confidencialitat, la integritat i la disponibilitat de la informació gestionada pel Centre, el CED aplica un conjunt integrat de mesures de seguretat de caràcter **físic, lògic i organitzatiu**, que s'implementen de manera proporcionada segons la sensibilitat i la classificació de cada conjunt de dades.

- **Mesures físiques i d'infraestructura:** L'emmagatzematge central i les infraestructures crítiques s'ubiquen en instal·lacions autoritzades i amb accés físic restringit del CED (sales de servidors amb control d'entrada). L'accés físic es regula segons polítiques establertes i només està autoritzat per personal amb responsabilitats concretes.
- **Mesures lògiques i tècniques:** En l'àmbit lògic, s'apliquen controls d'accés basats en el principi de mínim privilegi i la separació de tasques, gestionats per usuari i per rol. Les dades sensibles es protegeixen amb xifrat en repòs i en trànsit; les connexions

externes i les comunicacions segueixen protocols segurs (TLS/HTTPS). Les defenses perimetrals inclouen tallafocs i sistemes de detecció/prevenició d'intrusions; les credencials es gestionen amb polítiques robustes (longitud mínima, complexitat, rotació periòdica) i, per als perfils amb accés a dades sensibles, s'exigeix autenticació de múltiples factors. La infraestructura actual del CED incorpora, entre altres components, un servidor Synology per a emmagatzematge, controls d'accés gestionats per la Unitat d'Informàtica i la Unitat de Banc de Dades, i el tallafocs del Campus UAB; aquests elements s'integren sota polítiques i procediments comuns de seguretat.

- **Còpies de seguretat i recuperació:** Les còpies de seguretat són redundants i s'executen amb la periodicitat i el retenció definides segons la categoria de les dades. Existeixen protocols documentats de restauració i de recuperació davant d'incidents, amb proves periòdiques per verificar la validesa de les còpies i la capacitat de recuperació.
- **Traçabilitat i auditories:** El CED manté registres d'accés i d'auditories que assegurin la traçabilitat de les operacions sobre dades. Els logs es custodien amb control d'integritat i s'utilitzen en auditories periòdiques per revisar el compliment de les polítiques i detectar incidents.
- **Gestió d'incidents:** S'estableix un procediment institucional per a la gestió d'incidents de seguretat i de violacions de dades personals que contempla la detecció, la contenció, l'anàlisi, la notificació i la remediació. Quan una violació de dades personals pugui suposar un risc per als drets i llibertats de les persones, el CED notificarà l'incident a l'Agència Espanyola de Protecció de Dades (AEPD) en els terminis establerts per la normativa i, si escau, comunicarà la violació a les persones afectades, d'acord amb les exigències del RGPD.
- **Responsabilitat i revisió:** Les mesures de seguretat són objecte d'actualització periòdica per respondre a l'evolució de les amenaces, la tecnologia i la normativa. El/la Responsable de Gestió de Dades, el/la Delegat/da de Protecció de Dades (DPO) i la Unitat d'Informàtica coordinen la implementació i la revisió d'aquestes mesures.
- **Protecció de dades personals i mesures ètiques:** El CED assumeix el compromís de respectar la normativa de protecció de dades i les obligacions ètiques associades a la recerca. En aquest marc, el Centre:
 - aplica mesures tècniques i organitzatives adequades per garantir el compliment del Reglament (UE) 2016/679 (RGPD) i de la Llei Orgànica 3/2018, de Protecció de Dades Personals i garantia de drets digitals (LOPDGDD);
 - respecta la normativa sobre *secret estadístic* conforme a l'àmbit de l'estadística oficial, tal com estableix la normativa estatal aplicable;
 - implanta tècniques d'anonimització i pseudonimització adequades al risc i al context, i utilitza tècniques de divulgació segura quan es requereix evitar riscos de reidentificació;

- garanteix la recollida de consentiment informat clar i suficient quan sigui la base legal aplicable, indicant usos previstos, períodes d'emmagatzematge i possibles transferències o reutilitzacions;
- sotmet a revisió ètica tots els projectes amb risc per a les persones; i
- respecta les condicions d'accés imposades per proveïdors o convenis contractuals.

Accés, llicències i reconeixement

El CED treballa sota la premissa que la màxima disponibilitat i reutilització de les dades, sempre que sigui compatible amb la protecció de persones i amb compromisos contractuals, augmenta el valor científic i social dels resultats de recerca. Per això, la política fomenta l'ús de llicències que afavoreixen la reutilització i el reconeixement, com ara Creative Commons Attribution (CC-BY), sempre que la naturalesa de les dades i les restriccions legals ho permetin.

Requeriments per a la difusió.

Quan sigui procedent alliberar conjunts de dades, el CED promou que els datasets públics es dipositin en repositoris reconeguts i subministrin, com a mínim:

1. **Identificadors persistents** (per exemple, DOI).
2. **Informació per a la citació:** autor(s) amb ORCID quan sigui possible, títol, any, DOI i versió.
3. **Metadades estàndard:** ús d'esquemes reconeguts (DataCite, Dublin Core), amb informació bàsica i complements tècnics que permetin la interpretació (cobertura temporal i espacial, definicions de variables, procediments de recollida i processament).
4. **Documentació d'acompanyament:** README i codebook que expliquin l'estructura del conjunt de dades, les decisions de neteja i transformació, els scripts d'anàlisi disponibles i les limitacions d'ús.

Dades de tercers i restriccions contractuals.

Quan les dades procedeixen de tercers o estan subjectes a condicions contractuals, el CED respectarà estrictament aquelles condicions: no es publicarà ni es posarà a disposició pública informació contrària als termes acordats, i es formalitzaran mecanismes d'accés restringit quan calgui (acords de confidencialitat, Data Use Agreements - DUA).

Política d'embargament i condicions d'accés.

El PGD ha d'indicar, amb justificació, qualsevol període d'embargament proposat. Per a dades restringides o confidencials s'aplicaran procediments formals de sol·licitud i avaluació d'accés, amb la formalització d'un DUA que especifiqui finalitats, durada de l'accés, mesures de seguretat i obligacions del sol·licitant. Quan s'accedeixi a dades sensibles, l'accés

preferentment s'efectuarà en entorns tècnics segurs (Data Safe Haven) amb registre d'activitats i controls d'auditoria.

Reconeixement i responsabilitat de l'usuari.

Tots els usuaris que accedeixin a conjunts de dades del CED s'han d'obligar formalment a reconèixer l'origen de les dades mitjançant la cita corresponent i a respectar les condicions d'ús establertes. El CED pot exigir la remissió de resultats o publicacions derivades com a condició de l'accés quan així ho requereixi el DUA

Suport, formació i recursos

L'eficàcia d'aquesta política depèn de la capacitat i el suport que es posin a disposició dels equips de recerca. El CED ofereix assistència tècnica i formativa en les següents dimensions: elaboració i revisió de Plans de Gestió de Dades (PGD/DMP), procediments de dipòsit i conservació, requisits tècnics per a l'emmagatzematge segur i pràctiques d'anonimització, així com assessorament referent a la normativa de protecció de dades i a les condicions contractuals de tercers.

Les activitats de capacitat inclouen sessions formatives periòdiques, materials pràctics (guies, checklists, plantilles) i suport individual o per projecte a través dels responsables i contactes del RGD, la Unitat de Banc de Dades i la Unitat d'Informàtica. El Centre es compromet, a més, a reservar recursos pressupostaris per garantir l'emmagatzematge i la preservació a llarg termini de les dades i per mantenir el suport tècnic necessari.

Alineament institucional, referències normatives i obligacions operatives derivades

Alineament institucional.

Aquesta política s'ha redactat amb l'objectiu d'assegurar coherència amb els marcs internacionals i nacionals rellevants: directrius de la Comissió Europea en matèria de Ciència Oberta i gestió de dades, la normativa espanyola i catalana aplicable, l'estratègia i les orientacions de la Institució CERCA, així com els requisits específics establerts pels finançadors. L'enfocament adoptat busca conjugar l'aposta per l'obertura i la reutilització amb el respecte estricte a la privacitat, la seguretat i les obligacions contractuals.

La política s'entén com una **guia viva**: s'ha concebut per ser clara, accessible i pràctica, i per garantir que les dades del CED siguin valuoses avui i mantenibles demà. Les mesures i procediments que descriuen han de preservar tant la dignitat i els drets de les persones participants en la recerca com el rigor i la traçabilitat científica que exigeix la comunitat acadèmica.

Notes legals i referències normatives

Aquesta política es desplega en el marc jurídic i reglamentari europeu i espanyol aplicable a la protecció de dades i a la confidencialitat estadística. Les referències fonamentals són les següents:

- **Reglament (UE) 2016/679 del Parlament Europeu i del Consell**, del 27 d'abril de 2016, relatiu a la protecció de les persones físiques pel que fa al tractament de dades personals i a la lliure circulació d'aquestes dades (Reglament General de Protecció de Dades — *RGPD*). Aquest Reglament estableix els principis de licitud, limitació de la finalitat, minimització de dades, exactitud, limitació de la conservació, integritat i confidencialitat, i imposa obligacions com la designació del Delegat de Protecció de Dades, la realització d'Avaluacions d'Impacte sobre la Protecció de Dades (DPIA) quan correspongui, i la notificació de violacions de seguretat en els terminis establerts (72 hores).
- **Llei Orgànica 3/2018**, del 5 de desembre, de Protecció de Dades Personals i garantia de drets digitals (*LOPDGDD*), que desenvolupa i complementa el RGPD en l'ordenament jurídic espanyol.
- **Normativa sobre estadística oficial i secret estadístic**: normativa estatal que regula la funció estadística pública i el deure de confidencialitat de les dades estadístiques, amb les obligacions corresponents en matèria de tractament i reserva d'informació (normativa bàsica sobre la funció i l'organització estadística, i les disposicions que estableixen el caràcter de secret estadístic per a dades proporcionades amb finalitat estadística). El CED assegura el compliment de les obligacions de secret estadístic en els casos en què aquesta normativa sigui d'aplicació.
- **Directrius i polítiques europees de Ciència Oberta i gestió de dades** (Comissió Europea, Horizon Europe, i orientacions sobre pràctiques FAIR), que promouen la gestió responsable de dades i l'obertura en condicions que preservin els drets fonamentals i la seguretat.
- **Autoritats de control i orientacions**: Agència Espanyola de Protecció de Dades (AEPD) i, en l'àmbit autonòmic, l'Autoritat Catalana corresponent, així com les guies i recomanacions de l'European Data Protection Board (EDPB) i altres organismes competents.

Obligacions operatives derivades de la normativa.

En l'aplicació d'aquest marc legal (punt 11.2), el CED:

- manté un registre d'activitats de tractament conforme al RGPD;
- designa un/a Delegat/da de Protecció de Dades quan la naturalesa de les activitats ho exigeixi i assegura la seva participació en qüestions rellevants;
- realitza Avaluacions d'Impacte sobre la Protecció de Dades (DPIA) prèvies per a tractaments amb alt risc per als drets i llibertats de les persones (p. ex. tractaments extensius de dades sensibles o combinació de conjunts de dades amb possible re-identificació);

- formalitza acords de encarregat de tractament (Data Processing Agreements) amb tercers que gestionin dades en nom del CED;
- notifica les violacions de seguretat a l'AEPD i, si escau, als interessats, d'acord amb els terminis i criteris del RGPD; i
- adopta mesures tècniques i organitzatives per garantir la protecció de dades «per disseny i per defecte» (article 25 del RGPD).

Secret estadístic. Quan les dades s'han recollit o s'utilitzen amb finalitat estadística i queden afectades per la normativa de secret estadístic, el CED s'ajusta a les obligacions de confidencialitat, a les limitacions d'accés i a les sancions corresponents i aplica mesures addicionals de protecció per impedir usos aliens a la finalitat estadística autoritzada.

Cerdanyola del Vallès, 18 de desembre de 2024

ANNEXOS

ANNEX 1: Plantilla de Pla de Gestió de Dades

Indicacions d'ús: emplena aquesta plantilla abans de començar el projecte i actualitza-la periòdicament. Sigues concís/sa en la pàgina principal i amplia en annexos tècnics.

1. Identificació del projecte

- **Títol del projecte:**
- **Codi / Referència (si escau):**
- **Investigador/a responsable (IP):** (nom, càrrec, correu, ORCID)
- **Equip/projecte:** (altres responsables, tècnics)
- **Període del projecte:** data d'inici / finalització
- **Finançador(s):** (nom + referència del conveni)
- **Versió del PGD:** data i autor de l'actualització

2. Resum del projecte (150–300 paraules)

Breu descripció del projecte i dels seus objectius associats a les fonts de dades que s'utilitzaran i es generaran (ja siguin primàries o secundàries), finalitat científica i valor potencial de reutilització.

3. Tipus de dades generades i formats

Per a cada tipus de dades (ex.: dades primàries o secundàries, enquestes, registres administratius, dades administratives, fitxers multimèdia, codis, resultats anàlisis) indica:

- **Nom del dataset / element:**
- **Tipus:** (numèric, text, geoespacial, imatge, etc.)
- **Format preferent:** (p.ex. CSV UTF-8, JSON, NetCDF, TIFF)
- **Mida estimada:** (GB)
- **Freqüència de generació / actualització:**

4. Metadades i estàndards

- **Esquema de metadades a utilitzar:** (DataCite, Dublin Core, esquema específic)
- **Ús d'identificadors:** DOI per datasets finals; ORCID per autors.
- **Vocabularis controlats:** (p.ex. ISO3166, nomenclatures INE, no aplica, descripció al README)
- **Fitxa mínima de metadades inclosa al dataset:** títol, autors, abstract, data, format, identificador, cobertura temporal i geogràfica, variables clau, llicència, contactes.

5. Privacitat i aspectes ètics

- **Conté dades personals?** (Sí / No)
- **Categoria de risc:** (Justifica)
- **Base legal (RGPD):** (consentiment, interès públic, execució d'una missió pública...)

- **Consentiment informat:** (incloure text o referència al formulari)
- **Mesures d'anonimització/ pseudonimització:** (mètode, responsable)

6. Emmagatzematge i seguretat

- **On s'emmagatzemaran les dades durant el projecte:** Servidor institucional / Altra ubicació (indicar)
- **Còpies de seguretat:** freqüència, ubicacions.
- **Xifrat:** (sí/no) quin nivell (AES-256...).
- **Control d'accés:** rols i permisos (qui pot llegir/editar/eliminar).
- **Entorns segurs per a dades sensibles:** (sí/no) descripció.

7. Control de qualitat i versions

- **Controls aplicats:** validació, detecció de valors extrems, integritat.
- **Sistema de versionat:** (p.ex. etiquetatge semàntic, Git per a codi).
- **Documentació de transformacions:** scripts, logs, fitxers de provenència.

8. Compartició i accés

- **Nivell proposat:** obert / restringit / confidencial / tancat (justificació)
- **Repositori suggerit per dipòsit final:** (CORA-RDR, Dataverse, Zenodo, etc.)
- **Embargament (si escau):** durada i motiu.
- **Condicions per a l'accés restringit:** DUA, revisió de sol·licituds, etc.

9. Preservació i retenció

- **Durada mínima de retenció:** (p.ex. 5 anys)
- **Repositori de preservació:** (institucional / extern)
- **Formats per preservació:** (formats oberts preferits)
- **Procés d'eliminació segura:** descriure.

10. Responsables i recursos

- **Responsable de gestió de dades (RGD):** nom, contacte.
- **Responsable tècnic d'emmagatzematge:** nom, contacte.
- **Pressupost estimat per PGD (recursos d'emmagatzematge/preservació):** €

11. Annexos (listar documents tècnics)

- Formularis, models de consentiment, scripts, fitxers de metadades detallats.

12. Signatura

- **Investigador/a Responsable (IP):** nom, data, signatura
- **Responsable de Gestió de Dades (RGD):** nom, data, signatura

ANNEX 2: Formulari de sol·licitud d'accés a dades del CED

Instruccions: persona/entitat que sol·licita dades completa aquest formulari; RGD verifica i remetrà a CAD.

1. Identificació sol·licitant

- Nom i cognoms:
- Institució / Organització:
- Càrrec:
- Correu electrònic institucional:
- Telèfon:
- ORCID (si escau):

2. Projecte i finalitat

- Títol projecte:
- Resum de l'objectiu i justificació de la necessitat de les dades (max 500 paraules):
- Durada prevista de l'accés (dates):
- Resultats esperats / productes (articles, informes, etc.):

3. Dades sol·licitades

- Identificació del dataset (títol / DOI / codi intern):
- Parts del dataset demanades (variables/periods/mostres):
- Nivell d'accés sol·licitat: visualització / descàrrega / accés en entorn segur:
- Raó per la qual l'accés obert no és suficient:

4. Requeriments tècnics i seguretat

- Entorn d'anàlisi proposat (Own workstation / Secure remote environment / Data Safe Haven):
- Mesures de seguretat que es prendran (descripció):

5. Altres compromisos

- Acceptes signar un DUA amb condicions i termes d'ús? (Sí / No)
- Declaració de no intentar la re-identificació de persones: (Signatura / reconeixement)

6. Declaracions i signatura

- Data i lloc:
- Signatura del sol·licitant:

ANNEX 3: Model de Data Use Agreement (DUA) (model orientatiu)

ACORD D'ÚS DE DADES (Data Use Agreement)

Entre: **Centre d'Estudis Demogràfics (CED)** (en endavant "CED") i **[Nom Entitat / Persona Sol·licitant]** (en endavant "Sol·licitant").

Preàmbul

CED és titular / gestor del dataset **[Títol / DOI / Codi]**. El Sol·licitant sol·licita accés amb la finalitat de **[descriure]**. Ambdues parts acorden les condicions següents:

1. Objecte

Permetre l'accés a **[descripció del dataset]** sota les condicions i limitacions definides en aquest acord.

2. Durada

L'accés és autoritzat per període de **[mesos/anys]**, des de **[data inici]** fins **[data final]**. Renovació per escrit.

3. Condicions d'ús

3.1. El Sol·licitant s'obliga a:

- No intentar la reidentificació de persones.
- Utilitzar les dades únicament per a la finalitat aprovada.
- No transferir les dades a tercers sense autorització prèvia i per escrit.
- Mantenir mesures de seguretat tècniques i organitzatives (MFA, xifrat, control d'accés).

3.2. Publicacions i reconeixement:

- En qualsevol publicació s'ha d'esmentar i citar el dataset amb el DOI i incloure la frase: "Dades facilitades per Centre d'Estudis Demogràfics (CED) i subjectes a DUA."
- Compartir còpia dels resultats i publicacions amb CED.

4. Responsabilitats legals

- El Sol·licitant és responsable de l'ús conforme a RGPD. Si es produeix una violació de dades, notificar al RGD i DPO del CED en 72 hores.

5. Retorn i destrucció

A la finalització de l'accés, el Sol·licitant ha de:

- Destruir totes les còpies de dades (incloses còpies de seguretat) o retornar les dades al CED com s'indica al PGD.
- Presentar un informe breu d'ús.

6. Incompliment

El incompliment d'aquest acord pot comportar la revocació de l'accés, responsabilitats legals i sancions contractuals.

7. Legislació aplicable i jurisdicció

L'acord es regeix per la legislació espanyola i la jurisdicció dels tribunals corresponents (o altra acordada).

Signatures

Pel CED:
(nom, càrrec, data)

Pel Sol·licitant:
(nom, càrrec, data)

ANNEX 4: Checklist per al dipòsit de datasets

Descripció	Sí / No / N/A	Observacions
Existeix un README/fitxa del dataset?		
Metadades mínimes (DataCite, etc.) completes?		
Formats de dades segons especificació (CSV, pdf, etc.)?		
Scripts d'extracció/transformació disponibles i documentats?		
Versions del codi (repo Git) amb DOI/etiquetat?		
Controls de qualitat aplicats (valors estranys, missings)?		
Consentiment i bases legals correctament documentades?		
Llicència d'ús indicada (CC-BY, CC0, etc.)?		
Nivell d'accés i procediment d'accés definits?		
Pla de preservació i repositori assignat?		
Document d'anotacions (codebook, README) adjunt?		
Data i responsable de la revisió QC		

ANNEX 5: Full de metadades DataCite

Instruccions: emplenar per cada dataset; aquest full servirà per alimentar el registre institucional o el repositori.

1. Identificació i accés

- **Title (títol):**
- **Creators (autors):** (llista amb ORCID si disponible)
- **Publisher:** Centre d'Estudis Demogràfics (CED)
- **PublicationYear:** YYYY
- **Identifiers:** (DOI o codi intern)
- **ResourceType:** Dataset / Code / Other

2. Descripció

- **Abstract / Description:** (300–500 paraules)
- **Subject:** (keywords, temàtiques)
- **Dates:** coverage temporal (from/to)

3. Coverage

- **Spatial coverage:** (llocs geogràfics, coordenades si escau)
- **Temporal coverage:** (període cobert pel dataset)

4. Technical metadata

- **Format:** (CSV UTF-8, JSON...)
- **Size:** (GB o MB)
- **Checksum:** (SHA-256)
- **Version:** (v1.0, etc.)
- **Software environment / dependencies:** (scripts, containers)

5. Rights & Licenses

- **Rights:** (RGPD notes, restriccions)
- **License:** (CC-BY 4.0 / CC0 / restricció per DUA)

6. Related identifiers

- **Related publications:** (articles amb DOI)
- **Related datasets / code:** (DOIs, URLs)

7. Contact & curation

- **Contact person:** nom, correu, rol
- **Custodian / Data manager:** nom, contacte
- **Preservation repository:** (CED institucional / CORA RDR / Dataverse)

8. Access policy

- **Access level:** Obert / Restringit / Confidencial / Tancat
- **Procedure to request access:** URL o contacte

9. Additional notes

- **Embargo:** si escau, data de finalització
- **Other important metadata:** (notes de qualitat, advertències)

ANNEX 6: Glosari

Ciència oberta

És un model de recerca col·laboratiu, transparent i accessible, on es millora l'accés als recursos científics, incloses les publicacions, les dades, la metodologia, el codi, entre d'altres, de manera que aquestes siguin accessibles a tots els nivells de la societat, tant per a aficionats com per a professionals, especialment d'aquell coneixement científic generat amb fons públics.

Accés obert

(Open Access –OA-) és l'accés immediat, sense requeriments de registre, subscripció o pagament –és a dir, sense restriccions- a material digital educatiu, acadèmic, científic o de qualsevol altre tipus

Gestió de dades:

La Gestió de Dades de Recerca comprèn el procés actiu del maneig de les dades generades o recopilades durant totes les fases de recerca. Aquest procés ha de planificar-se des de l'inici del projecte i abasta totes les decisions relacionades amb el tractament de les dades al llarg del seu cicle vital.

Cicle de Vida de les Dades de Recerca:

El cicle de vida de les dades comença amb la planificació, continua durant l'execució del projecte fins a la difusió dels resultats i conclou amb la preservació de les dades. Aquest procés garanteix que les dades siguin precises, completes, fiables, accessibles i reutilitzables al llarg del temps.

Protecció de dades i ètica:

La recerca demogràfica està subjecta a nombroses regulacions degut a la naturalesa de les seves dades. És essencial identificar i documentar tots els requisits legals aplicables.

Llicències i propietat intel·lectual:

Llicències d'ús de les dades segons la seva naturalesa, en general només les fonts secundàries podran ser compartides.

Dataset: fitxer de dades i metadades

Conjunt de fitxers que han d'incloure un README o fitxer de metadades, en els quals es descriuen els seus continguts: carpetes i subcarpetes, codis utilitzats, descripció de les variables, etc. i a què es refereixen a una acció de recerca determinada, com un article, una contribució a un congrés o un apartat del projecte de recerca.

Dades FAIR:

Sigles angleses de: FINDABLE : Assegura que les dades puguin ser fàcilment localitzades dins de les infraestructures de dades; ACCESSIBLE: Garanteix l'accés als conjunts de dades per als usuaris autoritzats; INTEROPERABLE: Garanteix l'accés als conjunts de dades per als usuaris autoritzats; REUSABLE: Facilita la reutilització de les dades per a futures investigacions

Dades obertes

Posicionar les dades en accés obert, de manera que puguin ser accessibles tenint en compte les restriccions de llicències d'ús.

Repositoris:

Per maximitzar la visibilitat i reutilització de les dades, cal escollir repositoris adequats segons criteris com l'especialització (repositoris temàtics, multidisciplinars o institucionals), la visibilitat científica (presència en la comunitat investigadora), l'adjudicació d'Identificadors persistents-DOI (per garantir la traçabilitat), la dotació mecanismes de control d'accés a les dades (per protegir dades sensibles, comerciabilitats, etc.) i la disposició d'eines de cerca que permetin facilitar la recuperació i reutilització de dades.

Nivells de compartició i accés a les dades:

Obert: dades anonimitzades sense restriccions ètiques o contractuales. Reutilització obert.

Restringit: dades amb risc moderat o amb condicions per part de proveïdors/participants (acords d'ús, embargo). Accés sota sol·licitud i acord.

Confidencial: dades amb risc elevat d'identificació; sol·licituds avaluades per comitè i amb mesures tècniques (xifrat, entorns segurs).

Tancat: dades no compartibles per contracte o legalment restringides.

Vocabularis controlats

Glossaris reconeguts per variables demogràfiques (p.ex. codis INE, ISO 3166 per països, etc.).