

Guia per elaborar un pla de gestió de dades per a doctorands

(Doc.CO20/01) (4 RDM\Plans Gestio Dades\PlansdeGestiodeDades_Doctorands-CA.docx, 29.01.20)

Aquest document vol servir de suport als doctorands a l'hora de crear els seus Plans de Gestió de Dades (*Data Management Plan, DMP*).

Aquest document ha estat elaborat pel Grup de treball de Suport a la Recerca del CSUC que està format per representants de les universitats següents: Universitat de Barcelona, Universitat Autònoma de Barcelona, Universitat Politècnica de Catalunya, Universitat Pompeu Fabra, Universitat de Girona, Universitat de Lleida, Universitat Rovira i Virgili, Universitat Oberta de Catalunya, Universitat de Vic-Universitat Central de Catalunya, Universitat Ramon Llull, Universitat de les Illes Balears i Universitat Jaume I.

Aquest document està subjecte a la llicència de Reconeixement de Creative Commons (<http://creativecommons.org/licenses/by/4.0/>).

Versió digital: <http://hdl.handle.net/2072/370509>

NOTA: El punt 6. Dades sensibles/personals ha estat elaborat per la Universitat de Lleida i és necessari per al tractament de dades de les tesis doctorals de la UdL.

Sobre la teva recerca

Nom i adreça electrònica	
Director/s de tesi	
Títol provisional de la tesi	
Describeu la teva recerca	<i>Aproximadament 50 paraules</i>
Durada de la teva recerca	Inici: <i>DD-MM-AAAA</i> Fi: <i>DD-MM-AAAA</i>
Projecte vinculat	<i>Es tracta d'una tesi vinculada a un projecte? Quin?</i>
Finançament	<i>Has rebut finançament per a fer la tesi? De qui?</i>

Sobre aquest pla de gestió de dades

Data de creació	<i>DD-MM-AAAA</i>
Última actualització	<i>DD-MM-AAAA</i>
Versió i data	<i>Fes una nova versió cada vegada que hi hagi canvis significatius (nous datasets, canvis en la teva investigació o altres factors)</i>

Dades sensibles/personals	<p><i>Si treballes amb <u>dades de caràcter personal o sensible</u>, tens l'obligació legal de tractar-les d'acord amb la normativa vigent. Les dades personals són qualsevol informació que permet identificar una persona (nom, adreça, localització, etc.).</i></p> <p><input type="checkbox"/> No treballaré amb dades personals</p> <p><input type="checkbox"/> Treballaré amb dades personals [vegeu punt 2 i 6]</p>
---------------------------	--

1. Captura de dades

Descriu les dades que crearàs/recolliràs

1.1 Durant la teva recerca utilitzaràs dades ja existents? Si no és així, indica la procedència de les dades que utilitzaràs

- No
- Dades pròpies o dades del grup de recerca en que participes
- Col·laboradors acadèmics
- Col·laboradors de caràcter comercial
- Bases de dades/arxius disponibles de manera pública
- Proveïdors comercials de dades
- Altres (indica quines):

1.2 Descripció de les dades

Descriu les dades que crearàs o les dades de tercers que reutilitzaràs i indica:

- si usaràs protocols o estàndards usats en la teva àrea de recerca
- quines eines, instruments, equipament, maquinari o programari faràs servir

Si reutilitzes dades de tercers, assegura't que disposes dels permisos adequats i de conèixer les condicions d'ús d'aquestes dades.

1.3 Tipus i formats de les dades

Tingues en compte:

- la tipologia de dades: per exemple, si treballaràs amb mesures, simulacions, observacions, text (text, MS Word), imatges, audiovisuals o mostres, estadístiques (fulls de càlcul), amb models computacionals, amb dades d'una enquesta qualitativa (qüestionaris), gravacions (àudio, vídeo), programari (codi), etc.
- la longevitat dels formats de fitxer: utilitza preferentment estàndards oberts perquè les dades puguin ser llegides per diversos programes facilitant la preservació i la compartició amb altres usuaris.

1.4 Indica el volum de dades

- < 10 GB
- 10-30 GB
- 30-50 GB
- 50 GB-250 GB
- 250 GB-500 GB
- 500 GB-2 TB
- 2 TB

2. Emmagatzematge de dades i seguretat

Assegura't que totes les teves dades de recerca s'emmagatzemen de manera segura i que es fan còpies de seguretat periòdicament

2.1 Indica les restriccions (comercials, ètiques o de confidencialitat) que poden afectar les teves dades

- Obligacions contractuals
- Obligacions legals: protecció de dades personals ([LOPDGDD](#), [RGPD](#)...) [vegeu 4.1]
- Obligacions legals: copyright, propietat intel·lectual [vegeu 4.1]
- Restriccions ètiques [vegeu 4.1]
- Aspectes comercials (p. ex. patentabilitat)
- Estàndards formals de seguretat
- Sense obligacions
- Altres, indica:

Explica breument les restriccions

2.2 Riscos principals en la seguretat de les dades

Indica els principals riscos, per exemple: esborrat accidental de les dades, pèrdua o robatori de les dades. Descricu les conseqüències d'una possible pèrdua de dades

2.3 Mesures a prendre per mitigar els riscos d'una possible pèrdua de dades

- Restriccions d'accés
- Encriptació
- Processament de dades
- Pseudoanonimització
- Anonimització
- Còpies de seguretat periòdiques
- Altres, indica:

Indica també els procediments que usaràs per assegurar la confidencialitat de les dades personals

2.4 On emmagatzemaràs les dades?

En cas de tractar i emmagatzemar dades personals contesta el punt: 6.6 On emmagatzemaràs les dades en cas de tractar dades personals?

- A la xarxa del teu departament o grup de recerca
- A la xarxa de la universitat
- Emmagatzematge físic (p. ex. USB, disc dur extern)
- Servei al núvol (p. ex. DropBox)
- Altres, indica:

Descriu breument les condicions d'emmagatzematge i còpies

3. Documentació de les dades

Documenta les dades per facilitar-ne la comprensió i la reutilització

3.1 Nom i estructura dels fitxers i carpetes

Describeix com organitzaràs i anomenaràs les carpetes i fitxers de dades

3.2 Control de versions

Describeix breument com controlaràs les versions. Indica també què faràs en cas d'eliminar dades

- Sense control de versions (p. ex. els fitxers originals se sobreescriven)
- Programari amb control de versions, indica'l:
- Programari amb opció de seguiment de canvis
- Número de versió i data en el nom del fitxer o carpeta
- Fent una còpia de l'script amb el qual es processen les dades
- Altres, indica:

3.3 Quins estàndards de metadades utilitzaràs?

Vegeu "[Disciplinary metadata standards](#)" del DCC o "[Metadata standards](#)" a la Wikipedia.

- No usaré cap estàndard (especifica les metadades necessàries per entendre les dades)
- Esquema de metadades genèric (p. ex. Dublin Core)
- Esquema de metadades automàtic de Windows (p.ex. de Word, Excel)
- Esquema de metadades temàtic especialitzat, indica:
- Un altre esquema de metadades, indica:

Indica com es crearan (en un fitxer "readme", un full de càlcul, incrustades a les dades) i quina documentació elaboraràs per tal que les dades siguin entenedores per altres

4. Accedir, compartir i reutilitzar les dades

4.1 Tens restriccions per compartir les dades en relació al reglament vigent ([General Data Protection Regulation](#)) o altres (ètiques, comercials, de seguretat, propietat intel·lectual o copyright)?

Indica quines.

4.2 Quins són els usuaris potencials de les teves dades i com les descobriran?

Describeix breument qui podria estar interessat en la teva recerca i quina difusió en faràs (p .ex. dades al repositori, pàgina web, publicacions en congressos, , etc.)

4.3 Especifica quines llicències aplicaràs a les dades per permetre la seva reutilització al màxim

Es recomana l'ús de llicències Creative Commons (CC - BY o CC Zero) o GNU

5. Dipòsit i preservació de les dades

Tingues en compte, també, que totes les tesis doctorals avaluades favorablement han de ser publicades als respectius dipòsits institucionals d'accés obert.

5.1 Quins criteris faràs servir per seleccionar les dades a preservar a llarg termini?

- Tipus de dades (raw, processades) i facilitat en la generació
- Rellevància del contingut per altres
- Facilitat de reutilització del format per altres
- Dades vinculades a una publicació
- Verificació de la recerca
- Temps disponibles
- Recursos econòmics disponibles
- Altres, indica:

5.2 Quant de temps preservaràs les dades?

Diferents estàndards internacionals recomanen un mínim de 10 anys.

5.3 En quin repositori dipositaràs les teves dades?

Tingues en compte els requeriments específics en quant a format, metadades, mida, cost, etc. que poden tenir per dipositar dades.

El repositori per dipositar dades a la UdL és [CORA-Repositori de Dades de Recerca](#). És un repositori de dades federat i multidisciplinar que permet a les universitats catalanes, centres de recerca CERCA i altres entitats que fan recerca la publicació de conjunts de dades de recerca en mode FAIR i seguint les directrius de l'EOSC.

- Repositori institucional
- Repositori temàtic (internacional), indica:
- Repositori multidisciplinar (p. ex. Zenodo, Figshare, Dryad)
- Altres, indica:

6. Dades sensibles / personals

6.1 Treballaràs amb dades personals?

Una dada personal és aquella relativa a una persona física identificada o identificable. En aquest sentit, s'hi poden considerar dades com el nom o DNI d'una persona que permetrien identificar-lo de forma directa, així com un conjunt de dades d'una persona que tot i no saber el seu nom permeten identificar-la.

Per exemple, una adreça IP o una adreça de correu electrònic són considerades dades personals o saber la data de naixement, etc .

En cas de treballar amb dades personals especifica amb quines dades.

- No
- Sí, amb dades personals de menors d'edat de menys de 14 anys
- Sí, amb dades personals de menors d'edat entre 14 i 17 anys
- Sí, amb dades personals de majors d'edat

6.2 Treballaràs amb dades especialment protegides, com ara dades de salut, biomètriques o genètiques, dades d'origen racial o ètnic, conviccions filosòfiques o religioses, idees polítiques o altres dades amb risc elevat per a les persones afectades?

En cas afirmatiu, indica amb quines dades.

6.3 En la meua investigació és imprescindible identificar persones concretes?

- No, perquè les dades personals són anònimes (vegeu apartat 6.4)
- Sí, (vegeu apartat 6.5)

6.4 Per què les dades personals són anònimes?

- Perquè no es capten ni es desen cap identificador de les persones afectades (adreces de correu, números d'IP de dispositius connectats a internet, número de telèfons mòbils, noms i cognoms o DNI, etc.).
- Perquè la participació en el projecte només genera un codi per a cada participant, la qual cosa fa impossible a l'equip investigador de la UdL identificar la seva identitat.
- Perquè en cap cas no es realitzen enregistraments de la veu ni de la imatge de les persones afectades.
- Perquè de manera raonable i sense esforços desproporcionats no és possible deduir, dels resultats de l'execució del Projecte (contingut de les enquestes i entrevistes, contingut de les analítiques mèdiques, etc.), la identitat de les persones.
- Perquè, tot i que es capten o es desen identificadors de les persones afectades, s'han implementat les mesures tècniques indispensables per evitar la reidentificació ("anonimització" a través del xifrat de les dades o encriptació).

6.5 Per què les dades personals identifiquen persones concretes?

- Perquè es capten o desen identificadors de les persones afectades i no s'han implementat mesures tècniques que evitin la reidentificació.
- Perquè es realitzen enregistraments de la veu o de la imatge de les persones afectades.
- Perquè de manera raonable i sense esforços desproporcionats és possible deduir, dels resultats de l'execució del Projecte, la identitat de les persones.
- Perquè s'utilitzen sistemes de pseudonimització de les dades personals.

6.6 Confirma quines eines utilitzaràs per a la gestió de les dades personals.

En cas de seleccionar la segona opció cal que contesteu el punts:

6.7 Quines eines alienes a la UdL utilitzaràs per la gestió de dades personals? Per què són imprescindibles?

6.8 Confirmo que tot i utilitzar eines alienes a la UdL per a la gestió de dades personals, aquestes dades també les emmagatzemaré en un entorn de la UdL.

- Confirmo la meua obligació d'utilitzar només eines de la UdL, o de les quals la UdL ha adquirit la llicència, sota el control del meu director de tesi.
- Confirmo que, malgrat conec l'obligació d'usar només eines de la UdL, o de les quals la UdL ha adquirit la llicència, esdevé imprescindible per a la meua investigació utilitzar-ne d'altres d'addicionals i compto amb el vistiplau del meu director de tesi.

6.7 Quines eines alienes a la UdL utilitzaràs per la gestió de dades personals? Per què són imprescindibles?

Per exemple plataformes específiques per a enquestes, com ara Google forms, EUSurvey, etc.

6.8 Confirmo que tot i utilitzar eines alienes a la UdL per a la gestió de dades personals, aquestes dades també les emmagatzemaré en un entorn de la UdL.

- Confirmo que, en qualsevol cas, les dades restaran emmagatzemades en un entorn (preferentment electrònic) de la UdL.