LA GESTIÓ DE LES DADES DE RECERCA A LA UPC

Compartir les dades de recerca és un dels mandats de la ciència oberta, que té per objectiu incrementar la cooperació i la transparència de la recerca en tots els seus estadis i augmentar-ne la seva visibilitat i impacte. El resultat és una recerca més robusta i replicable que permet incrementar els retorns dels fons públics dedicats a la investigació. Alhora, afavoreix la disciplinarietat creuada i l'obertura de nous camps de recerca. En alguns casos, l'accés o l'ús de les dades es pot restringir per salvaguardar la privadesa de les persones, protegir la confidencialitat o permetre l'explotació comercial dels resultats de recerca. Però sempre cal fer una correcta gestió de les dades de recerca seguint els principis FAIR, gestió que cal planificar prèviament.

1.- Què entenem per dades de recerca?

Les dades de recerca són els registres numèrics, textuals, imatges o sons que s'utilitzen com a fonts primàries per a la investigació científica i que són comunament acceptats per validar les conclusions i els resultats d'una investigació per la comunitat científica. Les dades de recerca poden ser dades experimentals, dades d'observació, dades operatives, dades de tercers, dades del sector públic, dades de seguiment, dades processades o dades reutilitzades. Font: Codata

2.- Com hem de gestionar aquestes dades? El Pla de Gestió de Dades o Data Management Plan

El Pla de Gestió de Dades (abreviat com PGD o DMP en anglès) és un document viu que cal elaborar a l'inici de la nostra recerca i anar actualitzant durant la vida del projecte. Inclou aspectes com:

- Els tipus de dades que recollirem, el format, el nom dels arxius i la gestió de les diferents versions
- On s'emmagatzemaran les dades i com es compartiran
- Les mesures de seguretat de les dades
- Qui serà el responsable
- Els criteris per la seva selecció i preservació
- Com garantim l'aplicació dels principis FAIR
- Aspectes ètics i de propietat intel·lectual

En el cas dels projectes finançats, és molt possible que al moment de la sol·licitud ja haguem d'esbossar com serà el nostre PGD, tot i que no s'haurà d'acabar de redactar de forma íntegra fins a l'inici del projecte, com un lliurable més. És important que a les memòries de sol·licituds quedi constància que es farà una bona gestió de dades seguint els principis FAIR.

Existeixen eines i aplicatius que ajuden a configurar el PGD amb tots els apartats necessaris com CORA-eiNa DMP. Les biblioteques de la UPC donen suport en l'elaboració del pla de gestió de dades.

3.- Principis FAIR. Què són i com els puc complir.

Els principis de dades FAIR, sigles en anglès de (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable - **T**robables, **A**ccessibles, **I**nteroperables i **R**eusables), són un conjunt de característiques que han de tenir els objectes de dades (metadades i dades), per poder ser recuperables i reutilitzables per l'ésser humà i les màquines. Font: Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

Per complir els principis FAIR les dades han de ser:

Findables

El conjunt de dades tenen metadades prou riques per descriure les dades i disposen d'un identificador únic i persistent (DOI, handle,....) que fa que les dades i metadades puguin ser fàcils de trobar per altres.

Accessible

Les dades són recuperables per persones i màquines mitjançant un protocol de comunicació estandarditzat, amb autenticació i autorització quan sigui necessari. Les metadades han de ser accessibles fins i tot quan les dades ja no estan disponibles.

Interoperable

Les dades i metadades associades utilitzen un "llenguatge formal, accessible, compartit i àmpliament aplicable per a la representació del coneixement". Això implica utilitzar idiomes, formats i vocabularis acceptats per a la comunitat a les dades i metadades. Les metadades haurien de fer referència i descriure les relacions amb altres dades, metadades i informació mitjançant identificadors.

Reusable

Les metadades associades proporcionen informació precisa i rellevant i les dades inclouen una llicència d'ús clara i informació detallada de la procedència. Utilitzar estàndards de metadades i dades específiques de l'àrea temàtica.

Cal remarcar que els principis FAIR s'han d'aplicar al conjunt de dades. Totes les dades prèvies han de ser correctament gestionades per arribar a assolir un conjunt de dades FAIR, però no s'han de complir obligatòriament sempre amb tots els principis.

4.- Recollida i tractament de dades personals

I si la meva recerca inclou el tractament de dades personals, entenent com a tal "*qualsevol operació o conjunt d'operacions realitzades sobre dades personals o conjunts de dades personals, ja sigui per procediments automatitzats o no, com ara la recollida, registre, organització, estructuració, conservació, adaptació o modificació, extracció, consulta, utilització, comunicació per transmissió, difusió o qualsevol altra forma d'habilitació d'accés, confrontació o interconnexió, limitació, supressió o destrucció*"?

Hi ha diferents situacions en les quals podem necessitar recollir dades personals al llarg del desenvolupament del nostre projecte de recerca. En aquests casos, haurem d'especificar també al PGD quines mesures prenem per salvaguardar els drets i llibertats de les persones i garantir el compliment, tant del Reglament General de Protecció de Dades (abreviat com RGPD o GDPR en anglès) com la Llei orgànica 3/2018, de 5 de desembre, de protecció de dades personals i garantia dels drets digitals.

Situació 1: Recollida de dades personals únicament a efectes de comunicació i disseminació dels projectes.

En aquest cas, no es tracta pròpiament de dades de recerca que haurem de tractar amb criteris FAIR, ni tampoc generarem datasets a publicar, però sí que haurem de definir molt clarament quin ús en farem i com s'emmagatzemaran.

Quan recollim dades personals, cal demanar sempre el consentiment exprés de les persones indicant quina és la finalitat del tractament que es farà, qui serà el responsable del seu tractament i com la persona interessada podrà exercir els seus drets.

En aquest enllaç trobareu models en català, castellà i anglès de la informació que cal afegir en els vostres formularis de recollida de dades.

D'entre els diferents codis que ha establert la Universitat especificats en el RAT (registre d'activitats de tractament de dades de la UPC) heu de triar quina és la finalitat del tractament de les dades. En aquest cas (situació 1) podem triar per exemple F05.12 Actuacions per promoure la Recerca i la Innovació.

I un cop recollides les dades. Com les haig de tractar?

Les dades personals recollides hauran de conservar-se de manera que es pugui garantir el dret de les persones a l'accés, rectificació, supressió i limitació del tractament de les seves dades. El seu accés ha d'estar limitat a les persones que han de fer ús legítim, i protegit per contrasenya. Es poden fer servir únicament per a la finalitat que s'ha indicat a la seva recollida.

Què faig un cop finalitzat el tractament de dades personals?

Un cop finalitzada l'activitat per la qual s'estaven tracten dades personals, el responsable del tractament estarà obligat a bloquejar les dades. El bloqueig de les dades consisteix en la identificació i reserva d'aquestes, adoptant mesures tècniques i organitzatives, per impedir-ne el tractament, incloent-hi la visualització, excepte per a la posada a disposició de les dades

als jutges i tribunals, el Ministeri Fiscal o les Administracions públiques competents, en particular de les autoritats de protecció de dades, per a l'exigència de possibles responsabilitats derivades del tractament i només pel termini de prescripció d'aquestes. Transcorregut aquest termini s'ha de procedir a la destrucció de les dades.

Situació 2: La meva recerca inclou la recollida de dades mitjançant enquestes o participació en experiments, però no és necessari tractar les dades de les persones que hi participen.

Es recomana fer la recollida de les dades de recerca sense que calgui especificar les dades que permetin identificar les persones. Es poden recollir grups d'edat, i altres dades que permetin mantenir l'anonimat dels participants. En aquest cas, no cal seguir cap procediment especial. Cal recollir sempre les mínimes dades necessàries.

En cas que sigui necessari recollir les dades personals, caldrà seguir un procediment d'anonimització de les dades o de pseudoanonimització abans de procedir al seu tractament. Pots trobar informació i tutorials sobre anonimització en aquest enllaç. La millor pràctica és realitzar l'anonimització en el moment de la recollida. Així no es tracten dades personals ni s'aplica el GDPR.

És necessari recollir un consentiment informat, que consta de dues parts, un full d'informació amb dades sobre el projecte, la seva finalitat i quines implicacions pot tenir per la persona participant, i d'altra banda el full de consentiment. Podeu trobar més informació i models a la web del comitè d'ètica de la UPC.

Un cop anonimitzades les dades, ja no ens resulta d'aplicació el GDPR, però si encara conservem les dades originals, aquestes s'han de tractar amb els protocols de seguretat necessaris.

Situació 3: La meva recerca necessita tractar les dades personals per assolir els seus objectius, però no contenen dades de categoria especial (salut, polítiques, sexuals, ètniques..) o de menors o persones vulnerables.

En aquest cas és imprescindible que abans de fer la recollida de dades, feu la sol·licitud de revisió ètica del projecte al comitè d'ètica de la UPC. Sense la seva validació prèvia, no podreu en cap cas publicar els resultats de la teva recerca. Pots trobar informació addicional sobre el procediment i documentació a presentar al comitè a la seva web.

Situació 4: La meva recerca necessita tractar les dades personals per a assolir els seus objectius, incloent-hi dades de categoria especial (salut, polítiques, sexuals, ètniques..) o de menors o persones vulnerables.

A més de la revisió del projecte per part del comitè d'ètica de la UPC, en aquest cas cal extremar les mesures de seguretat en l'emmagatzematge de les dades. S'hauran de detallar totes les mesures tècniques i organitzatives que es faran servir per a garantir la seva seguretat.

Situació 5: Reutilització de dades personals recollides de projectes de recerca previs o d'accés públic.

En el cas de dades recollides per un altre projecte, s'ha indicar quina va ser la metodologia utilitzada per a la recollida original, el procediment del consentiment informat, i tenir el permís del propietari del dataset.

Si es fan servir dades disponibles públicament, ens hem de donar detalls de la font i assegurar-nos de que es poden fer servir per a recerca. El fet de que les dades siguin públiques, no vol dir que no hi hagi límits per al seu ús. Hem d'assegurar-nos que el nostre ús compleix amb els termes i condicions publicats pel Data Controller.

En cas de dubte, consultar sempre al delegat de protecció de dades de la UPC Data Protection Officer de la UPC.

Finalment, aquest arbre de decisió elaborat per la Comissió Europea t'ajudarà a decidir quins són els protocols necessaris a seguir en cada cas.

5.- Publicació de dades en accés obert.

La ciència oberta inclou com una de les seves premisses, la publicació en accés obert de les dades que validen els resultats de la recerca. D'acord amb el lema fonamental que està sota els principis FAIR "*tan obert com sigui possible i tan tancat com sigui necessari*" la publicació no és obligatòria però sempre cal seguir els principis FAIR. A més, si la recerca ha estat finançada, s'ha de complir amb les obligacions que estableix l'organisme finançador respecte l'accés obert. Consulta la teva biblioteca per veure quin és el repositori més recomanable per al teu cas. Existeixen repositoris multidisciplinars com el CORA.RDR del CSUC o Zenodo i d'altres d'especialitzats.

Per les dades de projectes Horizon Europe és obligatori fer servir una llicència *Creative Commons Attribution International Public Licence* (CC BY) a la seva darrera versió, o una que atorgui drets equivalents.

ANNEXOS

a. Documentació per a la fase de proposta del projecte: Informació a incloure sobre el Pla de Gestió de Dades

b. Model de consentiment Informat (per dades personals): Document a l'emplaçament del Comitè d'Ètica.

c. Més models de consentiments, a l'apartat "Models i Formularis" de l'emplaçament del Comitè d'Ètica.

c. Informació sobre anonimització i pseudo-anonimització de l'Agència Espanyola de Protecció de Dades.