Proposta per establir una política d'accés obert a les dades de recerca a les Universitats de Catalunya

(Doc. 16/30) (B6SR\GDR\1603Politiques_rdm-v8, GTSR 11.07.16)

1. Introducció

1.1 L'accés obert a les publicacions científiques i a les dades de recerca

En els darrers anys, els mecanismes tradicionals pels quals es comunicaven els resultats de la recerca han viscut canvis profunds deguts a l'Accés Obert (OA). Per accés obert s'entén la pràctica de proporcionar un accés en línia a la informació científica de forma gratuïta i reutilitzable per a l'usuari final. El moviment de l'accés obert s'ha desenvolupat sobretot a partir de les declaracions de Budapest (2002) i de Berlín (2003) i s'ha referit fins fa poc als articles publicats a revistes revisades, a '*working papers*' i a comunicacions a congressos.

Arreu del món les entitats promotores de la recerca, les universitats i els centres d'investigació han aprovat polítiques que afavoreixen l'accés obert a les publicacions en qualsevol de les seves opcions (en revistes o en repositoris). Els fonaments de l'accés obert són dos:

- La recerca és majoritàriament finançada amb fons públics i, per això, l'accés als resultats ha de ser públic, i
- La ciència millora com més fàcil sigui accedir i reutilitzar els seus resultats.

El concepte d'accés obert s'ha ampliat darrerament fins incloure també les dades creades o utilitzades pels investigadors durant la recerca. El concepte de dades de recerca és molt ampli, però l'aplicació de l'accés obert s'adreça, principalment, a aquelles en les quals es basen els resultats publicats en articles, *working papers* o ponències.

Posar les dades de recerca en obert implica almenys dues coses:

- Que les dades estiguin dipositades en algun repositori d'accés públic acompanyades de la seva descripció i un identificador permanent.
- Que les dades s'ofereixin en formats reutilitzables i mitjançant llicències que no imposin restriccions legals.

A Europa el programa Horizon 2020 ja fa obligatori dipositar en obert totes les publicacions científiques resultat de projectes que financia. Així mateix, l'Horizon 2020 ha iniciat l'Open Research Data Pilot (ORD Pilot), amb la finalitat de millorar i maximitzar l'accés i reutilització de les dades de recerca generades pels projectes. Actualment, el pilot ORD afecta només alguns àmbits científics, però és voluntari per a la resta. No serà així a partir de l'any 2017, quan s'estendrà a tots els projectes. Aquesta iniciativa pilot no només requereix que les dades de recerca és publiquin en obert sinó que inclou l'obligació de crear plans de gestió de dades per descriure quines dades es generaran o s'utilitzaran, com es gestionaran, quines es publicaran, com se'n facilitar la reutilització i com es conservaran.

Aquest document vol facilitar l'adopció d'estratègies encaminades a establir una política per posar en obert les dades de recerca, sota el supòsit que aviat serà indispensable tenir-ne i que iniciar-les de forma anticipada i conjunta ha de fer més fàcil desenvolupar-les.

1.2 Treballs conjunts de les universitats de Catalunya en suport a la recerca

Les universitats de Catalunya tenen antics i amplis antecedents de treball conjunt: el sistema d'informació UNEIX, els repositoris cooperatius TDX (de tesis) i RECERCAT (d'articles i treballs de recerca), l'acord del CIC per establir polítiques de foment de l'accés obert, la decisió d'usar l'identificador ORCID, el portal de la recerca i acords de bones pràctiques per al compliment mandats d'accés obert de Horizon 2020 i la Llei espanyola de la ciència.

Aquestes actuacions conjuntes no han disminuït la capacitat de cada universitat d'establir polítiques pròpies i, en canvi, han permès establir-ne amb més facilitat. Això ha fet que els resultats en aquests àmbits siguin avui millor a Catalunya que a la resta de l'estat i que estiguin, a nivell europeu, en una bona posició.

Sempre i quan permetin definir a cada universitat els graus i mètodes d'aplicació, els treballs conjunts per establir polítiques de recolzament de l'activitat investigadora tenen els avantatges següents:

- En l'establiment de polítiques, faciliten l'encert, redueixen la feina a fer i fan que la seva adopció sigui més fàcil.
- En la creació d'infraestructures, disminueixen el cost.

1.3 Posicionaments (recomanacions, mandats i exemples)

A nivell internacional, existeixen recomanacions, mandats i universitats pioneres pel que fa les dades de recerca en obert que poden servir de referent a les Universitats de Catalunya, com són:

Recomanacions, directrius internacionals per a les dades de recerca

- *ERAC Opinion on Open Research Data* [en línia]. Brussels : European Research Area and Innovation Committee, February 2016 [consulta: 20 de maig de 2016]. Disponible a: <http://www.earto.eu/fileadmin/content/Website/ERAC_Opinion_on_Open_Research_Data.PDF>.
- Erway, Ricky. *Starting the Conversation: University-wide Research Data Management Policy* [en línia]. Dublin, Ohio : OCLC Research, 2013 [consulta: 20 de maig de 2016]. Disponible a: <http://www.oclc.org/content/dam/research/publications/library/2013/2013-08.pdf>.
- *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020, Version 2.1* [en línia]. European Comission, February 2016 [consulta: 20 de maig de 2016]. Disponible a: <https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf>.
- *LERU roadmap for research data* [en línia]. Leuven : LERU, League of European Research Universities, 2014 [consulta: 20 de maig de 2016]. Disponible a: <http://www.leru.org/files/publications/AP14_LERU_Roadmap_for_Research_data_final.pdf>.
- *Policy RECommendations for Open Access to research Data in Europe (RECODE)* [en línia]. RECODE Project Consortium, 02.02.2015 [consulta: 20 de maig de 2016]. Disponible a: <http://recodeproject.eu/wp-content/uploads/2015/02/RECODE-D5.1-POLICY-RECOMMENDATIONS-_FINAL.pdf>.

Mandats institucionals de dades de recerca:

- *University College of London Research Data Policy* [en línia]. Londres : UCL Research Data and Network Services, 2013 [consulta: 20 de maig de 2016]. Disponible a: <https://www.ucl.ac.uk/isd/services/research-it/documents/uclresearchdatapolicy.pdf>.
- *Policy on Research Data Management* [en línia]. [Nijmegen]: Radboud University, 2013 [consulta: 20 de maig de 2016]. Disponible a: <http://www.ru.nl/research-information-services/institutional-policy/policy-research-data-management/>.

Serveis de suport a la publicació en obert de dades de recerca:

- *Data management* [en línia]. Massachusetts : Massachusetts Institute of Technology [consulta: 20 de maig de 2016]. Disponible a: <http://libraries.mit.edu/data-management/>.
- *Managing research data* [en línia]. Melbourne : Monash University, 29.06.2015 [consulta: 20 de maig de 2016]. Disponible a: <https://www.monash.edu/library/researchdata>.
- *Research data* [en línia]. London : University College of London, 2016 [consulta: 20 de maig de 2016]. Disponible a: <https://www.ucl.ac.uk/research-it-services/about/research-data-service>.
- *Research data: get going* [en línia]. Utrecht : Utrecht University, 04.05.2016 [consulta: 20 de maig de 2016]. Disponible a: <http://libguides.library.uu.nl/data>.
- *Research data management* [en línia]. Edimburgh: University of Edimburgh, 13.08.2015 [consulta: 20 de maig de 2016]. Disponible a: <http://www.ed.ac.uk/information-services/research-support/data-management>.

2. Elements a tenir en compte a l'hora de promoure la gestió i l'Accés Obert de les dades de recerca

A l'hora d'establir una estratègia institucional per a la gestió de dades de la recerca encaminada a promoure'n l'accés obert, el primer que cal fer és identificar els agents implicats dins i fora de la institució. Els agents interns són, principalment, els òrgans de govern, els investigadors i totes les unitats de suport a la recerca de la universitat, incloent-hi la biblioteca i àrea de tecnologies. Externament cal comptar amb altres universitats, amb consorcis com el CSUC, i amb les entitats finançadores de la recerca. Amb tots ells caldrà establir un entorn de diàleg i col·laboració.

Un cop identificats aquests agents, cal avaluar l'estat de la qüestió:

- conèixer les necessitats dels investigadors com a generadors o compiladors de dades, a partir d'enquestes o entrevistes i adreçant-se especialment a aquells investigadors que participen en projecte H2020 o que tenen interès en publicar en obert les dades de recerca.
- determinar els requeriments de les polítiques emergents dels finançadors de la recerca, especialment les d'entitats europees com la Comissió Europea, el Wellcome Trust.
- analitzar la infraestructura de la qual es disposa, tant les infraestructures de les Universitats (servidors on s'allotgen dades) com les externes (Zenodo, Fighshare, Dataverse, etc.).

- detectar els serveis que poden intervenir en el procés de gestió de les dades durant tot el cicle de la recerca (gestors de recerca, biblioteques, informàtics, etc.) per tal de donar suport als investigadors en matèria d'assessorament, de formació, de descripció, de preservació i de publicació de dades.

Posteriorment, cal determinar quina és la infraestructura necessària per donar suport a aquesta estratègia, tenint en compte no només els aspectes tècnics sinó també els de personal.

Finalment cal avaluar els costos d'implementació de la gestió de dades de recerca: el cost de la infraestructura, el cost dels serveis de suport, la formació al personal de suport i als investigadors i les campanyes de difusió i sensibilització.

Tota aquesta anàlisi ha d'anar encaminada a establir una política institucional.

3. Proposta de redactat per una política de gestió de les dades de recerca a les universitats catalanes

Les universitats catalanes adoptaran una política de gestió de dades de recerca que inclogui els aspectes següents:

Accés Obert per defecte: accés obert de les dades de recerca com a requeriment estàndard, proporcionant el suport adequat. La política ha de ser d'obligat compliment amb possibilitat de tancar dades en casos en què es puguin demostrar temes ètics, legals, de confidencialitat, de seguretat, etc. La política demanarà que les dades de recerca tinguin assignades llicències que permetin la seva reutilització.

Responsabilitats: definició de la responsabilitat de tots els agents implicats dins la institució: els òrgans de govern, els investigadors i tot el personal de suport a la gestió i publicació de les dades de recerca. Els investigadors gestionaran les seves dades de recerca seguint estàndards i procediments específics i la institució assumirà la responsabilitat d'oferir els serveis necessaris per a què ho puguin fer (infraestructura, assessorament, formació, etc.).

Lloc de dipòsit: determinació del lloc de dipòsit de les dades, podent ser el repositori institucional o altres repositoris de confiança (llista de repositoris o de criteris que els investigadors puguin utilitzar per triar un repositori adequat), en el cas que no es disposi d'una infraestructura adient.

Termini de dipòsit: publicació de les dades de recerca en obert tan aviat com sigui possible. La política contemplarà períodes de temps raonables entre l'obtenció de les dades i la seva publicació en obert. Per exemple, les dades associades a una publicació hauran de ser accessibles en el mateix moment de la publicació, o un cop finalitzat el projecte de recerca.

Pla de gestió de dades: elaboració d'un pla de gestió de dades com a requeriment estàndard, en les activitats de recerca de la institució. Aquest requeriment es podrà implementar de manera progressiva.

Costos: establiment de mecanismes oportuns per repercutir els costos de la infraestructura de la gestió de les dades. Per exemple, garantint un espai estàndard sense cost a tots els investigadors i establint tarifes per a requeriments específics.

Preservació i conservació: compromís de la institució per preservar i conservar les dades dipositades durant un període de temps determinat (10-20 anys, per exemple), establint mecanismes per garantir-ne l'accés i la seva integritat.

Acompanyament i seguiment: compromís de la institució de dotar d'un servei d'assessorament i suport als investigadors en la gestió de dades de recerca i l'establiment d'un pla de formació.

La política també ha d'incloure mecanismes de seguiment i d'avaluació de la seva eficàcia i possibles incentius per als investigadors.

4. Proposta de pla d'acció per a les universitats de Catalunya per impulsar una política d'accés obert a les dades de recerca

L'Horizon 2020 incorpora el concepte de 'Ciència Oberta' com a estratègia per millorar la disseminació del coneixement i, per tant, la innovació. Malgrat que la idea de l'Accés Obert a la recerca científica està consolidada entre la comunitat científica, el nivell de solucions pràctiques per fer-ho efectiu varia segons a l'àmbit al que s'apliqui. Pel que fa a posar les dades de recerca en obert, en aquests moments ho estan fent els països millors en recerca, i, per tant, sembla raonable iniciar-lo també a Catalunya. Per una altra, cal dir que hi ha encara poques pràctiques ben establertes, i que, per tant, cal fer aquests primers passos amb prudència.

A continuació es proposen uns acords per establir de forma conjunta entre les universitats de Catalunya un full de ruta en relació les dades de recerca. Aquest tindria un marc temporal curt, al final del qual caldria avaluar els avanços assolits i establir noves fites o mesures.

La proposta es composa de les següents mesures:

- Les universitats es comprometen a identificar els elements necessaris per iniciar aquest suport mencionat anteriorment i a, efectivament, iniciar-lo, donant prioritat als projectes de recerca amb finançament del programa Horizon 2020 (i en especial en els inclosos al programa pilot).
- Les universitats es comprometen a establir els serveis i els recursos necessaris per donar un suport institucional de gestió de les dades de recerca. Això suposarà estudiar les alternatives existents per dipositar les dades i fer una proposta de la o de les més convenients, la qual cosa pot suposar la creació d'una infraestructura conjunta. Pel que fa al repositori on dipositar les dades, de forma provisional, els serveis de suport a la gestió de dades de recerca, recomanaran en cada cas el repositori més adient d'entre els existents.
- En el termini d'un any, les universitats es comprometen a adoptar una política interna de gestió de dades de recerca que inclogui els elements mencionats anteriorment (accés obert per defecte, responsabilitats, lloc i termini de dipòsit, pla de gestió de dades, costos, preservació i conservació i acompanyament i seguiment).
- Aquests acords seran revisats al cap d'un any de la seva aprovació amb la intenció de modificar-los, concretar-los o ampliar-los en funció de l'experiència obtinguda.