

Publicat a *blok de bid* (<http://www.ub.edu/blokdebid>)

[Inici](#) > [Allau de dades](#)

Allau de dades

Enviat per [Isabel Bernal](#) ^[1] el Dc, 08/01/2014 - 11:52 [Blok de BiD - cat](#) ^[2] [dades de recerca](#) ^[3] [Isabel Bernal](#) ^[4] [polítiques](#) ^[5]

Isabel Bernal

**Unidad de Recursos de Información Científica para la Investigación
CSIC**

Wilma Mossink, Magchiel Bijsterbosch, and Joeri Nortier. *European Landscape Study of Research Data Management*. Utrecht: SURF, 2013.

<http://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20vs1%204%2014.08.13.pdf> ^[6]



SIM4RDM (*Support Infrastructure Models for Research Data Management*, <http://www.sim4rdm.eu/> ^[7]) és un projecte europeu iniciat a l'octubre de 2011 i finançat pel 7è Programa Marc de la Comissió Europea que s'enquadra en el creixent nombre d'iniciatives que, davant l'allau de dades científiques, promouen bones pràctiques i estàndards en la seva gestió per garantir el progrés científic i econòmic mitjançant el seu accés, reutilització i preservació digital. En concret, SIM4RDM es marca com a objectiu facilitar a la comunitat científica l'ús efectiu d'infraestructures emergents per a dades de recerca i parteix de la consideració prèvia del relatiu desconeixement pels investigadors de metodologies, instruments i capacitats necessàries. Per a això, el consorci SIM4RDM es beneficia d'experiències acumulades per les organitzacions participants en el projecte, entre les quals destaquen JISC (Regne Unit) i SURF (Holanda).

La pedra angular de l'informe *European Landscape Study of Research Data Management*, publicat al maig de 2013, gira al voltant dels resultats -desiguals en la seva cobertura geogràfica i representativitat de les respostes- dels qüestionaris enviats a 4 grans grups involucrats en la gestió de dades científiques a Europa (agències de finançament, organismes nacionals, institucions de recerca i editors) per definir les seves accions perquè la comunitat científica millori el maneig de les seves dades un cop *finalitzats* els seus projectes de recerca. L'informe es completa amb una introducció força exhaustiva que recorre, a partir de bones pràctiques existents, els elements que han de formar part de les polítiques i els plans de gestió de dades, i amb un parell de seccions annexes a l'anàlisi de les respostes als qüestionaris: d'una banda, entrevistes a un grup d'investigadors sobre les seves necessitats i un paquet de recomanacions generals i específiques per a cada grup enquestat.

La secció introductòria ofereix una molt bona panoràmica de les claus per dissenyar polítiques i plans de gestió de dades a partir de l'estudi de diversos documents de referència, com l'informe *Riding the Wave* (2010), i la seva continuació *Surfboard for Riding the Wave* (2011), els Principis i Guia de l'OCDE, i diversos exemples concrets de polítiques i plans de gestió, principalment del món anglosaxó, com són els de la *National Science Foundation* (2011) la *US National Academy of Sciences* (1995) als Estats Units, les recomanacions de l'informe de *JISC Dealing with Data* (2007), els principis de polítiques de dades dels *Research Councils* al Regne Unit (2012) i el pla de gestió de *Australian National University* (2010). Entre les moltes consideracions, podem destacar les relatives als incentius, a la formació i al suport per als investigadors, la inclusió de plans de gestió de dades en la preparació de propostes, el desenvolupament d'infraestructures i de codis de conducta, la creació d' perfils professionals específics, la definició de responsabilitats entre les parts implicades (autors de dades, agències finançadores i institucions científiques) i les polítiques de preservació digital i d'utilització de les dades. Una part especialment interessant en la introducció es refereix a estudis sobre pràctiques en la gestió de dades en disciplines científiques concretes (ex. EUDAT) i sobre les possibles reticències i mancances de la comunitat científica per compartir les seves dades (ex. el projecte PARSE).

El gruix de l'informe està compost per l'anàlisi de les respostes de les enquestes sobre polítiques de gestió de dades i de les entrevistes a investigadors preparades per l'equip de SIM4RDM. Malgrat la seva vocació europea, els resultats dels qüestionaris són relativament desiguals i en alguns casos -especialment, l'enquesta a editors- han impedit oferir conclusions prou argumentades. No obstant això, l'exercici en conjunt reflecteix tendències generals en matèria de planificació de gestió i reutilització de dades a diferents nivells. Les enquestes amb el major nombre i varietat de respostes són les dirigides a agències finançadores i a organismes d'investigació. Pel que fa a les primeres, cal destacar el creixent interès i la posada en marxa de polítiques per al finançament i la gestió de dades (gairebé un 50 % dels enquestats), tot i que la xifra cau pel que fa a l'obligatorietat i l'avaluació de plans de gestió de dades en les propostes de projectes. En general, s'evidencia que la garantia de preservació digital és una altra àrea relativament poc desenvolupada per agències finançadores. D'altra banda, l'anàlisi dels resultats de l'enquesta sobre polítiques institucionals és la més rica en qüestions i en cobertura geogràfica (participació de 15 països) i reflecteix que més d'un terç de les institucions enquestades disposen d'una política, si bé sol ser de molt recent creació. Un altre aspecte a ressaltar és la precarietat de la majoria de les moltes iniciatives en marxa per produir eines, recursos, guies i suport per millorar la gestió de les dades, i en aquest sentit potser es pugui establir una relació amb el baix nombre de respostes recollides (fonamentalment de Finlàndia, Holanda i Regne Unit) en l'enquesta a organismes nacionals encarregats de crear codis de conducta i de coordinar activitats -per a l'accés, la formació, la preservació i creació d'infraestructures de dades i el suport a agències finançadores.

Entre les principals motivacions de les institucions de recerca per desenvolupar plans d'acció

destaquen possibles requeriments legals d'agències financeres o institucionals (sobretot pel que fa al dipòsit, gestió i preservació de les dades). L'exhaustivitat dels plans de gestió varia considerablement entre les institucions enquestades encara que en major o menor mesura en la meitat dels casos solen incloure línies d'acció per a la captura, manteniment i retenció de dades, l'aplicació d'estàndards tècnics i bibliogràfics, i qüestions de confidencialitat i de seguretat i de llicències d'ús. D'altra banda, el suport de les institucions a la reutilització de les dades sol canalitzar mitjançant la creació de repositoris, però els sistemes d'incentius perquè els investigadors dipositin les seves dades continuen sent minoritaris, amb només un quart de les institucions enquestades oferint serveis de suport com la inclusió de les dades en les llistes de publicacions institucionals, la gestió de dades en l'avaluació i auditories de projectes, l'assignació de DOIs, la traducció de les metadades a l'anglès perquè tingui més visibilitat, el finançament parcial per aplicar formats estàndards etcètera. Les parts menys detallades de l'informe corresponen a l'enquesta als editors i les entrevistes a investigadors encara es deixen entreveure qüestions de rellevància, com la relativa manca de polítiques editorials per al dipòsit i l'accés a les dades i la manca de consens en la citació de les dades i la seva integració amb els articles corresponents. Per la seva banda, les 12 entrevistes a investigadors deixen en evidència les qüestions que ells consideren prioritàries a l'hora de gestionar les seves dades i les necessitats més urgents. En aquest sentit, cal esmentar l'existència d'infraestructures de dipòsit fàcils d'usar, el reconeixement de l'autoria de les dades, el suport institucional, els incentius per al seu dipòsit i formació.

Finalment, l'informe tanca amb una sèrie de conclusions generals i recomanacions perquè els organismes nacionals, agències financeres, institucions de recerca i editors promoguin una millor gestió de dades per part de la comunitat científica. El treball futur de SIM4RDM (en vigor fins a la primavera de 2014) preveu la inclusió de més stakeholders en l'àmbit de les dades científiques, com ara més òrgans nacionals i organitzacions internacionals, centres de dades i proveïdors d'infraestructures (biblioteques), comitès editorials de revistes científiques i societats acadèmiques.

Informes com aquest són una prova de l'interès que desperta aquest tema, i estan encoratjats en gran mesura per l'auge de l'accés obert, la multiplicació de mandats sobre la difusió dels resultats d'investigació i pels desenvolupaments tecnològics d'última hora. L'impuls de les agències financeres és clau per canviar de manera fonamental les pràctiques en la gestió i la comunicació d'aquests resultats de recerca i les *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* ^[8], acabades de publicar a meitat de desembre, tindran sens dubte un efecte positiu en el panorama europeu. No obstant això, hi ha diversos interrogants entorn de la gestió d'aquest tros molt considerable de ciència, començant per la necessitat d'abordar des d'un consens internacional una definició estàndard del que entenem per "dades d'investigació" i no repetir els errors derivats de la relativa manca d'harmonització entre els múltiples models de mandats d'accés obert a les publicacions científiques -en aquest sentit, l'aplicació d'estàndards internacionals i la identificació d'excepcions són importants. D'altra banda, la creixent facilitat tecnològica per recuperar informació sobre el seu impacte i el trànsit web que generen, per analitzar mitjançant tècniques de mineria i les perspectives de integrar-los en exercicis d'avaluació ha obert la porta a noves iniciatives amb una vocació fortament comercial (bases de dades bibliogràfiques, repositoris de pagament, serveis d'allotjament massiu, etc.) que podrien tenir un efecte no desitjat sobre l'objectiu final de descobriment, obertura i reutilització en nom del progrés científic i socioeconòmic. Altres qüestions que queden per a treball futur es refereixen a la sempre delicada qüestió de la gestió i preservació de la integritat de les dades, la protecció de privacitat i les dificultats d'atribuir adequadament aquestes obres, de forta empremta col·lectiva en moltes disciplines, als seus corresponents autors.

URL d'origen: <http://www.ub.edu/blokdebid/ca/node/441>

Enllaços:

- [1] <http://www.ub.edu/blokdebid/ca/users/isabel-bernal>
- [2] <http://www.ub.edu/blokdebid/ca/category/butllet%C3%AD/blok-de-bid-cat>
- [3] <http://www.ub.edu/blokdebid/ca/category/etiquetes/dades-de-recerca>
- [4] <http://www.ub.edu/blokdebid/ca/category/ressenyadors/isabel-bernal>
- [5] <http://www.ub.edu/blokdebid/ca/category/etiquetes/politiques>
- [6] http://www.sim4rdm.eu/sites/default/files/uploads/documents/SIM4RDM%20landscape%20report%20vs1%204_14.08.13.pdf
- [7] <http://www.sim4rdm.eu/>
- [8] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf