

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

[Inici](#) > La conservació de dades: l'estat actual i el futur més proper

La conservació de dades: l'estat actual i el futur més proper

Enviat per [Alexandre López...](#) [1] el Dc, 04/12/2013 - 17:58 [Alexandre López-Borrull](#) [2]
[Blok de BiD - cat](#) [3] [conservació de dades](#) [4] [data curation](#) [5] [openAIRE](#) [6]
[preservació digital](#) [7]

Alexandre López-Borrull

Estudis de Ciències de la Informació i de la Comunicació [professor]

Universitat Oberta de Catalunya (UOC)

National Information Standards Srganisation (NISO) (2013). *Information Standards Quarterly*. EEUU, MA: NISO. Vol. 25. Núm. 3. 44 p. [Consulta: 2/12/2013]. Disponible a: <http://www.niso.org/apps/group_public/download.php/11593/isqv25no3.pdf [8]>. ISSN 1041-0031.



El darrer número que la [National Information Standards Organization](#) [9] (NISO) ha publicat està centrat en la [Data Curation](#) [10], que podem traduir com a *Conservació de dades*¹ i que pren una actualitat molt gran atesa la tendència i el debat obert en la comunitat científica per tal de poder preservar, enllaçat a ser possible amb els articles científics, les dades de la investigació per a la seva posterior reutilització. Per a abordar la temàtica, a banda de l'editorial hi ha quatre articles que fan una aproximació a

la temàtica en diversos àmbits (Química, Arqueologia, Ciències de la Terra) i l'experiència i el coneixement generat a [OpenAIRE](#) [11].

El document comença amb una editorial escrita per Sarah Callaghan, on a la vegada que presenta el número tot resumint cada una de les aportacions, fa una defensa de la necessitat d'estàndards que permetin adequadament la conservació digital. Així, considerant les dades de projectes de recerca, publicacions i literatura gris, metodologies i resultats de laboratori. Tantes dades, defensa Callaghan tan sols poden ser conservades si el procés s'automatitza, i això només s'aconsegueix si hi ha estructures i ontologies estandarditzades.

El primer dels articles, *Data Curation: Issues in the Chemical Sciences*, escrit per Colin B.

Bird, Cerys Willoughby, Simon J. Coles i Jeremy G. Frey, comença tot fent palès que la Ciència és de fet, fortament dependent de la preservació i l'addició de valor a la recerca feta anteriorment, i això inclou les dades, tant les dades en brut (*raw data* [12]) com les derivades. Per tant, defensen, la Ciència, és fortament dependent de la conservació. Al llarg de l'article fan esment a diversos documents que fan referència a la temàtica, algun dels quals escrit pels propis autors.

Dins la conservació de dades, per als autors és bàsic com a ingredient les metadades, i sobretot des de la vessant no només de la recuperació, l'intercanvi o la identificació, sinó sobretot de la captura del context, que identifiquen com una necessitat bàsica en el cicle de recerca. En aquest sentit, diuen, el repte per als químics serà comprendre i assumir la importància del context i la captura apropiada i amb suficient descripció (metadades) en el temps adequat. Des del meu punt de vista, és destacable el rol que reserven als bibliotecaris i professionals de la informació en l'assistència en el procés, tant directament com assistint en la formació dels químics per a saber conservar ells mateixos els seus resultats.

Considerant la conservació a la pràctica, els autors esmenten diversa bibliografia que dona suport a la idea que encara recentment, els químics: 1. empen documents en paper (la llibreta de laboratori) com a document de treball habitual; 2. empen molt programari propietari; 3. en la recerca en col·laboració són dependents d'intercanvi de correus persona a persona. El repte, doncs, és el canvi d'aquestes formes de treballar que semblen incompatibles amb l'adequada conservació digital. Això, afirmen, no succeirà evidentment de la nit al dia, i caldrà formació i promoció. I el paper, novament, dels professionals de la informació serà central. Finalment, considerem addient esmentar com en diverses part de l'article es fa èmfasi dels punts crítics associat a el cost de la conservació (traducció lliure de *burden of curation*) en la doble vessant de càrrega i cost.

Quant al segon article, *Data Curation in the OpenAIRE Scholarly Communication Infrastructure*, escrit per Jorchen Schirwagen, Paolo Manghi, Natalia Manola, Lukasz Bolikowski, Najla Rettberg i Birgit Schmidt, fan una presentació de la iniciativa *OpenAire* (*Open Access Infrastructure for Research in Europe*), pensada per a la gestió de publicacions millorades (traducció per a *enhanced publication*), és a dir, les publicacions en comunicació acadèmic que venen acompanyades de diversa informació (projectes de finançament, fonts de dades, citacions, publicacions similars, mètriques) que li confereixen una major profunditat així com la capacitat de ser millor entesa i reutilitzada.

Així, es presenta la tasca duta a terme dins el projecte, i dona algunes de les ingents estadístiques, com ara els vuit milions de registres bibliogràfics. A la vegada, però també esmenta els principals reptes amb els quals ha de bregar, com són la diversitat de dades científiques i la diversitat d'aproximacions culturals. Sens dubte, allò que ha permès la creació de l'infraestructura, afirmen, ha estat la capacitat de poder oferir unes directrius clares per a la interoperabilitat i la posterior conservació de les dades. Tot descrivint el seu funcionament, els autors forneixen una via clara per a una futura infraestructura més ambiciosa.

Ray Moore i Tim Evans, per la seva banda, en el seu article *Preserving the Grey Literature Explosion: PDF/A and the Digital Archive*, en el qual s'inclou, en primer lloc, una presentació de l'*Archaeological Data Service (ADS)* [13], creat el 1996 amb la intenció de proporcionar suport per a la preservació de la documentació en Arqueologia, però també de la creació i guiatge de bones pràctiques per a fer-ho possible. En una disciplina amb una gran quantitat de literatura grisa, també s'encarrega de la seva preservació i posada a disposició per a la reutilització.

En el mateix article es descriu com el format PDF (*Portable Document Format*) esdevé un estàndard i l'ús que en fan en el seu projecte del *PDF/A (archival)* [14] com a format de

preservació en les seves diverses versions, PDF/A-1, A-2 i A-3, els avantatges que els ha donat i les diverses dificultats per al futur que consideren.

Finalment, el darrer article, *Ensuring the Long Term Impact of Earth Science Data through Data Curation and Preservation* escrit per Esther Conway, Sam Sepler, Wendy Garland, David Hooper, Fulvio Marelli, Luca Liberti, Emanuela Piervitali, Katrin Molch, Helen Glaves i Lucio Badiali, explica els avantatges de la conservació i preservació de dades en les Ciències de la Terra. Aquestes provenen tant de la capacitat de poder estudiar les dades com de poder predir comportaments.

Així, de la seva experiència dins la Science Data Infrastructure for Preservation (SCIDIP-ES) esmenten els guanys en el coneixement en desastres naturals, temes de salut, energia, clima, aigua, ecosistemes i agricultura, tot contemplant el retorn en la inversió que ha suposat, la innovació en la interoperabilitat, la compartició de despeses i com han evitat la duplicació de recursos i l'augment de la responsabilitat social i ambiental.

Per tant, es tracta d'un número monogràfic per a comprovar les potencialitats i els reptes que suposa la conservació de dades, en un moment de molta actualitat per quant es tracta d'una temàtica molt present avui dia a la literatura. Recomanat també per gestors de repositoris i experts en preservació digital.

¹Mentre no existeixi una terminologia acceptada pel TERMCAT, emprarem Conservació de dades per a *Data Curation* i Conservació Digital per a *Digital Curation*.

URL d'origen: <http://www.ub.edu/blokdebid/ca/content/la-conservacio-de-dades-lestat-actual-i-el-futur-mes-proper>

Enllaços:

- [1] <http://www.ub.edu/blokdebid/ca/users/alexandre-lopez-borrull>
- [2] <http://www.ub.edu/blokdebid/ca/category/ressenyadors/alexandre-lopez-borrull>
- [3] <http://www.ub.edu/blokdebid/ca/category/butllet%C3%AD/blok-de-bid-cat>
- [4] <http://www.ub.edu/blokdebid/ca/category/etiquetes/conservacio-de-dades>
- [5] <http://www.ub.edu/blokdebid/ca/category/etiquetes/data-curation>
- [6] <http://www.ub.edu/blokdebid/ca/category/etiquetes/openaire>
- [7] <http://www.ub.edu/blokdebid/ca/category/etiquetes/preservaci%C3%B3-digital>
- [8] http://www.niso.org/apps/group_public/download.php/11593/isqv25no3.pdf
- [9] <http://www.niso.org/home/>
- [10] http://en.wikipedia.org/wiki/Data_curation
- [11] <https://www.openaire.eu/>
- [12] http://en.wikipedia.org/wiki/Raw_data
- [13] http://en.wikipedia.org/wiki/Archaeology_Data_Service
- [14] <http://es.wikipedia.org/wiki/PDF/A>