

# Item

## Editorial

### Dades: oportunitats i nous escenaris

Fa més d'una dècada que la nostra professió viu sota l'embat periòdic de *noves* oportunitats laborals en forma de *nous* perfils professionals: arquitecte de continguts, expert en usabilitat, *webmaster*, gestor del coneixement... i altres més recents com *content curator* i *community manager*. Perfils aparentment allunyats de la biblioteca per a una professió que, fa aproximadament cinc anys, va renunciar fins i tot a la seva denominació tradicional, almenys en la seva etapa formativa, amb l'aparició dels estudis d'Informació i Documentació.

En altres ocasions, les oportunitats no tenen com a origen un nou perfil professional, sinó una nova tendència o realitat. Aquestes oportunitats no ho són només per als professionals, sinó també per a les mateixes biblioteques, per tal de recuperar un espai rellevant en les institucions o administracions de les quals depenen. Alguns exemples que s'han anat produint podrien ser la digitalització, l'emmagatzematge dels objectes digitals en repositoris, la seva gestió i l'obertura en forma d'*open data*, etc.

Aquest número de l'*Item* està més centrat en aquest segon tipus d'oportunitat. No es parla tant de nous perfils com d'un **nou centre d'atenció i possible escenari de noves oportunitats professionals**: les dades.

Gestionar dades o *amb* dades ni és nou ni aliè a les tasques del bibliotecari-documentalista. No obstant això, com ens diuen en Genís Roca i l'Óscar Marín a les seves entrevistes, s'està produint un nou fenomen, la producció d'un nou flux de dades provinent de la nostra expansiva participació a les xarxes socials o fins i tot amb un origen molt més inanimat, com les dades que pot recollir i facilitar un telèfon mòbil, un terminal de lectura de targetes de crèdit o les mateixes ciutats convertides en «intel·ligents» recol·lectores de dades. I aquest fenomen és ja conegut com a *big data*.

Per altra banda, l'Ignasi Labastida i en Llorenç Arguimbau ens parlen de les dades de recerca, més nombroses i obertes que mai, com un escenari ple d'incògnites i a la vegada d'oportunitats, especialment per a les biblioteques de recerca, universitàries i centres de recursos.

Però també coneixerem nous perfils, ja que, com ens expliquen la Karma Peiró i en Javier Guallar, l'aplicació de conceptes com *big data* o *open data* en un sector com el periodisme ha facilitat l'aparició del periodista de dades, una nova pràctica professional que els bibliotecaris-documentalistes haurien, com a mínim, de seguir de prop. I l'Óscar Marín, per la seva banda, ens explicarà què són un enginyer de dades i un científic de dades.

Queden per una altra ocasió altres aspectes relacionats i no menys rellevants, com la visualització de dades (tant de bo puguem veure durant els propers anys professionals de la informació i la documentació treballant en aquest sector o especialitzant-se en alguns dels estudis de màster o postgrau sobre aquesta temàtica) o el *linked open data*, que reclama ja un full de ruta i projectes d'envergadura a Catalunya, com els que podem veure en d'altres comunitats i països.

Si és cert que els canvis propicien oportunitats, no ho és menys que les oportunitats s'han d'aprofitar i sovint requereixen també adaptació i millora de les pròpies habilitats com a professionals.

Quan la feina del documentalista perd l'aixopluc de la biblioteca, l'administració pública o la universitat, i s'entra en un terreny compartit amb altres professionals, sembla necessari mantenir un esperit obert a la recerca d'eines, estàndards i pràctiques professionals que poden ser diferents dels emprats habitualment, i que poden facilitar la incorporació dels professionals de la Informació i la Documentació a l'àmbit privat i empresarial.

Això comporta una especial atenció a la pròpia formació tècnica, que contempli especialitzacions resultants de la combinació dels estudis d'Informació i Documentació amb d'altres disciplines (comunicació, periodisme, informàtica, estadística, visualització, etc.).

La qual cosa inclou, també, l'assistència i participació en jornades i congressos d'àmbits no sempre directament relacionats amb la nostra professió. Només sobre el tema

que ara ens ocupa, hem tingut l'oportunitat de participar recentment en esdeveniments sobre periodisme de dades, visualització, *big data*, etc., i com a col·lectiu que reclama presència en tots els àmbits laborals relacionats amb la Informació i Documentació, no ens podem permetre estar al marge.

Aquest no serà el darrer *nou* escenari que es posarà de moda, i molt probablement continuaran apareixent *nous* perfils que ens semblaran inspirats en les nostres pràctiques i coneixements. Que les dades, la informació i la comunicació siguin valors a l'alça, en tot cas sembla una bona notícia.

# Introducció al periodisme de dades. Característiques, panoràmica i exemples

## **Karma Peiró**

Periodista digital freelance i professora de periodisme a la URL  
*karmapeiro@mac.com*

## **Javier Guallar**

Professor de documentació i comunicació a la UB, URL i UOC  
*jguallar@gmail.com*

### **Resum:**

Introducció al periodisme de dades, que es pot considerar una de les tendències més rellevants en el periodisme en els últims anys. Es mostra una proposta de definició i de tipologia de productes de periodisme de dades, així com la seva relació amb altres termes com, entre d'altres, periodisme de precisió, documentació periodística, infografia, *open data* i *big data*. Es traça una panoràmica sobre la seva situació actual a nivell nacional i internacional amb una variada selecció d'exemples destacats.

### **Paraules clau:**

periodisme, dades, periodisme de dades, periodisme de precisió, documentació periodística, infografia, dades obertes, grans dades, professió, periodistes, informàtics, infografistes, documentalistes

## Introducción al periodismo de datos. Características, panorámica y ejemplos

### Resumen:

Introducción al periodismo de datos, que se puede considerar una de las tendencias más relevantes en el periodismo en los últimos años. Se muestra una propuesta de definición y de tipología de productos de periodismo de datos, así como su relación con otros términos como, entre otros, periodismo de precisión, documentación periodística, infografía, *open data* y *big data*. Se traza una panorámica sobre su situación actual a nivel nacional e internacional con una variada selección de ejemplos destacados.

### Palabras clave:

periodismo, datos, periodismo de datos, periodismo de precisión, documentación periodística, infografía, datos abiertos, grandes datos, profesión, periodistas, informáticos, infografistas, documentalistas

## Introduction to data journalism. Features, overview and examples

### Abstract:

Introduction to data journalism, which can be considered as one of the most important trends in journalism in recent years. The article proposes a definition and a typology of data journalism products, as well as shows its relationship to other terms such as, among others, precision journalism, journalistic documentation, graphics, open data and big data. It draws an overview of its current situation nationally and internationally with a wide selection of outstanding examples.

### Keywords:

journalism, data, data journalism, precision journalism, journalistic documentation, graphics, open data, big data, profession, journalists, computer programmers, graphic designers, news librarians

## 1. Introducció

En els darrers anys sona cada cop més en els mitjans de comunicació i entre els periodistes i altres professionals de la informació l'expressió «data journalism» o «periodisme de dades», i probablement es pugui considerar la tendència més rellevant en el sector dels mitjans digitals en els anys 2011-12<sup>1</sup>.

Malgrat que l'així anomenat «periodisme de dades» s'ha estès considerablement en poc temps a països com els EUA (per exemple, a capçaleres com *The New York Times* o *Los Angeles Times*), el Regne Unit (amb el cas emblemàtic de *The Guardian*) o diversos països d'Amèrica Llatina (els diaris *La Nación* a Argentina i a Costa Rica són un bon exemple), el moviment ha tingut fins ara escassa presència en els mitjans catalans i espanyols.

Però el 2013 pot ser l'any del seu definitiu enlairament a Espanya. Alguns indicadors d'aquesta afirmació poden ser l'aparició del primer programa de televisió que assumeix plenament aquesta tendència, *El objetivo*, de la periodista Ana Pastor, o l'èxit de participació de les I Jornades de Periodisme de Dades i Open Data (#13JPD)<sup>2</sup>, celebrades a Barcelona i Madrid els dies 24-26 de maig del 2013 i organitzades per les periodistes Mar Cabra i Karma Peiró (coautora d'aquest article) amb el suport de l'Open Knowledge Foundation.

Però tot i que l'expressió «periodisme de dades» està en boca de molts professionals i ha començat a arribar al gran públic, hi ha molts dubtes, desconeixement i imprecisions sobre el seu abast i significació.

«Vivim en un món digital, on gairebé qualsevol fet pot ser descrit amb números: la carrera professional, 300.000 documents confidencials, les persones que formen el teu cercle d'amics; fotos, vídeo i àudio; assassinats, malalties,



©istockphoto/DeshaCAM

vots polítics, corrupció i mentides, etc.», es diu al manual *The Data Journalism Handbook*<sup>3</sup>, escrit de manera col·laborativa per periodistes de diversos països, entusiastes d'aquesta especialitat.

El periodisme de dades es presenta en aquests moments per als periodistes com una gran oportunitat per injectar qualitat a la professió, perduda en la majoria de mitjans a causa de l'acceleració de la informació i de la retallada de personal a les redaccions a conseqüència de la crisi econòmica.

D'altra banda, des del sector professional de la documentació, al qual es dirigeix particularment aquesta revista, s'ha generat també força interès per l'oportunitat professional que pugui representar, alhora que inseguretat per l'abast de la pròpia participació. Des d'aquest àmbit sovint es pregunta: què és periodisme de dades?, té relació amb la documentació periodística de tota la vida actualitzada?, hi tenen un paper a jugar els documentalistes?

1. Guallar, 2013.

2. I Jornades de Periodisme de Dades i Open Data #JPD13 ([periodismodatos.okfn.es](http://periodismodatos.okfn.es)).

3. Gray, *et al.*, 2012.

## 2. Definició i característiques

### 2.1. Definició

Segons Flores i Salinas<sup>4</sup>, el periodisme de dades «consiste en usar herramientas estadísticas y de visualización para contar mejor las viejas historias y descubrir nuevas historias que contar».

Nosaltres proposem la següent definició:

*Periodisme de dades és l'especialitat o pràctica del periodisme que utilitza com a matèria primera i fonamental dades (i no altres fonts d'informació), per tal de crear productes informatius diversos, com ara: articles amb dades, infografies o visualitzacions interactives amb dades i conjunts de dades en obert.*

L'èmfasi de la definició està –clarament– en el terme «dades». Aquestes són la font i l'origen de la informació que es publica i són també les protagonistes destacades del producte informatiu que s'ofereix al lector. La pròpia varietat d'aquests productes és en si una altra característica del fenomen.

### 2.2. Relació amb altres especialitats

Per completar la significació del concepte «periodisme de dades» és interessant també veure la seva vinculació amb altres especialitats o pràctiques, com són: el periodisme d'investigació, el periodisme de precisió, la documentació periodística, la infografia, el *fact checking* o els moviments *open data* i *big data*. Amb tots aquests conceptes hi té relació i segurament qualsevol visió del periodisme de dades quedaria incompleta sense contemplar aquest conjunt de relacions.

**Periodisme d'investigació.** Per a algunes veus el periodisme de dades és bàsicament un tipus de de «periodis-

me d'investigació» en el qual s'ha posat l'èmfasi en l'ús de tècniques d'extracció i anàlisi de dades.

**Periodisme de precisió.** Hi ha molts autors, particularment en l'àmbit acadèmic, que consideren que el periodisme de dades és l'evolució de l'anomenat «periodisme de precisió», especialitat periodística que va tenir una important rellevància els anys setanta i vuitanta del segle passat. El seu pare i fundador va ser el periodista Philip Meyer, autor de *Precision Journalism*, que investigava amb dades estadístiques, i pretenia l'aplicació dels mètodes de recerca científics –quantitatius, afegiríem– a la pràctica del periodisme. Sobre periodisme de precisió hi ha una abundant bibliografia i és per tant una sòlida base sobre la qual es pot fonamentar des del punt de vista acadèmic l'actual periodisme de dades, tot i que sigui encara aviat per poder-ho establir així.

**Documentació periodística.** La documentació aplicada al sector dels mitjans de comunicació es coneix habitualment com a documentació periodística o documentació informativa. Com a professió ha tingut fins ara una important presència a Espanya, especialment en el cas de les empreses periodístiques grans i mitjanes, que solen comptar amb professionals especialitzats en documentar les notícies que es publiquen, localitzar tot tipus d'informació (especialment retrospectiva) i gestionar sistemes documentals. La localització, gestió i publicació de dades formen part, per tant, de les funcions associades tradicionalment a la documentació periodística. El periodisme de dades, segons aquesta visió, hi tindria molta relació.

**Infografia.** Els professionals de la infografia que treballen en els mitjans estan participant generalment de manera molt activa i entusiasta en el moviment. Cal remarcar, però, que infografia no és igual a periodisme de dades; tot i que sovint la presència important de productes de periodisme de dades que són visualitzacions interactives pugui fer confondre una cosa amb l'altra. Per tant, sí que hi ha infografies que són excel·lents exemples de periodisme

4. Flores; Salinas, 2012: 7; també citat per Chaparro, 2012: 5.



©istockphoto/Stepan Popov

de dades, però aquestes no són sinó un dels diversos productes que es poden dur a terme periodísticament amb dades, com veurem més endavant.

**Fact checking.** Una especialitat concreta del periodisme relacionada amb l'ús de dades és el *fact checking* o comprovació de fets o dades. Molt poc conegut al nostre país (tot just es comencen a veure els primers exemples en aquest any 2013), ja fa uns anys que triomfa als EUA i l'Amèrica Llatina. Es basa en la comprovació o verificació d'informacions, especialment de tipus polític<sup>5</sup>. Alguns dels millors exemples actuals de periodisme de dades són alhora fantàstiques mostres de *fact checking*, com el mitjà dels EUA, *Politifact*<sup>6</sup>.

**Open data i big data.** La denominació «open data» o «dades en obert» fa referència a la capacitat de capturar, tractar, modificar, reutilitzar o difondre dades; en definitiva, de gestionar dades sense restriccions. Té relació (o es pot considerar com una petita part) d'una tendència molt més àmplia, el gran moviment «open» que promou el lliure ac-

... des d'un punt de vista periodístic, les dades per si soles no expliquen històries: potser ofereixen context, però no donen respostes si no són traduïdes, si no són explicades. Serà necessària, per tant, la labor del professional que les relacioni, i les relati.

cés al coneixement, a la ciència, al software, a les dades... Per la seva banda, l'expressió *big data* o «grans dades» fa referència als programes, processos o tècniques per gestionar grans volums de dades, una cosa que només ha estat possible des de fa pocs anys. La irrupció d'aquests dos moviments han situat en aquests moments les dades en el centre del paradigma de la informació. Així, el periodisme de dades no seria altra cosa que l'aplicació d'aquests dos moviments al camp del periodisme. En aquest context, haurem de tenir present que, des d'un punt de vista periodístic, les dades per si soles no expliquen històries: potser ofereixen context, però no donen respostes si no són traduïdes, si no són explicades. Serà necessària, per tant, la labor del professional que les relacioni, i les relati.

### 2.3. Productes de periodisme de dades

De quina manera, com o en què es concreta el periodisme de dades? Encara que ens trobem en una fase inici-

5. Guallar, 2012.

6. *Politifact* [www.politifact.com](http://www.politifact.com).

al d'aquest moviment entès com a tal, sí que es poden distingir, de manera exploratòria, al menys tres tipus de productes de periodisme de dades, que estan presents a la proposta de definició esmentada abans. Una característica de la diferenciació entre aquests productes és si el format principal és el text, el multimèdia o les dades en brut. És així com podem distingir entre:

- Articles basats en dades o que les continguin
- Infografies i visualitzacions basades en dades
- Bases de dades i conjunts de dades en obert

**Articles amb dades.** És un lloc comú considerar que un article periodístic perquè tingui qualitat ha de comptar amb un bon suport documental, i en aquest sentit no és cap novetat l'existència d'un article amb aportació de dades o basat en dades. La novetat més aviat rau en la proliferació d'aquests articles (que no són un fet aïllat o puntual) i, especialment, en la seva ambició, al basar-se, cada cop més, en l'anàlisi de grans volums de dades en obert. Així, un producte característic del fenomen que estem estudiant és l'article periodístic amb un aportació destacada de dades fruit d'un procés de recerca per part del periodista o equip periodístic.

**Infografies i visualitzacions.** Una infografia, visualització o reportatge multimèdia basat en dades és un dels productes més coneguts i característics actualment de l'anomenat periodisme de dades. Els antecedents són les infografies publicades per la premsa impresa al llarg dels anys noranta i primera dècada de 2000 pels diaris més importants; uns productes liderats per les seccions d'infografia dels mitjans, i en els quals sovint han col·laborat informàtics, documentalistes i periodistes. Amb l'expansió de la premsa digital, els gràfics es fan interactius, i algunes grans capçaleres com *The New York Times* o *The Guardian*, per exemple, meravellen amb alguns d'aquests multimèdies. La irrupció amb força en els darrers anys dels *big data* i l'*open data* fa que molts d'aquests gràfics integrin o es basin en importants volums de dades, que es mostren de manera molt atractiva visualment i amb diverses possibilitats d'interacció per part de l'usuari. Comunicar visua-

litzant informació és ara una tendència a l'alça per mostrar grans quantitats de dades, que si no fos així, serien difícils d'interpretar i entendre per l'audiència.

**Conjunts de dades en obert.** El moviment *open data* aplicat al periodisme implica també la publicació en obert de dades perquè els lectors les puguin no només llegir o visualitzar (com als casos anteriors), sinó també usar, tractar, manipular. Perquè això sigui possible s'han d'oferir en un format obert, és a dir, accessible i reutilitzable. Alguns d'aquests productes periodístics el que fan és mostrar dades que les administracions públiques del seus països no fan accessibles als ciutadans (això passa, per exemple, a Amèrica Llatina). En aquests casos el periodisme de dades afegeix a la missió periodística d'informar una indubtable funció de servei públic. En ocasions, es poden presentar els conjunts de dades (*datasets*) que s'han utilitzat per construir una informació publicada i en altres casos es mostren de manera independent. També es pot construir una base de dades, amb diverses possibilitats de consulta de les dades per part de l'usuari.

Per últim, cal tenir present que també trobem exemples de confluència dels tres tipus de productes esmentats: per exemple, grans reportatges que contenen articles basats en dades acompanyats per visualitzacions i la publicació de les dades en obert que permeten la seva consulta i ús. Un bon exemple és *Offshore Leaks*, que es detalla al capítol següent.

## 2.4. Equips mixtes

Una altra de les característiques del periodisme de dades és la necessitat de treballar en equips mixtes on els periodistes es coordinen amb perfils més tècnics. Aquí tenen un important paper a jugar altres professionals, com ara informàtics, programadors, infografistes i documentalistes. La relació entre periodistes i programadors la descriu molt clarament a les Jornades #13JPD de Barcelona Manuel Aristarán, informàtic del diari argentí *La Nación*: el periodista necessita d'algú que l'ajudi a explicar gràficament la història i a analitzar les dades, i el programador ne-



cessita qui trobi la història enmig de les dades. Aristarán va explicar que cada vegada més els periodistes s'estan apropiant a aquesta disciplina de treballar amb dades: «sobretot quan veuen com millora la qualitat de la notícia en incorporar una visualització del tema tractat, i els periodistes vénen als tècnics per demanar-nos ajuda de com treballar un conjunt de dades».

### 3. El periodisme de dades a nivell internacional

#### 3.1. Bon exemples internacionals

L'especialitat s'ha anat obrint pas en els mitjans de molts països en els darrers anys, i potser són els EUA el lloc on se'n poden trobar més mostres. Diversos diaris nord-americans estan dedicant actualment espais específics dins del seus webs al periodisme de dades, en molts casos sota l'explícita denominació «Data» i en altres en format de blogs, tot i que no sempre trobarem aquests apartats clarament diferenciats i els productes de dades també els podem trobar com a reportatges especials o inclosos dins de la secció de Multimèdia o Gràfics.

Entre d'altres, es poden citar les seccions «**Maps & Apps**»<sup>7</sup> del *Chicago Tribune*, «**Data Desk**»<sup>8</sup> de *Los Angeles Times*, «**Data Post**»<sup>9</sup> de *The Washington Post*, el blog «**Open**»<sup>10</sup> de *The New York Times* o «**Data**»<sup>11</sup> del *Texas Tribune*. No podem oblidar *Politifact*<sup>12</sup>, referent de l'especialitat concreta del *fact checking* que abans hem esmentat.

7. Maps & Apps. *Chicago Tribune*. <<http://www.chicagotribune.com/news/data/>>.

8. Data Desk. *Los Angeles Times*. <<http://datadesk.latimes.com/>>.

9. Data Post. *The Washington Post*. <<http://www.washingtonpost.com/wp-srv/metro/data/>>.

10. Open. *The New York Times*. <<http://open.blogs.nytimes.com/>>.

11. Data. *The Texas Tribune*. <<http://www.texastribune.org/library/data/>>.

12. *Politifact*. [www.politifact.com](http://www.politifact.com).

13. Data. *La Nación*. <<http://blogs.lanacion.com.ar/data/>>.

14. «La evolución patrimonial del Gabinete», *La Nación*. <<http://www.lanacion.com.ar/1546303-los-bienes-de-los-funcionarios-en-la-primera-news-application-de-la-nacion>>.



Figura 1. *La Nación*, «La evolución patrimonial del gabinete»

Sense deixar el continent americà, no podem deixar de parlar del ric moviment de periodisme de dades a l'Amèrica Llatina. Aquí cal destacar el diari argentí *La Nación*, que s'ha convertit amb el blog «**Data**»<sup>13</sup> en un dels mitjans capdavanters del moviment a nivell mundial. El seu equip de dades ha dut a terme reportatges tan rellevants com «**La evolución patrimonial del Gabinete**»<sup>14</sup>, que mostra l'evolució al llarg dels anys del patrimoni de membres del govern i alts càrrecs.

El subtítol dels apartats de Data és ben explícit: «Facts are sacred», tota una declaració de principis, que pren com a referència la frase de l'editor de *The Guardian* als anys 20, Charles Prestwich Scott, «Comment is free, but facts are sacred»

A Europa el referent és *The Guardian*. Les seves seccions de dades són «**Data Store**»<sup>15</sup>, on es recullen tots els productes de periodisme de dades del diari, i el popular «**Datablog**»<sup>16</sup>, el blog de l'especialitat que és una referència obligada. El subtítol dels apartats de Data és ben explícit: «Facts are sacred», tota una declaració de principis, que pren com a referència la frase de l'editor de *The Guardian* als anys 20, Charles Prestwich Scott, «Comment is free, but facts are sacred»<sup>17</sup>. Al capdavant hi és la figura d'en Paul Rogers, segurament l'exponent principal del periodisme de dades en aquests moments a nivell internacional, qui és a més autor de diversos articles i d'un llibre sobre la matèria.

Un dels molts exemples que es podrien posar del diari anglès és el reportatge de l'anàlisi dels rumors a Twitter durant els altercats de l'estiu del 2011, «**Reading the Riots**»<sup>18</sup>, pel qual es van analitzar 2,6 milions de tuits, i que mostra com neixen i moren aquest rumors, quantes hores duren o des de quins barris s'intensifiquen.

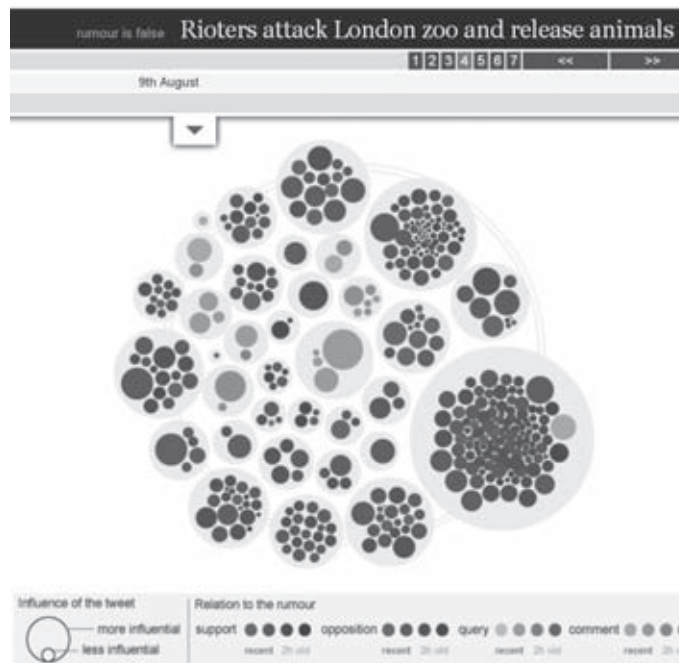


Figura 2. *The Guardian*, «Reading the Riots»

### 3.2. El cas *Offshore Leaks*

Volem comentar amb més detall un dels exemples més notables de periodisme de dades fins al moment. L'abril de 2013 el Consorci Internacional de Periodistes d'Investigació (ICIJ, en les seves sigles en anglès) va donar a conèixer una de les més grans filtracions de dades aconseguida i analitzada fins ara per un mitjà de comunicació. «**Offshore Leaks**»<sup>19</sup> és el títol d'un ampli reportatge multimèdia on es posen al descobert les xarxes d'empreses, polítics, recaptadors i celebritats d'arreu del món que tenen societats en paradisos fiscals. A través de mapes, vídeos i infografies, el lector és capaç d'entendre l'entramat de

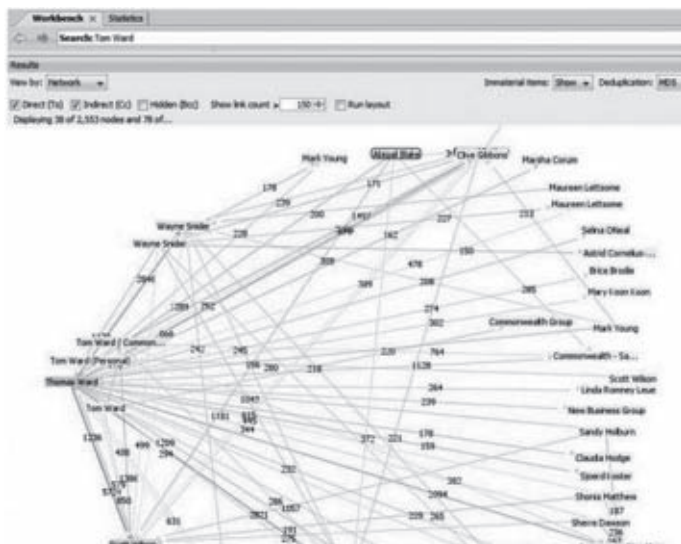
15. Data store. *The Guardian*. <<http://www.guardian.co.uk/data>>.

16. Datablog. *The Guardian*. <<http://www.guardian.co.uk/news/datablog>>.

17. «Comment is free, but facts are sacred», *The Guardian*, 29 November 2002. <<http://www.guardian.co.uk/commentisfree/2002/nov/29/1>>.

18. «Reading the Riots», *The Guardian*. <<http://www.guardian.co.uk/uk/interactive/2011/dec/07/london-riots-twitter>>.

19. «Offshore Leaks», *The International Consortium of Investigative Journalists*. <<http://www.icij.org/offshore>>.



**Figura 3.** *Offshore Leaks*, xarxa de contactes en el món dels paradisos fiscals

relacions i les seves milionàries transaccions de diners al llarg dels anys.

Per presentar aquest macroprojecte que posa al descobert els paradisos fiscals, l'ICIJ va treballar durant un any i mig amb 86 periodistes de 46 països, els quals van manejar 260 gigabytes de dades. Duncan Campbell, membre fundador de l'ICIJ, resumeix la investigació d'aquesta manera: «S'han revisat més 130.000 noms, prop de 12.000 intermediaris (o agents) i unes 22.000 empreses implicades. Les operacions investigades corresponen a 170 països. Mai abans no s'havia aconseguit res d'igual»<sup>20</sup>.

Després de mesos publicant històries de personatges rellevants en algunes de les principals capçaleres medià-

tiques de tot el món, l'ICIJ i el diari *La Nación*<sup>21</sup> de Costa Rica van crear una base de dades consultable que permet buscar informació de 100.000 companyies. Qualsevol persona pot introduir el nom d'algú implicat, realitzar la cerca a partir del seu país o ciutat i veure qui evadeix impostos.

Les dades mostrades –que cobreixen un període de 30 anys fins al 2010– són una petita part d'una filtració de 2,5 milions d'arxius que l'ICIJ va analitzar. La base de dades inclou entitats constituïdes en en paradisos fiscals: Illes Verges Britàniques, Illes Caiman, Illes Cook, Singapur, Hong Kong, Samoa, Seychelles, Maurici, Labuan i Malàisia. S'ha eliminat aquella informació que podria implicar problemes de privacitat, com adreces de correu i números de telèfon, passaports, comptes bancaris, béns o imatges, i tot el visible està contrastat pels periodistes involucrats. Des d'aquesta organització també s'ha animat els ciutadans a verificar els resultats amb altres fonts d'informació, i si troben res fals o rellevant a assenyalar, es poden posar en contacte amb el nucli que ha portat la batuta d'aquesta recerca.

La periodista espanyola Mar Cabra ha estat la *Data Research Manager* del projecte. Durant any i mig es va encarregar de servir milers de documents als 86 periodistes dels 46 països implicats. Els periodistes rebien les dades, les contrastaven i construïen les històries que van ajudar a explicar, des de les principals capçaleres internacionals, com polítics, recaptadors i celebritats d'arreu del món tenien interessos en paradisos fiscals. «Per exemple, de França es va obtenir informació de Jean-Jacques Augier, tresorer de l'última campanya de François Hollande. El periodista de *Le Monde* ho va publicar el mateix dia i a la mateixa hora que la resta de periodistes treien a la llum altres històries», explicava Cabra a les Jornades #JPD13.

20. Campbell, 2013.

21. Data. *La Nación*. <<http://www.nacion.com/data/>>.

## 4. El descobriment del periodisme de dades a casa

### 4.1. Situació a Espanya

L'Open Knowledge Foundation<sup>22</sup> (OKFN) és una organització internacional dedicada a promoure el coneixement obert en el seu sentit més ampli (*open education, open government, open data, data journalism, etc.*). El mes d'octubre de 2012 es va crear el grup espanyol de l'OKFN, integrat pels principals professionals de diferents disciplines que promouen activitats que ajudin a entendre la importància de totes aquestes branques del coneixement que s'estan desplegant, en aquests moments, gràcies a les facilitats de comunicació d'Internet i també a les eines gratuïtes i digitals que es troben a la xarxa.

Sota el paraigües del grup espanyol de l'OKFN es van organitzar les I Jornades de Periodisme de Dades i Open Data a Barcelona #JPD13 celebrades del 24 al 26 de maig del 2013 a Barcelona, Madrid, Donostia i Sevilla alhora, i que hem anat esmentat al llarg d'aquest article. Aquesta ha estat la primera vegada que es realitzava un esdeveniment sobre periodisme de dades d'aquestes dimensions a Espanya, amb conferències, *barcamp*, tallers i *hackathon*. Tot i que la programació més forta es va centrar a la Ciutat Comtal, amb l'acollida de ponents internacionals i nacionals, l'interès pel periodisme de dades va quedar manifest, a través de les xarxes socials, des de tots els punts de l'Estat.

Abans hem parlat del retard del periodisme de dades a Espanya respecte a altres països. Segurament el motiu és el fet que encara no hi ha una Llei de Transparència de

la Informació<sup>23</sup>. Som el país de la Unió Europea amb més d'un milió d'habitants, que no té una regulació que permeti als ciutadans accedir lliurement a les dades públiques. «Tenim un projecte de Llei, però la legislació de Rwanda que s'ha aprovat aquest any és molt millor que el nostre esborrany», en paraules de la periodista Mar Cabra.

L'esdeveniment de periodisme de dades celebrat a Barcelona va servir també per ser conscients que, malgrat la novetat d'aquesta disciplina a Espanya i els inconvenients per aconseguir informació pública, ja existeixen exemples destacables de periodisme de dades que s'han convertit en un referent. Veiem alguns dels més destacats.

### 4.2. Bons exemples a Espanya

«El periodisme actual, basat en globus sonda, declaracions, en comentaris de tertulians i filtracions interessades no està ajudant la professió perquè actua d'altaveu dels governs i poders de torn, mentre que les dades són objectives», va dir la periodista Eva Belmonte a les Jornades #JPD13. Belmonte és l'autora de ***El BOE nuestro de cada día***<sup>24</sup>, un blog on diàriament aboca informacions publicades al Butlletí Oficial de l'Estat i que, a la majoria dels ciutadans, ens poden passar desapercebudes. Aquests dades, relacionades, contextualitzades i visualitzades amb eines digitals, ens permeten entendre millor les accions del Govern.

Un altre exemple és ***Tu Derecho a Saber***<sup>25</sup>, una iniciativa de la Fundació Ciudadana Civio<sup>26</sup>, dirigida per l'informàtic David Cabo, en la qual col·labora també Belmonte. *Tu Derecho a Saber* és una plataforma que recull les peticions d'informació dels ciutadans a les institucions públiques. Tot i que més de la meitat d'aquestes sol·licituds (el 53%) no reben cap resposta per part de les administracions im-

22. Open Knowledge Foundation ([okfn.org](http://okfn.org))

23. Vicepresidencia del Gobierno. Ministerio de la Presidencia. Ley de Transparencia, Acceso a la Información Pública y Buen Gobierno. <<http://www.leydetransparencia.gob.es/index.htm>>.

24. El BOE nuestro de cada día ([elboenuestrodecadadia.com](http://elboenuestrodecadadia.com)).

25. Tu Derecho a Saber ([www.tuderechoasaber.es](http://www.tuderechoasaber.es)).

26. Fundació Ciudadana Civio ([www.civio.es](http://www.civio.es)).



Figura 4. «Tu derecho a saber»

plicades, el projecte continua endavant amb força. «En els últims mesos, sembla que els ajuntaments vulguin obrir les seves dades, però els formats en els quals les publiquen (PDF), les converteixen en documents inútils», va explicar Eva Belmonte a les Jornades.

La Fundació Ciudadana Civio també és la impulsora de l'**Indultómetro**<sup>27</sup> i **Dónde van mis Impuestos**<sup>28</sup>. El primer exemple mostra, a partir de les informacions publicades al BOE durant més d'una dècada, la quantitat d'indults que hi ha hagut a Espanya, relacionats amb els diferents governs. El segon és una visualització pràctica –creada per David Cabo– perquè tothom, des d'una pàgina web, pugui comparar els diferents pressupostos que destina el Govern en cadascuna de les partides que administra. Es pot consultar des de 2008 fins al moment actual, i veure les diferències invertides, per exemple, en educació o serveis socials respecte d'altres anys. Les dades són dels Presupostos Generals de l'Estat. Un altre projecte de Civio

27. Indultómetro ([www.civio.es/proyectos/el-indultometro](http://www.civio.es/proyectos/el-indultometro)).

28. Dónde van mis impuestos ([www.dondevanmisimpuestos.es](http://www.dondevanmisimpuestos.es)).

29. España en llamas ([www.espanaenllamas.es](http://www.espanaenllamas.es)).

30. Tu tasa de paro ([www.tusasadeparo.es](http://www.tusasadeparo.es)).

31. Kaleidos ([www.tusasadeparo.es/about](http://www.tusasadeparo.es/about)).



Figura 5. «El indultómetro»

que cal destacar és **Espanya en llamas**<sup>29</sup>, que permet visualitzar bona part dels 187 mil incendis que han cremat més d'un milió d'hectàrees a Espanya durant una dècada (2001 -2011).

Una visualització senzilla i similar a *Dónde van mis impuestos* és la que han realitzat els joves de **Tu tasa de paro**<sup>30</sup>, a partir de les dades que l'Institut Nacional de Estadística publica sobre l'Encuesta de Población Activa. La pàgina –creada per l'equip de Kaleidos<sup>31</sup>– permet comparar la taxa d'atur que hi ha a Espanya per províncies, edat, sexe i nivell d'estudis. D'aquesta manera es poden treure conclusions de certes zones segons les variables escollides.

Són representacions senzilles però que ajuden molt a entendre el moment actual. En aquest sentit cal destacar la labor de la periodista Vicky Bolaños, juntament amb un equip d'informàtics que fa anys evolucionen en el seu pro-



Figura 6. «Tu tasa de paro»

jecte **Qué hacen los Diputados**<sup>32</sup>. Des de juny de 2011, segueixen al dia l'activitat parlamentària. «Existeix molta informació disponible sobre els diputats però és difícil de trobar, entendre i interpretar. Vam pensar que podríem fer servir les eines digitals per monitoritzar la tasca dels polítics. Volem ajudar a millorar la transparència», expliquen des de la seva pàgina. Actualment, a través de *crowdfunding*, han aconseguit diners per construir una eina que permeti a moltes mans actualitzar diàriament la informació del Congrés dels Diputats.

No ens podem oblidar de **Data'n'Press**<sup>33</sup>, una empresa jove catalana muntada pel periodista Eduard Martín-Borregón, el seu germà, enginyer de telecomunicacions, David Martín-Borregón, i l'informàtic Arnau Udina. Ells són els responsables de diverses visualitzacions per a *El Periódico de Catalunya* i *Vilaweb*, a més de monitoritzar debats polítics a través de Twitter.

La pàgina (...) permet comparar la taxa d'atur que hi ha a Espanya per províncies, edat, sexe i nivell d'estudis. D'aquesta manera es poden treure conclusions de certes zones segons les variables escollides.

Del *hackathon* de les #JPD13 van sortir també projectes tan interessants com **Vidas contadas**<sup>34</sup>, que es va endur tres dels quatre premis que es podien aconseguir. El projecte vol posar nom i explicar la història de les persones afusellades, desaparegudes i represaliades durant la guerra civil espanyola i el franquisme. «És un intent d'aplicar la tecnologia i les seves possibilitats periodístiques per recuperar la memòria», expliquen els seus impulsors. **Cerrado por reformas**<sup>35</sup> (que va aconseguir un accèssit) és una visualització d'un mapa d'Espanya que marca els locals comercials que s'han anat tancant amb motiu de la crisi econòmica. I el tercer projecte guanyador, **¿Dónde van las subvenciones europeas?**<sup>36</sup>, són unes visualitzacions per donar a conèixer en què es gasten (o no) els diners rebuts de la Unió Europea per a les subvencions agrícoles d'Espanya. Els periodistes Antonio Villarreal García i Adrián Blanco Ramos ja han començat a investigar i han descobert camps de golf, parròquies, aparcaments privats, etc., en lloc de camps per conrear.

32. Qué hacen los diputados ([quehacenlosdiputados.net](http://quehacenlosdiputados.net)).

33. Data'n'Press ([www.datanpress.com](http://www.datanpress.com)).

34. *Periodismo de datos*. «Vidas contadas». <<http://periodismodatos.okfn.es/vidas-contadas-memoria-historica>>.

35. Cerrado por reforma ([cerradoporreforma.org](http://cerradoporreforma.org)).

36. *Periodismo de datos*. «¿Dónde van las subvenciones europeas?». <<http://periodismodatos.okfn.es/category/hackathon>>.

Una de les diferències entre el periodisme de dades que es fa a nivell internacional i el que es fa a Espanya és que el primer es cou des de les redaccions de les principals capçaleres mediàtiques, amb equips mitjans o grans, on es barregen perfils tècnics i periodístics. En el cas espanyol, tots els exemples mencionats fins ara han estat creats per periodistes que en les seves hores lliures inverteixen, de manera voluntària, el seu temps sense esperar –en la majoria dels casos– cap remuneració.

Sí que es poden citar, per acabar aquest apartat, algunes mostres de periodisme de dades publicades per mitjans espanyols, com ara alguns dels multimèdies de *Lainformacion.com* (sens dubte el diari més destacat a Espanya en visualització d'informacions en els darrers anys), que es poden consultar a la secció **Gráficos**<sup>37</sup>, com per exemple l'especial «**Patrimonio de los diputados**»<sup>38</sup>. També exemples puntuals com «**Todos los papeles de Bárcenas**»<sup>39</sup> a *El País*, iniciatives d'*El Periódico de Catalunya* com la de publicar en obert les dades de les seves enquestes d'opinió (el «**Baròmetre**»<sup>40</sup>), o aquest «**Mapa electoral**»<sup>41</sup> dels ja esmentats *Data'n'Press* a *Vilaweb*. Però no veiem a cap mitjà espanyol una secció de *data* comparable a la de moltes capçaleres importants d'altres països.

## 5. Present i futur

El moment actual de crisi econòmica a Espanya –i com a conseqüència la reducció dràstica de periodistes– és l'excusa perfecta perquè els mitjans no inverteixin temps ni planifiquin complementar les notícies, almenys, amb visu-

alitzacions senzilles. Ni mencionar introduir-se a explorar tot el que les dades obertes que corren per Internet poden ajudar la professió.

Però arreu del món s'estan obrint noves vies d'aprendre el necessari per treballar amb dades de manera gratuïta i col·laborativa. Per exemple, la *School of Data*<sup>42</sup>, organitzada per l'OKFN, és un dels casos més coneguts. Ells són els responsables, juntament amb l'European Journalism Centre<sup>43</sup>, del primer manual de periodisme de dades, *The Data Journalism Handbook*, ja esmentat. I, recentment, s'està plantejant la creació de La escuela de datos per als professionals llatinoamericans i espanyols.

Després dels exemples que hem vist, queda clar que el periodisme de dades pot actuar a Espanya, en aquests moments, com un antídote per a la professió periodística, per democratitzar-la i per evitar que caigui en el futur lúgubre que molts li auguren. I també pot ser un revulsiu per a tots aquells professionals que treballen conjuntament amb els periodistes en la creació de productes informatius: informàtics, infografistes... i, per descomptat, documentalistes. Quan la Llei de Transparència entri en vigor, malgrat les restriccions inicials d'accés a la informació que els experts auguren, hi haurà una oportunitat enorme perquè tant els professionals com els ciutadans en general puguin recuperar dades públiques, es puguin analitzar, contextualitzar i visualitzar. Tot aquest poder en mans ciutadanes ajudarà, sens dubte, a entendre millor la realitat del moment i actuar en conseqüència.

Les previsions sobre què ens depararà el futur, a nivell global, quan tothom estigui acostumat a viure *amb* i *per* les dades és inimaginable en aquests moments. Internet és

37. Gráficos. *Lainformacion.com*. <<http://graficos.lainformacion.com>>.

38. *Lainformacion.com*. «Especial Patrimonio de los diputados». <<http://especiales.lainformacion.com/espana/patrimonio-diputados>>.

39. *El País*. «Todos los papeles de Bárcenas». <[http://elpais.com/especiales/2013/caso\\_barcanas/todos\\_los\\_papeles.html](http://elpais.com/especiales/2013/caso_barcanas/todos_los_papeles.html)>.

40. *El Periódico de Catalunya*. «El baròmetre de Catalunya d'EL PERIÓDICO, en obert al complet». <<http://www.elperiodico.cat/ca/noticias/politica/barometre-catalunya-del-periodico-obert-complet-2413825>>.

41. *Vilaweb*. «Mapa electoral». <<http://www.vilaweb.cat/mapa-electoral>>.

42. School of Data ([schoolofdata.org](http://schoolofdata.org)).

43. European Journalism Centre ([ejc.net](http://ejc.net)).

un ecosistema que accelera a velocitats insospitades per a la humanitat. El que han generat les comunicacions a les xarxes socials no ho podríem haver imaginat fa cinc anys. I, si algun guru l'havia pronosticat, era vist per les masses com una teoria poc consistent o una utopia. De la mateixa manera, què ens donaran o trauran les dades durant la propera dècada ens és desconegut hores d'ara perquè depèn de les accions de milions de persones actuant en temps real i des de qualsevol racó del planeta.

No obstant, per acabar aquest article, potser són bones les paraules de l'antropòleg Eudald Carbonell, que fa poc va comentar que un dels trets que definirà el futur serà la democratització de la ciència: «Hi haurà una consciència crítica de la nostra espècie. Vindrà una transhumanitat, la del meu fill, on els ciutadans ja no diran: "M'han dit o m'han explicat que..." Diran: "Sé que..."».

## 6. Bibliografia

CAMPBELL, Duncan. «How ICIJ's Project Team Analyzed the Offshore Files» [En línia], *ICIJ*, 3 abril 2013. <<http://www.icij.org/offshore/how-icijs-project-team-analyzed-offshore-files>> [Consulta: 12/07/2013].

CHAPARRO-DOMÍNGUEZ, María Ángeles. «La evolución del periodismo de precisión: el blog de *The Guardian* sobre periodismo de datos» [En línia]. *I Congreso Internacional de Comunicación y Sociedad Digital*, Logroño: Universidad Internacional de la Rioja, enero 2013. <<http://dialnet.unirioja.es/servlet/articulo?codigo=4247831>> [Consulta: 12/07/2013].

FLORES VIVAR, Jesús; Salinas Aguilar, Cecilia. «Sinergias en la construcción del Nuevo Periodismo derivadas del Data Journalism y el Transmedia Journalism» [En línia]. *III Congreso Internacional Comunicación 3.0*. Salamanca, 10-11 octubre 2012. <[http://www.2ip.es/wpcontent/uploads/2013/01/sinergias\\_construccion\\_nuevoperiodismo.pdf](http://www.2ip.es/wpcontent/uploads/2013/01/sinergias_construccion_nuevoperiodismo.pdf)> [Consulta: 12/07/2013].

GRAY, Jonathan; BOUNEGRU, Liliana; CHAMBERS, Lucy (ed). *The Data Journalism Handbook*. Open Knowledge Foundation, 2012 [En línia]. <<http://www.datajournalismhandbook.com/>> [Consulta: 12/07/2013].

[versió en castellà:] <<http://interactivos.lanacion.com.ar/manual-data/index.html>> [Consulta: 12/07/2013].

GUALLAR, Javier. «Comprobadores de datos o fact checkers» [En línia], *#documedios*, 22 octubre 2012. <<http://documedios.wordpress.com/2012/10/22/comprobadores-de-datos-o-fact-checkers/>> [Consulta: 12/07/2013].

GUALLAR, Javier. «Prensa digital en 2010-11», *Anuario ThinkEPI*, vol. 7, 2013, p. 194-199.

I Jornades de Periodisme de Dades i Open Data #JPD13 [En línia]. <<http://periodismodatos.okfn.es>> [Consulta: 12/07/2013].

MEYER, Philip. *Precision Journalism. A reporter's introduction to social science methods*. 1970.

PEIRÓ, Karma. «El por qué del éxito de las I Jornadas de Periodismo de Datos y Open Data #jpd13» [En línia], *Karma Peiró. Historias de una periodista digital*, 6 junio 2013. <<http://karmapeiro.wordpress.com/2013/06/06/el-por-que-del-exito-de-las-i-jornadas-de-periodismo-de-datos-y-open-data-jpd13-cabralens/>> [Consulta: 12/07/2013].

ROGERS, Simon. «Data journalism at the Guardian: what is it and how do we do it?» [En línia], *The Guardian Data Blog*, 28 July 2011. <<http://www.guardian.co.uk/news/data-blog/2011/jul/28/data-journalism>> [Consulta: 12/07/2013].

ROGERS, Simon (ed.) *Facts are sacred. The power of data*. Londres: Guardian Books, 2011.



# Les dades de recerca: una oportunitat professional per als gestors d'informació

**Llorenç Arguimbau**

Observatori de la Recerca (OR-IEC). Institut d'Estudis Catalans (IEC)

*llarguimbau@iec.cat*

## **Resum:**

En un context d'obertura de quantitats massives de dades, aquestes són considerades com actius amb valor afegit. Per tant, les dades científiques s'han de gestionar, compartir i preservar a llarg termini, des de la seva creació fins a la conservació definitiva, i mantenir l'enllaç amb les publicacions científiques derivades. L'article s'estructura en cinc apartats: marc conceptual, infraestructures de dades, agents involucrats en la comunicació científica, documentació científica i gestió de dades, i experiències pràctiques (directoris internacionals i iniciatives espanyoles). Per acabar, s'extreuen algunes conclusions, tot i que es tracta d'una àrea emergent amb nombroses qüestions a resoldre, on els gestors d'informació hi poden jugar un paper destacat.

## **Paraules clau:**

accés obert, ciència, comunicació científica, dades, dipòsits digitals, gestió, infraestructures, preservació, publicacions, recerca, reutilització

## Los datos de investigación: una oportunidad profesional para los gestores de información

### Resumen:

En un contexto de apertura de cantidades masivas de datos, éstos son considerados como activos con valor añadido. Por lo tanto, los datos científicos han de ser gestionados, compartidos y preservados a largo plazo, desde su creación hasta la conservación definitiva, y mantener el enlace con las publicaciones científicas derivadas. El artículo se estructura en cinco apartados, que corresponden a sus objetivos básicos: marco conceptual, infraestructuras de datos, agentes involucrados en la comunicación científica, documentación científica y gestión de datos, y experiencias prácticas (directorios internacionales e iniciativas españolas). Por último, se extraen algunas conclusiones, aunque se trata de un área emergente con numerosas cuestiones por resolver, donde los gestores de información pueden jugar un papel destacado.

### Palabras clave:

acceso abierto, ciencia, comunicación científica, datos, repositorios digitales, gestión, infraestructuras, preservación, publicaciones, investigación, reutilización

## Research data: professional opportunity for information managers

### Abstract:

In the context of huge amounts of data becoming publicly available, this data is considered as an asset with added value. There should therefore be long term management dissemination and preservation of scientific data from its creation to its ultimate storage. Moreover, it should be referenced in scientific publications in which it is used. The article is divided into five sections which correspond to its basic objectives: conceptual framework; data infrastructures; agents involved in scientific communication; scientific documentation and data management; case studies (international directories and Spanish initiatives). Finally, some conclusions are drawn, all be it that this is an emerging field with various questions still to be addressed. It is in this context that information managers could play an important role.

### Keywords:

open access, science, scientific communication, data, digital repositories, management, infrastructures, preservation, publications, re-search, reuse



©istockphoto/AlexRaths

## Introducció

Les tecnologies de la informació i de la comunicació (TIC) i el moviment d'accés obert (*open access*) al coneixement han tingut un impacte evident en el món de la comunicació científica, sobretot en les publicacions. Però, què succeeix amb les dades creades o recollides per qualsevol investigació científica? Quin paper hi poden jugar els professionals de la gestió d'informació? Aquestes dues preguntes configuren l'eix central al voltant del qual gira l'article. Hem de tenir en compte que el tractament de dades científiques és una àrea emergent que presenta encara nombroses qüestions per resoldre.

«Durant els últims anys s'han incrementat extraordinàriament el volum i la rapidesa amb què aquestes dades [científiques] s'obtenen i es comparteixen gràcies al perfeccionament dels instruments de mesura i a la utilització

d'Internet per a la comunicació entre els investigadors. Com a resultat, la cadena de la comunicació científica s'ha ampliat [...] per incorporar la captura, descripció i reutilització de les dades primàries en què es basen aquestes publicacions»<sup>1</sup>. Per tant, també cal administrar el cicle global de les dades, des de la creació fins a la conservació definitiva, i mantenir l'enllaç amb les publicacions científiques que se n'han derivat.

Hi ha un debat intens dins de la comunitat científica sobre la necessitat i els procediments per administrar, compartir i reutilitzar les dades<sup>2</sup>. En els darrers anys, aquesta temàtica ha despertat l'interès de tots els agents involucrats en la comunicació científica: investigadors, institucions científiques, organismes internacionals, governs i agències financeres, editorials i bibliotecaris-documentalistes. Un article científic pot quedar obsolet, però no les dades en sí, perquè poden ser reutilitzades, recombinades o repro-

1. Àngel Borrego, «Enriquir les publicacions amb dades empíriques» [En línia], *Blok de BiD*, 11/04/2012. <<http://www.ub.edu/blokdebid/ca/content/enriquir-les-publicacions-amb-dades-emp%C3%ADriques-0>> [Consulta: 06/06/2013].
2. Daniel Torres-Salinas; Nicolás Robinson-García; Álvaro Cabezas-Clavijo, «Compartir los datos de investigación en ciencia: introducción al *data sharing*», *El Profesional de la Información*, vol. 21, núm. 2, p. 174.

cessades amb una altra tecnologia o simplement vistes des d'una altra perspectiva<sup>3</sup>.

Tot i que el fet de compartir dades entre els científics és consubstancial a la recerca, cal assenyalar algunes fites destacades de les darreres dècades. L'any 1971 es va crear Protein Data Bank (PDB)<sup>4</sup> (Figura 1), el primer banc de dades científiques. Així mateix, l'any 1983, *Journal of Biological Chemistry* va ser una revista pionera a exigir les dades per a la publicació d'articles<sup>5</sup>. A partir de mitjans dels anys noranta, algunes agències anglosaxones encarregades del finançament de la recerca van començar a considerar que les dades s'han de compartir, sobretot si s'han obtingut amb diners públics. L'any 2003, la declaració de Berlín sobre l'accés obert al coneixement en ciències i humanitats contemplava també les dades.



Figura 1. Protein Data Bank (PDB)

L'article vol ser una introducció general al tema i, per tant, la metodologia està fonamentada en la revisió bibliogràfica i en la identificació de fonts d'informació de referència.

## 1. Marc conceptual

Quantitats ingents de dades són presents arreu de la nostra societat, mitjançant les aplicacions web 2.0, les administracions públiques, els mitjans de comunicació, les institucions científiques, etc. *Big data* és un concepte habitual que s'utilitza per definir grans paquets de dades, els més coneguts dels quals estan vinculats amb l'anàlisi de l'ús de xarxes socials i de motors de cerca a Internet. A la vegada, totes aquestes dades es troben cada cop més lligades als moviments *open*<sup>6</sup>. D'aquesta manera, creix la consciència que cal posar-les a disposició de tothom i preservar-les com a actius d'utilitat social i comuna<sup>7</sup>.

Segons l'Open Knowledge Foundation (OKFN)<sup>8</sup>, una dada oberta és qualsevol «dada que pot ser utilitzada, reutilitzada i redistribuïda lliurement per qualsevol persona, tan sols subjecte, com a màxim, a l'exigència d'atribució i de compartir en les mateixes condicions». Així, les propietats principals de les dades obertes són les següents:

- Disponibilitat i accés: han d'estar disponibles en el seu conjunt, amb un cost raonable de reproducció (preferentment, mitjançant la descàrrega a Internet) i en una forma adequada i modificable.
- Reutilització i redistribució: s'han de proporcionar en condicions que facilitin el reaprofitament, fins i tot la combinació amb altres conjunts de dades (*datasets*).
- Participació universal: tothom ha de ser capaç d'utilitzar, reutilitzar i redistribuir les dades obertes sense discriminar cap camp d'acció, persona o grup (per exemple, sense restriccions a usos comercials o limitacions per a finalitats educatives).

3. Remedios Melero, «Una plenamar de dades» [En línia], *Blok de BiD*, 24/11/2010. <<http://www.ub.edu/blokdebid/ca/content/una-plenamar-de-dades>> [Consulta: 06/06/2013].
4. Protein Data Bank (PDB) ([www.rcsb.org/pdb](http://www.rcsb.org/pdb)) [Consulta: 06/06/2013].
5. Daniel Torres-Salinas, «Compartir datos (*data sharing*) en ciencia: contexto de una oportunidad», *Anuario ThinkEPI*, vol. 4 (2010), p. 259.
6. Antonia Ferrer-Sapena; Enrique A. Sánchez-Pérez, «Open data, big data: ¿hacia dónde nos dirigimos?» [En línia], *Notas ThinkEPI*, 1 febrero 2013. <<http://www.thinkepi.net/open-data-big-data-hacia-donde-nos-dirigimos>> [Consulta: 06/06/2013].
7. Alicia García-García; Josep-Manuel Rodríguez-Gairín; Tomás Saorín; Luis-Millán González; Xavi García-Massó; Antonia Ferrer-Sapena; Fernanda Peset, «ODiSEA: International Registry on Research Data» [En línia], *BiD: textos universitaris de biblioteconomia i documentació*, núm. 29 (2012). <<http://www.ub.edu/bid/29/garcia2.htm>> [Consulta: 06/06/2013].
8. OKFN, *The Open Data Handbook*, Cambridge: OKFN, 2012. <<http://opendatahandbook.org>> [Consulta: 06/06/2013].

Tim Berners-Lee<sup>9</sup> ha proposat un esquema de cinc estrelles per a les dades obertes: s'atribueix una estrella a les dades en qualsevol format amb una llicència, i cinc estrelles a les dades enllaçables o *linked data*.

Pel que fa a la definició de dada de recerca, els National Institutes of Health (NIH) (Figura 2) dels Estats Units d'Àmerica afirmen que es tracta de qualsevol «material factual enregistrat que és acceptat de forma generalitzada per la comunitat científica per validar els resultats d'investigació»<sup>10</sup>. A la vegada, aquesta definició és el punt de partida a l'elaborada posteriorment per l'Organització per a la Cooperació i el Desenvolupament Econòmic (OCDE)<sup>11</sup>. Ambdues organitzacions no consideren dades de recerca les notes de laboratori, les anàlisis preliminars, els esborranys d'articles científics o els projectes d'investigacions futures.

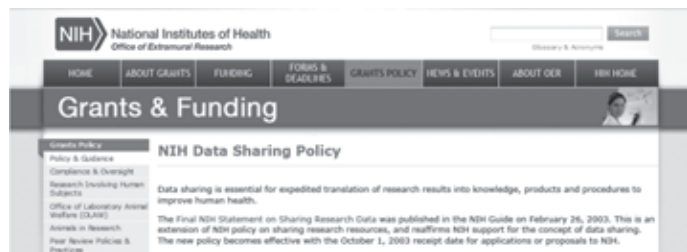


Figura 2. NIH Data Sharing Policy

Un informe de la Research Information Network (RIN)<sup>12</sup> ha elaborat una classificació de les dades de recerca a partir de tres criteris no excloents:

- Procés d'obtenció: experiments científics (per exemple, seqüències genètiques o experiments de física de partícules); simulacions (moviments astronòmics o models econòmics); o observacions de fenòmens específics (per exemple, enquestes sociològiques o dades climatològiques).
- Finalitat de recollida: projecte específic, d'abast mitjà o d'interès general.
- Tractament: dades primàries creades o recollides en experiments o observacions (*raw data*), o bé derivades a partir del tractament d'aquestes dades (*final research data*).

Ara bé, aquest marc conceptual no s'ajusta a totes les especialitats del coneixement: per exemple, un investigador d'humanitats pot treballar amb qualsevol registre fruit de l'experiència humana<sup>13</sup>. Per tant, el concepte de dada de recerca varia de manera considerable segons la disciplina científica i el mètode d'investigació. «La naturaleza de los datos de investigación es muchísimo más variable y dependiente de la disciplina y de su particular ciclo de vida. Además, los requisitos técnicos y legales para garantizar el acceso son más complejos que los de las publicaciones»<sup>14</sup>.

El *data sharing* consisteix a compartir en accés obert les dades científiques amb la finalitat que puguin ser reutilitzades per altres investigadors per generar coneixement. El potencial de la reutilització com a eina de validació de la ciència i com a motor d'innovació és inqüestionable<sup>15</sup>.

9. Tim Berners-Lee, *5 Star Open Data* [En línia]. <<http://5stardata.info>> [Consulta: 06/06/2013].

10. NIH, *NIH Data Sharing Policy* [en línia]. Bethesda, Maryland: NIH, 2007. <[http://grants.nih.gov/grants/policy/data\\_sharing](http://grants.nih.gov/grants/policy/data_sharing)> [Consulta: 06/06/2013].

11. OCDE, *OECD Principles and Guidelines for Access to Research Data from Public Funding* [En línia], Paris: OCDE, 2007. <<http://www.oecd.org/science/sci-tech/38500813.pdf>> [Consulta: 06/06/2013].

12. Alma Swan; Sheridan Brown, *To Share or not to Share: publication and Quality Assurance of Research Data Outputs* [En línia], London: Research Information Network, 2008. <<http://www.rin.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs>> [Consulta: 06/06/2013].

13. Torres-Salinas; Robinson-García; Cabezas-Clavijo, *op. cit.*, p. 176.

14. Cristina González Copeiro; Jordi Serrano-Muñoz (coord.), *La conservación y reutilización de los datos científicos en España: informe del grupo de trabajo de buenas prácticas* [En línia], Madrid: FECYT, 2012. <[http://www.recolecta.net/buscador/documentos/informe\\_datos\\_cientificos\\_en\\_esp.pdf](http://www.recolecta.net/buscador/documentos/informe_datos_cientificos_en_esp.pdf)> [Consulta: 06/06/2013].

15. Fernanda Peset; Antonia Ferrer-Sapena; Imma Subirats-Coll, «Open data y linked open data: su impacto en el área de bibliotecas y documentación», *El Profesional de la Información*, vol. 20, núm. 2 (2011), p. 165-173.

En aquest sentit, sembla evident que les dades obtingudes amb finançament públic haurien de ser compartides amb tothom. Nombroses agències de finançament i institucions científiques ja disposen de polítiques per impulsar l'accés obert al coneixement.

Amb el propòsit d'incentivar el moviment per compartir i reutilitzar dades de recerca, cal promocionar els seus beneficis:

- Progrés científic: nous estudis que combinen diversos conjunts de dades, els quals s'analitzen des de perspectives innovadores<sup>16</sup>.
- Eficiència: millor rendiment dels recursos econòmics invertits, eliminació de projectes duplicats i estalvi de costos<sup>17</sup>.
- Transparència: lluita contra el frau científic<sup>18</sup>.
- Visibilitat: augment del nombre de citacions dels treballs amb dades obertes<sup>19</sup>.

Malgrat els avantatges, el *data sharing* no és una pràctica generalitzada entre els científics. Segurament hi ha massa obstacles i incògnites que encara cal superar:

- Marc inestable: manca de reconeixement i protecció dels investigadors que comparteixen<sup>20</sup>.
- Mal ús: utilització fraudulenta de les dades o crítica, més o menys justificada, a la investigació realitzada<sup>21</sup>.
- Capacitació insuficient: desconeixement i manca de

formació dels investigadors per garantir un tractament adequat de les dades obertes<sup>22</sup>.

- Dificultats tècniques: la implantació i gestió de polítiques i d'infraestructures s'enfronta a costos elevats, manca d'estandardització de les dades, formats múltiples, etc. En aquest sentit, els *datasets* poden ser objectes compostos i heterogenis, amb diferents elements o tipus de dades: textos, fulls de càlcul, gràfics, imatges, etc.<sup>23</sup>

## 2. Infraestructures de dades

Les dades s'han d'administrar i preservar mitjançant infraestructures fiables i duradores. «De los numerosos procesos que tienen lugar durante la gestión de todo el ciclo de vida de los datos científicos, el almacenamiento en depósitos cobra una gran importancia por sus implicaciones en cuanto a la preservación para el futuro, la citación, el reconocimiento de la autoría, la utilización de enlaces persistentes...»<sup>24</sup>.

La pregunta sobre com es poden gestionar, conservar i compartir dades científiques presenta diverses respostes. Una solució són els bancs de dades (*data banks*), infraestructures específicament dissenyades per a aquesta tasca i, per tant, l'objectiu ideal de qualsevol disciplina científica<sup>25</sup>. Ara bé, també cal tenir en compte els dipòsits digitals<sup>26</sup>, on les dades conviuen amb les publicacions i altres

16. A. Ramasamy; A. Mondry; C.C. Holmes; D.G. Altman, «Key Issues in Conducting a Meta-Analysis of Gene Expression Microarray Datasets» [En línia], *PLoS Medicine*, vol. 5, núm. 9, e184. <<http://dx.doi.org/10.1371/journal.pmed.0050184>> [Consulta: 06/06/2013].

17. García-García *et al.*, 2012, *op. cit.*

18. Torres-Salinas, *op. cit.*, p. 259.

19. Heather A. Piwowar; Roger S. Day; Douglas B. Fridsman, «Sharing Detailed Research Data Is Associated with Increased Citation Rate» [En línia], *PLoS ONE*, vol. 2, núm. 3, e308. <<http://dx.doi.org/10.1371/journal.pone.0000308>> [Consulta: 06/06/2013].

20. Christine L. Borgman, «Data, disciplines, and scholarly publishing», *Learned Publishing*, vol. 21, núm. 1 (2008), p. 29-38.

21. Torres-Salinas; Robinson-García; Cabezas-Clavijo, *op. cit.*, p. 181.

22. Torres-Salinas, *op. cit.*, p. 259.

23. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 15.

24. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 37.

25. Torres-Salinas; Robinson-García; Cabezas-Clavijo, *op. cit.*, p. 176.

26. Silvia Arano; Gemma Martínez; Marina Losada; Marta Villegas; Anna Casaldàliga; Núria Bel, «La comunidad "Recursos y datos primarios" de la Universitat Pompeu Fabra: los repositorios institucionales como infraestructuras científicas: estudio de caso», *Revista Española de Documentación Científica*, vol. 34, núm. 3 (2011), p. 387.

Per ara, els bancs de dades s'adrecen a comunitats científiques molt concretes: astronomia, ciències de la terra, gnòmica i proteòmica, neurociència, oncologia, etc.

materials, i les revistes<sup>27</sup>, en les quals les dades actuen com a material suplementari dels articles. D'altra banda, els canals informals (peticions concretes als autors, llocs web d'investigadors o de grups de recerca, etc.) són una alternativa poc aconsellable perquè dificulten la localització i reutilització de dades degut a diferents factors: multiplicitat de formats, manca de normalització, desconeixement del context de creació, etc.

Similars als dipòsits de literatura científica, els *data banks* estan orientats específicament a la gestió de dades especialitzades en una disciplina científica. Per tant, es disposa d'una gran varietat de solucions i resulta difícil consensuar directrius o orientacions. En aquest sentit, són eines complexes per al dipòsit i recuperació de dades que requereixen d'especialistes i que provoquen costos elevats de manteniment (cal garantir-ne la viabilitat econòmica). Per ara, els bancs de dades s'adrecen a comunitats científiques molt concretes: astronomia, ciències de la terra, gnò-



©istockphoto/Denysov Dmytro

mica i proteòmica, neurociència, oncologia, etc. Per exemple, GenBank<sup>28</sup> és un model internacional en la publicació en accés obert de dades genètiques des de l'any 1982.

En relació als dipòsits digitals, generalment s'han orientat cap a la producció científica, deixant de banda formats menys habituals com els materials multimèdia, els programes d'ordinador o les dades<sup>29</sup>. Com a prova d'això, el maig de 2013 el directori OpenDoar té registrats 2.279 dipòsits, dels quals tan sols 84 (un 3,7%) recopilen conjunts de dades.

Una altra possibilitat per compartir dades de recerca és fer-ho com a material suplementari als articles de revista. En aquesta línia, Borrego<sup>30</sup> comenta un informe del projecte europeu *Opportunities for Data Exchange* (ODE) que examina el paper de les editorials. L'informe identifica diversos mecanismes perquè les editorials facin disponibles els fitxers de dades vinculades amb els articles que publiquen, ja sigui mitjançant la publicació de materials

27. Borrego, 11/04/2012, *op. cit.*

28. GenBank ([www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)) [Consulta: 06/06/2013].

29. Arano *et al.*, *op. cit.*, p. 387.

30. Borrego, 11/04/2012, *op. cit.*

suplementaris (en el servidor de l'editorial o en un dipòsit extern) o, més recentment, mitjançant l'aparició de revistes dedicades a la publicació d'articles descriptius sobre conjunts de dades disponibles en un dipòsit (vegeu, per exemple, les revistes *Earth System Science Data* o *Giga Science*). Tanmateix, l'informe detecta dificultats rellevants perquè les editorials es responsabilitzin adequadament de la revisió, publicació i preservació de dades.

Pel que fa a les qüestions de caire més tècnic, el format i la descripció dels fitxers de dades són factors essencials per garantir la consulta i preservació a llarg termini<sup>31</sup>. És recomanable usar formats oberts i que permetin la indexació del contingut. També resulta necessari un control de versions dels fitxers, que han d'estar ben organitzats amb una estructura ordenada i una nomenclatura que permeti la identificació dels continguts. A més, la descripció mitjançant metadades facilita la localització, comprensió i utilització de les dades. A la vegada, les metadades s'estructuren en registres segons esquemes normalitzats, que són prioritaris per a la interoperabilitat dels sistemes. També cal tenir en compte que els conjunts de dades s'han d'associar a un identificador digital únic i permanent. Finalment, és recomanable fer còpies de seguretat en diferents tipus de suports físics, garantir la integritat dels fitxers i prevenir l'obsolescència tecnològica.

### 3. Agents de la comunicació científica

Una gran diversitat d'actors estan involucrats en el cicle de les dades científiques, des dels reguladors (governos, agències finançadores i organismes internacionals), pas-

sant pels gestors (editorials i biblioteques), fins als productors (investigadors i institucions científiques).

Des de mitjans dels anys noranta, cada cop més organismes encarregats del finançament de la recerca científica van començar a prendre consciència de la necessitat de compartir les dades, sobretot les que s'havien obtingut a partir de fons públics<sup>32</sup>. En aquesta línia, es dissenyaren polítiques, estàndards, directrius i recomanacions des d'entitats com, per exemple, l'Economic and Social Research Council (ESRC) del Regne Unit, la National Science Foundation (NSF) dels Estats Units o l'Australian National Data Service (ANDS). Degut al seu caràcter pioner, cal destacar la política dels NIH nord-americans, que des de l'any 2003 obligava a compartir les dades dels projectes d'investigació finançats amb diners públics. Així, les sol·licituds presentades a les seves convocatòries havien d'incorporar un pla de gestió de dades i, a més, els NIH impulsaren diversos dipòsits digitals. Per tant, les agències de finançament han de tenir un paper clau per garantir la sostenibilitat de les dades, a partir de l'exigència d'un pla de tractament i preservació<sup>33</sup>.

Pel que fa als organismes internacionals, la UNESCO manifestà la necessitat de prendre mesures per a «encoratjar les universitats i altres organismes de recerca, tant públics com privats, per garantir la conservació de les dades d'investigació»<sup>34</sup>. D'altra banda, l'OCDE impulsà el *data sharing*, mitjançant una guia per estimular l'accés a les dades procedents de finançament públic<sup>35</sup>.

La Unió Europea (UE) considerà les dades científiques com un factor de progrés econòmic, ja que la reutilització de dades prèvies proporcionava una oportunitat d'innovació a un cost raonable<sup>36</sup>. D'aquesta manera, la directiva

31. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 17-23.

32. Torres-Salinas; Robinson-García; Cabezas-Clavijo, *op. cit.*, p. 178.

33. Alice Keefer, «La preservació de les dades de recerca i les agències de finançament de la R+D» [En línia], *Blok de BiD*, 11/10/2011. <<http://www.ub.edu/blokdebid/ca/node/129>> [Consulta: 06/06/2013].

34. UNESCO, *Charter on the Preservation of Digital Heritage* [En línia], Paris: UNESCO, 2003. <[http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html)> [Consulta: 06/06/2013].

35. OCDE, 2007.

36. Fernanda Peset, «Opinions del sector científic sobre la preservació d'informació» [en línia], *Blok de BiD*, 19/09/2012. <<http://www.ub.edu/blokdebid/ca/content/opinions-del-sector-cient%C3%ADfic-sobre-la-preservaci%C3%B3-de-la-informaci%C3%B3>> [Consulta: 06/06/2013].



2003/98/EC sobre reutilització de dades advocava per una major disponibilitat i incorporava les dades de recerca<sup>37</sup>. A més, una comunicació oficial de l'any 2007 sobre l'accés, difusió i preservació de la informació científica a l'era digital equiparava les publicacions i les dades de recerca<sup>38</sup>. Una comunicació posterior refermava la confiança en el potencial de les dades obertes per al creixement econòmic, l'avenç social, la transparència governamental i, naturalment, el progrés científic<sup>39</sup>.

A la vegada, la UE ha donat suport actiu a les dades científiques obertes mitjançant els seus programes marc d'R+D, ja sigui a través de projectes d'investigació o de cofinançament d'infraestructures i dipòsits. En aquest darrer sentit, cal destacar les iniciatives OpenAIRE<sup>40</sup> i OpenAIREplus, que recullen entre les seves finalitats el tractament de dades científiques, l'enllaç amb les publicacions i la creació de serveis addicionals per als científics. Així, el mes de maig de 2013 ha entrat en funcionament Zenodo<sup>41</sup> (Figura 3), un dipòsit digital obert impulsat per OpenAIRE i l'European Organization for Nuclear Research (CERN), amb el suport de la UE. Zenodo exemplifica una nova generació de dipòsits que impulsa la integració entre la producció científica, els conjunts de dades i la informació de finançament. Es tracta d'un servei senzill i innovador que permet compartir, distribuir, accedir i reutilitzar els resultats de recerca de tots els camps científics.

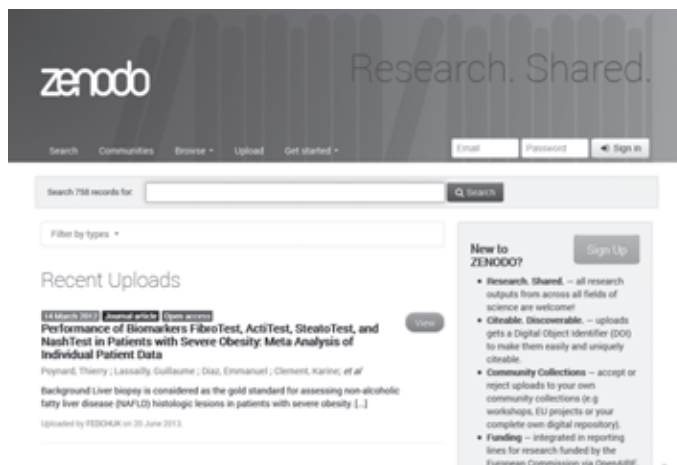


Figura 3. Zenodo

*Riding the wave: how Europe can gain from the rising tide of scientific data*<sup>42</sup> és un informe europeu sobre els beneficis i els costos de la posada en funcionament d'una infraestructura global i multidisciplinària per treballar, compartir i conservar dades científiques. A l'informe, els experts es plantegen preguntes sobre com incentivar la contribució dels diferents agents (sobretot, cal implicar els investigadors), com garantir la preservació a llarg termini, com establir el context i origen de dades o, per últim, com sufragar la infraestructura.

37. UE, «Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information» [En línia], *Official Journal of the European Union*, 31/12/2003. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF>> [Consulta: 06/06/2013].

38. Comissió Europea (CE), *Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee on scientific information in the digital age* [En línia]: *Access, dissemination and preservation*, COM(2007) 56 final, Brussels: Commission of the European Communities, 2007. <[http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/communication-022007\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/communication-022007_en.pdf)> [Consulta: 06/06/2013].

39. Comissió Europea (CE), *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* [En línia]: *Open data, an engine for innovation, growth and transparent governance*, COM(2011) 882 final, Brussels: European Commission, 2011. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>> [Consulta: 06/06/2013].

40. OpenAIRE: *Open Access Infrastructure for Research in Europe* ([www.openaire.eu](http://www.openaire.eu)) [Consulta: 06/06/2013].

41. Zenodo ([www.zenodo.org](http://www.zenodo.org)) [Consulta: 06/06/2013].

42. *Riding the wave: how Europe can gain from the rising tide of scientific data* [En línia], [S.l.]: European Union, 2010. <<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>> [Consulta: 06/06/2013].

L'any 2012 es van publicar els resultats d'una enquesta sobre el paper de la UE en l'accés i la preservació digital de la informació científica<sup>43</sup>. Entre d'altres elements, l'enquesta posava de manifest algunes barreres per accedir a les dades científiques, com ara la falta d'inversió en la creació i manteniment d'infraestructures, pocs incentius als investigadors per fer disponibles les dades i manca de polítiques nacionals.

A l'Estat espanyol, l'article 37 de la Ley 14/2011, de 1 de junio, de la ciencia, la tecnología y la innovación, impulsa el desenvolupament de dipòsits digitals d'accés obert per a les publicacions. Així mateix, es fixa l'obligació de dipositar en accés obert la versió final dels continguts publicats que hagin estat finançats majoritàriament amb fons públics, però enlloc s'esmenten explícitament les dades científiques. Tot i això, sembla evident que les agències finançadores han de començar a implantar mandats obligatoris d'accés obert a les dades obtingudes amb diner públic, un cop els protocols i les infraestructures tecnològiques estiguin prou consolidades<sup>44</sup>. En aquesta línia de treball, l'any 2012 la Fundación Española para la Ciencia y la Tecnología (FECYT) creà un grup de treball sobre «Depósito y Gestión de Datos en Acceso Abierto» en el marc del projecte RE-COLECTA, la primera iniciativa estatal per generar una infraestructura que impulsi la ciència oberta (*open science*).

Pel que fa a les editorials, les instruccions per als autors de revistes internacionals contempen sovint la necessitat, i a vegades l'obligació, de compartir les dades de les recerques publicades<sup>45</sup>. Una opció habitual és indicar a quin banc s'han de dipositar les dades, quins estàndards cal seguir o com s'han d'identificar unívocament els *datasets*. Si no existeixen dipòsits especialitzats, la mateixa revista pot facilitar l'emmagatzematge (per exemple, *Science* i, properament, *Nature* amb la plataforma *Scientific Data*). Si això tampoc no és possible, la revista té l'alternativa de reclamar



©istockphoto/Josef Kubicek

una disposició favorable per part dels autors a respondre peticions concretes. Ara bé, cal tenir present que l'adopció de polítiques editorials favorables al *data sharing* comporta algunes dificultats (per exemple, dades sensibles sobre pacients o fàrmacs), fet que pot repercutir en una rebaixa dels nivells d'exigència als autors. D'altra banda, els productors comercials de bases de dades comencen a incorporar grans conjunts de dades a la seva oferta de continguts. Així, Thomson Reuters va posar en funcionament l'octubre de 2012 el producte *Data Citation Index*<sup>46</sup>.

Respecte a les institucions científiques, la seva responsabilitat principal radica a fixar una política interna de dades, amb la creació de dipòsits corporatius, l'establiment de guies de bones pràctiques i estàndards, la formació del personal o la coordinació dels diferents serveis implicats (informàtica, biblioteca i serveis científics)<sup>47</sup>.

43. Comissió Europea (CE), *Online Survey on Scientific Information in the Digital Age* [En línia], Luxembourg: Publications Office of the European Union, 2012. <[http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/survey-on-scientific-information-digital-age\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf)> [Consulta: 06/06/2013].

44. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 42.

45. Torres-Salinas; Robinson-García; Cabezas-Clavijo, *op. cit.*, p. 180-181.

46. Data Citation Index ([wokinfo.com/products\\_tools/multidisciplinary/dci](http://wokinfo.com/products_tools/multidisciplinary/dci)) [Consulta: 06/06/2013].

47. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 8.

Quina és la posició de la comunitat científica respecte a *data sharing*? Com a productors i usuaris de les dades, un 60% dels investigadors es declaren favorables a compartir-les, però la major part no ho porta a la pràctica: menys del 20% emmagatzema dades en un dipòsit digital o les envia als editors de revista conjuntament amb els manuscrits<sup>48</sup>. Segons els investigadors, el principal obstacle per compartir és que no existeix una disponibilitat general de dades. A més, si estan disponibles, resulten difícils de localitzar o hi manca informació addicional per interpretar-les.

El projecte *Data Audit Framework* (DAF), finançat pel Joint Information Systems Committee (JISC) del Regne Unit, va pretendre esbrinar quin era el tractament de dades per part dels científics i com se'ls podia donar suport<sup>49</sup>. La iniciativa detectà que els científics no disposaven d'una política clara per conservar les dades, les quals es trobaven disperses en diversos suports físics i formats informàtics, sense catàlegs de col·leccions de dades. Això comportava pèrdues irreparables, cerques per localitzar dades o identificar versions, etc.

Com a primer pas, el projecte DAF plantejava implementar polítiques corporatives per crear i conservar dades, amb el suport de guies per afavorir bones pràctiques: com tractar diferents tipus de dades, com fer còpies a llarg termini, com crear plans de gestió, etc. Altres propostes eren crear dipòsits institucionals i formar els nous investigadors en tècniques d'administració i preservació de dades. Ara bé, una de les principals dificultats derivava del complex cicle de vida d'un projecte de recerca, en el qual investigadors diferents podien recopilar dades de fonts diverses i en formats múltiples<sup>50</sup>.

Segons els resultats del projecte DAF, en general els investigadors donaven poca importància a la reutilització posterior dels seus *datasets*. Malgrat això, cal tenir en compte que ells són els encarregats de proporcionar la informació bàsica per a la descripció de dades. Així, una descripció mínima ha de contemplar elements com el context de les dades (resum, objectius, metodologia de la investigació), descripció (contingut, estructura, terminologia, dates de creació i de modificació, versions, responsables), tecnologia (formats, estructura i nom dels fitxers, programes informàtics, sistemes d'emmagatzematge i seguretat) i, per últim, aspectes legals i polítiques d'accés<sup>51</sup>.

«Encara que hi ha una certa manca en els marcs tècnics i institucionals per regular la normalització de l'accés obert a les dades de recerca, de manera gradual estan emergint iniciatives internacionals que afavoreixen poder-les compartir»<sup>52</sup>. És necessària una anàlisi profunda sobre la identificació complexa dels drets d'autor d'objectes digitals que no són publicacions. Amb les possibilitats actuals de publicació en accés obert, els investigadors han d'establir drets i tipus de llicència perquè les dades puguin ser reutilitzades<sup>53</sup>: Creative Commons, Scientific Commons, Open Data Commons, etc.

## 4. Documentació científica i gestió de dades

En el futur serà imprescindible comptar amb dipòsits de dades de recerca on dipositar i compartir els resultats de les investigacions pagades amb fons públics. «¿Quién se encargará de normalizar los datos, de describirlos, de man-

48. Borrego, 11/04/2012, *op. cit.*

49. Daniel Torres-Salinas, «Hacia la gestión de datos de investigación en las universidades: la *Data Asset Framework*», *Anuario ThinkEPI*, vol. 4 (2010), p. 262-265.

50. Àngel Borrego, «Els reptes de la gestió de dades de recerca» [en línia], *Blok de BiD*, 06/11/2012. <<http://www.ub.edu/blokdebid/ca/content/els-reptes-de-la-gesti%C3%B3-de-dades-de-recerca>> [Consulta: 06/06/2013].

51. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 16.

52. García-García *et al.*, 2012, *op. cit.*

53. Arano *et al.*, *op. cit.*, p. 401.

tener y preservar esos repositorios de datos, de investigación, gubernamentales...? ¿También dejaremos que sean otros? ¿No es nuestra profesión la gestión de información, sea en bibliotecas, en centros de documentación o en empresas?»<sup>54</sup>. En aquest context, les qualitats organitzatives i de preservació dels bibliotecaris-documentalistes són essencials per afrontar l'allau de dades científiques. La idea de treballar amb informació factual a les biblioteques no és una novetat, ja que des dels anys seixanta del segle XX hi havia *data librarians*, sobretot als Estats Units d'Amèrica<sup>55</sup>. Aquest perfil presenta possibilitats excel·lents per establir relacions sòlides entre biblioteques i científics.

En un editorial de l'any 2009, la revista *Nature* ja reclamava l'equivalent digital de les biblioteques actuals per desenvolupar el *data sharing*, i apuntava directament cap a les biblioteques universitàries<sup>56</sup>. D'altra banda, la Ligue des Bibliothèques Européennes de Recherche - Association of European Research Libraries (LIBER) impulsà un *Working Group on e-Science*. L'informe final del grup reuneix deu recomanacions per a les biblioteques per començar les tasques vinculades amb les dades científiques<sup>57</sup>.

En general, la contribució de les biblioteques pot situar-se al voltant de tasques de descripció segons estàndards internacionals, polítiques de preservació digital i qüestions relacionades amb la reutilització i privacitat de dades<sup>58</sup>. Per tant, s'obren bones perspectives de futur professional en aquest camp, on els nous rols i funcions dels gestors d'informació poden ser els següents:

- Des de les biblioteques: vetllar per la qualitat de la informació associada als conjunts de dades, implementar estàndards i metadades per a l'intercanvi d'informa-

En general, la contribució de les biblioteques pot situar-se al voltant de tasques de descripció segons estàndards internacionals, polítiques de preservació digital i qüestions relacionades amb la reutilització i privacitat de dades

ció i ontologies que ajudin a organitzar les dades per a la posterior recuperació, impulsar la interoperabilitat dels sistemes, garantir la preservació adequada dels dipòsits digitals.

- En relació als investigadors: localitzar conjunts de dades rellevants per a les investigacions, actuar com a intermediaris amb els dipòsits, formar els investigadors, redactar i difondre guies de bones pràctiques, promoure l'ús de llicències alternatives al copyright.
- Respecte als governs, entitats científiques i organismes internacionals: col·laborar en l'elaboració de polítiques, directrius o normes; donar suport a la implementació i manteniment d'infraestructures.

En aquest sentit, els professionals han de gestionar de forma integral el cicle de vida de les dades d'investigació

54. Tony Hernández-Pérez, «Datos abiertos y repositorios de datos: ¿nuevo reto para los bibliotecarios» [en línia], *Notas ThinkEPI*, 4 febrero 2013. <<http://www.thinkepi.net/datos-abiertos-repositorios-datos-nuevo-reto-bibliotecarios>> [Consulta: 06/06/2013].

55. Luis Martínez-Urbe; Stuart Macdonald, «Un nuevo cometido para los bibliotecarios académicos: data curation», *El Profesional de la Información*, vol. 17, núm. 3, p. 275.

56. Torres-Salinas, *op. cit.*, p. 261.

57. Birte Christensen-Dalsgaard (coord.), *Ten recommendations for libraries to get started with research data management* [en línia], The Hague: LIBER, 2012. <[http://www.libereurope.eu/sites/default/files/WGSC\\_20120801.pdf](http://www.libereurope.eu/sites/default/files/WGSC_20120801.pdf)> [Consulta: 06/06/2013].

58. Miquel Plaza-Navas; Agnès Ponsati Obiols; Isabel Bernal Martínez; Mercedes Baquero Arribas, «La biblioteca de recerca i la seva aportació al desenvolupament de l'activitat R+D+i», *Item*, núm. 56 (2012), p. 22.

amb la finalitat de garantir la seva preservació, localització i reutilització (*data curation*)<sup>59</sup> (Figura 4). El cicle ha de contemplar la recollida, selecció, integració, anàlisi, publicació i conservació de dades en un dipòsit. El desplegament acadèmic de la matèria, així com dels projectes de recerca relacionats, són factors determinants.

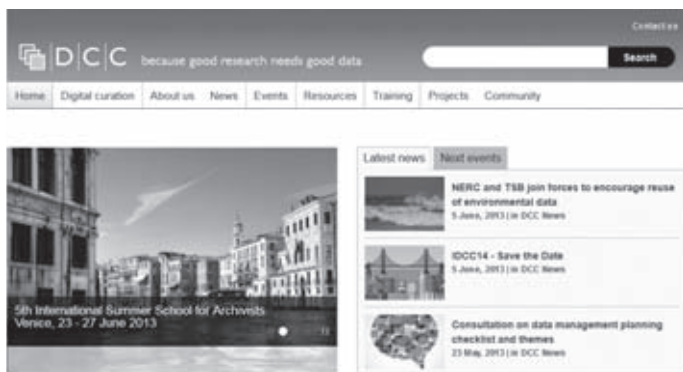


Figura 4. Digital Curation Centre (DCC)

La major part de les biblioteques considera la conservació de dades com a part del seu treball, tot i que encara són una minoria les que l'accepten per a l'emmagatzematge i la preservació<sup>60</sup>. D'altra banda, el projecte europeu ODE ha estudiat la problemàtica associada a la citació de les dades, les quals permeten atorgar crèdits als autors dels conjunts, conservar evidències empíriques de recerques prèvies i facilitar als investigadors la localització i reutilització de dades<sup>61</sup>. Mitjançant una enquesta a bibliotecaris, el projecte afronta qüestions com, per exemple, quines metadades cal recopilar a la citació o com s'han de citar les actualitzacions.

A l'Estat espanyol es detecta una manca de sol·licitud de serveis bibliotecaris en relació a les dades científiques<sup>62</sup>. Per tant, s'imposa la formació adequada dels bibliotecaris, seguida de la promoció intensiva de les seves capacitats per generar confiança entre els investigadors.

## 5. Experiències en gestió de dades científiques

### 5.1. Directoris internacionals

En el camp emergent de la gestió i preservació de dades científiques, resulta evident la necessitat de comptar amb instruments d'identificació sistemàtica dels dipòsits temàtics i institucionals. Per tant, es requereixen directoris per identificar de manera fàcil i fiable els dipòsits especialitzats amb el propòsit de facilitar la descoberta de nous conjunts de dades, la recombinació dels quals, des d'una perspectiva multidisciplinària, fomentarà la innovació i la rendibilitat de la inversió en ciència<sup>63</sup>.

**Databib**<sup>64</sup> (Figura 5) es defineix com un catàleg, registre, directori i bibliografia que s'orienta a identificar i localitzar dipòsits de dades de recerca. Originalment impulsat per l'Institute of Museum and Library Services (IMLS) dels Estats Units, el projecte es fonamenta en un sistema d'editors globals, en el qual no hi ha representants nacionals. La major part dels 575 dipòsits actuals corresponen a ciències biològiques (25,2%), ciències mediambientals (12,7%), ciències de la terra (10,6%), matemàtica i física (9,6%) i ciències mèdiques i de la salut (8,9%). Per àrees geogràfiques, es constata un domini aclaparador dels Es-

59. Digital Curation Centre (DCC) ([www.dcc.ac.uk](http://www.dcc.ac.uk)) [Consulta: 06/06/2013].

60. Borrego, 11/04/2012, *op. cit.*

61. Borrego, 06/11/2012, *op. cit.*

62. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 42.

63. García-García *et al.*, 2012, *op. cit.*

64. Databib ([databib.org](http://databib.org)) [Consulta: 06/06/2013].

tats Units (324 dipòsits) i del Regne Unit (67). La fitxa de cada dipòsit enumera, entre d'altres informacions, l'entitat responsable, país, data d'inici, resum, matèries i condicions d'accés, reutilització i dipòsit de dades.



Figura 5. Databib

**re3data.org:** Registry of Research Data Repositories<sup>65</sup> està finançat per la German Research Foundation (DFG). Es tracta d'un registre global amb la finalitat de promoure la cultura de compartir, donar accés i millorar la visibilitat dels dipòsits de dades científiques. La descripció de cada un dels 338 projectes actuals és exhaustiva, amb informació general, organitzacions responsables, polítiques, aspectes legals, estàndards, comentaris i versions.

**ODISEA:** International Registry on Research Data<sup>66</sup> és un projecte conjunt de cinc universitats espanyoles. Recull dipòsits internacionals que conserven conjunts de dades, materials addicionals d'articles i documents gràfics i multimèdia<sup>67</sup>. La classificació dels 177 dipòsits actuals es basa en les àrees de coneixement de la base de dades Essential Science Indicators (ESI), de Thomson Reuters. Les disciplines amb més dipòsits són biologia molecular i genètica, biologia i química, farmacologia i toxicologia i,

Les disciplines amb més dipòsits són biologia molecular i genètica, biologia i química, farmacologia i toxicologia i, finalment, ciències de les plantes i els animals.

finalment, ciències de les plantes i els animals. L'anàlisi geogràfica també confirma el predomini clar dels països anglosaxons. El registre de cada dipòsit conté camps de directori i d'anàlisi: entitat responsable, àrea geogràfica, tipus i quantitat de dades, format i grau de compliment del protocol OAI-PMH.

## 5.2. Iniciatives a l'Estat espanyol

A l'Estat espanyol les infraestructures dedicades a les dades de recerca encara es troben en un estat inicial i els projectes en funcionament són escassos. A continuació es comenten quatre de les iniciatives més destacades.

Des de l'any 2004, el Centro Nacional de Datos Polares (CNDP)<sup>68</sup> (Figura 6) recull les dades en brut i les de calibratges instrumentals que els investigadors espanyols

65. re3data.org: Registry of Research Data Repositories ([www.re3data.org](http://www.re3data.org)) [Consulta: 06/06/2013].

66. ODISEA: International Registry on Research Data ([odisea.ciepi.org](http://odisea.ciepi.org)) [Consulta: 06/06/2013].

67. García-García *et al.*, 2012.

68. Centro Nacional de Datos Polares ([hielo.igme.es](http://hielo.igme.es)) [Consulta: 06/06/2013].

subvencionats amb projectes estatals han d'enviar-hi obligatòriament en el termini màxim de tres mesos. El CNDP disposa d'un protocol de remissió, emmagatzematge i difusió de dades, les quals s'acaben integrant en el Global Change Master Directory (GCMD)<sup>69</sup>.



Figura 6. Centro Nacional de Datos Polares (CNDP)

**Digital.CSIC**<sup>70</sup> és el dipòsit digital institucional del Consejo Superior de Investigaciones Científicas (CSIC), que compta amb 72.642 registres (juny de 2013), dels quals 15 són conjunts de dades procedents de 9 instituts d'investigació de camps científics ben diferenciats: Estación Experimental de Aula Dei (EEAD), Instituto de Ciencias Marinas de Andalucía (ICMAN), Instituto de Estudios Documentales sobre Ciencia y Tecnología (IEDCYT), Instituto de Historia (IH), Instituto de Lenguas y Culturas del Mediterráneo y Oriente Próximo (ILC), Instituto de Química Física Rocasolano (IQFR), Instituto Mediterráneo de Estudios Avanzados (IMEDEA), Institut Milà i Fontanals (IMF) i, per últim, Instituto Pirenaico de Ecología (IPE).

En el camp de l'ensenyament superior, cal destacar la iniciativa de la Universitat Pompeu Fabra (UPF) d'integrar conjunts de dades a l'e-Repository<sup>71</sup>, el dipòsit digital de la institució<sup>72</sup>. Actualment presenta sis conjunts de dades procedents de l'Institut Universitari de Lingüística Aplicada (IULA) i tres del Departament de Ciències Polítiques i Socials.

La Biblioteca de Datos del Centro de Estudios Avanzados en Ciencias Sociales (CEACS)<sup>73</sup>, que forma part de l'Institut Juan March de Estudios e Investigaciones, proporciona una col·lecció extensa de dades i d'informació agregada i geogràfica. També proporciona serveis especialitzats i formació als investigadors sobre temes d'opinió pública, sistemes polítics, eleccions, enquestes socials, etc.

## 6. Conclusions

En un context general d'obertura de quantitats massives de dades, considerades com actius amb valor afegit, és imprescindible una gestió adequada per garantir la seva disponibilitat i preservació a llarg termini. Amb la finalitat de generar nous coneixements, «una óptima gestión de [...] datos favorece la innovación y el desarrollo de la misma, puesto que permitiría la explotación de datos de alta calidad (compartir-reutilizar)»<sup>74</sup>. Es tracta d'un domini emergent, amb bones perspectives de futur però també amb nombroses incògnites a resoldre. Per tant, els bibliotecaris-documentalistes han de col·laborar amb la resta d'agents implicats amb l'objectiu d'impulsar polítiques, directrius, infraestructures, estàndards, guies de bones pràctiques i accions de formació.

69. Global Change Master Directory ([gcmd.nasa.gov](http://gcmd.nasa.gov)) [Consulta: 06/06/2013].

70. Digital.CSIC ([digital.csic.es](http://digital.csic.es)) [Consulta: 06/06/2013].

71. e-Repository ([repositori.upf.edu](http://repositori.upf.edu)) [Consulta: 06/06/2013].

72. Arano et al., *op. cit.*, p. 387.

73. Biblioteca de Datos del Centro de Estudios Avanzados en Ciencias Sociales ([www.march.es/ceacs/biblioteca/datalib](http://www.march.es/ceacs/biblioteca/datalib)) [Consulta: 06/06/2013].

74. González Copeiro; Serrano-Muñoz, *op. cit.*, p. 15.



©istockphoto/Sergey Skleznev

En el context del *big data*, l'àmbit de la investigació científica és un dels principals productors de dades. Les dades científiques tenen caràcter heterogeni i s'utilitzen com a matèria primera dels resultats, sobretot publicats en forma de documents. Cal mantenir el lligam de les dades amb les publicacions i possibilitar futures investigacions a partir de la reutilització i combinació de múltiples *data-sets*. En general, els investigadors es mostren favorables al *data sharing* però no el porten a la pràctica degut a la manca d'incentius, d'instruments tecnològics i de coneixements i habilitats sobre la matèria.

«Junto a una investigación de excelencia debe fomentarse también una gestión de los datos de excelencia, que permita su acceso óptimo y garantice su reutilización futura»<sup>75</sup>. Per tant, en un futur immediat els bibliotecaris-documentalistes hauran de tenir un paper rellevant en la gestió integral del cicle de les dades científiques, des de la creació fins a la preservació. El desplegament acadèmic i professional del *data curation* facilitarà establir una relació

directa amb els investigadors i impulsar la promoció dels serveis bibliotecaris. En darrer terme, l'impacte del *data sharing* s'haurà de traduir en avenços científics i en més transparència i eficiència del sistema de recerca, però també en una major consolidació dels serveis d'informació i documentació científica.

## Agraïments

A Mercè Martí i Robert Latorre, companys de l'Observatori de la Recerca (OR-IEC), per la revisió del text. A Netta Azgad, per la col·laboració en les traduccions.

75. Torres-Salinas, *op. cit.*, p. 265.



## Bibliografía

ARANO, Silvia; MARTÍNEZ, Gemma; LOSADA, Marina; VILLEGAS, Marta; CASALDÁLIGA, Anna; BEL, Núria. «La comunidad “Recursos y datos primarios” de la Universitat Pompeu Fabra: los repositorios institucionales como infraestructuras científicas: estudio de caso». *Revista Española de Documentación Científica*, vol. 34, núm. 3 (2011), p. 385-407.

BERNERS-LEE, Tim. 5 *Star Open Data* [En línea]. <<http://5stardata.info>> [Consulta: 06/06/2013].

Biblioteca de Datos del Centro de Estudios Avanzados en Ciencias Sociales [En línea]. <<http://www.march.es/ceacs/biblioteca/datalib>> [Consulta: 06/06/2013].

BORGMAN, Christine L. «Data, disciplines, and scholarly publishing». *Learned Publishing*, vol. 21, núm. 1 (2008), p. 29-38.

BORREGO, Àngel. «Els reptes de la gestió de dades de recerca» [En línea]. *Blok de BiD*, 06/11/2012. <<http://www.ub.edu/blokdebid/ca/content/els-reptes-de-la-gesti%C3%B3-de-dades-de-recerca>> [Consulta: 06/06/2013].

BORREGO, Àngel. «Enriquir les publicacions amb dades empíriques» [En línea]. *Blok de BiD*, 11/04/2012. <<http://www.ub.edu/blokdebid/ca/content/enriquir-les-publicacions-amb-dades-emp%C3%ADriques-0>> [Consulta: 06/06/2013].

Centro Nacional de Datos Polares [En línea]. <<http://hielo.igme.es>> [Consulta: 06/06/2013].

CHRISTENSEN-DALSGAARD, Birte (coord.). *Ten recommendations for libraries to get started with research data management* [En línea]. The Hague: LIBER, 2012. <[http://www.libereurope.eu/sites/default/files/WGSC\\_20120801.pdf](http://www.libereurope.eu/sites/default/files/WGSC_20120801.pdf)> [Consulta: 06/06/2013].

COMISSIÓ EUROPEA (CE). *Communication from the Commission to the European Parliament, the Council and the Euro-*

*pean Economic and Social Committee on scientific information in the digital age* [En línea]: *Access, dissemination and preservation*. COM(2007) 56 final. Brussels: Commission of the European Communities, 2007. <[http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/communication-022007\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/communication-022007_en.pdf)> [Consulta: 06/06/2013].

COMISSIÓ EUROPEA (CE). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* [En línea]: *Open data, an engine for innovation, growth and transparent governance*. COM(2011) 882 final. Brussels: European Commission, 2011. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>> [Consulta: 06/06/2013].

COMISSIÓ EUROPEA (CE). *Online Survey on Scientific Information in the Digital Age* [En línea]. Luxembourg: Publications Office of the European Union, 2012. <[http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/survey-on-scientific-information-digital-age\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/survey-on-scientific-information-digital-age_en.pdf)> [Consulta: 06/06/2013].

Databib [En línea]. <<http://databib.org>> [Consulta: 06/06/2013].

Data Citation Index [En línea]. <[http://wokinfo.com/products\\_tools/multidisciplinary/dci](http://wokinfo.com/products_tools/multidisciplinary/dci)> [Consulta: 06/06/2013].

Digital.CSIC [En línea]. <<http://digital.csic.es>> [Consulta: 06/06/2013].

Digital Curation Centre (DCC) [En línea]. <<http://www.dcc.ac.uk>> [Consulta: 06/06/2013].

e-Repository [En línea]. <<http://repositori.upf.edu>> [Consulta: 06/06/2013].

FERRER-SAPENA, Antonia; SÁNCHEZ-PÉREZ, Enrique A. «Open data, big data: ¿hacia dónde nos dirigimos» [En línea]. *Notas ThinkEPI*, 1 febrero 2013. <<http://www.thinkepi.net/open-data-big-data-hacia-donde-nos-dirigimos>> [Consulta: 06/06/2013].

GARCÍA-GARCÍA, Alicia; RODRÍGUEZ-GAIRÍN, Josep-Manuel; SAORÍN, Tomás; GONZÁLEZ, Luis-Millán; GARCÍA-MASSÓ, Xavi; FERRER-SAPENA, Antonia; PESET, Fernanda. «ODISEA: International Registry on Research Data» [En línia]. *BiD: textos universitaris de biblioteconomia i documentació*, núm. 29 (2012). <<http://www.ub.edu/bid/29/garcia2.htm>> [Consulta: 06/06/2013].

GenBank [En línia]. <<http://www.ncbi.nlm.nih.gov/genbank>> [Consulta: 06/06/2013].

Global Change Master Directory [En línia]. <<http://gcmd.nasa.gov>> [Consulta: 06/06/2013].

GONZÁLEZ COPEIRO, Cristina; SERRANO-MUÑOZ, Jordi (coord.). *La conservación y reutilización de los datos científicos en España: informe del grupo de trabajo de buenas prácticas* [En línea]. Madrid: FECYT, 2012. <[http://www.recolecta.net/buscador/documentos/informe\\_datos\\_cientificos\\_en\\_esp.pdf](http://www.recolecta.net/buscador/documentos/informe_datos_cientificos_en_esp.pdf)> [Consulta: 06/06/2013].

KEEFER, Alice. «La preservació de les dades de recerca i les agència de finançament de la R+D» [En línia]. *Blok de BiD*, 11/10/2011. <<http://www.ub.edu/blokdebid/ca/node/129>> [Consulta: 06/06/2013].

HERNÁNDEZ-PÉREZ, Tony. «Datos abiertos y repositorios de datos: ¿nuevo reto para los bibliotecarios» [En línia]. *Notas ThinkEPI*, 4 febrero 2013. <<http://www.thinkepi.net/datos-abiertos-repositorios-datos-nuevo-reto-bibliotecarios>> [Consulta: 06/06/2013].

MARTÍNEZ-URIBE, LUIS; MACDONALD, Stuart. «Un nuevo comedido para los bibliotecarios académicos: data curation». *El Profesional de la Información*, vol. 17, núm. 3, p. 273-280.

MELERO, Remedios. «Una plenamar de dades» [En línia]. *Blok de BiD*, 24/11/2010. <<http://www.ub.edu/blokdebid/ca/content/una-plenamar-de-dades>> [Consulta: 06/06/2013].

NATIONAL INSTITUTES OF HEALTH (NIH). *NIH Data Sharing Policy* [En línia]. Bethesda, Maryland: NIH, 2007. <[http://grants.nih.gov/grants/policy/data\\_sharing](http://grants.nih.gov/grants/policy/data_sharing)> [Consulta: 06/06/2013].

ODISEA: International Registry on Research Data [En línia]. <<http://odisea.ciepi.org>> [Consulta: 06/06/2013].

OpenAIRE: Open Access Infrastructure for Research in Europe [En línia]. <<http://www.openaire.eu>> [Consulta: 06/06/2013].

ORGANITZACIÓ PER A LA COOPERACIÓ I EL DESENVOLUPAMENT ECONÒMIC (OCDE). *OECD Principles and Guidelines for Access to Research Data from Public Funding* [En línia]. Paris: OCDE, 2007. <<http://www.oecd.org/science/sci-tech/38500813.pdf>> [Consulta: 06/06/2013].

OPEN KNOWLEDGE FOUNDATION (OKFN). *The Open Data Handbook*. Cambridge: OKFN, 2012. <<http://opendata-handbook.org>> [Consulta: 06/06/2013].

PESET, Fernanda. «Opinions del sector científic sobre la preservació d'informació» [En línia]. *Blok de BiD*, 19/09/2012. <<http://www.ub.edu/blokdebid/ca/content/opinions-del-sector-cient%C3%ADfic-sobre-la-preservaci%C3%B3-de-la-informaci%C3%B3>> [Consulta: 06/06/2013].

PESET, Fernanda; FERRER-SAPENA, Antonia; SUBIRATS-COLL, Imma. «Open data y linked open data: su impacto en el área de bibliotecas y documentación». *El Profesional de la Información*, vol. 20, núm. 2 (2011), p. 165-173.

PIWOWAR, Heather A.; DAY, Roger S.; FRIDSMA, Douglas B. «Sharing Detailed Research Data Is Associated with Increased Citation Rate» [En línia]. *PLoS ONE*, vol. 2, núm. 3, e308. <<http://dx.doi.org/10.1371/journal.pone.0000308>> [Consulta: 06/06/2013].

Protein Data Bank (PDB) [En línia]. <<http://www.rcsb.org/pdb>> [Consulta: 06/06/2013].

RAMASAMY, A.; MONDRY, A.; HOLMES, C.C.; ALTMAN, D.G. «Key Issues in Conducting a Meta-Analysis of Gene Expression Microarray Datasets» [En línia]. *PLoS Medicine*, vol. 5, núm. 9, e184. <<http://dx.doi.org/10.1371/journal.pmed.0050184>> [Consulta: 06/06/2013].

re3data.org: Registry of Research Data Repositories [En línia]. <<http://www.re3data.org>> [Consulta: 06/06/2013].

PLAZA-NAVAS, Miquel; PONSATI OBIOLS, Agnès; BERNAL MARTÍNEZ, Isabel; BAQUERO ARRIBAS, Mercedes. «La biblioteca de recerca i la seva aportació al Desenvolupament de l'activitat R+D+i». *Item*, núm. 56 (2012), p. 13-32.

*Riding the wave: how Europe can gain from the rising tide of scientific data* [En línia]. [S.l.]: European Union, 2010. <<http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>> [Consulta: 06/06/2013].

SWAN, Alma; BROWN, Sheridan. *To Share or not to Share: publication and Quality Assurance of Research Data Outputs* [En línia]. London: Research Information Network, 2008. <<http://www.rin.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs>> [Consulta: 06/06/2013].

TORRES-SALINAS, Daniel. «Compartir datos (*data sharing*) en ciencia: contexto de una oportunidad». *Anuario ThinkEPI*, vol. 4 (2010), p. 258-261.

TORRES-SALINAS, Daniel. «Hacia la gestión de datos de investigación en las universidades: la *Data Asset Framework*». *Anuario ThinkEPI*, vol. 4 (2010), p. 262-265.

TORRES-SALINAS, Daniel; ROBINSON-GARCÍA, Nicolás; CABEZAS-CLAVIJO, Álvaro. «Compartir los datos de investigación en ciencia: introducción al data sharing». *El Profesional de la Información*, vol. 21, núm. 2, p. 173-184.

UNESCO. *Charter on the Preservation of Digital Heritage* [En línia]. Paris: UNESCO, 2003. <[http://portal.unesco.org/en/ev.php-URL\\_ID=17721&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html)> [Consulta: 06/06/2013].

UNIÓ EUROPEA (UE). «Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information» [En línia]. *Official Journal of the European Union*. 31/12/2003. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF>> [Consulta: 06/06/2013].

Zenodo [En línia]. <<http://www.zenodo.org>> [Consulta: 06/06/2013].

# Les dades de la recerca: de la foscor a la claror

## **Ignasi Labastida**

CRAI-Unitat de Recerca, Universitat de Barcelona  
*ilabastida@ub.edu*

### **Resum:**

Un cop consolidades les polítiques d'accés obert en els resultats de la recerca en publicacions periòdiques, han començat a sorgir accions per obrir altres resultats de recerca com ara les dades obtingudes o emprades. Tradicionalment, aquestes dades només eren accessibles a un grup reduït d'investigadors, però ara es planteja fer-les públiques i oferir-les a qualsevol que les vulgui utilitzar per reproduir la recerca o verificar-ne els resultats. Tanmateix, aquesta obertura de les dades comporta molts dubtes i moltes preguntes per part dels investigadors i les institucions de recerca. No només cal superar les reticències tradicionals, sinó que també cal determinar quines dades s'han de compartir, com s'ha de fer, on s'han de dipositar... En aquest procés, les biblioteques de recerca tenen una nova oportunitat per ajudar a respondre aquestes preguntes i entomar el repte de col·laborar a difondre els resultats de la recerca en obert.

### **Paraules clau:**

dades de recerca, accés obert, biblioteques de suport a la recerca, dades obertes

## Los datos de la investigación: de la oscuridad a la claridad

### Resumen:

Una vez consolidadas las políticas de acceso abierto en los resultados de la investigación en publicaciones periódicas, han empezado a surgir acciones para abrir otros resultados de la investigación como los datos obtenidos o utilizados. Tradicionalmente, estos datos sólo eran accesibles a un grupo reducido de investigadores, pero ahora se plantea hacerlos públicos y ofrecerlos a cualquiera que los quiera utilizar para reproducir la investigación o verificar los resultados. Sin embargo, esta apertura de los datos comporta muchas dudas y muchas preguntas por parte de los investigadores y las instituciones de investigación. No sólo hay que superar las reticencias tradicionales, sino que también hay que determinar qué datos hay que compartir, cómo hay que hacerlo, dónde hay que depositarlos... En este proceso, las bibliotecas de investigación tienen una nueva oportunidad para ayudar a responder estas preguntas y asumir el reto de colaborar a difundir los resultados de la investigación en abierto.

### Palabras clave:

datos de investigación, acceso abierto, bibliotecas de apoyo a la investigación, datos abiertos

## Research data: from darkness to clarity

### Abstract:

Once open access policies on the publication of the research results have been consolidated, further actions to open more results arisen from research have begun to appear. Among those actions we find policies towards data used or obtained in a research process. Traditionally, those data were only available to a small group of researchers, but now it has been considered to provide them to the public and therefore anyone could use them to reproduce or verify the results. Nevertheless, the opening of data brings many doubts and many questions from researchers and from research institutions. Not only to overcome the traditional reluctance, but also to determine which data to share, how to share them, where to post them ... In this new opening process, research libraries have a new opportunity to help to answer all those questions and take the challenge to collaborate in disseminating research results openly.

### Keywords:

data research, open access, research libraries, open data

## 1. Introducció

Les dades que es fan servir en l'àmbit de la recerca han estat tradicionalment accessibles només a un grup reduït de persones. Els investigadors han estat molt gelosos de les seves dades i normalment les han tingut tancades [1]. Les raons per no compartir-les han estat variades: des de la por a perdre la primícia o que siguin mal interpretades fins a la necessitat de dedicar temps a polir-les o preparar-les per fer-les públiques<sup>1</sup>. Tanmateix, els darrers anys aquesta tendència està canviant lentament. Els investigadors han vist com durant la primera dècada del segle XXI s'ha estès el moviment de l'accés obert que aposta per fer accessible de manera gratuïta la publicació dels resultats de la recerca i permetre'n la difusió sense barreres tecnològiques ni legals<sup>2</sup>. Les institucions públiques i privades que financen la recerca han establert una sèrie de polítiques<sup>3</sup> imposant que aquests resultats es posin a l'abast de tothom en un temps no superior a un any després de ser publicats en revistes de referència. El pas següent ha estat implementar aquestes polítiques a altres resultats de la recerca: les dades. Per tant, aquest canvi de tendència pel que fa a la difusió de les dades no ha vingut principalment per part dels investigadors sinó per qui en finança la recerca o per part de les mateixes publicacions. El nombre de revistes que demanen als investigadors la difusió de les dades obtingudes o emprades en una recerca abans de decidir si aproven o no la publicació d'un resultat ha augmentat els darrers anys [2]. Aquests canvis comporten l'aparició de dubtes i preguntes sobre quines dades cal oferir, on s'han d'ubicar, com s'han de citar o si hi haurà algun reconeixement específic per als investigadors que ho facin com en el cas de les publicacions, o quins beneficis comporta. Sobre aquest darrer punt, els beneficis, s'han escrit diversos estudis. N'hi ha alguns que apunten que compartir les dades augmenta el nombre de



©istockphoto/sonyae

citacions de les publicacions relacionades [3] i d'altres que apunten també als beneficis de la institució per aconseguir noves col·laboracions o nous fons [4].

En aquest article revisarem els aspectes a tenir en compte a l'hora de difondre les dades de recerca de manera oberta. Veurem què s'està fent en l'àmbit internacional i en l'estatal. I finalment, analitzarem el paper que hi pot jugar una biblioteca o un centre de recursos d'una institució de recerca. Comencem per veure què entenem per dada de recerca.

### 1.1 Les dades de la recerca

No és fàcil parlar de dades de recerca perquè cadascú pot tenir la seva interpretació del que és una dada segons la disciplina. Per a un meteoròleg una dada pot ser la tempe-

1. Hi ha un document dinàmic a la xarxa que intenta donar resposta a alguna d'aquestes raons: Christopher Gutteridge, Alexander Dutton, «Concerns about opening up data, and responses which have proved effective» [En línia]. <[https://docs.google.com/document/d/1nDtHpnIDTY\\_G32EMJni-XaOGBufjHCck4VC9WGO7jK4/edit#](https://docs.google.com/document/d/1nDtHpnIDTY_G32EMJni-XaOGBufjHCck4VC9WGO7jK4/edit#)> [Consulta: 1/7/2013].
2. «Iniciativa de Budapest per a l'accés obert» [En línia], febrer de 2002. <<http://www.budapestopenaccessinitiative.org/translations/catalan-translation>> [Consulta: 1/7/2013].
3. Es poden consultar aquestes polítiques a [www.sherpa.ac.uk/juliet](http://www.sherpa.ac.uk/juliet) [Consulta: 1/7/2013].

ratura en un determinat lloc i en un determinat moment, però per a un metge una dada pot ser una imatge obtinguda en un aparell de ressonància magnètica. De fet, si consultem el *Gran Diccionari de la Llengua Catalana* ens trobem com a primera definició de dada: «Qualsevol informació que facilita la formació d'una idea, d'un judici, d'una conclusió»<sup>4</sup>.

Aquesta varietat de significats té la seva rellevància per tal d'establir plans de gestió de dades o per determinar l'estat legal d'una dada, com veurem a continuació.

Podem classificar les dades d'acord amb la seva procedència, i llavors parlarem de dades primàries i dades secundàries. Les dades primàries són aquelles dades que obtenim per l'observació o l'experimentació; en canvi, les dades secundàries són aquelles que han estat tractades i que poden ser fruit de la combinació de diversos conjunts de dades primàries.

El que tractarem en aquest article serà bàsicament les dades primàries, que podem separar en dades observacionals fenomenològiques, experimentals o computacionals, segons quin sigui el seu origen [5].

## 1.2 Estat legal de les dades

Com ja s'ha dit anteriorment, l'estat legal de les dades dependrà del que entenguem com a dada. En el cas que considerem com a dada un document, una imatge o un registre sonor, haurem de tenir en compte que poden gaudir de protecció per la propietat intel·lectual i, per tant, caldrà tenir en compte qui n'és el titular i quins usos poden estar permesos. El titular o beneficiari dels drets de propietat intel·lectual pot ser l'investigador, el centre de recerca o

Quan la dada no tingui la consideració per ser protegida per la propietat intel·lectual, com és el cas de moltes dades primàries, el fet de fer-ne una compilació o una base de dades sí que genera uns drets.

el finançador de la recerca, segons estigui estipulat per llei, per conveni o acord, o segons l'establert en l'ajut per realitzar la recerca. Cal establir clarament qui és o serà el titular d'aquestes dades des de bon principi per poder determinar qui decidirà com es difondran<sup>5</sup>.

Quan la dada no tingui la consideració per ser protegida per la propietat intel·lectual, com és el cas de moltes dades primàries, el fet de fer-ne una compilació o una base de dades sí que genera uns drets. Aquesta consideració pot variar segons les lleis vigents a cada país [6]. A la Unió Europea, una base de dades pot ser protegida com una obra artística, científica o literària sempre que la selecció dels continguts o la seva disposició sigui original. I en el cas que no ho sigui, el sol fet d'invertir temps, diners o energia en la creació de la base dades també atorga un dret al seu creador. Aquest dret és l'anomenat dret *sui generis*<sup>6</sup>, que permet al creador de la base dades prohi-

- «Dada» a Grup Enciclopèdia Catalana, *Gran Diccionari de la Llengua Catalana* [En línia]. Barcelona: 1a edició, 1998. <<http://www.diccionari.cat/lexicx.jsp?GECART=0157101>> [Consulta: 1/7/2013].
- Un exemple de política de dades el trobem al projecte IceCube ([icecube.wisc.edu/science/data](http://icecube.wisc.edu/science/data)) [Consulta: 1/7/2013].
- Introduït per la *Directiva 96/9/CE del Parlamento Europeu y del Consejo, de 11 de marzo de 1996, sobre la protección jurídica de las bases de datos* [En línia]. <[http://eur-lex.europa.eu/smartapi/cgi/sga\\_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=en&numdoc=31996L0009&model=guichett](http://eur-lex.europa.eu/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=en&numdoc=31996L0009&model=guichett)> [Consulta: 10/8/2013].

bir-ne l'extracció o la reutilització d'una part substancial o la totalitat dels continguts. Per tant, com en el cas anterior, on els continguts sí que eren protegibles, també caldrà establir qui té la titularitat d'aquests drets.

## 2. Aspectes a tenir en compte a l'hora de difondre i compartir les dades

A l'hora de compartir les dades de la recerca no només cal tenir en compte què entenem per dades i quin és el seu estat legal, sinó que també cal avaluar altres aspectes com són: la selecció de les dades, la descripció de les dades, com indicar-ne l'obertura, la seva conservació i preservació i, finalment, la citació de les dades

### 2.1 Selecció de les dades

El nombre de dades que tenim al nostre abast creix de manera desmesurada dia a dia. Ens trobem davant de l'anomenat «diluvi» o «allau de dades» [7] i, per tant, és lògic plantejar-se quines dades cal compartir. Cal fer una preselecció de les dades amb les quals s'ha treballat al llarg d'una investigació? O s'haurien d'exposar totes les dades per tal que qualsevol persona pogués verificar-les i reutilitzar-les? No hi ha una resposta única per a tots els casos i totes les disciplines. A més, la determinació de quines dades cal difondre pot estar fixada per una política de la institució, agència o administració que finança la recerca. Tanmateix, es poden establir algunes pautes per discernir quines dades cal difondre [8]. Entre els criteris que es poden seguir hi ha la importància de les dades i del context on han estat obtingudes o observades; el grau de reutilització de les dades, és a dir, si tenen un format que en permetrà la reutilització posterior; la relació amb altres dades importants, o el fet que les dades són úniques i

difícilment es podran tornar a recollir. També hi ha un altre punt clau, el de la dimensió de les dades i el cost que pot tenir conservar-les. Per aquesta raó, és important també determinar si les dades seleccionades mantindran la seva validesa i importància al llarg del temps.

És important seleccionar les dades a priori, és a dir, determinar al principi de la recerca quines dades seran les que es difondran i es compartiran. Aquest punt serà l'inici d'un pla de gestió de dades que és necessari elaborar en qualsevol recerca si al final volem tenir-les en obert. Tampoc no podem oblidar que en alguns àmbits s'utilitzen dades personals o confidencials i, per tant, caldrà tenir previst un protocol o procediment per anonimitzar les dades o excloure-les de la difusió.

### 2.2 Descripció de les dades

Un cop determinades les dades que es compartiran, caldrà descriure-les perquè posteriorment es puguin citar, identificar o conèixer com han estat obtingudes. Hi ha diferents models per descriure les dades, però potser caldria veure aquells que han estat creats especialment per a les dades com, per exemple, el model proposat per la Data Documentation Initiative (DDI)<sup>7</sup> per al camp de les ciències socials.

Com a informació bàsica caldria donar un títol, el nom de les persones que han obtingut les dades, un identificador, un nombre de matèries, la situació legal, les dates més rellevants, com ara la d'obtenció o processament de les dades, i el format. També caldrà indicar, si és possible, qui ha finançat la recerca, on s'han obtingut les dades o quina metodologia s'ha seguit. El nombre de metadades que cal per descriure un conjunt de dades vindrà donat molts cops pel lloc on s'allotjaran i que marcarà quines són obligatòries i quines són recomanables, a més d'oferir estàndards ja establerts.

7. Data Documentation Initiative ([www.ddialliance.org](http://www.ddialliance.org)) [Consulta: 1/7/2013].



## 2.3 L'obertura de les dades

Per obrir les dades caldrà indicar-ho, ja que el marc legal actual estableix que no dir res equival a reservar-se tots els drets<sup>8</sup> i, per tant, haver de demanar permís per utilitzar l'obra o creació, llevat dels casos previstos en la llei. Aquests casos són els límits o excepcions dels drets de propietat intel·lectual als quals podem recórrer en determinades condicions en el marc de la recerca, però són molt detallats.

Així doncs, si volem compartir les dades caldrà indicar-ho mitjançant algun tipus d'avís legal. Possiblement el mitjà més estès actualment són les llicències per a continguts oberts. Aquestes llicències són textos legals mitjançant els quals els autors o titulars d'una obra o creació, de la qual ostenten els drets, només se'n reserven alguns i permeten, a priori, determinats usos amb algunes condicions. De llicències d'aquest tipus n'hi ha moltes, però possiblement el sistema més estès actualment és el de Creative Commons. Quan les dades que vulguem obrir puguin ser protegides per propietat intel·lectual com una obra artística, científica o literària podem fer servir alguna de les sis llicències que ofereix Creative Commons. No obstant això, si volem complir estrictament la definició d'accés obert caldria utilitzar la llicència de Reconeixement (CC BY)<sup>9</sup>, que permet qualsevol tipus d'ús sense restriccions sempre que se citi l'autoria i sigui visible la indicació de la llicència.

En el cas que l'objecte protegible per propietat intel·lectual no siguin les dades sinó la base de dades, les sis llicèn-

cies estàndard de Creative Commons només serviren en el cas que la base de dades fos original, perquè en el text legal hi ha una renúncia explícita a qualsevol *dret sui generis*<sup>10</sup>. Aquest fet implica que no té sentit posar una llicència d'aquest tipus en moltes de les bases de dades de recerca que es generen, ja que no tenen el grau d'originalitat suficient per ser protegides més enllà d'aquest *dret sui generis* aplicable. Així doncs, decidir quina llicència s'adapta millor a una base de dades no és fàcil i, de fet, a Creative Commons es va obrir una discussió sobre això<sup>11</sup> que va desembocar en l'elaboració d'una nova llicència, CC Zero (CC0)<sup>12</sup>, i a l'establiment d'un protocol per implementar l'accés obert a dades<sup>13</sup>. La llicència CC0 és, de fet, una renúncia a exercir qualsevol dret per part del titular de la propietat intel·lectual. Aquesta llicència va més enllà de la llicència de Reconeixement (CC BY) perquè ni tan sols requereix el reconeixement d'autoria, segons la legislació aplicable, sinó que demana que se segueixin les normes de citació de la comunitat científica. Aquesta llicència o renúncia és la que actualment es fa servir per les metadades del projecte Europeana<sup>14</sup>.

La fundació Open Knowledge també ofereix unes eines legals per obrir les dades mitjançant el seu projecte Open Data Commons<sup>15</sup>. Hi ha tres llicències dissenyades per a bases de dades, i que són compatibles amb el protocol de Creative Commons esmentat anteriorment. Al contrari que les sis llicències genèriques de Creative Commons, aquestes llicències han estat creades per tenir en compte les peculiaritats de les bases de dades, com per exemple el *dret sui generis* aplicable a la Unió Europea. Una de les tres llicències és molt similar a la llicència CC0, mentre

8. L'article 1 de l'actual llei de propietat intel·lectual vigent a l'Estat espanyol estableix que no cal cap tràmit per gaudir dels drets: «La propietat intel·lectual d'una obra literària, artística o científica correspon a l'autor per l'únic fet de la creació». Real Decreto Legislativo 1/1996, de 12 de abril, por el que se aprueba el Texto Refundido de la Ley de Propiedad Intelectual, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia, *Boletín oficial del Estado*, 22 de abril de 1996, n.º. 97, p. 14369-14396.
9. OASPA, «Why CC BY?» [En línia]. <<http://oaspa.org/why-cc-by>> [Consulta: 1/7/2013].
10. Cal tenir en compte que als EUA el *dret sui generis* per a bases de dades no està previst.
11. Creative Commons, «Data» [En línia]. <<http://wiki.creativecommons.org/Data>> [Consulta: 1/7/2013].
12. Creative Commons, «About CC0 - "No Rights Reserved"» [En línia]. <<http://creativecommons.org/about/cc0>> [Consulta: 1/7/2013].
13. Science Commons, «Protocol for Implementing Open Access Data» [En línia]. <<http://sciencecommons.org/projects/publishing/open-access-data-protocol>> [Consulta: 1/7/2013].
14. Europeana, «Data Exchange Agreement» [En línia], juliol 2012. <<http://pro.europeana.eu/web/guest/data-exchange-agreement>> [Consulta: 10/8/2013].
15. Open Data Commons ([opendatacommons.org](http://opendatacommons.org)) [Consulta: 1/7/2013].



©istockphoto/tommytucker

que les altres dues inclouen el reconeixement d'autoria, en un cas, i la condició *copyleft* en l'altre, que requereix mantenir la llicència en qualsevol adaptació que es faci de la base de dades, seguint la filosofia del programari lliure.

Un altre referent per entendre com obrir les dades són els anomenats *Principis de Panton*<sup>16</sup>, un document elaborat l'any 2010 al Regne Unit per un grup d'experts en dades obertes, i que estableix els principis per considerar unes dades com a obertes.

## 2.4 Conservació i preservació

Quan les dades ja han estat seleccionades, identificades i obertes cal col·locar-les en un lloc on estiguin disponibles i localitzables de manera perdurable. Generalment no és convenient deixar-les en pàgines personals, sinó que cal ubicar-les en repositoris dedicats a aquest efecte o que al-

menys permetin la interconnexió amb d'altres mitjançant protocols estàndard.

A l'hora de decidir en quin repositori cal dipositar les dades reapareix el debat sobre el repositoris locals i els globals. Cal crear repositoris locals de dades? O bé cal treballar a nivells superiors? Una primera opció seria dipositar les dades en el repositori institucional on s'ha dut a terme la recerca<sup>17</sup>. Tanmateix, el cost de preservar aquestes dades i el nombre creixent de les dades pot fer inviable aquesta primera solució i, per tant, s'hauria d'optar per opcions consorciades a un nivell superior [9]. En aquest segon cas ens trobem amb infraestructures temàtiques, és a dir, repositoris per a àmbits o disciplines de recerca que permeten una adaptació de les metadades a les necessitats de cada camp. Una altra possibilitat és crear repositoris nacionals o continentals que estiguin connectats amb altres de similars i que permetin una visió més global amb subdivisions temàtiques especialitzades.

16. El text original en anglès i la traducció al català es troben disponibles a Panton Principles, «Principis per obrir les dades científiques» [En línia]. <<http://pantonprinciples.org>> [Consulta: 1/7/2013].

17. Segons el registre de repositoris d'accés obert (*roar.eprints.org*), a principis de juliol de 2013 hi havia uns 3.400 repositoris al món.

Algunes d'aquestes solucions ja s'han posat en marxa a diferents nivells i actualment ja es poden dipositar les dades en repositoris d'accés públic. Per ajudar a prendre una decisió es pot consultar un parell de directoris de repositoris de dades<sup>18</sup>. Cal dir que el nombre de revistes que requereixen que les dades estiguin dipositades en algun repositori públic i, per tant, siguin accessibles va en augment i moltes aconsellen consultar aquests directoris per trobar el millor lloc on hostatjar-les<sup>19</sup>.

## 2.5 Citació

Finalment, cal veure com se citen les dades. Generalment, aquest tema està vinculat amb la ubicació de les dades. Quan les dades es troben en un repositori normalment hi ha la possibilitat d'utilitzar el sistema Handle<sup>20</sup>, que és el més estès en aquest tipus de plataformes. No obstant això, molts repositoris que ofereixen un servei gratuït d'hostatjament de dades, de vegades vinculat a la publicació en revistes, disposen del sistema d'identificació DOI, el mateix que s'utilitza en la majoria de publicacions periòdiques internacionals. Entre aquests repositoris que utilitzen el DOI hi ha Dryad<sup>21</sup> o Figshare<sup>22</sup>. Aquests identificadors estan pensats per garantir un accés permanent a l'objecte associat. A més, hi ha altres mètodes per identificar unívocament un conjunt de dades com, per exemple, l'empremta numèrica universal o UNF (Universal Numerical Fingerprint) [10]. Aquest identificador es crea mitjançant un algoritme per preservar la integritat del conjunt de dades i poder verificar que no ha canviat encara que hagi variat el lloc on ha estat emmagatzemat o allotjat.

I un cop les dades són citables es poden construir mètriques per avaluar-ne les citacions, com s'ha fet durant

anys amb les publicacions [11]. Hi ha diversos models proposats, però encara no hi ha un estàndard adoptat. En el que tothom coincideix és que cal buscar un sistema per recompensar i incentivar els investigadors més enllà de la bona voluntat de compartir les dades [12].

## 3. Situació de l'obertura de dades de recerca en el context internacional

En l'àmbit internacional hi ha tres zones on s'està treballant de manera més activa per obrir les dades de recerca. Per una banda hi ha els EUA, on diferents agències i instituts que financen la recerca tenen polítiques de gestió de les dades de recerca [13] en què s'inclouen els períodes durant els quals els investigadors poden gaudir de l'exclusivitat d'accés a les dades i els terminis per fer públiques les dades primàries i secundàries. Cal tenir en compte que la llei de propietat intel·lectual dels EUA, Copyright Act, estableix que el que crea el Govern federal passa automàticament al domini públic, és a dir, no té cap mena de protecció de drets<sup>23</sup>. Aquesta situació afecta moltes d'aquestes agències o instituts nacionals. En moltes biblioteques i centres de recursos es poden trobar serveis especialitzats adreçats a donar suport als investigadors en la gestió de dades<sup>24</sup>.

Un segon referent és Austràlia. De fet, aquest país va ser un dels primers a adoptar polítiques per obrir les dades públiques i, per tant, és lògic que aquestes polítiques també s'apliquen en les dades de recerca. El Servei Nacional de Dades Australià ha establert un repositori de dades nacio-

18. El Registre de Repositoris de Dades de Recerca o re3data.org ([www.re3data.org](http://www.re3data.org)) i el Databib ([databib.org](http://databib.org)) [Consulta: 1/7/2013].

19. Per exemple, la política editorial de SpringerPlus ([www.springerplus.com/about#editorialpolicies](http://www.springerplus.com/about#editorialpolicies)) [Consulta: 1/7/2013].

20. Handle ([www.handle.net](http://www.handle.net)) [Consulta: 1/7/2013].

21. Dryad és un repositori genèric per hostatjar dades relacionades amb publicacions científiques ([datadryad.org](http://datadryad.org)) [Consulta: 10/8/2013].

22. Figshare s'ofereix com un espai alternatiu on publicar dades i altres resultats de la recerca al que ofereix l'actual sistema de comunicació científica ([figshare.com](http://figshare.com)) [Consulta: 10/8/2013].

23. Secció 105 de la *Copyright Act* [En línia], <<http://www.law.cornell.edu/uscode/text/17/105>> [Consulta: 1/7/2013].

24. Un exemple és el del MIT i el seu servei «Manage your data» ([libraries.mit.edu/guides/subjects/data-management/index.html](http://libraries.mit.edu/guides/subjects/data-management/index.html)), que ofereix guies, seminaris, atenció personalitzada a grups i investigadors [Consulta: 1/7/2013].

nal, Research Data Australia<sup>25</sup>, que el juliol de 2013 ja disposava de més de 56.000 conjunts de dades de recerca. Les dades allotjades estan etiquetades amb metadades que inclouen les condicions d'accés i la llicència d'ús, si existeix.

Finalment, el tercer focus és la Unió Europea, on s'està establint una política conjunta de cara al nou programa marc d'ajuts per a la recerca anomenat *Horizon 2020*. En aquest programa, que s'iniciarà l'any 2014, es preveu una acció pilot sobre dades de recerca que requerirà que en determinats àmbits s'hagin de fer públiques les dades amb les quals s'ha treballat en el marc d'un projecte finançat per aquest programa<sup>26</sup>. Si mirem els desenvolupaments que s'han fet a Europa a nivell estatal possiblement destaca el Regne Unit, on el Govern ha apostat clarament per l'accés obert després de la publicació de l'informe *Finch*<sup>27</sup>. Com en el cas australià, cal destacar que el Regne Unit és un dels estats més actius en l'obertura de les dades públiques després de la campanya iniciada el març de 2006 reclamant l'alliberament d'aquestes dades<sup>28</sup>. Pel que fa a les accions respecte les dades de recerca, cal destacar l'informe elaborat per la Royal Society [14] on es recullen una sèrie de recomanacions adreçades a investigadors, centres de recerca, societats científiques, editors, governs i agències de finançament per tal de facilitar la ubicació i la difusió en obert de les dades recollides en qualsevol recerca. Com en el cas dels EUA, la majoria d'universitats disposen de serveis de suport a la gestió d'aquestes dades<sup>29</sup>.

No voldria acabar aquest apartat sense esmentar l'Aliança per a les Dades de Recerca (Research Data Alliance)<sup>30</sup>, una

organització internacional impulsada pels organismes de finançament de la recerca a Austràlia, els EUA i la Unió Europea. L'objectiu d'aquesta organització és facilitar i accelerar l'intercanvi de dades de recerca mitjançant la implementació de la infraestructura tecnològica i les connexions socials necessàries. Aquesta aliança s'organitza mitjançant grups de treball i grups d'interès oberts a la participació de qualsevol persona interessada i en trobades semestrals plenàries.

## 4. Situació de les dades a l'Estat espanyol

A finals de l'any passat, la FECYT (Fundación Española para la Ciencia y la Tecnología) va publicar un informe del grup de treball sobre dipòsit i gestió de dades en accés obert [14]. Possiblement aquest document és un bon inici per establir una política de gestió de dades en qualsevol institució de recerca. S'hi apunten algunes iniciatives estatals, com ara el projecte ODISEA<sup>31</sup>. Aquest projecte té com a objectiu crear un inventari internacional dels repositoris que admeten el dipòsit de dades de recerca. A més d'aquest projecte, s'esmenten les publicacions d'autors estatals sobre el tema de les dades de recerca, es fa un repàs de les sessions, jornades o conferències organitzades per difondre l'estat de les dades de recerca i conscienciar els investigadors, i finalment se citen projectes de recerca relacionats amb les dades.

25. Research Data Australia ([researchdata.andis.org.au](http://researchdata.andis.org.au)) [Consulta: 1/7/2013].

26. «Towards better access to scientific information: Boosting the benefits of public investments in research» [En línia]. Communication from the European Commission, 17 de juliol 2012. <[http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/era-communication-towards-better-access-to-scientific-information\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf)> [Consulta: 1/7/2013].

27. Resposta del govern britànic a la publicació de l'informe: Letter to Dame Janet Finch on the Government Response to the Finch Group Report: «Accessibility, sustainability, excellence: how to expand access to research publications» [En línia]. <<http://www.bis.gov.uk/assets/biscore/science/docs/l/12-975-letter-government-response-to-finch-report-research-publications.pdf>> [Consulta: 1/7/2013].

28. Free Our Data: Make taxpayers' data available to them ([www.freeourdata.org.uk](http://www.freeourdata.org.uk)) [Consulta: 1/7/2013].

29. Un exemple d'això és el servei ofert per la Universitat de Cambridge: Support for Managing Research Data ([www.lib.cam.ac.uk/dataman](http://www.lib.cam.ac.uk/dataman)) [Consulta: 1/7/2013].

30. Research Data Alliance ([rd-alliance.org](http://rd-alliance.org)) [Consulta: 1/7/2013].

31. ODISEA, International Registry on Research Data ([odisea.ciepi.org](http://odisea.ciepi.org)) [Consulta: 1/7/2013].



©istockphoto/Steven Frame

Tanmateix, no hi ha cap política adoptada per una institució de recerca o una administració pública que requereixi cap tipus de dipòsit de dades. Tampoc no hi ha cap iniciativa per construir un repositori estatal. En l'informe tampoc no s'esmenten serveis de suport a la recerca com els que hi ha als EUA o al Regne Unit. Això no vol dir que no hi hagi interès a crear-los, sinó que possiblement encara no estan prou consolidats per posar-los com a exemple.

## 5. El paper d'una biblioteca o un centre de recursos

Des de sempre les biblioteques o els centres de recursos han ofert serveis de suport als investigadors, des de l'accés a les publicacions fins a l'ajuda per fer cerques bibliogràfiques. Aquests centres sempre han estat un dels principals punts d'informació als quals els investigadors s'han adreçat. Els darrers anys, un dels temes més tractats ha estat l'accés obert i tot el que comporta: repositoris, revistes, ajudes, aspectes legals... La majoria d'institucions de

Des de sempre les biblioteques o els centres de recursos han ofert serveis de suport als investigadors, des de l'accés a les publicacions fins a l'ajuda per fer cerques bibliogràfiques.

recerca han construït unitats de suport a la recerca constituïdes principalment per bibliotecaris i documentalistes, sense oblidar altres perfils. Cal no ignorar tampoc les connexions necessàries amb altres unitats de la institució com ara les oficines de gestió de la recerca o de transferència, els centres d'informàtica o els serveis de publicacions. Sense una estratègia global i coordinada, la institució no podrà oferir un suport eficient als investigadors. Ara, el repte que es presenta és col·laborar i donar suport a la gestió i la difusió de les dades de recerca [15]. En aquest camp, les biblioteques o els centres de recursos poden oferir tota la seva expertesa en la gestió de la informació i l'experiència adquirida en la darrera dècada amb el moviment de l'accés obert. Un bon punt de partida per establir un suport és seguir les recomanacions de LIBER [16]. En aquestes recomanacions s'inclou la possibilitat d'oferir suport específic per a la gestió de les dades amb ajudes centrades en els temes legals o en la descripció dels conjunts de dades mitjançant metadades, la creació de llocs de treball o perfils específics dedicats a la gestió i conservació de les dades, la participació en les polítiques institucionals, l'oferiment de col·laboracions amb investigadors i grups de recerca, i l'oferiment d'espais on allotjar

les dades, com per exemple els repositoris institucionals, o l'aliança amb altres serveis o institucions per construir infraestructures comunes.

Un dels serveis més citats en el suport que ofereixen les biblioteques o centres de recursos anglosaxons és ajudar en l'elaboració dels plans de gestió de dades. Cada cop més, qui finança la recerca demana l'elaboració d'aquests plans i previsiblement serà així en la nova convocatòria d'ajuts europeus. Un pla de gestió de dades hauria d'incloure la tipologia de dades, les mostres i els materials que s'utilitzaran i es produiran al llarg de la recerca, així com el programari o altres eines que es faran servir. Cal incloure-hi els estàndards que se seguiran pel que fa a la descripció de les dades mitjançant metadades i els formats utilitzats. S'hi han d'afegir les polítiques d'accés i reutilització sense oblidar l'estat de la propietat intel·lectual o la protecció de les dades per temes de seguretat, confidencialitat o privacitat. I finalment, cal establir un pla d'arxivament i preservació de les dades a llarg termini.

A més, des de les biblioteques o centres de recursos es poden realitzar campanyes de sensibilització i informació per tal d'arribar als investigadors de la institució.

## 6. Conclusions

Com a conclusió m'agradaria tornar a remarcar que l'obertura de les dades de recerca ofereix moltes oportunitats per a investigadors, institucions de recerca, centres de recursos, emprenedors i ciutadans. Per una banda l'investigador pot aconseguir una major visibilitat dels resultats de la seva recerca i, per tant, augmentar el seu reconeixement i les seves citacions. La institució on treballa pot aconseguir més prestigi i establir noves col·laboracions amb centres de recerca d'arreu del món. Per a la biblioteca o el centre de recursos és una oportunitat per adaptar-se als canvis, establir nous perfils, entomar nous reptes... Els emprenedors poden desenvolupar nous projectes a partir de les dades obertes. I, finalment, al ciutadà li oferim la

possibilitat d'acostar-lo a la recerca que es desenvolupa al seu entorn i que en la majoria de casos ell mateix està finançant. Ha arribat el moment que les dades de la recerca vegin la llum!

## Referències

- [1] TENOPIR, C.; ALLARD, S.; DOUGLASS, K.; AYDINOGLU, A. U.; WU, L.; *et al.* (2011) «Data Sharing by Scientists: Practices and Perceptions». PLoS ONE 6(6): e21101. DOI:10.1371/journal.pone.0021101.
- [2] STODDEN, V.; GUO, P.; MA, Z. (2013) «Toward Reproducible Computational Research: An Empirical Analysis of Data and Code Policy Adoption by Journals». PLoS ONE 8(6): e67111. DOI:10.1371/journal.pone.0067111.
- [3] PIWOWAR, H. A.; DAY, R. S.; FRIDSMA, D. B. (2007) «Sharing Detailed Research Data Is Associated with Increased Citation Rate». PLoS ONE 2(3): e308. DOI:10.1371/journal.pone.0000308
- [4] WOLSKI, M.; RICHARDSON, J.; REBOLLO, R.; «Shared benefits from exposing research data», 32nd Annual IATUL Conference Proceedings, 2011.
- [5] TORRES-SALINAS, D.; ROBINSON-GARCÍA, N.; CABEZAS-CLAVIJO, A. «Compartir los datos de investigación: introducción al data sharing». *El profesional de la información*, 2012, marzo-abril, vol. 21, núm. 2, p. 173-184. DOI:10.3145/epi.2012.mar08.
- [6] DE COCK BUNING, M.; VAN DINTHER, B.; JEPPESEN DE BOER, C. G.; Ringnald, A. «The legal status of research data in the Knowledge Exchange partner countries» [En línia], Centre for Intellectual Property Law (CIER), Molengraaff Institute for Private Law, Utrecht University, The Netherlands, octubre de 2011. <<http://www.knowledge-exchange.info/Default.aspx?ID=461>> [Consulta: 1/7/2013].

- [7] HEY, Tony; TREFETHEN, Anne. «The Data Deluge: an e-Science Perspective» dins *Grid Computing: Making the Global Infrastructure a Reality*, Fran Berman, Geoffrey C. Fox, and Anthony J. Hey (eds.), 2003, p. 808-822.
- [8] TJALSMA, Heiko; ROMBOUTS, Jeroen. *Selection of Research Data. Guidelines for appraising and selecting research data. A report by DANS and 3TU.Datacentrum* [En línia], Surf Foundation, 2010. <[http://www.surf.nl/nl/themas/openonderzoek/cris/Documents/SURFshare\\_Collectioneren\\_Selection%20of%20Research%20Data\\_DANS\\_3TU\\_DEFtt.pdf](http://www.surf.nl/nl/themas/openonderzoek/cris/Documents/SURFshare_Collectioneren_Selection%20of%20Research%20Data_DANS_3TU_DEFtt.pdf)> [Consulta: 1/7/2013].
- [9] BEAGRIE, N.; CHRUSZCZ, J.; LAVOIE, B. *Keeping Research Data Safe: A Cost Model and Guidance for UK Universities* [En línia]. JISC, 2008. <<http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf>> [Consulta: 1/7/2013].
- [10] ALTMAN, M., KING, G. A. «Proposed Standard for the Scholarly Citation of Quantitative Data» [En línia], *D-Lib Magazine*, 2007, vol.13, núm. 3/4. <<http://www.dlib.org/dlib/march07/altman/03altman.html>> [Consulta: 1/7/2013].
- [11] COSTAS, R.; MEIJER, I.; ZAHEDI, Z.; WOUTERS, P. *The Value of Research Data - Metrics for datasets from a cultural and technical point of view. A Knowledge Exchange Report* [En línia]. Leiden: 2013. <<http://www.knowledge-exchange.info/datametrics>> [Consulta: 1/7/2013].
- [12] NIU, Jinfang. «Reward and Punishment Mechanism for Research Data Sharing», *IASSIST Quarterly*, No. Winter, 2006.
- [13] DIETRICH, D.; ADAMUS, T.; MINER, A.; STEINHART, G. «De-Mystifying the Data Management Requirements of Research Funders», *Issues in Science and Technology Librarianship*, 70, 2012. DOI:10.5062/F44M92G2.
- [14] The Royal Society. *Science as an open enterprise*. London: The Royal Society Science Policy Centre report, 02/12, June 2012. ISBN: 978-0-85403-962-3.
- [15] Grupo de Trabajo de Depósito y Gestión de datos en Acceso Abierto del proyecto RECOLECTA. *La conservación y reutilización de los datos científicos en España. Informe del grupo de trabajo de buenas prácticas* [En línia]. Madrid: Fundación Española para la Ciencia y la Tecnología (FECYT), (2012). <<http://www.fecyt.es>> [Consulta: 1/7/2013].
- [16] MARTÍNEZ-URIBE, L.; MACDONALD, S. «Un nuevo cometido para los bibliotecarios académicos: data curation» [En línia]. *El profesional de la información*, 2008, vol. 17, núm. 3, p. 273-280. <<http://www.elprofesionaldelainformacion.com/contenidos/2008/mayo/03.pdf>> [Consulta: 1/7/2013].
- [17] CHRISTENSEN-DALSGAARD, Birte; *et al.* «Ten recommendations for libraries to get started with research data management» [En línia]. Final report of the LIBER working group on E-Science / Research Data Management, juliol 2012. <<http://www.libereurope.eu/sites/default/files/The%20research%20data%20group%202012%20v7%20final.pdf>> [Consulta: 1/7/2013].

Aquest text està subjecte a la llicència de Reconeixement 3.0 de Creative Commons disponible a:  
<<http://creativecommons.org/licenses/by/3.0/es/>>.

# Web 2.0, identitat i identitat digital. La vida, el temps, la mort, la memòria i la identitat

## **Ernest Benach**

Expresident del Parlament de Catalunya.  
*@ebenach* (Twitter)

## **Miquel Pueyo**

Professor de Periodisme i Comunicació a la Universitat de Lleida.  
*mediterranicontinental.wordpress.com*

### **Resum:**

Fins i tot en la societat 2.0, el més important són les relacions humanes, no la tecnologia. Cadascú de nosaltres construeix la seva identitat en la vida real i la seva identitat digital, a Internet. Cal tenir en compte algunes recomanacions per gestionar la identitat digital, com ara la vulnerabilitat, la influència dels altres, la tendència a convertir-se en eterna (de qualsevol cosa que pengem), l'oportunitat de relacionar-nos, la nostra credibilitat i reputació, i la inevitabilitat de la nostra desaparició física. La nostra identitat digital, en l'hora de la nostra mort, hauria de poder ser regulada a través del nostre testament digital o de la utilització de determinats recursos existents a Internet.

Per a comunicadors, periodistes i documentalistes, la web 2.0 també proporciona noves oportunitats professionals i metodològiques. En el cas dels documentalistes i els serveis de documentació, la socialització de la xarxa els afectarà, i aquells que entenguin el funcionament d'aquesta Internet relacional n'adquiriran avantatges personals i professionals.

### **Paraules clau:**

Internet, mort digital, testament digital, societat xarxa, identitat



## **Web 2.0, identidad e identidad digital. La vida, el tiempo, la muerte, la memoria y la identidad**

### **Resumen:**

Incluso en la sociedad 2.0, lo más importante son las relaciones humanas, no la tecnología. Cada uno de nosotros construye su identidad en la vida real y su identidad digital, en Internet. Hay que tener en cuenta algunas recomendaciones para gestionar la identidad digital, como la vulnerabilidad, la influencia de los demás, la tendencia a convertirse en eterna (de cualquier cosa que colgamos), la oportunidad de relacionarnos, nuestra credibilidad y reputación, y la inevitabilidad de nuestra desaparición física.

Nuestra identidad digital, en la hora de nuestra muerte, debería poder ser regulada a través de nuestro testamento digital o de la utilización de determinados recursos existentes en Internet.

Para comunicadores, periodistas y documentalistas, la web 2.0 también proporciona nuevas oportunidades profesionales y metodológicas. En el caso de los documentalistas y los servicios de documentación, la socialización de la red les afectará, y aquellos que entiendan el funcionamiento de esta conexión relacional adquirirán ventajas personales y profesionales.

### **Palabras clave:**

Internet, muerte digital, testamento digital, sociedad red, identidad

## **Web 2.0, identity and digital identity. Life, time, death, memory and identity**

### **Abstract:**

Even at 2.0 society, human relationships are the most important thing, not technology. Each one of us builds his identity in real life and his digital identity in the internet. Keep some suggestions in mind in order to manage digital identity, such as vulnerability, other people's influence, tendency to become eternal (anything we post), the opportunity to relate to people, our credibility and reputation, and the inevitability of our physical disappearance.

At the time of our death, we should be able to regulate our digital identity destination, preferably through our digital will or the use of certain Internet resources.

Web 2.0 also provides new professional and creative chances for communicators, journalists and documentalists makers. In the case of the documentalists and documentation services, they will be concerned by the socialization of the net, and those who will be able to understand the internet functioning will obtain personal and professional benefits.

### **Keywords:**

Internet, digital death, digital will, network society, identity

«L'element més important no és la tecnologia, sinó les relacions humanes»



©istockphoto/mediaphotos

## 1. Societat, identitat i identitat digital

### 1.1 Tecnologia o relacions humanes?

«L'element més important no és la tecnologia, sinó les relacions humanes»<sup>1</sup>. Aquesta és la divisa de l'obra de Howard Rheingold, professor de la Universitat d'Stanford i crític i assagista especialitzat en les implicacions culturals, socials i polítiques de les noves tecnologies de la informació. Ell va encunyar l'expressió *virtual community* i ha afirmat sempre que les comunitats virtuals són un instrument d'afirmació de la democràcia de base i que és necessari que existeixi una comunitat *on line* global. També va ser Rheingold el primer a parlar de multituds intel·ligents (*smartmobs*), que serien una forma d'organització social nascuda de i estructurada a l'entorn de les noves tecnologies de la informació i la comunicació. Rheingold no només creu en les NTIC (Noves Tecnologies de la Informació i la Comunicació), sinó que aposta decididament per les

NTIC com a element nuclear de les nostres vides i de la nostra societat, la qual cosa no obsta perquè continuï posant les relacions humanes per damunt de la tecnologia.

### 1.2 La mort per correu electrònic

La mort ha estat durant segles un misteri per a les societats humanes, però només des de mitjans del segle XX s'ha convertit en un tabú. Ara, a començaments del segle XXI, la nostra societat adquireix un ritme accelerat al qual intentem adaptar-nos, i la gestió de la mort constitueix un capítol no del tot resolt de les relacions humanes. La nova societat digital, que tantes transformacions ha provocat i provocarà en la nostra manera de viure, també modificarà la manera com entenem i gestionem la mort, i com actuem enfront d'aquest esdeveniment indestruïble de la vida.

Rafael Bartolozzi era un important artista plàstic del nostre país i va ésser també alcalde de Vespella de Gaià, al Tarrago-

1. *ArtNodes. Intersections between Arts, Sciences and Technology*. «Entrevista a Howard Rheingold» [En línia]. <<http://www.uoc.edu/artnodes/espai/cat/art/rheingold.html>> [Consulta: 26/07/2013].

nès. El 22 de març de 2009 molts dels seus amics van rebre un correu electrònic des de l'adreça de l'artista, amb un assumpte on es podia llegir: «Rafael Bartolozzi ha muerto». Era el correu de Bartolozzi que comunicava la seva mort. Alguns dels receptors del correu van pensar que es tractava de l'anunci d'un nou muntatge de l'artista; alguna nova instal·lació, tal vegada. D'altres van creure que es referia a la mort de l'artista com a tal; és a dir, Bartolozzi ja no faria més art. Els que estaven al corrent de la seva malaltia van deduir ràpidament el sentit d'aquell correu electrònic. Senzillament, la seva família havia aprofitat el correu electrònic de l'artista per comunicar la seva mort. En aquells moments, rebre un correu d'una persona anunciant la seva mort resultava estrany, però cada cop, sigui a través del correu o d'alguna de les xarxes socials, serà més habitual. D'una banda, la identitat física desapareix, mor, i de l'altra, és justament la seva identitat virtual la que ens ho comunica.

### 1.3 La construcció de la identitat

Això ens duu a la necessitat d'aclarir què és la identitat i què és la identitat digital. El sentit de la identitat personal és inseparable d'allò que hem denominat la construcció del jo. El tecnòleg Anil Dash, emprenedor i bloguer conegut ens ho adverteix: «Jo sóc la primera i principal font d'informació sobre mi mateix»<sup>2</sup>. En la mesura que el terme identitat no fa referència únicament a la imatge o a les característiques i la personalitat d'un individu, la identitat també pot ser definida com la capacitat de l'individu per mantenir constantment la seva personalitat o com el conjunt de característiques que fan que una persona sigui ella mateixa, en un sentit molt ampli, donat que constitueix una realitat extraordinàriament complexa i que no només es basa en les circumstàncies i les característiques d'una persona, sinó també en la visió que en tenen els altres.

La nostra identitat es troba inserida en un context cultural determinat. L'edat, el gènere, la classe social, la religió, la professió, la nacionalitat, l'opció sexual... són alguns dels

elements que determinen la nostra identitat. L'entorn també hi contribueix força. La família, les amistats, l'estatus econòmic, la nostra trajectòria professional i vital integren igualment aquest conjunt de factors. El cos, la nostra gestualitat i les nostres actituds corporals i estètiques també participen en la projecció d'una identitat concreta. Podem anar rapats, dur una llarga cua, corbata, una brusa i uns pantalonets curts...

Al llarg de la vida, la construcció d'aquest jo s'enriqueix i es modula mitjançant la interacció amb els altres éssers socials. La nostra identitat s'impregna amb les actituds i expectatives que tenen –o que creiem que tenen– la nostra família, els amics, els coneguts, els nostres professors o els companys de feina. En cadascuna d'aquestes múltiples relacions, es projecten identitats més o menys fragmentàries que, en conjunt, constituïran el nostre jo general i complet.

### 1.4 La identitat digital

Al mateix temps, en el món del primer quart del segle XXI, la identitat digital està convertint-se en alguna cosa gairebé tan important com la identitat real. Nosaltres la definirem com el conjunt de trets propis d'un individu que, a Internet, el diferencien d'altres identitats i que permeten determinar qui és i verificar que efectivament és qui assegura ser. També pel que fa a la identitat digital, allò decisiu no és només el que nosaltres fem, expressem o compartim a la xarxa, sinó el que altres expressen sobre nosaltres. En resum, la suma de les característiques d'una persona (intrínseques o adquirides al llarg del temps), que permeten diferenciar-la de qualsevol altra i que poden ésser verificades constitueixen la seva identitat digital.

En una conversa telefònica som capaços d'assegurar que la persona que està a l'altre costat de la línia es qui diu ser perquè reconeixem la seva veu. En canvi, en una comuni-

2. Anil DASH, *A blog about making culture* [Blog]. <<http://dashes.com/ani>> [Consulta: 26/07/2013].

cació remota a través d'Internet, el que rebem de l'altre son bytes, processats per les aplicacions corresponents i presentats en forma d'imatge, so, textos, etcètera.

D'entrada, atorguem un certa confiança als processos de generació i transmissió de dades per tal de poder creure que ens adrecem a una determinada persona i no a una altra. Tanmateix, garantir la identitat digital a Internet és un procés complex i no exempt de riscos, a pesar de les eines existents per generar propietats digitals identitàries, associar-les a entitats concretes i autenticar-les a través de certificats digitals, targetes intel·ligents, recursos biomètrics, contrasenyes dinàmiques, etcètera.

La conformació de la nostra identitat digital a Internet és quelcom que anem fent a poc a poc, en la mesura que ens connectem en xarxa, ens donem a conèixer i interactuem amb altres identitats. És aquesta identitat digital la que fa que els altres ens identifiquin i reconeixin, i ens prenguin en consideració dins aquesta societat xarxa que té unes normes de convivència que es van consensuant entre els seus usuaris. La socialització virtual és tan important com la que es produeix en el món físic i, en alguns sentits, encara més simple. Està basada en expectatives com l'increment de la transparència i de les relacions entre individus, la compartició de recursos i coneixements i la lliure circulació d'informació. Naturalment, cal conèixer les oportunitats i els riscos de la nostra presència a la xarxa i adquirir un domini bàsic de les formes i les eines amb què podem gestionar la nostra identitat, la nostra visibilitat i la nostra reputació. El correu electrònic, els formularis, la imatge, la geolocalització, els avatars, les adreces i els *nicknames* constitueixen elements molt importants en la configuració de la nostra identitat.

### 1.5 Algunes recomanacions per gestionar la nostra identitat digital

1. Existir i interactuar en el ciberespai posa al nostre abast oportunitats de tota mena, però la força i la visibilitat característiques de la xarxa també ens poden convertir en un objectiu més accessible i vulnerable.
2. Podem crear i gestionar la nostra identitat digital, però no preveure ni controlar allò que altres individus opinaran o compartiran. Per tant, hem de ser conscients que la nostra identitat es pot veure afectada per opinions, informacions i comentaris de fonts i persones diverses.
3. És important tenir tanta cura de la nostra identitat digital com de la nostra identitat en el món físic. No podem saber qui acabarà veient aquella fotografia de carnet toltos que vam penjar ni qui repescarà aquell comentari irat que vam fer, a propòsit d'un col·lega.
4. La creació d'una identitat digital és una oportunitat d'explicar-se a un mateix i de connectar amb persones amb les quals compartim interessos, objectius o punts de vista similars.
5. Cal ser conscients que tot allò que publiquem o pengem a Internet quedarà per a la posteritat i pot tenir conseqüències sobre la nostra identitat.
6. L'espècie humana sent una debilitat molt intensa pels relats, relacionada amb l'empatia i la capacitat de cooperar, però transmetre aquesta emoció requereix habilitats comunicatives específiques per als recursos digitals.
7. Una decisió important que hem de prendre és si volem o no mantenir aïllats (amb diverses identitats, més o menys obertes) els diversos àmbits de la nostra personalitat: treball, família, amistat, ideals, etcètera.
8. El rastre de la nostra activitat a la xarxa, llevat que prenguem precaucions extraordinàries, és molt més visible del que potser desitjaríem i també contribueix a configurar la nostra identitat digital.
9. Cal decidir quin límit posem a la informació que estem disposats a compartir. Això significa que hem de prendre precaucions a l'hora de gestionar les nostres contrasenyes i noms d'usuari, i que no és una bona idea facilitar adreces familiars, comentar qüestions confidencials relacionades amb la feina o divulgar indicis sobre la nostra situació econòmica.



©istockphoto/fotocelia

10. La nostra credibilitat a la xarxa dependrà en gran mesura de la nostra manera d'actuar, del grau de reciprocitat de la nostra activitat i de la solvència de la informació que compartim.
11. Encara que no ens vingui gaire de gust pensar en la nostra desaparició física, és recomanable manifestar als nostres familiars, amics o marmessors què desitgem que passi amb la nostra identitat digital, si un dia desapareixem. Que sigui esborrada? Que quedi surant en el núvol digital? Que sigui reconvertida en un espai memorial?

## 2. La mort digital. Com es pot gestionar?

La mort és el desenllaç natural de la història del nostre cos i de les seves funcions. És la nostra fi com a éssers humans i també de la nostra identitat legal, sobretot des que existeixen registres. La dicotomia viu/mort és d'aque-

Encara que no ens vingui gaire de gust pensar en la nostra desaparició física, és recomanable manifestar als nostres familiars, amics o marmessors què desitgem que passi amb la nostra identitat digital, si un dia desapareixem.

lles que no admeten matisos: no es pot estar «una mica mort» ni «mig viu». La mort anul·la la nostra identitat legal. El carnet d'identitat i el de conduir perden la seva validesa, els nostres estalvis, els nostres béns mobles i immobles, accions i tota mena de drets i propietats passen a mans dels nostres hereus, en un procés legalment tipificat.

Què passa, en canvi, amb tot allò que constitueix la nostra identitat digital i que inclou comptes de correu, perfils en xarxes socials, àlbums de fotografies, vídeos penjats a Vimeo o Youtube, blogs, cançons comprades a iTunes, etcètera? La resposta a aquesta pregunta no és senzilla, donat que no és el mateix allò que pugui succeir en el cas del correu electrònic, els perfils, els blogs; donat que no totes les companyies que proporcionen aquests serveis mantenen les mateixes polítiques, pel que fa a la gestió dels traspàs dels seus clients i a la relació amb els seus hereus.

Facebook, per exemple, la xarxa social més popular, va topar ben aviat amb el problema de la gent que moria i deixava un perfil orfe. Un dels directius de la companyia va perdre una persona molt propera i la companyia va optar

per una solució molt pragmàtica. Mitjançant un formulari i la demostració de la defunció d'un familiar, un tràmit que es pot realitzar a través de la xarxa, el perfil pot ésser eliminat o transformat en una pàgina en memòria del traspasat. D'altra banda, la mateixa companyia ha posat en marxa l'aplicació *ifidie*<sup>3</sup>, amb la qual qualsevol usuari pot preparar un comiat, sigui en vídeo o text, i designar tres dels teus amics a la xarxa com a fiduciaris de la seva mort. Després de la confirmació de la mort a través dels fiduciaris, el missatge de *d'ifidie* és publicat i arriba a tots els amics de Facebook del difunt.

Twitter, la popular xarxa de microblogging, també ha definit la seva política en relació a la mort de persones que hi tenen un perfil. D'entrada, n'esborra el perfil de seguidor (*follower*), però admet la continuïtat temporal del perfil com un espai d'homenatge. Els hereus ho han de comunicar a través de *privacy@twitter.com* però l'empresa no lliura, com tampoc no ho fa Facebook, cap contrasenya d'accés. Per contra, si es demana, Twitter empaqueta les piulades de l'usuari i les envia als seus hereus.

Què hauríem de fer amb aquest creixent volum d'informació que deixarem enrere? Com hauríem de resoldre la paradoxa de viure en una societat on la incineració i la dispersió de les cendres es va imposant com un mètode resolutiu per eliminar qualsevol vestigi del cos físic, mentre que al mateix temps deixem enrere enormes repositoris d'informació sobre l'antic propietari d'aquest cos desaparegut?

S'ha reflexionat poc sobre aquesta qüestió<sup>4</sup>, però comencen a aparèixer recomanacions perquè les persones es plantegin la conveniència de deixar algun tipus de testament digital o instruccions per garantir que els seus descendents tinguin ocasió de fer allò que aquestes persones voldrien que s'esdevingués, ja sigui eliminant del tot la

seva identitat digital o reconvertint-la temporalment o definitivament. En el mes d'abril de 2012, el govern dels Estats Units va publicar una recomanació als ciutadans amb activitat a les xarxes socials perquè deixin a la família o als amics instruccions explícites (i les contrasenyes!) sobre allò que volen que es faci amb els seus perfils digitals; i estats com Connecticut, Rhode Island, Indiana, Oklahoma o Idaho han aprovat disposicions legals per permetre a marmessors i famílies l'accés a comptes de xarxes i de correu electrònic. Així mateix, la Comissió Europea fa temps que estudia la necessitat d'ordenar allò que ja s'anomena l'oblit digital o dret a l'oblit digital, i alguns governs hi estant reflexionant, donat que instruments com la Llei de protecció de dades de 1999, reformada l'any 2007 i vigent a Espanya, són insuficients o inadequats.

### 3. El testament digital

Un document d'aquestes característiques –ja sia davant notari o advocat, o mitjançant un document d'últimes voluntats o una simple llista lliurada a familiars o amics– hauria d'incloure les voluntats del seu titular, tant correus electrònics com perfils en xarxes socials, informació financera, documents sonors i visuals, contractes virtuals..., i, naturalment, els noms d'usuari i totes les contrasenyes actualitzades.

A la xarxa, també comença a ser freqüent trobar-se amb recursos, iniciatives, empreses i organitzacions que comparteixen la voluntat de gestionar la mort, el dol i la memòria en l'àmbit digital, i donar resposta a la voluntat dels difunts i dels seus hereus. Es poden agrupar en quatre àmbits, dels quals esmentem alguns exemples: memorialis (*www.epilogable.com*, *www.duelia.org*); serveis relacionats amb la missatgeria electrònica (*www.deadmansswitch.net*); nota-

3. *ifidie* (*ifidie.net*).

4. Els signants d'aquest article ens permetem de suggerir-vos, si us interessa ampliar la vostra informació sobre aquest tema, la consulta de la segona part de l'obra de la qual són autors, i que fa referència als interrogants que plantegen la identitat digital, i la mort, el dol i la memòria 2.0: BENACH, E. PUEYO, M. *Mort certa, hora incerta. De l'edat mitjana a la societat digital*. Lleida: Pagès editors, 2013. També podeu visualitzar: *Entrevista a «Els matins» de TV3 26/02/2013* [En línia]. <<http://www.youtube.com/watch?v=6VUa-YxnOMw>> [Consulta: 12/07/2013].

ries digitals (*www.assetlock.net*), i el que podríem anomenar gestors de l'eternitat digital (*www.chronicleoflife.com*, *www.lifenaut.com*).

## 4. Gestió de la identitat digital i de la posteritat digital. Comunicadors, periodistes i documentalistes

Molt abans d'Internet, el poeta T. S. Eliot, en un dels seus comentaris al descens a l'infern del Dant, va escriure: «L'infern és un lloc on res no connecta amb res». Una definició que en plena eclosió de la web 2.0 o web social sembla forçosament premonitòria. En aquesta nova societat hiperconnectada, local i global al mateix temps, no exempta de riscos però dotada de la força que li atorga la seva condició de policèntrica i social, creiem que més enllà dels recursos tecnològics que la converteixen en un artefacte tecnosocial que només ha començat a explotar les seves possibilitats, continua bevent en les mateixes fonts del gust o de la necessitat pel coneixement, la comunicació, la cooperació, el relat i la memòria que han caracteritzat les societats humanes de tots els temps.

En aquest sentit, reivindiquem la necessitat de madurar el tractament de la nostra identitat digital, de la que anomenem la mort digital i de les activitats relacionades amb el relat, el dol i la memòria, i de superar la falsa dicotomia entre una hipotètica fredor tecnològica de la xarxa i la suposada calidesa superior de la interacció cara a cara, i ho fem per tres raons.

La primera, perquè la distància entre Internet i el món real és cada cop més relativa. La segona, perquè la qualitat i la intensitat de les emocions no és pas menor en un cas que en l'altre. La tercera, perquè de la mateixa manera que durant segles hem anat integrant a la nostra vida innovacions

i recursos que inicialment foren considerats suspectes o perillosos (des de la ferradura a l'anestèsia, passant per la lectura en veu alta o la impremta), estem convençuts, pel que fa a la xarxa, que si algun risc planteja és precisament el de quedar-se'n al marge.

Així mateix, creiem que Internet, primer, i ara la web social constitueixen un repte professional i metodològic per a totes aquelles professions i activitats relacionades amb la informació i la comunicació, així com amb la gestió, la creació, l'ordenació i la selecció de continguts, i amb la fascinació que la nostra espècie sent per les històries, en general. Tant a nivell personal com col·lectiu, ja sia com a instrument o com a creació artística, es produeix una dialèctica molt interessant entre el poder de la mort que destrueix l'ésser viu i el temps que extingeix la memòria, i l'esforç de les generacions per preservar la memòria dels morts (a través de tombes, memorials i històries) i ara també de construccions i vestigis identitaris que romandran a la xarxa, i d'iniciatives empresarials que ens proposen gestionar la nostra memòria o garantir l'aplicació del nostre testament digital<sup>5</sup>.

### 4.1 Els professionals i els recursos 2.0

En aquest escenari, els professionals dels mitjans de la comunicació es troben davant per davant d'un nou model d'accés a la informació i als continguts a Internet, en el qual les xarxes i les relacions assoleixen un protagonisme creixent. El periodista multimèdia, per exemple, ja no només redacta les informacions i se n'oblida, sinó que s'ha de preocupar de distribuir-les, adaptar-les als diversos suports, comentar-les en les xarxes socials i interactuar amb els seus lectors i seguidors.

Els documentalistes i els serveis de documentació dels mitjans també s'enfronten a aquesta nova situació. Acostumats al seu tradicional rol d'intermediació entre les fonts

5. Per a un estat de la qüestió relativament recent i per ampliar el punt de vista de tot allò que pot ésser inclòs dins l'anomenat testament digital, us recomanem: CARROLL, E.; ROMANO, J. *Your digital Afterlife. When Facebook, Flickr and Twitter Are Your Estate*. Berkeley: New Riders, 2011.

La possibilitat d'establir contacte amb l'audiència que ofereixen els blogs es multiplica si també es fa una utilització activa de les xarxes socials...



©istockphoto/Kevin Winter

o els documents, d'una banda, i els periodistes, de l'altra, havien estat gairebé sempre allunyats de la relació directa amb l'audiència, i mantenien un perfil més aviat discret, sense assumir cap protagonisme. La socialització de la xarxa també els afectarà, com a tots els professionals relacionats amb la comunicació, i els plantejarà noves exigències, nous reptes i noves oportunitats. Els documentalistes de mitjans hauran de ser competents en aquesta nova manera de fer circular la informació i d'utilitzar amb seguretat els recursos per gestionar-la en l'àmbit 2.0.

Les xarxes socials i els *social media* en general són extraordinàries fonts d'informació, i un ús professional eficient de serveis com Twitter o Delicious, per part d'un periodista o un documentalista de mitjans, constitueix un avantatge competitiu per a ell i la seva empresa. A més de la utilització de les xarxes com a fonts informatives, cada cop seran més útils per difondre el treball dels documentalistes.

#### 4.1.1 Blogs de documentació

L'eclosió dels blogs en el periodisme digital no ha estat un fenomen regular ni general, almenys a Catalunya i a Espanya. Per contra, en altres països els blogs dels departaments de documentació o similars s'estan convertint en un producte periodístic flexible i de gran diversitat temàtica, per bé que predominen els reportatges i els continguts retrospectius o relacionats amb el fons d'arxiu: efemèrides, cronologies, reportatges històrics. Vegem-ne alguns exemples interessants: *The Guardian*<sup>6</sup>, *The Times Tribune*<sup>7</sup>, *The Columbus Dispatch*<sup>8</sup>, *San Antonio Express*<sup>9</sup>, *La Vanguardia*<sup>10</sup>.

#### 4.1.2 Perfils en xarxes socials

La possibilitat d'establir contacte amb l'audiència que ofereixen els blogs es multiplica si també es fa una utilització

6. *The Guardian* ([www.guardian.co.uk/theguardian/series/from-the-archive](http://www.guardian.co.uk/theguardian/series/from-the-archive)).
7. *The Times Tribune* ([blogs.thetimes-tribune.com/pages/Pages from the past](http://blogs.thetimes-tribune.com/pages/Pages%20from%20the%20past)).
8. *The Columbus Dispatch* ([www.dispatch.com/content/blogs/a-look-back/index.html](http://www.dispatch.com/content/blogs/a-look-back/index.html)).
9. *San Antonio Express* ([blog.mysanantonio.com/vault](http://blog.mysanantonio.com/vault)).
10. *La Vanguardia* ([blogs.lavanguardia.com/hemeroteca](http://blogs.lavanguardia.com/hemeroteca)).



activa de les xarxes socials, amb perfils propis per al departament de documentació o l'hemeroteca, en els serveis més populars. Suggerim com a mostra, en el cas de Facebook, l'Hemeroteca de *La Vanguardia*<sup>11</sup>, així com els comptes de Twitter que mantenen *The Guardian Library*<sup>12</sup>, *Times Archive*<sup>13</sup> i *Christian Science Monitor Library*<sup>14</sup>.

Més enllà del perfil del departament o del servei de documentació, és també el perfil personal del professional allò que aquesta Internet *relacional*<sup>15</sup> està reclamant, i aquells que ho entenguin i posin en pràctica assoliran un avantatge. Alguns professionals de la documentació periodística ja ho estan fent, i utilitzen Internet *relacional* per compartir, a més d'altres informacions, els seus processos documentals a l'hora de construir una notícia o l'enllaç a una peça de documentació periodística publicada. Per exemple, a Twitter ho fan regularment Iris López<sup>16</sup>, documentalista de TVE; Anna Cornet<sup>17</sup>, documentalista de premsa; Virginia M.<sup>18</sup>, documentalista de Barcelona Televisió; Carlos Martín<sup>19</sup>, documentalista de la Sexta; José Arenedo<sup>20</sup>, documentalista, i Federico Ayala<sup>21</sup>, documentalista d'ABC.

## 5. Conclusions

Allò més important en la societat 2.0 no és la tecnologia, sinó les relacions humanes, com en diverses ocasions ha afirmat Howard Theingold. Qualsevol persona que vulgui gestionar conscientment la seva identitat digital ha de tenir en compte algunes recomanacions bàsiques.

Tant si ens agrada pensar-hi com si no, la mort és el desenllaç natural de la vida de tota persona i és recomanable que aquesta persona manifesti als seus amics, familiars o marmessors què desitja que passi amb la seva identitat digital, el dia del seu traspàs. La confecció d'un testament digital constitueix la millor mesura per assumir un cert control sobre la posteritat digital de l'individu. Bo i reconeixent que aquest és encara un marc a definir millor, des d'un punt de vista legal, ens sembla essencial a l'hora d'assegurar el predomini de la voluntat del finat i garantir, entre d'altres opcions, l'anomenat oblit digital.

Per a un periodista, comunicador o documentalista, a banda de la seva dimensió personal, pot trobar en la web 2.0 noves oportunitats i avantatges professionals, i desplegar millors oportunitats de cooperació, debat i divulgació, en l'àmbit de la seva professió.

## 6. Bibliografia

ARIÈS, Philippe. *L'homme devant la mort*. Paris: Éditions du Seuil. Traducció a l'espanyol: *El hombre ante la muerte*. Madrid, Taurus, 1983.

BENACH, E.; PUEYO, M. *Mort certa, hora incerta. De l'edat mitjana a la societat digital*. Lleida: Pagès editors, 2013.

CARROLL, E.; ROMANO, J. *Your digital Afterlife. When Facebook, Flickr and Twitter Are Your Estate*. Berkeley: New Riders, 2011.

11. Facebook de *La Vanguardia* ([www.facebook.com/pages/La-Hemeroteca-de-La-Vanguardia/291917499474](http://www.facebook.com/pages/La-Hemeroteca-de-La-Vanguardia/291917499474)).

12. Twitter de *The Guardian Library* ([twitter.com/guardianlibrary](https://twitter.com/guardianlibrary)).

13. Twitter de *Times Archive* ([twitter.com/TimesArchive](https://twitter.com/TimesArchive)).

14. Twitter de *Christian Science Monitor Library* ([twitter.com/CSMlibrary](https://twitter.com/CSMlibrary)).

15. La Internet relacional, com el màrqueting social o el màrqueting de relacions, està basada en el foment de l'intercanvi beneficiós per a totes les parts implicades, ja sigui des del punt de vista de l'anàlisi de costos o del principi de l'intercanvi social.

16. Twitter d'Iris López ([twitter.com/Iris\\_Lopez](https://twitter.com/Iris_Lopez)).

17. Twitter d'Anna Cornet ([twitter.com/annacornet](https://twitter.com/annacornet)).

18. Twitter de Virginia M. ([twitter.com/virhouse](https://twitter.com/virhouse)).

19. Twitter de Carlos Martín ([twitter.com/MrCarlosMartin](https://twitter.com/MrCarlosMartin)).

20. Twitter de José Arenedo ([twitter.com/josearenedo](https://twitter.com/josearenedo)).

21. Twitter de Federico Ayala ([twitter.com/Federico\\_Ayala](https://twitter.com/Federico_Ayala)).

# Europa, *cloudcomputing* i protecció de dades de caràcter personal

## **Ramon Miralles**

Coordinador d'Auditoria i Seguretat de la Informació. Autoritat Catalana de Protecció de Dades  
*sramon.miralles@gencat.cat*

### **Resum:**

El *cloudcomputing* és un model de prestació i/o aprovisionament de serveis de tecnologies de la informació i la comunicació. La seva adopció tant a nivell empresarial com particular, i inclús per part del sector públic, s'està produint a una extraordinària velocitat, com tot allò que té a veure amb les tecnologies i la xarxa; els beneficis del seu ús són evidents, però no està exempt de riscos, alguns d'ells molt rellevants, especialment en relació al tractament de les dades personals. Segons els analistes, en els propers anys això anirà a més, tant pel que fa als usuaris individuals de la xarxa i serveis en línia, com per les organitzacions, i en tots dos casos afectarà la seva manera d'utilitzar les TIC. En relació als usuaris de la xarxa, la informàtica en núvol té molts punts de connexió amb la web 2.0 i, addicionalment, per a les empreses, està estretament lligada als processos d'externalització dels serveis de tecnologies de la informació i la comunicació.

### **Paraules clau:**

*cloudcomputing*, protecció de dades, dades personals, legislació, TIC, riscos tecnològics, economia, Europa, mercat digital, informàtica en núvol

## Europa, *cloudcomputing* y protección de datos de carácter personal

### Resumen:

El *cloudcomputing* es un modelo de prestación y/o aprovisionamiento de servicios de tecnologías de la información y la comunicación. Su adopción tanto a nivel empresarial como particular, e incluso por parte del sector público, se está produciendo a una extraordinaria velocidad, como todo lo que tiene que ver con las tecnologías y la red; los beneficios de su uso son evidentes, pero no está exento de riesgos, algunos de ellos muy relevantes, especialmente en relación al tratamiento de los datos personales.

Según los analistas, en los próximos años esto irá a más, tanto por lo que respecta a los usuarios individuales de la red y servicios en línea, como para las organizaciones, y en ambos casos afectará a su manera de utilizar las TIC. En relación a los usuarios de la red, la informática en la nube tiene muchos puntos de conexión con la web 2.0 y, adicionalmente, para las empresas, está estrechamente ligada a los procesos de externalización de los servicios de tecnologías de la información y la comunicación.

### Palabras clave:

*cloudcomputing*, protección de datos, datos personales, legislación, TIC, riesgos tecnológicos, economía, Europa, mercado digital, computación en la nube

## Europe, *cloud computing* and personal data protection

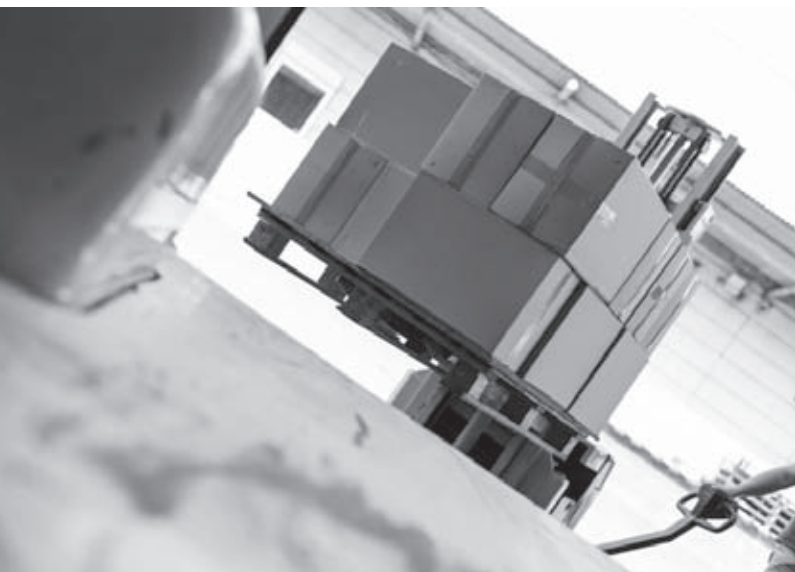
### Abstract:

The *cloud computing* is a model of delivery and/or provisioning service of information technology and communication. Its adoption both by business and individual users, and even by the public sector, is occurring very quickly, like everything that has to do with the technology and Internet. The benefits of its use are obvious, but is not without risks, some of them very important, especially in relation to the processing of personal data.

According to analysts in the coming years this will be go further, both with respect to individual users of the network and online services, and for the organizations, and in both cases it will affect how they use ICT. In relation to the users of the network, cloud computing has many points in common with Web 2.0 and, in addition to businesses, is closely linked to the processes of outsourcing of services of information technology and communication.

### Keywords:

cloud computing, data protection, personal data, law, ICT, technological risks, economy, Europe, digital market



©istockphoto/AzmanL

## 1. *Cloudcomputing*: una breu introducció

### 1.1 Punt de partida

Abans d'entrar a connectar el *cloudcomputing* amb el dret a la protecció de les dades de caràcter personal em sembla d'utilitat fer una breu aproximació a la història i evolució de la informàtica en núvol, a la que em referiré a partir d'ara com a *cloudcomputing*.

Des de l'òptica de les telecomunicacions, s'entén per «cloud» o «núvol» el conjunt de dispositius i infraestructures de comunicacions pels quals, de manera «imprevisible», passa la informació quan es vol transmetre d'un punt a l'altre de la xarxa Internet. Aquesta falta de predicció afecta tant a la quantitat com al tipus de dispositius; de fet, tots els elements que es troben al mig d'aquest intercanvi d'informacions tradicionalment s'han situat gràficament en un «núvol».

El punt d'inici de la comunicació és conegut (un usuari o procés inicia una transacció), i el de destí també (un servi-

El punt d'inici de la comunicació és conegut (...), i el de destí també (...), i així successivament. Però el camí que seguirà la informació transportada entre tots dos punts respon a unes regles que, si bé estan fixades per un protocol tècnic (TCP), tenen resultats impredecibles a priori...

dor dona resposta a la transacció), i així successivament. Però el camí que seguirà la informació transportada entre tots dos punts respon a unes regles que, si bé estan fixades per un protocol tècnic (TCP), tenen resultats impredecibles a priori; d'aquesta manera, a la pràctica podem intuir per on passarà la informació, però sense poder-ne estar del tot segurs.

Cal afegir un altre element que, tot i que en essència aquests dispositius es dediquen a gestionar el trànsit de la informació entre origen i destí, no se'ns pot escapar que durant aquest trànsit la informació es pot sotmetre a tractaments que vagin més enllà de facilitar la transmissió de paquets d'informació.

En aquest sentit, resulta rellevant fer referència a l'ús de tecnologies d'inspecció de paquets (*Deep Packet Inspection*, DPI), especialment per part dels estats, és a dir, la capacitat que tenen alguns equipaments de xarxa, que no són punt final de comunicacions, de tractar no només les capçaleres dels paquets que ha de retransmetre per la xarxa, sinó també, convenientment configurats, d'analitzar-ne el contingut.

De fet, la circumstància que les dades personals es puguin tractar exclusivament a efectes de trànsit per la xarxa està prevista a la Llei orgànica 15/1999, de 13 de desembre, de protecció de dades personals (en endavant LOPD). Quant a la determinació de l'àmbit territorial d'aplicació, a l'article 2.1.c<sup>1</sup> es preveu que la LOPD serà aplicable quan el responsable del tractament, tot i no estar establert en el territori de la Unió Europea, utilitzi mitjans situats en territori espanyol, llevat si aquests mitjans s'utilitzen únicament amb la finalitat de trànsit.

Ara bé, en el moment que el núvol deixa de ser exclusivament un mitjà de transport de la informació per passar a tenir capacitat de processament de la informació, se li afegeix el *computing*; tot i que en realitat la capacitat de processament no recau exactament en el núvol, sinó en aplicacions, plataformes i infraestructures disponibles a la xarxa i que, en alguns aspectes, es comporten com els dispositius del núvol a què ja he fet referència.

## 1.2 Cloudcomputing. Una definició

El *cloudcomputing* implica que en el processament d'informació hi concorren una sèrie de característiques que tenen com a conseqüència directa que l'origen, i especialment el destí, d'una transacció deixa de tenir uns valors absoluts, per passar a ser relatius: «la informació no sempre és on realment sembla i no sempre es tracta allà on sembla que s'està processant».

La definició de *cloudcomputing* que generalment serveix de base per dotar de contingut aquest concepte és la del

NIST<sup>2</sup> (*Institut Nacional d'Estàndards i Tecnologies*, una agència del U.S. Department of Commerce, creada el 1901). L'última versió d'aquesta definició de *cloudcomputing* elaborada pel NIST és de setembre de 2011<sup>3</sup>.

Segons aquesta definició, hi ha cinc característiques bàsiques que defineixen el *cloudcomputing*:

<b>Autoservei</b>	L'usuari pot utilitzar capacitats addicionals de processament o emmagatzemament de la informació, sense demanar-ho expressament al proveïdor del servei (per exemple, demanar més espai per emmagatzemar informació).
<b>Ampli accés a la xarxa</b>	Es pot accedir a la xarxa des de diferents dispositius i xarxes (típicament l'accés als serveis i informacions en el núvol es pot fer tant des d'ordinadors, com des de <i>tablets</i> o <i>smartphones</i> ).
<b>Agrupació i reserva de recursos</b>	Hi ha un conjunt de recursos compartits pels usuaris, d'acord amb les seves necessitats puntuals, que implica que en cada moment els recursos reservats poden ser diferents (l'economia d'escala és un dels aspectes fonamentals en els serveis <i>cloud</i> , es comparteixen els recursos amb d'altres usuaris, de manera que si jo no utilitzo, per exemple, capacitat de processament l'utilitza un altre).
<b>Rapidesa i elasticitat</b>	Es pot accedir als nous recursos de forma immediata aparentment il·limitada (les noves necessitats de serveis; per exemple, si una empresa de comerç electrònic necessita més ample de banda durant una campanya de Nadal hi pot accedir de manera immediata i després deixar anar aquests recursos, de manera que només pagarà per allò realment utilitzat).
<b>Servei mesurable i supervisat</b>	Se'n controla l'ús i en tot moment es pot conèixer, de manera transparent, el nivell de recursos utilitzats en cada moment (això és especialment rellevant de cara a la posterior facturació del servei; pagarem en funció del consum de recursos TIC, i per calcular aquest consum cal que sigui mesurable).

1. Artículo 2. Ámbito de aplicación.

1. La presente Ley Orgánica será de aplicación a los datos de carácter personal registrados en soporte físico, que los haga susceptibles de tratamiento, y a toda modalidad de uso posterior de estos datos por los sectores público y privado.

Se regirá por la presente Ley Orgánica todo tratamiento de datos de carácter personal:

a) ...

b) ...

c) Cuando el responsable del tratamiento no esté establecido en territorio de la Unión Europea y utilice en el tratamiento de datos medios situados en territorio español, salvo que tales medios se utilicen únicamente con fines de tránsito.

2. NIST ([www.nist.gov](http://www.nist.gov)).

3. Vegeu: U.S. Department of Commerce. *The NIST Definition of Cloud Computing* [En línia]. <<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>>.

### 1.3 Per què el núvol?

Aquesta capacitat de procés en el núvol està connectada amb la tendència d'externalització dels serveis TIC de les organitzacions i la reconversió d'aquests serveis en el que s'anomena una *utility computing*.

En relació a aquesta qüestió, recomano la lectura de l'article de Nicholas Carr<sup>4</sup> publicat en el *MIT Sloan Management Review* (Massachusetts Institute of Technology), d'abril de 2005, que amb el títol «The End of Corporate Computing»<sup>5</sup> descriu els motius que han de portar les organitzacions a deixar de considerar les TIC com un actiu de la seva propietat, per passar a tractar-les com un servei que compren.

Al seu article, Carr fa un paral·lelisme entre el procés de transformació de l'ús de les TIC que han de seguir les organitzacions per ser competitives, amb la transformació que es va produir a principis del segle XX, quan les empreses industrials va començar a tancar i a desmantellar les fonts de energia utilitzades per les seves indústries i que eren de la seva propietat (rodes d'aigua, màquines de vapor, generadors elèctrics, etc.).

Aproximadament a partir de 1880, la producció comercial d'electricitat va començar a ser possible; el 1902, als Estats Units hi havia unes 50.000 plantes de generació privada d'energia i només 3.600 estacions podien vendre energia a d'altres, amb moltes limitacions i inicialment amb un preu alt. Però entre 1907 i 1920 la quota de producció d'energia elèctrica per ser comercialitzada va passar del 40% al 70%, i el 1930 ja arribava al 80%.

Els motius d'aquesta ràpida adopció d'un nou model de subministrament d'energia per a les indústries eren senzills: uns costos menors i una complexitat de gestió menor, que permetia que les indústries es poguessin centrar en el seu negoci.

Per a Nicholas Carr, al seu article de 2005, hi ha tres avenços tecnològics clau en la transformació de les TI: la virtualització, el *grid computing* i els serveis web que, combinats amb l'augment de capacitat de les xarxes de comunicacions i la fibra òptica, donen com a resultat un escenari idoni per dur a terme aquesta transformació, de manera que el major obstacle no serà la tecnologia, sinó l'actitud de les organitzacions a l'hora d'assumir aquest nou model d'ús de les TIC en els seus negocis.

Certament el *cloudcomputing* ja és una realitat per als usuaris de la xarxa, a títol individual. Els principals casos d'ús del *cloudcomputing* impliquen companyies i serveis com Facebook, Amazon, Nasdaq o Google. En un informe del Pew Research Center<sup>6</sup>, ja en un estudi de setembre de 2008, s'assenyalava, en relació a l'ús del *cloudcomputing*, que el 69% dels usuaris d'Internet dels Estats Units emmagatzema dades o utilitza aplicacions basades en serveis de *cloudcomputing*.

L'explosió real del *cloudcomputing* vindrà provocada per aquest model d'ús de les TIC per part de les empreses<sup>7</sup> i de les administracions públiques. Tal i com el mateix NIST inclou a la seva definició, el *cloudcomputing* és un paradigma que encara està en evolució.

4. Nicholas Carr és autor del *bestseller* de 2008 del *Wall Street Journal* titulat *El gran canvi: cablejant el món, des d'Edison a Google*, que és considerat com un dels llibres més influents en relació al *cloudcomputing*. Vegeu informació sobre Nicholas Carr a [www.nicholasgcarr.com/info.shtml](http://www.nicholasgcarr.com/info.shtml).
5. «The End of Corporate Computing» [En línia], *MIT Sloan Management Review*, 2005. <<http://sloanreview.mit.edu/the-magazine/articles/2005/spring/46313/the-end-of-corporate-computing>>.
6. Un informe d'aquest centre d'investigació (juny de 2010) recull l'opinió dels experts que el 2020 la majoria d'usuaris d'Internet utilitzarà aplicacions basades en *cloudcomputing*, en lloc de les aplicacions d'escriptori <<http://pewinternet.org/Reports/2010/The-future-of-cloud-computing.aspx>>.
7. En aquest sentit, resulta d'especial interès la tasca de Salesforce en relació a l'ús empresarial del *cloudcomputing* [www.salesforce.com/es/cloudcomputing](http://www.salesforce.com/es/cloudcomputing).

«la informació no sempre és on realment sembla i no sempre es tracta allà on sembla que s'està processant».



©istockphoto/John Neff

## 2. Protecció de dades i estratègia europea de *cloudcomputing*

### 2.1 *Cloudcomputing*: la societat de la informació i el creixement econòmic a Europa

Cal tenir molt present que la rellevància del model dels serveis de processament de la informació basat en l'arquitectura *cloudcomputing* no se centra en l'element tecnològic, sinó en el seu impacte en el desenvolupament econòmic, situats com estem en un context de societat de la informació, i per tant d'ús intensiu de les tecnologies de la informació i d'Internet a tots els nivells.

D'aquí que les institucions europees consideren estratègica la manera en què Europa adopti i desenvolupi tot el que envolta el *cloudcomputing*.

La meua intenció es connectar la regulació del dret a la protecció de dades a Europa amb el posicionament de les institucions europees respecte del *cloudcomputing*, expressat mitjançant l'anomenada «estratègia europea de *cloudcomputing*».

Per fer-ho, en primer lloc convé, a banda de les definicions formals fetes a la introducció, contextualitzar allò a què ens referim d'una manera genèrica com a *cloudcomputing*, que en realitat aglutina situacions fàctiques molt diverses i heterogènies.

En aquest sentit, no cal explicar els aspectes positius de l'ús de les TIC per al processament de la informació relacionada amb l'activitat d'empreses, administracions i societat en general, ni com han estat motor de canvis substancials en la «manera de fer les coses», des de que en la dècada del setanta es va començar a generalitzar l'ús de sistemes informàtics. Tampoc cal que entri en detalls de com, a partir dels noranta, Internet ens ha canviat la vida, a nivell personal i professional.

Ja des de la dècada dels setanta es va generar una percepció, que s'ha mantingut des de llavors, que la «informàtica» era «complicada» i «costosa», de manera que sempre ha existit un cert anhel que les TIC fossin un recurs senzill d'incorporar a l'activitat de les organitzacions i que el seu cost estigués estretament vinculat al seu consum.

Pel que fa a la simplificació en l'ús de les TIC, resulta evident que el fet que els serveis basats en el *cloudcomputing* ens permetin «agafar» i «deixar», segons ens convinguin i de manera autònoma (sense intervenir el proveïdor), recursos de processament de la informació (autoservei), uns recursos que aparentment se'ns presenten com il·limitats, ja que sembla que mai es puguin acabar (elasticitat) i que les aplicacions i la informació siguin accessibles des de qualsevol xarxa o dispositiu, fan més senzill el seu ús.

Respecte l'eficiència econòmica del model *cloud*, hem de pensar que aquesta se sustenta en la compartició de recursos per al processament de la informació, per tant aplicant economies d'escala; és per això que una característica essencial dels serveis en *cloud* és que allò que consumim és mesurable, i per tant pagarem per allò que realment consumim.

Els aspectes econòmics favorables també es manifesten de manera global. Una part del futur creixement econòmic i d'ocupació a la Unió Europea passa per aquest model de negoci, com reflexa un estudi de la consultora IDC de finals del 2012: en el 2020 el volum de consum d'aquests serveis podria arribar als 250.000 milions d'euros, afavorint la creació de fins a 3,8 milions de llocs de treball; sempre i quan Europa sigui proactiva a donar resposta a les incerteses que el núvol genera en les organitzacions. Es plantegen dubtes i preocupacions respecte de qüestions relacionades amb la jurisdicció i legislació aplicable, la localització de les dades, la seguretat de la informació o el nivell de responsabilitat dels proveïdors de serveis *cloud*, que han estat identificades com les principals barreres per al desplegament d'aquests serveis a Europa.

## 2.2 Riscos: pèrdua de control i les dades de caràcter personal

Ara bé, en l'ús de serveis *cloud* també es produeix el que podríem anomenar «un dany col·lateral principal»: la pèrdua o cessió del control sobre la informació i el seu processament.

Una pèrdua de control que addicionalment es veu amplificada per una generalitzada falta de transparència en els termes i condicions dels contractes de serveis *cloud*.

En la meua opinió, aquesta pèrdua de control es manifesta al menys en tres dimensions:

- a. Perdem el control directe sobre la informació (les dades): qüestió d'especial rellevància en el cas de dades de caràcter personal (cal recordar que, com veurem, el responsable del tractament manté intactes les seves obligacions).
- b. Perdem el control sobre els processos de tecnologies de la informació: el rendiment i disponibilitat de les tecnologies de la informació i els aspectes tecnològics de la seguretat de la informació passen a dependre de tercers, i tot amb una dependència absoluta de les telecomunicacions.
- c. Perdem el control en relació a la capacitat contractual: tret de comptades ocasions (molt excepcionalment, per tant), no podem negociar els termes i condicions contractuals.

No hem d'oblidar que quan el Tribunal Constitucional Espanyol va considerar al dret a la protecció de dades com de caràcter fonamental, a les sentències 290/2000<sup>8</sup> i 292/2000<sup>9</sup>, de 30 de novembre de 2000, l'autodetermina-

8. Sentència 290/2000 del Tribunal Constitucional d'Espanya [En línia]. <[http://www.agpd.es/portalwebAGPD/canaldocumentacion/sentencias/tribunal\\_constitucional/common/pdfs/Sentencia2901.PDF](http://www.agpd.es/portalwebAGPD/canaldocumentacion/sentencias/tribunal_constitucional/common/pdfs/Sentencia2901.PDF)>.

9. Sentència 292/2000 del Tribunal Constitucional d'Espanya [En línia]. <[http://www.agpd.es/portalwebAGPD/canaldocumentacion/sentencias/tribunal\\_constitucional/common/pdfs/Sentencia292.pdf](http://www.agpd.es/portalwebAGPD/canaldocumentacion/sentencias/tribunal_constitucional/common/pdfs/Sentencia292.pdf)>.



ció informativa formava part central del seu argumentari, de manera que:

- Primer, «el dret a l'autodeterminació informativa és un dret actiu de control sobre el conjunt d'informacions relatives a una persona».
- En segon lloc, «el dret a la protecció de dades garanteix als individus un poder de disposició sobre les seves dades personals».
- I en tercer lloc, que aquest poder de disposició sobre les pròpies dades personals no val res si l'afectat desconeix quines dades seves són les que tenen tercers en el seu poder, qui són aquests tercers i amb quina finalitat tenen les seves dades.

Per tant, l'escenari proposat pels models de serveis basats en *cloudcomputing* impacta directament en el contingut de l'autodeterminació informativa, i en conseqüència del dret fonamental a la protecció de dades de caràcter personal.

### 2.3 Desequilibri en les condicions d'ús i en els contractes de serveis *cloud*

Les pràctiques contractuals més comunes del mercat de *cloud* es fonamenten en la proposta de clàusules estàndard per part dels proveïdors; i aquí voldria fer referència a com expressa un dictamen del Grup de Treball de l'Article 29 sobre la Protecció de Dades (Dictamen 1/2010 sobre els conceptes de «responsable del tractament» i «encarregat del tractament», de 16 de febrer de 2010<sup>10</sup>): «el desequilibri en la capacitat de negociar els contractes per part de petits i mitjans responsables de tractaments, respecte dels grans proveïdors de serveis de *cloudcomputing*, no justifica la acceptació de clàusules i condicions que no compleixin amb la legislació de protecció de dades.»

Hi ha una certa tendència a aglutinar sota el concepte de *cloudcomputing* situacions molt heterogènies que simplifiquen en excés un escenari que és força complex, no ja tecnològicament, sinó per les diferents relacions contractuals o de servei que es poden arribar a generar:

- a. Contractació de serveis a un proveïdor de *cloud* per part d'una empresa, en el context de l'externalització dels seus serveis de TIC (es tracta d'una situació força habitual, en què una part, o inclús tots els serveis de tecnologies de la informació són prestats per un tercer; un dels serveis més habituals és el de correu electrònic o el d'allotjament de llocs web).
- b. Però aquest proveïdor a la vegada pot tenir contractada, totalment o parcialment, la seva infraestructura TIC a un altre proveïdor de serveis *cloud* (sovint el proveïdor principal té una part del seu servei subcontractada a una altra empresa; per exemple, els servidors o les telecomunicacions).
- c. Si l'entitat que contracta el serveis *cloud* pertany al sector públic, caldrà afegir d'altres consideracions al servei contractat (normes de contractació del sector públic o especificacions concretes en matèria de seguretat, derivades per exemple de l'Esquema Nacional de Seguretat<sup>11</sup>).
- d. Si l'objecte del contracte implica el tractament de dades de caràcter personal, tenim una nova variable: ens trobarem amb tota la casuística derivada de l'aplicació del marc jurídic derivat del dret a la protecció de dades de caràcter personal.
- e. A títol individual, ja sigui per un ús domèstic o professional, podem optar per contractar serveis de *cloudcomputing* (per exemple, emmagatzemament Dropbox, ús de Google Apps o IDrive Apps de Microsoft, etc.).

10. Grupo del Artículo 29 sobre la Protección de Datos. Dictamen 1/2010 sobre los conceptos de «responsable del tratamiento» y «encargado del tratamiento» [En línia]. <[http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169\\_es.pdf](http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2010/wp169_es.pdf)>.

11. Boletín Oficial del Estado. Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica [En línia]. <<http://www.boe.es/buscar/pdf/2010/BOE-A-2010-1330-consolidado.pdf>>.

- f. A vegades, en situacions d'ús domèstic o privat de serveis de la societat de la informació, el proveïdor del servei és possible que estigui utilitzant serveis de *cloud*, propis o de tercers, sense que siguem conscients d'aquesta circumstància (xarxes socials, correu electrònic, etc.).
- g. I moltes altres casuístiques.

D'aquestes diverses casuístiques i de la seva combinació, m'he permès fer una certa agrupació dels tipus de *cloud* a què ens podem arribar a enfrontar:

<b>Cloudcomputing empresarial (inclou sector públic)</b>	Implica una decisió d'externalització de les TIC a mida; per exemple, el cas BBVA <sup>12</sup> .
<b>Cloudcomputing en cadena</b>	Hi ha diferents proveïdors de <i>cloud</i> implicats (aplicació, infraestructura, etc.).
<b>Cloudcomputing personal</b>	Quan individualment decidim utilitzar aquests serveis (tant per a ús particular com professional).
<b>Cloudcomputing invisible</b>	Es tracta d'una variant de l'anterior; simplement no sabem que estem contractant serveis de <i>cloud</i> .

En cap cas aquesta diversitat de supòsits pot suposar una disminució del nivell de protecció de les dades de caràcter personal: els principis, drets i obligacions que preveu la vigent *Directiva Europea de Protecció de Dades*, o els que preveu la Proposta de Reglament Europeu<sup>13</sup>, han de ser d'aplicació a totes i cadascuna d'aquestes casuístiques de tractament de dades personals, de manera que l'element clau serà de quina manera es formalitzen aquestes diferents situacions.

## 2.4 Les polítiques europees i el *cloud*: mercat únic digital

De fet, tal i com recorda el Supervisor Europeu de Protecció de Dades, en la seva opinió<sup>14</sup> sobre la estratègia europea de *cloud* (de novembre de 2012), no hi ha cap motiu per rebaixar el nivell de protecció en entorns *cloud*, respecte d'altres models de processament de la informació més convencionals: «not lead to a lowering of data protection standards as compared to those applicable to conventional data processing operations».

Les institucions europees, en multitud de documents, han manifestat la necessitat de generar un clima de confiança digital i seguretat jurídica com a base per al desenvolupament dels serveis de *cloud*. És per això que un dels aspectes en els quals cal aprofundir és en la qüestió de la contractació en el *cloud*, ja que si té una resposta eficaç (una bona solució), la principal problemàtica en aquest model de serveis, com ara la de determinació de les responsabilitats, es podrà substanciar més clarament.

L'estratègia de *cloudcomputing* té el seu origen en la comunicació de la Comissió d'11/01/2012, sobre «Un marc coherent per augmentar la confiança en el mercat únic digital del comerç electrònic i dels serveis en línia» –COM(2011) 942<sup>15</sup>–, en la qual una de les mesures previstes era l'adopció, durant el 2012, per part de la Comissió, d'una estratègia europea global sobre la computació en el núvol, que devia abordar la dimensió jurídica i tecnològica, així com la normalització, a fi d'estimular el sector del *cloud* i oferir la seguretat jurídica necessària a tots els agents participants en el desplegament del serveis *cloud*, i en especial als operadors econòmics.

12. «El BBVA se pasa a Google Apps» [En línia], *El País*, 11 de gener de 2012. <[http://tecnologia.elpais.com/tecnologia/2012/01/11/actualidad/1326276065\\_850215.html](http://tecnologia.elpais.com/tecnologia/2012/01/11/actualidad/1326276065_850215.html)>.
13. Proposta que ha rebut més de 3.000 esmenes, i que, per tant, la seva aprovació final serà complexa. Es pot consultar la proposta de gener de 2012 a [eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0011:FIN:ES:PDF](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0011:FIN:ES:PDF).
14. Opinion of the European Data Protection Supervisor on the Commission's Communication on «Unleashing the potential of Cloud Computing in Europe» [En línia]. <[https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-11-16\\_Cloud\\_Computing\\_EN.pdf](https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2012/12-11-16_Cloud_Computing_EN.pdf)>.
15. «Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Un marco coherente para aumentar la confianza en el mercado único digital del comercio electrónico y los servicios en línea» [En línia]. COM(2011) 942 final. Bruselas, 11.1.2012. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0942:FIN:es:PDF>>.



©istockphoto/leminuit

La Comissió considera que la «identificació i difusió de les millors pràctiques en matèria de models de condicions contractuals accelerarà l'acceptació de la computació en el núvol, a l'augmentar la confiança dels clients potencials». Òbviament, l'adopció de clàusules contractuals tipus també afavoreix un millor nivell de protecció de dades de caràcter personal.

En global, tot encaixa amb la voluntat de generar un marc de confiança digital, basat en un mercat únic digital, creant un adequat clima de seguretat i confiança, tant per als consumidors com per als proveïdors o comerciants, que afavoreixi l'adopció activa del *cloud* a Europa, connectant amb d'altres propostes de la Comissió, com la d'un reglament relatiu a una normativa comuna de compravenda europea, d'11/10/2011 COM(2011) 635, en què, per exemple, es regula el «contracte a distància».

Com deia, les institucions europees, en diferents manifestacions i documents, han identificat com una barrera al desplegament dels serveis *cloud* la manera en què es plantegen els termes i condicions contractuals.

En aquest sentit, s'estan duent a terme accions adreçades a millorar els mecanismes de contractació dels serveis *cloud*,

resultant d'especial rellevància la Comunicació de la Comissió Europea COM(2012) 529, del passat 27/09/2012, en la qual, amb el títol «Alliberar el potencial de la informàtica en núvol a Europa»<sup>16</sup>, s'estableix l'estratègia de *cloud computing* a Europa, i en la qual la Comissió planteja com a punt de partida la necessitat de «facilitar una adopció més ràpida de la informàtica en núvol en tots els sectors de l'economia, el que pot reduir els costos informàtics, així com, en combinació amb noves pràctiques empresarials digitals, impulsar la productivitat, el creixement i l'ocupació».

Segons aquesta comunicació, un dels àmbits en què cal prendre mesures és la contractació, a la qual es refereix com a «problemes amb els contractes», que es manifesten en preocupacions i inquietuds sobre:

- accés a les dades i la seva portabilitat,
- control de les modificacions,
- propietat de les dades,
- determinació de la responsabilitat en cas de fallada dels serveis (temps d'inactivitat o pèrdua de dades),
- resolució de conflictes, etc.

L'estratègia de *cloudcomputing* de la Comissió proposa tres accions específiques:

<b>Acció 1</b>	Establir un marc fiable de normes i estàndards (seguretat, interoperabilitat, portabilitat, etc.), recolzant-se en sistemes de certificació, i connectat amb la proposta de Reglament sobre la normalització europea de 01/06/2012, COM(2011) 315.
<b>Acció 2</b>	Condicions contractuals segures i justes. Elaborant «clàusules contractuals tipus» (per a finals del 2013) tenint en compte les peculiaritats d'àmbits i sectors com la compravenda, els consumidors, les petites i mitjanes empreses, els professionals, transferència de dades personals a tercers països, normes corporatives vinculants per als proveïdors de serveis <i>cloud</i> , etc.
<b>Acció 3</b>	Crear una Associació Europea d'Informàtica en núvol per impulsar la innovació i el creixement des del sector públic, per determinar els requisits del sector públic pels serveis en <i>cloud</i> , elaborar especificacions per a la contractació de TI o avançar cap a la contractació conjunta de serveis de informàtica en núvol pels organismes públics.

Ja he fet referència que en el desplegament dels serveis *cloud* té una especial transcendència la regulació del dret a la protecció de dades de caràcter personal, no hi ha dubte del valor econòmic de la informació, i en especial de les dades personals (en aquest sentit ha tingut èxit l'expressió que les dades són el «petroli de la societat de la informació»), i que, per tant, una bona part de la informació processada en el núvol és i serà de caràcter personal.

Tal i com va expressar la Sra. Kroes (European Commission Vice-President for the Digital Agenda), en una conferència a París, ja el novembre de 2010, «... a "cloud" without clear and strong data protection is not the sort of cloud we need»<sup>17</sup>.

En diferents estudis, la preocupació per la protecció de les dades personals es considera un dels més greus obstacles per a l'acceptació de la informàtica en núvol; és per això que la proposta de Reglament europeu de protecció de dades, a la qual ja he fet referència, és un instrument imprescindible per a aquesta estratègia, especialment pel que pot aportar el fet de disposar d'una sola regulació (en comptes de 27) o que es proposin mesures de simplificació de les transferències internacionals:

- clàusules tipus acordades per la Comissió o una Autoritat,
- normes corporatives vinculants (*Binding Corporate Rules*, BCR),
- o inclús en determinades circumstàncies, interès legítim de responsable o encarregat.

D'aquí que expressament en aquest document es destaquï la importància que el Consell i el Parlament treballin amb rapidesa per aconseguir l'adopció de la proposta de reglament el més aviat possible, en el 2013, tot i les dificultats que s'estan trobant.

A efectes pràctics, la situació actual de les condicions i els termes dels contractes de *cloud* han estat analitzats en detall des de la perspectiva de la protecció de dades de caràcter personal i per a casos concrets (Decisió de la agència de protecció de dades danesa<sup>18</sup>, en relació a la contractació dels serveis *Google Apps* per part d'un ajuntament de Dinamarca "Odense", o la consulta a l'APDCAT<sup>19</sup> per part d'un consell comarcal, també sobre la contractació del mateix servei). En tots dos casos la consulta plantejava si la contractació dels esmentats serveis, segons les condicions i clàusules contractuals proposades pel proveïdor, compleixen amb les previsions de la Directiva Europea de

16. «Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social Europeo y al Comité de las Regiones. Liberar el potencial de la computación en nube en Europa» [En línia]. COM(2012) 529 final. Bruselas, 27.9.2012. <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0529:FIN:ES:PDF>>.

17. Neelie Kroes, «Cloud computing and data protection» [En línia], Europa, 2010 <[http://europa.eu/rapid/press-release\\_SPEECH-10-686\\_en.htm](http://europa.eu/rapid/press-release_SPEECH-10-686_en.htm)>.

18. «Processing of sensitive personal data in a cloud solution» [En línia], Datatilsynet, 3 de Febrer de 2011. <<http://www.datatilsynet.dk/english/processing-of-sensitive-personal-data-in-a-cloud-solution>>.

19. «Dictamen en relació a la consulta formulada per un consell comarcal sobre la contractació del servei Google Apps for Business» [En línia], APDCAT. <[http://www.apd.cat/media/dictamen/ca\\_476.pdf](http://www.apd.cat/media/dictamen/ca_476.pdf)>.

... la consulta plantejava si la contractació dels esmentats serveis, (...), compleixen amb les previsions de la Directiva Europea de protecció de dades i amb les respectives legislacions nacionals.

protecció de dades i amb les respectives legislacions nacionals. En tots dos casos la resposta és molt similar. En síntesi, es diu que:

- Es detecten serioses dificultats per conèixer quin és l'entramat de termes i condicions aplicables a la contractació, el que «origina una evident incertesa sobre les condicions exactes en què les dades personals seran tractades per l'encarregat del tractament».
- Circumstància agreujada pel fet que es preveu contractualment que els termes i les condicions del contracte podran ser modificats unilateralment pel proveïdor dels serveis.
- No s'indica la localització dels centres de processament de dades: per motius de seguretat.
- No informen sobre les subcontractacions: referència genèrica al grup d'empreses i subcontractats.
- No indiquen el nivell de mesures de seguretat: a banda d'algunes vagues especificacions (auditoria SAS70) del tot insuficients, almenys en relació a la legislació que s'aplica a l'Estat espanyol.

- No hi ha previsions respecte del control que ha d'exercir el responsable del tractament (cal recordar que una de les principals obligacions és vetllar pel compliment de la normativa de protecció de dades per part d'aquell a qui contractem).
- I una de les qüestions més greus: no s'exclouen altres possibles tractaments, per a finalitats diferents per a les que les dades van ser recollides, o per a la pròpia finalitat del contracte (al contrari, es parla d'indexar continguts per a altres funcions, de millora de l'experiència del usuari).

Mentre s'aprova el Reglament, en opinió de la Comissió resulta rellevant el Dictamen del Grupo de Protección de Datos del Artículo 29 sobre la informàtica en núvol, de 1 de juliol de 2012<sup>20</sup>, ja que proporciona una bona base per a la transició en relació al *cloud*, des de la vigent Directiva de la UE sobre protecció de dades cap al nou Reglament de la UE sobre protecció de dades.

La preocupació respecte de l'impacte dels serveis *cloud* en les obligacions i els drets derivats de la protecció de les dades de caràcter personal s'ha manifestat des de diferents grups de treball sobre la matèria i des de les autoritats de control. Per exemple, resulten d'interès el *Sopot Memorandum*, sobre «Cloud Computing–Privacy and data Protection issues» del IWGDPT, d'abril de 2012, o la més recent «Resolution on cloudcomputing», de la 34a Conferència Internacional d'Autoritats de Protecció de Dades i Privacitat (Uruguai, 26 d'octubre de 2012). També s'han pronunciat durant el 2012 les autoritats de control a títol individual: CNIL, Garante Italiano, Autoritat Irlandesa, ICO, etc.

20. «Dictamen 5/2012 sobre la computació en nube» [En línia], Grupo de Protección de Datos del Artículo 29. <[http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196\\_es.pdf](http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2012/wp196_es.pdf)>.

...aplicar eines ja utilitzades de fa temps, amb el propòsit de fixar d'una manera clara aspectes relacionats amb la protecció de dades, igual que es fa amb els anomenats «acords de nivell de servei»...

### 3. Actuacions en desenvolupament connectades amb el *cloud* i la privacitat

Ja per finalitzar, voldria fer referència a algunes actuacions pràctiques que tenen a veure amb la problemàtica tractada al llarg de l'article. Per descomptat, donada la transcendència de la qüestió, hi ha molts projectes i experts implicats a estudiar les vies de solució per minimitzar els riscos a què he fet referència. En aquest sentit, jo només em referiré a dues iniciatives molt concretes, no perquè siguin especialment importants, sinó com una mostra dels diferents mecanismes que, tant des de la indústria com des dels poders públics, s'estan articulant per crear el clima de confiança en el *cloud* al qual he fet referència en diverses ocasions, i sobre les quals o bé he tingut una participació directa o bé he analitzat amb una certa profunditat. Per tant, em referiré a elles com uns exemples d'entre molts més sobre els quals s'està treballant.



©istockphoto/belekekin

Per una banda, la iniciativa de la Cloud Security Alliance<sup>21</sup> (recentment ha estat traduïda al castellà<sup>22</sup>) i la seva proposta d'acords de nivell sobre privacitat o sobre protecció de dades (PLA, *Privacy Level Agreement*) a incorporar als contractes de serveis *cloud*.

En realitat, es tracta d'aplicar eines ja utilitzades de fa temps, amb el propòsit de fixar d'una manera clara aspectes relacionats amb la protecció de dades, igual que es fa amb els anomenats «acords de nivell de servei» (SLA, *Services Level Agreement*), que regulen aspectes de prestació dels serveis TIC (temps de disponibilitat, aturades planificades, penalitzacions, etc.).

La funció del PLA és descriure amb claredat de quina manera el proveïdor té previst complir amb els requisits, ja siguin de caràcter normatiu o de bones pràctiques, relacionats amb les dades personals que seran objecte de tractament en virtut del contracte de serveis *cloud*.

21. Cloud Security Alliance ([cloudsecurityalliance.org/research/pla](http://cloudsecurityalliance.org/research/pla)).

22. «Acuerdo de Nivel de Privacidad» [En línia], Asociación Española para el Fomento de la Seguridad de la Información. <<https://www.ismsforum.es/noticias/noticia.php?idnoticia=460>>.

Tenen per objectiu:

- Proporcionar als clients dels serveis *cloud* una eina per avaluar i comparar el compromís d'un proveïdor per abordar la protecció de dades personals.
- Proporcionar als proveïdors una eina per divulgar d'una manera estructurada les seves pràctiques de protecció de dades.

I l'altre, el dret a la portabilitat, que s'incorpora com a nou dret en la proposta de Reglament europeu de protecció de dades.

Com deia l'article 18 de la Proposta de reglament, recull el dret a la portabilitat de les dades personals:

1. *Cuando se traten datos personales por vía electrónica en un formato estructurado y comúnmente utilizado, el interesado tendrá derecho a obtener del responsable del tratamiento una copia de los datos objeto de tratamiento en un formato electrónico estructurado y comúnmente utilizado que le permita seguir utilizándolos.*
2. *Cuando el interesado haya facilitado los datos personales y el tratamiento se base en el consentimiento o en un contrato, tendrá derecho a transmitir dichos datos personales y cualquier otra información que haya facilitado y que se conserve en un sistema de tratamiento automatizado a otro sistema en un formato electrónico comúnmente utilizado, sin impedimentos por parte del responsable del tratamiento de quien se retiren los datos personales.*
3. *La Comisión podrá especificar el formato electrónico contemplado en el apartado 1 y las normas técnicas, modalidades y procedimientos para la transmisión de datos personales de conformidad con lo dispuesto en el apartado 2. Dichos actos de ejecución se adoptarán con arreglo al procedimiento de examen contemplado en el artículo 87, apartado 2.*

Tal i com explicita l'exposició de motius de la Proposta de reglament, el dret a la portabilitat atribueix dues facultats a les persones interessades:

- Obtenir una còpia de les seves dades personals en un format electrònic estructurat i d'ús habitual.
- Transferir les seves dades, i d'altres informacions que hagi facilitat, d'un sistema de tractament electrònic a un altre.

El titular del dret a la portabilitat de les dades personals és la persona interessada, és a dir, aquella persona física a la qual es refereixen les dades personals sobre les quals es pretén exercir el dret a la portabilitat.

Tal i com vaig expressar en la meua participació en el congrés de la Red Derecho-TICS<sup>23</sup> el cert és que el dret a la portabilitat de les dades es pot entendre com una versió *premium*<sup>24</sup> del dret d'accés, que implica el deure jurídic del responsable del tractament que, a petició de la persona interessada, confirmi si tracta o no dades personals del sol·licitant, i en cas afirmatiu informi de la finalitat del tractament i origen de les dades, o si és el cas, si ha realitzat comunicacions de les dades a tercers, o té previst fer-ne.

Evidentment, l'exercici d'aquest dret comporta unes dificultats d'ordre pràctic; la seva necessitat no sembla que hagi de ser objecte de discussió, donat l'actual escenari digital, en què una ingent quantitat d'informació personal textual i multimèdia es mou per les xarxes socials o pels serveis de *cloudcomputing*, als quals confiem l'emmagatzemament i el processament massiu d'informació personal, cosa que obliga a preveure mecanismes que evitin, adduint raons tecnològiques, el «segrest» de la informació o acabar «captius digitals» d'un proveïdor concret.

23. «Libertad, transparencia y política en internet: ejercicio, amenazas y garantías», congrés organitzat per la Red Derecho-TICS, Madrid 18-19 octubre 2012 ([www.derechotics.com/congresos/2012-libertades-y-open-gov](http://www.derechotics.com/congresos/2012-libertades-y-open-gov)).

24. El mateix podríem dir del dret a l'oblit en relació als actuals drets de cancel·lació i oposició.

## 4. Conclusions

Els serveis *cloud* tenen per endavant un llarg recorregut; els aspectes funcionals, tecnològics i econòmics evidencien amb claredat que es tracta d'un model que, amb les evolucions que pertoquin, serà durant molt temps la base per a la prestació i obtenció de serveis relacionats amb les tecnologies de la informació i la comunicació.

Però han d'evolucionar en positiu si es vol que siguin motor de l'economia europea; per tant, cal que els desequilibris contractuals que es donen en aquests moments se solucionin mitjançant unes condicions i termes més clars i, en especial, més adaptats als requeriments que es deriven de la normativa europea de protecció de dades.

Si usuaris finals i empreses es troben segurs davant de les propostes contractuals i de serveis dels proveïdors, sens dubte aquests podran experimentar un ràpid creixement; tot i així, a banda dels aspectes contractuals també caldrà que la indústria treballi en la definició de solucions que donin resposta eficaç als compromisos contractuals.

## 5. Bibliografia

Addicionalment a les publicacions, articles i recursos a la xarxa a què s'ha fet referència al llarg de l'article, també poden resultar d'interès les següents publicacions:

CIERCO, David (coord.). *Cloud Computing: retos y oportunidades*. Madrid: Fundación IDEAS, 2011.

GUTWIRTH, S.; POULLET, Y.; DE HERT, P.; LEENES, R. (eds.). *Computers, Privacy and Data Protection: an Element of Choice*. Dordrecht: Springer, 2011.

TORRES, Jordi. *Empresas en la nube*. Barcelona: Llorenç Rubió, 2011.



# La reputació corporativa de les biblioteques en els mitjans socials

**Nieves González Fernández-Villavicencio**

Universidad P. de Olavide / Universidad de Sevilla

*ngonfer@upo.es*

## **Resum:**

La presència de les organitzacions en els mitjans socials està aconseguint un punt de maduresa, al mateix temps que augmenta la necessitat de gestionar la reputació digital, per l'impacte que produeixen aquests mitjans en un entorn digital i en xarxa.

Les biblioteques, com qualsevol altre tipus d'organització, han de tenir un pla en els mitjans socials que garanteixi una adequada gestió de la seva identitat digital i de la reputació de la marca, que augmenti el prestigi i l'autoria en el seu entorn d'influència, així com el nombre i fidelitat dels seus predictors, i estableixi les pautes per a la gestió de crisis, amb els recursos adequats.

En qualsevol cas, cal establir un sistema de mètriques que avaluï i monitoritzi els indicadors de reputació i influència, i permeti la presa de decisions.

## **Paraules clau:**

reputació digital, imatge de marca, biblioteques, indicadors, KPI, mètriques, mitjans socials, màrqueting digital

## La reputación corporativa de las bibliotecas en los medios sociales

### Resumen:

La presencia de las organizaciones en los medios sociales está alcanzando un punto de madurez, al mismo tiempo que aumenta la necesidad de gestionar la reputación digital, por el impacto que producen estos medios en un entorno digital y en red.

Las bibliotecas, como cualquier otro tipo de organización, deben tener un plan en los medios sociales que garantice una adecuada gestión de su identidad digital y de la reputación de la marca, que aumente su prestigio y autoría en su entorno de influencia, así como el número y fidelidad de sus predictores, y establezca las pautas para la gestión de crisis, con los recursos adecuados.

En cualquier caso, hay que establecer un sistema de métricas que evalúe y monitorice los indicadores de reputación e influencia, y permita la toma de decisiones.

### Palabras clave:

reputación digital, imagen de marca, bibliotecas, indicadores, KPI, métricas, medios sociales, marketing digital

## Corporate reputation of libraries in social media

### Abstract:

The presence of organizations in social media is reaching a degree of maturity and, at the same time, it increases the need to manage reputation online, due to the impact of social media in a digital and networked environment.

Libraries, as any other type of organizations, need to have a social media plan, in order to ensure proper management of digital identity and brand reputation, increase their prestige and influence in this environment –engagement and advocacy in stakeholders–, and establish the priorities for an eventual crisis by appropriate resources.

In any case, we have to articulate a serie of indicators that seeks to evaluate the library's reputation and influence for the purpose of decision-making.

### Keywords:

Online reputation, Brand image, libraries, KPI, metrics, social media, social media marketing

## 1. Ús dels mitjans socials per les organitzacions i conceptes relacionats amb la reputació digital corporativa. Els *stakeholders*

La Xarxa és, avui dia, un canal fonamental per a la interacció i la comunicació entre persones, empreses i institucions. Les xarxes socials han passat a formar part del dia a dia dels internautes espanyols. Els resultats de la 5ª Oleada Observatorio de RRSS de The Cocktail Analysis<sup>1</sup> mostren com el 93% dels usuaris d'Internet té almenys un compte actiu a les xarxes socials. Alhora, vivim un moment de desacceleració, amb la qual cosa podem dir que s'ha arribat a una fase de maduresa en la utilització d'aquests canals.

Els mitjans socials han vingut per quedar-se, per integrar-se com una via eficient de comunicació i atenció a usuaris, d'*engagement* (compromís), de conversa, de creació de continguts en comunitat, però també es tracta d'un indicador de la reputació digital i de la imatge corporativa, i del prestigi de la marca.

Avui dia, l'èxit d'una organització no dependrà exclusivament de la qualitat dels seus productes o serveis, ni de l'èxit de les seves campanyes de màrqueting, sinó que dependrà també de la seva responsabilitat amb l'entorn, de la transparència i la qualitat d'ús, de l'aportació a la societat i de l'existència de vincles sòlids, estables i duradors amb els seus grups d'interès. Internet està fomentant noves formes de relació entre les organitzacions i els seus *stakeholders* –els seus grups d'interès–, donant origen a un nou model d'influència i informació i a un canvi de direcció de la influència en les persones, en el seu paper de consumidors (Castelló, 2010).

Reputació *online* o digital, imatge corporativa i de marca, reputació corporativa, *e-reputació*... Què volen dir aquests termes i per què són importants avui per a les biblioteques?



©istockphoto/burakpekakan

La Viquipèdia defineix **Reputation management** o **reputació online**<sup>2</sup> com «...una tècnica per al control de la reputació i l'ús que es fa de les marques en el mitjans d'informació i zones d'influència informativa», en què «ja no val només el que digui l'empresa o el mitjà d'informació sinó que cal tenir molt en compte el que diu la societat sobre les marques i les companyies per petites o grans que siguin aquestes». La reputació digital és possiblement el valor principal d'una marca i inclou altres valors com notorietat, prestigi, fama, autoritat, confiança, estima o popularitat.

Per a Alloza (2012), el concepte de marca no depèn en exclusiva del que l'organització diu de si mateixa: «La marca corporativa es el balance que una empresa, institució o país consigue establecer entre las percepciones que comunica, es decir, las expectativas que crea y las realidades o experiencias que entrega a todos sus grupos de interés, tanto internos como externos», grups que ara es troben interconnectats. «Una marca es fuerte, creíble, generadora de confianza y de buena reputación, en la medida

1. Estudi disponible a <<http://the-cocktail.com/blog/posts/5-oleada-observatorio-de-rrss-el-70-de-los-usuarios-se-muestran-receptivos-a-la-presencia-de-marcas>>.  
2. Definició de reputació *online* a la Viquipèdia <[http://ca.wikipedia.org/wiki/Reputation\\_management](http://ca.wikipedia.org/wiki/Reputation_management)>.

*Si vols gaudir d'una bona reputació, preocupa't de ser el que aparentes ser.*

Sòcrates

en que no existe disonancia entre las percepciones o expectativas y las experiencias o realidades» (Alloza, 2012); i això es pot resumir en la frase «hacer lo que has dicho que ibas a hacer, hacerlo bien y comunicarlo» (Instituto de Análisis de los Intangibles, 2010).

Segons estudis, i des del punt de vista del risc reputacional, quan una empresa té bona reputació, els clients la recomanen fins a un 72,3%, i els detractors tan sols suposen un 3%; però quan aquesta reputació no és bona, només es recomana en un 9,4% dels casos, mentre que els detractors s'elevan al 40% (Alloza, 2012).

Per a Alfonso Alcántara, @Yoriento, especialista en la empresa 2.0: «Marca es el perfume que usas, reputación es el olor que dejas»<sup>3</sup>.

Avui dia, les organitzacions competeixen per aquests valors intangibles, la marca i la reputació corporativa, i estem assistint a un canvi de paradigma, un nou cicle econòmic que s'anomena «Economía de la reputación, que se basa



en indicadores no financieros como la reputación, la marca, la satisfacción de los clientes o el compromiso de los empleados» (Alloza, 2012).

Un altre concepte que hem esmentat i que té també un enorme impacte en el context de la reputació és la responsabilitat social corporativa (RSC): «Se denomina así al conjunto de obligaciones y compromisos con los grupos de interés, de tipo legal y ético, nacionales e internacionales, que se derivan del impacto que la actividad de la organización produce en el ámbito social, laboral, medioambiental y de los derechos humanos» (Valor, 2003).

El poder es trasllada, en el context actual, als grups d'interès, i el nou rol de les empreses i de les institucions és estar al seu servei. Es fa imprescindible mantenir una estreta relació amb els públics de l'organització, basada en el diàleg i la transparència de la comunicació. Són precisament els públics més propers els prioritaris per a una organització en aquest procés, ja que la confiança es construeix de dins cap a fora: «Empieza con los empleados que

3. URL del missatge de Twitter: <<https://twitter.com/Yoriento/status/245930786849685504>>.

recomiendan la empresa y sus productos a los clientes, después los clientes a los no clientes, y los no clientes a la sociedad en su conjunto» (Alloza, 2012).

La reputació de l'empresa entre els empleats es consolida com una de les claus i com a tendència de futur, però també la pròpia empremta digital dels treballadors influeix en la reputació corporativa. Poden arribar a ser els més grans ambaixadors de l'empresa o els seus principals detractors. Empreses com IBM tenen cura de la reputació digital dels seus treballadors, aportant-los formació i recomanacions per gestionar els seus perfils en els mitjans socials<sup>4</sup>.

D'altra banda, en l'àmbit digital emergeix amb força la figura del líder d'opinió o influenciador, tant en l'àrea de l'actualitat informativa com en el complex entramat de decisions sobre entreteniment, consum, tecnologia o tendències socials. La Word of Mouth Marketing Association (WOMMA) defineix la influència com l'habilitat per produir o contribuir a un canvi en l'opinió o el comportament de les persones; i l'influenciador és aquella persona o grup que posseeix un potencial d'influència superior a la mitjana a causa de diversos factors, com poden ser la freqüència de la comunicació, el poder de persuasió i la mida i la posició central a la xarxa<sup>5</sup>. A través de l'activitat en els mitjans socials dels influenciadors –publicació de piulades i entrades de blogs, actualització d'estat a les xarxes socials o recomanació de lectures i vídeos–, i de forma més o menys conscient, compleixen amb la funció de generar continguts, o de filtrar cap a altres usuaris sobre els quals tenen certa ascendència (Serrano-Puche, 2012).

Amb tot, la millor manera d'exercir una bona influència en el sector segueix sent oferir bons productes i serveis de qualitat, que faci que els clients/usuaris difonguin i recomanin la marca als altres, i arribar a ser una organització que funcioni bé, de tal manera que tant treballadors com col·laboradors (*stakeholders*), recomanin i defensin treballar amb l'organització.

Per tot això, tenir una bona salut a les xarxes socials és imprescindible. No obstant, encara hi ha empreses, organitzacions, a les quals els costa admetre un error o acceptar un suggeriment. Aquest tipus d'empreses veuen la xarxa com un atac incontrolable als seus punts febles, però es tracta d'una forma d'estar més a prop dels clients (Sánchez, 2012). La gestió de la seva reputació digital hauria de ser considerada una prioritat.

Les organitzacions, per tant, han de dur a terme una activitat d'alerta per aprendre a relacionar-se amb els diferents grups d'interès. Encara que la major dificultat consisteix a aconseguir que aquestes relacions es consolidin, és imprescindible dissenyar una estratègia que contempli les necessitats de cada grup. Per això, el gestor de reputació s'ha de concentrar a detectar les expectatives dels diferents *stakeholders* cap a l'organització, que van més enllà de les missions pròpies de l'activitat de qualsevol índole que pugui desenvolupar (Vaquer, 2012).

## 2. Com estan gestionant les organitzacions la seva reputació online

Les organitzacions es caracteritzen en general per una falta de planificació clara en aquests canals i per no donar resposta a la demanda de conversa per part dels usuaris. Aquest fet afecta negativament en diversos sentits, però sobretot afecta la capacitat de la institució per obtenir informació valuosa pel que fa a gustos i preferències dels seus usuaris. Les opinions o judicis de valor que tenen lloc en els diversos mitjans socials aporten gran part de la reputació corporativa digital, que ara és possible identificar i analitzar a través de les valoracions que els usuaris bolquen a la xarxa. No cal oblidar que quan una marca aconsegueix gran visibilitat en els mitjans, menys controla

4. Web d'IBM amb les recomanacions al seu personal: <[http://www03.ibm.com/employment/ibm\\_digital\\_reputation.html](http://www03.ibm.com/employment/ibm_digital_reputation.html)>.

5. Definicions del WOMMA: <<http://www.smmstandards.com/category/influence-relevance>>.

la seva imatge a la xarxa i que els usuaris comparteixen abans les males experiències amb les marques que les bones.

Encara que al principi les marques van utilitzar aquests canals com una forma de comunicació dels seus productes, serveis o notícies, amb el temps els usuaris els van exigir una resposta davant els seus problemes amb les marques. No obstant això, la realitat és que no es respon a l'usuari, tot i que se sap que una resposta ràpida de la marca produeix més satisfacció a l'usuari que la pròpia solució immediata del problema<sup>6</sup>. Nike, per exemple, és la marca que millor atén els seus clients a Twitter, amb una taxa de resposta del 73%<sup>7</sup>.

Un estudi de l'empresa de consultoria SocialBakers<sup>8</sup> indica que el 30% de les universitats del Regne Unit ignora la conversa que s'està produint a les seves pàgines de Facebook<sup>9</sup> o triguen massa a respondre als seus clients. Les universitats espanyoles, en general, tenen presència a la xarxa però no gestionen convenientment la seva imatge de marca. Destaquen, en aquest sentit, la Universitat de Girona o la de Granada, que compten amb un pla de màrqueting digital i un equip de persones que gestionen de manera eficient els mitjans socials i la reputació de la marca.

Qualsevol empresa o organització no pot oblidar que la seva imatge i reputació s'està configurant a la xarxa. Hi ha molts exemples de crisi que s'han resolt malament en els mitjans socials<sup>10</sup>, però també exemples d'una bona gestió d'aquesta reputació, com és el cas de la campanya *Engagement* de KLM<sup>11</sup>.

Per a Mut Camacho (2012), gestionar la reputació corporativa va més enllà d'accions de SEO (*Search Engine Optimization*), SEM (*Social Engine Marketing*), monitorització online, opinions personals en blogs o xarxes socials, o gestió de crisis als mitjans socials. És necessària la integració de la gestió *online* amb la gestió *offline* de la reputació d'una organització, i per tant, s'ha de gestionar sota els mateixos indicadors, que tindran en compte tant els valors corporatius com la jerarquizació de *stakeholders* (Vaquer, 2012). De tota manera, la gestió de la reputació digital institucional ha de formar part de l'estratègia de comunicació de l'organització i estar inclosa en els objectius de la institució.

Tampoc n'hi ha prou amb tenir un manual d'estil, d'identitat corporativa i una línia editorial, cosa que d'altra banda poques organitzacions tenen. Es tracta de definir i d'apropriar-se de la missió i els valors de l'organització, i «ser fieles a ellos con una estrategia que denote un conocimiento no sólo de los medios y de las herramientas, sino de los clientes/usuarios: por qué están ahí, qué necesitan, cómo empatizar con ellos más allá de los objetivos de negocio»<sup>12</sup>.

Per sobre de tot, la millor manera de gestionar la reputació *online* és tenir usuaris satisfets. Aquests usuaris satisfets seran els millors predictors de la marca, ja que el més difícil i car per a les empreses no és tenir clients, sinó mantenir-los, fidelitzar-los, i els mitjans socials juguen un paper fonamental en l'*engagement*, *advocacy*, atracció, implicació i defensa dels usuaris en relació a la marca.

Per gestionar les relacions amb els usuaris/clientes, els programes CRM (*Customer Relations Management*) haurien d'incloure tots aquells processos que ajuden a gestionar el cicle de vida amb els usuaris.

6. Més informació a: <<http://www.dimensionresearch.com>>.

7. Dades obtingudes a: <<http://www.puromarketing.com/42/16503/marcas-mejoran-atencion-cliente-traves-twitter.html>>.

8. Més informació a: <<http://www.puromarketing.com/42/16033/tardan-demasiado-marcas-responder-fans-redes-sociales.html>>.

9. Accessible a: <<http://www.socialbakers.com/blog/862-30-of-universities-in-the-uk-ignore-their-facebook-fans>>.

10. Exemples de crisis en els mitjans socials: <<http://www.lesechos.fr/entreprises-secteurs/tech-medias/actu/0202659203536-comment-les-entreprises-soignent-leur-e-reputation-554858.php>>.

11. Més informació a: <<http://blog.klm.com/klm%E2%80%99s-social-media-strategy-part-2/4700>>.

12. Dolores Vela. «¿Cuidamos nuestra imagen y reputación online?», *Social Media Strategies* [blog]. <<http://www.socialmediacm.com/cuidamos-nuestra-reputacion-online>>.

### 3. Reputació corporativa a les biblioteques

Com es pot gestionar la reputació digital corporativa de les biblioteques? Javier Celaya, en el V Congreso de Bibliotecas Públicas<sup>13</sup>, afirmava que les biblioteques compten amb un gran avantatge competitiu, una «muy buena imagen de marca», i per tant haurien d'utilitzar aquesta posició avantatjosa en la gestió de la seva reputació digital. No obstant això, comptar amb una bona imatge de marca, en aquest entorn digital i en xarxa, no és suficient.

Quina és la percepció que els usuaris tenen de les seves biblioteques? Segueixen considerant-les magatzems de llibres? S'ha adaptat la imatge de marca dels seus bibliotecaris a la nova societat digital? *LibraryScienceList.com* va publicar l'abril de 2013 una infografia<sup>14</sup> basada en dades d'un informe del centre d'investigació PewInternet, sobre les biblioteques del futur. Per il·lustrar-lo van utilitzar una imatge de bibliotecària massa tòpica –monyó, faldilla llarga i ulleres rodones–, encara que amb una tauleta digital a la mà, imatge que van canviar per una altra més actual davant les protestes generalitzades.

Hem vist recentment situacions difícils en què els mitjans socials s'han fet servir per organitzar la gent o respondre davant dels problemes, i aquesta flexibilitat i rapidesa de resposta és la que va contribuir més que cap altra acció a definir la imatge de marca de la biblioteca.

Un dels primers problemes a què s'enfronten les biblioteques és la seva falta de visibilitat i transparència en la gestió. Davant d'aquesta evidència, algunes biblioteques i serveis d'informació de tot el món, com l'Indianapolis Museum of Art (IMA)<sup>15</sup>, les biblioteques del NCSU<sup>16</sup> o la Harvard University<sup>17</sup>, estan publicant les dades d'ús de la



Figura 1. Web de l'IMA (Indianapolis Museum of Art)

biblioteca o dels mitjans socials a través de panells de control que es van actualitzant en els seus portals web.

La majoria de les biblioteques tenen presència a les xarxes socials i utilitzen aquests mitjans per aportar continguts i establir nous canals de comunicació amb els seus usuaris, per a activitats de màrqueting i promoció i per millorar el servei que ofereixen, però no estan monitoritzant les seves accions en un tant per cent molt elevat (González, 2013), i per tant tampoc estan gestionant convenientment la seva reputació digital. Les dades aportades es limiten al nombre de seguidors en els mitjans socials o la freqüència de l'activitat, però no s'han aportat estudis sobre les interaccions dels usuaris amb la marca biblioteca, o el seu nivell de satisfacció o compromís amb ella. En línies generals, les biblioteques s'obliden de cuidar la reputació i protegir-la, per la qual cosa és més que necessari treballar amb plans de reputació i de gestió de crisis en l'entorn digital.

13. Disponible a: <<http://travesia.mcu.es/portalsnb/jspui/handle/10421/4971>>.

14. Vegeu la infografia a: <<http://www.nievesglez.com/2013/04/infografia-sobre-las-bibliotecas-del.html>>.

15. Disponible a: <<http://dashboard.imamuseum.org>>.

16. Disponible a: <<http://www.lib.ncsu.edu/bythenumbers>>.

17. Disponible a: <<http://librarylab.law.harvard.edu/toolkit>>.

En un estudi de la Biblioteca de la Universitat de Saragosa, de finals de 2012, es fa referència a l'*índex Klout* per a mesurar la reputació de les biblioteques universitàries espanyoles de REBIUN<sup>18</sup>. Estudis d'aquesta mena poden servir de referència per contrastar posicions entre institucions similars.

Existeixen estudis més concrets referits a una biblioteca o xarxa de biblioteques dels quals se'n pot extreure algunes conclusions sobre la reputació digital.

Aquest és el cas de les biblioteques municipals de la Corunya. Per a Catuxa Seoane «La labor de planificación y mantenimiento de los perfiles institucionales en las Bibliotecas Municipales de A Coruña recae en un responsable de Servicios Web 2.0 que se encarga de coordinar un grupo de trabajo de Web 2.0 con representantes de cada una de las 8 bibliotecas del sistema. De este modo nos aseguramos que exista una persona de referencia que se responsabilice de los contenidos, de responder los comentarios y de poner en marcha nuevos servicios y por otra parte se vertebrará de manera coordinada la actividad de los diferentes miembros del grupo 2.0»<sup>19</sup>. Una experiència similar podem trobar-la a la xarxa de biblioteques de la ciutat de Barcelona.

Natalia Arroyo va presentar en el congrés de Fesabid 2013<sup>20</sup>, i per al cas de les biblioteques de la Fundació Germán Sánchez Ruipérez, una proposta de mesurament en els mitjans socials basada en una sèrie d'indicadors, entre els quals destaquem el seguiment o fidelització i la interacció, que es poden prendre com a indicadors per mesurar la reputació digital.

Com a model a seguir, recomanem la Biblioteca Pública de Nova York, que no només gaudeix d'una immillorable reputació digital, sinó que utilitza programes com *Hootsuite*

te, per obtenir analítiques i coordinar l'atenció a l'usuari a través d'un grup de bibliotecaris que responen a les seves preguntes<sup>21</sup>.

Finalment, l'autora d'aquest article està duent a terme un estudi sobre divuit biblioteques i xarxes de biblioteques espanyoles, de diferent tipus, i entre les quals n'hi ha de l'àmbit català: Biblioteques de Barcelona, la Biblioteca de la Universitat Politècnica de Catalunya i la Biblioteca de Escola Sant Jordi (Lleida). En aquest estudi es pretén obtenir els costos de la inversió i els beneficis del màrqueting digital a les biblioteques. Entre els ítems que s'estudien hi ha la reputació digital de les biblioteques mesurada en objectius com abast, freqüència de l'activitat, fidelització o trànsit a la web, influència (mencions), rellevància (interaccions) i retorn (conversió).

## 4. Pla de gestió de la reputació digital corporativa (RDC)

Desenvolupar un pla de gestió de la reputació digital no és altra cosa que establir un pla de màrqueting digital que contempli les actuacions a dur a terme per gestionar-la. Es tracta d'un procés mitjançant el qual es realitza un seguiment, s'informa i es reacciona a les opinions dels usuaris sobre l'empresa, el seu personal, productes i serveis. A la xarxa abunden les receptes per crear un Pla de Social Mèdia dedicat a la construcció de la marca. El pas previ és la definició d'una política d'ús de la web social en l'organització, que serveixi de marc per a qualsevol tipus d'actuació.

Basant-nos en les idees aportades per alguns autors com Julián Marquina (2012), Javier Leiva (2012), o empreses com Sysomos (2013), es proposen els següents punts a tenir en compte per a la creació del Pla:

18. Disponible a: <<http://blog.biblioteca.unizar.es/general/el-indice-klout-en-las-bibliotecas-universitarias-espanolas>>.

19. Disponible a: <<http://hdl.handle.net/10421/6623>>.

20. Disponible a: <<http://www.slideshare.net/fesabid/fesabid13-narroyo>>.

21. Disponible a: <<http://blog.hootsuite.com/8-tips-for-social-business-5>>.



1. Anàlisi de la marca i l'entorn. Monitoritzar les converses dels públics a la xarxa, al voltant de la marca, els productes i els serveis. Detectar àrees crítiques, riscos o usuaris insatsfets.
2. Monitoritzar l'activitat dels competidors més significatius, les seves relacions amb els seus grups d'interès.
3. Determinar quin és l'objectiu en relació a la marca, què es vol aconseguir. La diferència entre els valors que volem transmetre i els que arriben ens serveix d'orientació per marcar objectius.
4. Definir la imatge que es vol projectar, el valor i la missió davant els competidors, en què es diferencia de la competència, la seva identitat de marca, com comunica en línia. Sempre ha de ser consistent amb la imatge *offline*.
5. Decidir on es vol estar, en quina plataforma/es. Ha d'estar en concordança amb l'objectiu que s'hagi marcat. A més plataformes, més temps caldrà dedicar-hi, i aquesta és la inversió més gran que es farà: hores/persones. Cal esbrinar on són els usuaris que estan compartint la plataforma i parlant, per construir aquí la imatge de marca.
6. Assignar persones que es dediquen a aquestes tasques, experts en *community manager*, màrqueting digital, social media manager, però involucrant a tot el personal de l'organització en els mitjans socials. Quan s'assignen recursos personals adequats a aquesta tasca, els resultats són sorprenents.
7. Controlar l'atenció que s'està prestant als usuaris. Els mitjans socials són un canal de reclamació per als usuaris, que ben gestionats milloren la reputació de la marca. Definir un pla d'actuació a la xarxa en els espais de conversa.
8. Tenir un pla de crisi previst, que inclogui la reputació digital.
9. Construir relacions gràcies als mèdia: interactuar amb els clients/usuaris, acostar-s'hi, generar simpatia, deixar que els clients participin activament amb l'organització i ajudar a millorar el producte; admetre i corregir els errors.

10. Trobar els influenciadors, els líders d'opinió, dins i fora de l'organització. Detectar i saber actuar enfront dels trolls, en totes les seves variants<sup>22</sup>.
11. Establir mètriques que ens permetin realitzar un seguiment dels objectius que hem establert en els mitjans socials.
12. Monitorització detallada i anàlisi i avaluació de resultats; readaptació de l'estratègia.
13. Elaborar informes de resultats i comunicar, comunicar, comunicar als líders i als grups d'interès.

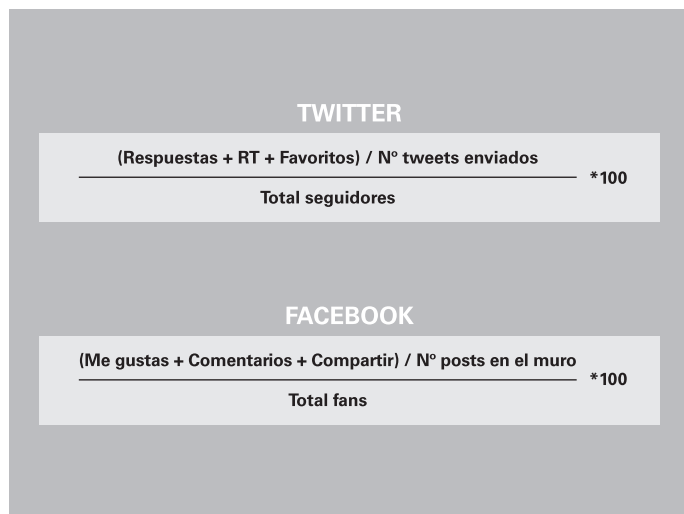
## 5. Mètriques i programes per monitoritzar la influència digital

Per mesurar la reputació i la influència digital s'utilitzen mètriques i els KPI –indicadors clau de rendiment–, variables que estan lligades a un objectiu i permeten monitoritzar-ne l'estat d'avanç o grau de compliment. Les mètriques i els KPI són tant qualitatives com quantitatives i, com ja hem vist anteriorment, estan relacionats amb l'activitat o el tràfic (visitants únics), mida (fans o subscriptors), influència (índexs com Klout), visibilitat (RT, abast), interacció (mencions, comentaris), o retorn (conversions). A més d'aquestes mètriques, se n'estan utilitzant d'altres que en combinen diverses, com la taxa d'*engagement*<sup>23</sup>, que fa referència a la interacció dels usuaris en els espais de la marca en els mitjans socials. Es tracta d'una mètrica que ha desenvolupat l'empresa d'anàlisi en mitjans socials, SocialBakers, i suma la interacció dels usuaris dividida pel nombre d'entrades (post) en aquests mitjans, dividit al seu torn pel nombre de fans, i multiplicat per 100. En el gràfic de la pàgina següent es pot veure aquesta mesura aplicada a Twitter i Facebook (Castelló, 2013).

Amb aquests indicadors podem calcular el retorn de la inversió, el ROI, un dels més grans objectius del màrque-

22. Dolores Vela. *Reputación de marca: Cómo gestionar Trolls* [En línia]. <<http://www.slideshare.net/doloresvelagarcia/100-sm>>.

23. Més informació a: <<http://www.socialbakers.com/blog/1427-engagement-rate-a-metric-you-can-count-on>>.



**Figura 2.** Taxa d'engagement, desenvolupada per SocialBakers

ting i l'estratègia de comunicació en xarxes socials, i que es tradueix en benefici menys inversió, dividit per la inversió. El ROI pot significar també retorn en influència, i amb l'aparició dels mitjans socials es parla de l'IOR (Impact On Relationship), com el càlcul de l'impacte de les relacions entre les marques i els seus seguidors (Castelló, 2013). Aquestes relacions poden mesurar-se basant-se en quatre variables:

1. L'**autoritat** del contingut de la marca (mencions de la marca a les xarxes socials o enllaços al contingut de la marca).
2. La **influència** de la marca en els mitjans socials (segons el nombre de seguidors de la marca en els mitjans socials i la seva evolució).
3. La **participació** i la interacció dels seguidors en els espais socials de la marca.
4. **Variables del trànsit** que genera el contingut de la marca des dels mitjans socials cap a la web de l'organització (es mesurarà l'evolució del trànsit que arriba

a la seu web des dels mitjans socials, percentatge de creixement, usuaris únics, mitjana de temps que s'hi passen o índex de rebot).

Per obtenir l'IOR és fonamental que l'empresa monitoritzi aquests valors en temps real, escoltant de manera activa, participant en les converses, oferint actualitzacions freqüents amb continguts rellevants i de qualitat, i responent de manera ràpida, demostrant transparència i confiança.

Tal i com indica Miguel del Fresno (2011), «la Reputación Online es posible identificarla, extraerla, clasificarla y analizarla, a partir de las opiniones que los usuarios diseminan en la web», pel que és necessari disposar de les eines adequades per localitzar aquests comentaris. Per quantificar i monitoritzar la reputació digital, s'estan utilitzant diferents eines com *BuzzMetrics*, *Swotti*, *Keotag*, *Addict-o-matic*, lectors de sindicació de continguts (RSS), el cercador *Technorati*, alertes amb la finalitat de detectar possibles crisis, com *Google Alerts* (avui dia una de les eines més comunes utilitzades per controlar la imatge d'una persona o empresa en línia)<sup>24</sup>. Un cop s'ha quantificat la reputació digital, podrà gestionar-se per millorar-la.

Alguns indicadors de reputació i influència són els que treballen amb algorismes que quantifiquen més d'una xarxa social, categoritzen el tipus de lideratge de l'usuari i determinen quina és la seva àrea d'influència. Aquest és el cas de les eines gratuïtes d'escolta i monitorització, *Klout*, *PeerIndex*, *Somes* i *SocialMention* (Serrano-Puche, 2012).

*Klout* mesura de 0 a 100 la influència personal de l'usuari en funció de la seva activitat fins en dotze xarxes socials. A la Universitat de l'Estat de Florida<sup>25</sup> utilitzen Klout perquè els alumnes aprenguin a guanyar influència. Aquest índex ofereix ara la possibilitat que alguns dels seus usuaris es converteixin en experts<sup>26</sup> i responguin a les preguntes de la seva àrea d'experiència, d'altres usuaris de l'índex.

24. Es pot consultar un llistat complet d'eines per monitoritzar en els mitjans digitals en el següent mapa de la web *Mindomo*: <<http://tinyurl.com/HerramientasSMM>>.

25. Més informació a: <<http://www.businessesgrow.com/2012/08/26/florida-state-university-class-using-klout-to-determine-student-grades>>.

26. Més informació a: <<http://klout.com/experts>>.

*PeerIndex* es defineix com una eina per a la mesura i compressió del «capital social» que una persona ha aconseguit en el mitjà *online*. És un concepte que s'assenta més en la noció d'autoritat i reputació, i menys en la d'influència (Serrano-Puche, 2012).

Un altre índex, *Somes*, mesura les incidències d'una marca en el cercador *Google* i en els principals mitjans socials. D'aquesta manera, construeix quatre índexs que es presenten en els resultats en temps real: *impacte* (capacitat de la marca per ser trobada ràpidament per un usuari a través del cercador *Google*); *presència* (accions de la marca en els mitjans socials); *conversa* (nivell de les converses sobre les marques que els usuaris generen i notícies aparegudes sobre la marca); i *índex Somes* (agregació de totes les dades obtingudes en els mitjans socials analitzats) (Ros-Diego, 2011).

L'eina *SocialMention* permet fer un seguiment d'un terme de cerca, una marca, en diferents espais digitals. Els resultats es presenten amb quatre indicadors: *strength*, que es refereix a la probabilitat que la selecció sigui centre de converses en mitjans socials; *sentiment*, que és la proporció de les mencions negatives respecte a les positives; *passion*, al·ludeix a les repeticions de la marca pels mateixos grups en les plataformes socials; i *reach*, una mesura del rang d'influència (Ros-Diego, 2011).

Mesurar la influència amb un algoritme és útil però no és definitiu. Es tracta de plataformes que són millorables i que no processen tots els usuaris, com és el cas de *Klout* o *PeerIndex*, ni tampoc el conjunt de totes les plataformes socials. A vegades fins i tot hi ha altres interessos lucratiu, derivats de la naturalesa de *Klout* i *PeerIndex* com a iniciatives comercials privades. El que resulta més útil d'aquestes mètriques és la seva evolució al llarg del temps, ja que poden detectar i analitzar aquelles accions que han fet que l'IOR s'incrementi o baixi, i marcar objectius periòdicament. També resulten d'interès per comparar els valors que ofereixen els índexs d'organitzacions o biblioteques de similars característiques.

Alguns indicadors de reputació i influència són els que treballen amb algoritmes que quantifiquen més d'una xarxa social, categoritzen el tipus de lideratge de l'usuari i determinen quina és la seva àrea d'influència.

## 6. Conclusions

Ens trobem en un moment de pressió econòmica i social en què cal demostrar la capacitat de les biblioteques per adaptar-se i reinventar-se o almenys adequar la seva identitat actual a les noves exigències de la societat. És fonamental, per tant, que aquest tipus d'organització sigui conscient de la necessitat d'adaptar les seves estratègies de comunicació interna i externa als nous requisits socials dels seus usuaris, de forma professional, per aprofitar al màxim les seves potencialitats i encaixar naturalment dins l'experiència d'un usuari que ja és digital.

La construcció de la pròpia identitat digital obliga a definir què, com i on es vol comunicar a Internet, i conèixer i utilitzar les diverses eines disponibles, construint aquesta presència poc a poc. Cal relacionar-se amb tercers que també són presents a Internet i comptar amb l'opinió que els altres generaran sobre la biblioteca i els bibliotecaris, que no controlem. «Podemos ayudar a encauzarla, podemos interactuar con ella, podemos incluso contestarla cuando sea claramente falsa. Lo que ya no podemos hacer es ignorarla ni impedir que se publique o que tenga impacto directo en nuestra reputación» (Alonso, 2011). És impres-



©istockphoto/fredpal

cindible, a més, tenir cura de la comunicació interna com a clau de reputació entre els empleats<sup>27</sup>, o de controlar les crisis reputacionals.

«La reputación corporativa y la marca serán inseparables» i «la reputación se redefinirá en base a los cambios sociales y políticos»<sup>28</sup>. Els seus líders són i seran els màxims guardians de la reputació corporativa, però aquesta dependrà dels grups d'interès –empleats, proveïdors, usuaris– i de la forma de gestionar-la.

## Bibliografía

ALONSO, J. «Identidad y Reputación digital». *Identidad digital y Reputación online. Cuadernos de comunicación Evoca*, 5, 2011.

ALLOZA LOSANA, Á. «Reflexiones sobre la reputación necesaria». *adComunica. Revista Científica de Estrategias, Ten-*

*dencias e Innovación en Comunicación*, 3, 2012, p. 27-47. DOI: <http://dx.doi.org/10.6035/2174-0992.2012.3.3> <<http://repositori.uji.es/xmlui/bitstream/handle/10234/42985/%C3%81ngel%20Alloza%20Losana.pdf?sequence=1>>.

CASTELLÓ MARTÍNEZ, A. *Estrategias empresariales en la Web 2.0. Las redes sociales online*. Alicante: Editorial ECU, 2010.

CASTELLÓ MARTÍNEZ, A. «El estudio del retorno de la inversión y el impacto en la relación de la comunicación empresarial y publicitaria en plataformas sociales: herramientas disponibles en el mercado». En: *Actas del 2º Congreso Nacional sobre Metodología de la Investigación en Comunicación*, Segovia, 2013. ISBN: 978-84-616-4124-6

DEL FRESNO, M. «Cómo investigar la reputación online en los medios sociales de la web 2.0». *Identidad digital y Reputación online. Cuadernos de comunicación Evoca*, 5, 2011.

27. Vegeu: <<http://blog.zyncro.com/2013/02/28/7-tendencias-en-reputacion-corporativa-para-2013>>.

28. Vegeu: <<http://blog.zyncro.com/2013/02/28/7-tendencias-en-reputacion-corporativa-para-2013>>.

GONZÁLEZ-FERNÁNDEZ-VILLAVICENCIO, N.; SEOANE GARCÍA, C.; SAN MILLÁN FERNÁNDEZ, E.; MENÉNDEZ, J. L. «Revisión y propuesta de indicadores (KPI) de la Biblioteca en los medios sociales». *Revista Española de Documentación Científica*, 36(1), 2013, p. 1-14. DOI: 10.3989/redc.2013.1.919.

Instituto de Análisis de los Intangibles. *La reputación de España en el Mundo*. Publicación propia (2008-2010).

LEIVA AGUILERA, J. *Gestión de la reputación online: crea fácilmente tu estrategia de presencia en la Red*. Barcelona, Editorial UOC, 2012.

MARQUINA-ARENAS, J. *Plan Social Media y Community Manager*. Barcelona, Editorial UOC, 2012.

MUT CAMACHO, M. «La desvirtualización del concepto reputación en el entorno virtual». *adComunica. Revista Científica de Estrategias, Tendencias e Innovación en Comunicación*, 3, 2012, p. 227-229. DOI: <<http://dx.doi.org/10.6035/2174-0992.2012.3.18>>.

ROS-DIEGO, V. J.; CASTELLÓ MARTÍNEZ, A. «La comunicación de la responsabilidad en los medios sociales». *Revista Latina de Comunicación Social*, 67, 2011, p. 47-67. DOI: 10.4185/RLCS-067-947-047-067.

SÁNCHEZ, E. «El social media en la estrategia de comunicación». *adComunica: revista científica de estrategias, tendencias e innovación en comunicación*, 3, 2012, p. 221-222. DOI: <<http://dx.doi.org/10.6035/2174-0992.2012.3.16>>.

SERRANO-PUCHE, J. «Herramientas web para la medición de la influencia digital: análisis de Klout y PeerIndex». *El Profesional de la Información*, 21(3), 2012, p. 298-303. DOI:10.3145/epi.2012.may.11.

Sysomos. *8 ways to build your brand using social media* (2013).

VALOR MARTÍNEZ, C.; DE LA CUESTA GONZÁLEZ, M. «Responsabilidad Social de la Empresa: Concepto, Medición y Desarrollo en España». *Boletín Económico ICE*, N° 2755, 2003, p. 7-20.

VAQUERO COLLADO, A. «La reputación online en el marco de la comunicación corporativa. Una visión sobre la investigación de tendencias y perspectivas profesionales». *adComunica: revista científica de estrategias, tendencias e innovación en comunicación*, 3, 2012, p. 49-63. DOI: 10.6035/2174-0992.2012.3.4.

# La gestió de la reputació digital a Biblioteques de Barcelona

## **Anna Bröll Nadal**

Directora tècnica de Coordinació i Serveis de Biblioteques de Barcelona  
*abroll@bcn.cat*

## **Anna Cabré Serra**

Responsable TIC i *community manager* de Biblioteques de Barcelona  
*acabres@bcn.cat*

## **Dolors Gándara Sanz**

Tècnic TIC i *community manager* de Biblioteques de Barcelona  
*dgandaras@bcn.cat*

### **Resum:**

Des de Biblioteques de Barcelona pensem que la gestió de la imatge i la reputació digital són tan importants com en l'entorn real tangible. La progressiva implantació de canals web 2.0 i de participació a les xarxes respon a l'aspiració de ser referents en entorns digitals, la qual cosa implica una adequada gestió de la identitat digital i la reputació de la marca. La marca Biblioteques de Barcelona té una bona reputació en l'entorn real tangible, només cal veure com any rere any som el servei municipal millor valorat pels ciutadans. Aquest prestigi ens ajuda a tenir una bona reputació a priori en l'entorn digital, i alhora ens obliga a mantenir aquesta reputació també en aquest entorn. Per això es tan important com gestionem el nostres canals digitals.

L'ús d'un canal de comunicació digital o un altre respon a dues variables: les característiques de cadascun i l'avaluació de l'ús i l'impacte. Podem diferenciar entre aquells canals que són contenidors de la informació original d'aquells que la distribueixen, com dels canals d'interacció amb l'usuari dels de comunicació i màrqueting.

El futur d'aquests canals dependrà de l'evolució tècnica de cada un dels canals i de les xarxes socials, així com de l'impacte que tenen en la societat i les estadístiques d'ús. L'anàlisi de tots aquests factors ens ajudarà a reflexionar, replantejar i millorar la gestió de la reputació digital de Biblioteques de Barcelona.

Estem atents a l'evolució del món digital i dels dispositius d'accés a la informació, adaptant-nos a les necessitats d'informació dels nostres usuaris i oferint-los nous serveis a través de nous canals.

### **Paraules clau:**

biblioteca pública, comunicació digital, identitat digital, marca, màrqueting digital, reputació digital, participació en línia, usuaris, web 2.0, xarxes socials

## La gestión de la reputación digital en Bibliotecas de Barcelona

### Resumen:

Desde Bibliotecas de Barcelona se cree que la gestión de la imagen y la reputación digital son tan importantes como en el entorno real tangible. La progresiva implantación de canales web 2.0 y de participación en las redes responde a la aspiración de ser referentes en entornos digitales, lo que implica una adecuada gestión de la identidad digital y la reputación de la marca.

La marca Bibliotecas de Barcelona tiene una buena reputación en el entorno real tangible, solo hay que ver cómo año tras año es el servicio municipal mejor valorado por los ciudadanos. Este prestigio ayuda a tener una buena reputación a priori en el entorno digital, y, al mismo tiempo, obliga a mantener esta reputación también en este entorno. Por eso es tan importante cómo se gestionan sus canales digitales.

El uso de un canal de comunicación digital u otro responde a dos variables: las características de cada uno y la evaluación del uso y el impacto. Se puede diferenciar entre los canales que son contenedores de la información original de los que la distribuyen, como de los canales de interacción con el usuario de los de comunicación y *marketing*.

El futuro de estos canales dependerá de la evolución técnica de cada uno de ellos y de las redes sociales, así como del impacto que tienen en la sociedad y las estadísticas de uso. El análisis de todos estos factores ayudará a reflexionar, replantear y mejorar la gestión de la reputación digital de Bibliotecas de Barcelona.

Asimismo, se atiende a la evolución del mundo digital y de los dispositivos de acceso a la información, adaptando las necesidades de información a los usuarios y ofreciéndoles nuevos servicios a través de nuevos canales.

### Palabras clave:

biblioteca pública, comunicación digital, identidad digital, marca, *marketing* digital, reputación digital, participación en línea, usuarios, web 2.0, redes sociales

## Digital reputation management in Biblioteques de Barcelona

### **Abstract:**

We at Biblioteques de Barcelona believe image and digital reputation management are as important as in the real, tangible environment. The steady introduction of web 2.0 channels and network participation responds to aspirations to become benchmarks in digital environments, which requires a suitable management of digital identities and brand reputations.

The Biblioteques de Barcelona brand has a good reputation in the real, tangible environment: you need only see how citizens have been rating it the best municipal service year after year. Such prestige helps us to enjoy a good a priori reputation in the digital environment and it also obliges us to maintain this reputation in this environment. That is why it is so important for us to manage our digital channels.

The use of one or other digital communications channel responds to two variables: the features of each and their use and impact assessment. We can distinguish between channels that are containers of original information and others that distribute it, as well as channels for interaction with users and communications and marketing channels.

The future of these channels will depend on the technical developments of each of the social media channels and networks, as well as the impact they have on society and user statistics. Analysing all these factors will help us to reflect, reconsider and improve Biblioteques de Barcelona's digital-reputation management.

We are constantly monitoring the development of the digital world and the devices for gaining access to information, adapting ourselves to the information needs of our users and offering them our services through new channels.

### **Keywords:**

public library, digital communication, digital identity, brand, digital marketing, digital reputation, online participation, users, web 2.0, social media networks





©bauenblog.info

## 1. Introducció

La reputació d'una organització, d'un individu o d'una marca es construeix a partir de l'opinió que tenen els altres sobre nosaltres<sup>1</sup>. En el cas de Biblioteques de Barcelona aquesta opinió ve determinada en gran part per la percepció que tenen els usuaris dels equipaments, de les col·leccions, dels serveis i del tracte personal rebut a les biblioteques de la xarxa de la ciutat. Però també els no usuaris tenen una opinió formada sobre nosaltres, basada en la informació i l'opinió que els arriba d'altres usuaris, i aquestes opinions també incideixen en la nostra reputació.

Biblioteques de Barcelona té una bona reputació en l'entorn real tangible, només cal veure com any rere any som el servei municipal millor valorat pels ciutadans. Aquest prestigi ens ajuda a tenir una bona reputació a priori en l'entorn digital, i alhora ens obliga a mantenir aquesta reputació també en aquest entorn.

Amb l'aparició de la web social, més participativa i més democràtica, l'apoderament dels usuaris i la seva capacitat de crear opinió són realment molt grans. Les organitzacions hem d'estar presents a les xarxes socials, crear la nostra identitat digital, comunicar els nostres valors i interactuar amb els usuaris. Si no ho fem, la nostra reputació *online* vindrà determinada exclusivament per les opinions dels altres<sup>2</sup>.

Des de la decisió de disposar de pàgina web des dels inicis de la creació de la xarxa urbana, de disposar d'un blog de recomanacions que busca la complicitat dels usuaris, *Bibarnabloc.cat*, fins la progressiva participació a les diferents xarxes socials, passant per l'edició del butlletí electrònic *Més Biblioteques* o el canal IRIS de suggeriments, consultes i reclamacions, tot plegat respon a una estratègia corporativa de crear una identitat digital i influir en la reputació de la marca Biblioteques de Barcelona.

1. LEIVA AGUILERA, Javier. *Gestión de la reputación online*. Barcelona: UOC, 2012.

2. ALONSO, Julio. «Identidad y reputación digital». *Cuadernos de comunicación Evoca*, 5, 2011, p. 5-9.

Imatge corporativa i identitat digital han d'anar de la mà, per això sempre utilitzem el logo de Biblioteques de Barcelona en tots els canals. És molt important que l'usuari ens reconegui i identifiqui en qualsevol mitjà digital i en totes les xarxes socials.

Cal anar construint la nostra presència posant informació a la xarxa, difonent-la i relacionant-nos amb els usuaris. No podem evitar els comentaris que els altres facin sobre nosaltres, però sí que podem mantenir un major coneixement sobre la nostra reputació i actuar des dels nostres canals d'informació, de comunicació i d'interacció amb els usuaris.

## 2. Informació, comunicació i interacció amb la comunitat

Per construir la nostra presència a la xarxa disposem de diferents canals digitals que hem classificat en tres categories: canals d'informació, canals de distribució i comunicació i canals d'interacció. Un mateix canal digital pot pertànyer a una o més categories, però els objectius varien depenent de la finalitat de cada mitjà.

### 2.1. Canals digitals d'informació

Són bàsicament els que contenen informació sobre la pròpia organització, les col·leccions, els serveis i les activitats.

És important la qualitat i la fiabilitat dels continguts que com a Biblioteques de Barcelona movem per les xarxes; per aquest motiu, la major part de la informació que distribuïm a través de les xarxes socials prové bàsicament dels nostres propis canals digitals d'informació, així com de fonts fiables com l'Ajuntament de Barcelona, la Diputa-

ció de Barcelona, mitjans de comunicació o de referents en l'entorn digital.

### El web de Biblioteques de Barcelona<sup>3</sup>

El web és el principal canal d'informació digital de Biblioteques de Barcelona. Des d'aquí es generen els continguts tant de serveis i col·leccions com d'agendes i notícies. També presentem continguts procedents del *Bibarnabloc*, del nostre canal de Youtube o els darrers tuits publicats.

Cada una de les 38 biblioteques que formen part de la xarxa disposa del seu propi web però formant part d'un únic *site*, per tal que aquelles informacions generals, com els continguts relatius als serveis de préstec, els serveis d'Internet, els tràmits en línia, les normatives i tarifes, horaris o qualsevol altre servei comú a totes les biblioteques siguin continguts únics per a tots els *minisites*. Així, les biblioteques es poden centrar en els continguts propis, com les agendes i notícies, o les especialitzacions, centres d'interès i projectes propis de cada centre.

D'aquesta manera, la identitat digital a nivell de xarxa es gestiona de forma centralitzada i cada centre gestiona la seva identitat particular, evitant així la duplicat i disparitat en els continguts comuns i propiciant la creació i el manteniment de continguts específics.

Tota la informació que posem al web és la que després distribuïm a través de les xarxes socials com ara Facebook i Twitter.

### *Bibarnabloc.cat*, el blog de les Biblioteques de Barcelona<sup>4</sup>

El *Bibarnabloc* és l'espai de participació i col·laboració entre els treballadors de Biblioteques de Barcelona i els usu-

3. Biblioteques de Barcelona ([www.bcn.cat/biblioteques](http://www.bcn.cat/biblioteques)).

4. *Bibarnabloc* ([bibarnabloc.cat](http://bibarnabloc.cat)).

El *Bibarnabloc* és l'espai de participació i col·laboració entre els treballadors de Biblioteques de Barcelona i els usuaris de les biblioteques.



aris de les biblioteques. Es va crear a mitjans de l'any 2011 amb l'objectiu de donar a conèixer el fons, sense oblidar que es tracta de l'element aglutinador de la nostra presència a la xarxa, donat que integra els ginys de les diferents xarxes socials on participem.

Al blog es realitzen un total de sis entrades setmanals, de dilluns a dissabte, sobre qualsevol temàtica de qualsevol document del fons de les biblioteques, així com de guies de lectura i activitats destacades, sempre associades, així sí, a documents disponibles a les biblioteques.

Un dels continguts més visitats són les videorecomanacions<sup>5</sup>, on un col·laborador recomana un document a través d'un videomuntatge allotjat al nostre canal de Youtube. Un cop l'any, coincidint amb el període previ a l'estiu, proposem que els col·laboradors realitzin una fotorecomanació, i el conjunt de totes les rebudes acaba essent, també, un vídeo, molt més desenfadat que les videorecomanacions.

Mensualment s'organitza un concurs per tal de fomentar la participació en el blog, ja que és imprescindible respondre a la pregunta i fer un comentari a un post del blog durant el mes en curs. A través de la Direcció de Programes i Cooperació, responsable de les activitats que s'organitzen a les biblioteques, se sol·licita a alguns autors que participen a xerrades un exemplar signat de la seva darrera obra.

Els concursos que organitzem al nostre blog i al FanPage de Biblioteques de Barcelona són una eina de màrqueting digital per tal d'augmentar les visites i les interaccions amb les nostres pàgines.

## Youtube<sup>6</sup>

El canal de Biblioteques a Youtube es va iniciar l'abril de l'any 2011, amb la intenció de recopilar tant el material que s'edita des de les pròpies biblioteques com des de diferents mitjans de comunicació o entitats.

5. Videorecomanacions de Biblioteques de Barcelona ([bibarnabloc.cat/tag/videorecomanacions](http://bibarnabloc.cat/tag/videorecomanacions)).

6. Canal de Youtube de Biblioteques de Barcelona ([www.youtube.com/bibliotequesbcn](http://www.youtube.com/bibliotequesbcn)).

El canal té diverses llistes de reproducció, depenent de la temàtica del vídeo pujat o de la font d'on ha compartit: tallers de creacions, vídeos cedits per BTV, formació AC-TIC, *Bibarnabloc*, serveis de les biblioteques, entrevistes amb autors, Sant Jordi, les biblioteques de Barcelona als mitjans, cursos Cibernàrium i presentacions de les biblioteques.

Col·laborem amb altres institucions com el Departament de Comunicació Web i Xarxes Socials de l'Ajuntament de Barcelona i Barcelona Activa per editar vídeos conjuntament per tal de compartir recursos i fer arribar la nostra marca a altres xarxes de la nostra ciutat.

L'estratègia digital de Biblioteques de Barcelona en aquest canal és la de publicar continguts audiovisuals per tal d'adaptar-nos al nostre *target* més jove, més acostumat a la cultura audiovisual que a l'escripta. Així, per exemple, s'elaboren videopresentacions de cadascuna de les 38 biblioteques per tal de fer publicitat dels nostres fons i serveis d'una manera més atractiva.

Amb l'objectiu de viralitzar els vídeos propis, els incrustem als webs i blogs, i els compartim a les xarxes socials. També en fem difusió a les pantalles informatives que hi ha a les biblioteques.

## Issuu<sup>7</sup>

Issuu és un servei en línia que permet la visualització en el navegador de material digitalitzat, amb el mateix aspecte que una publicació impresa.

Cada biblioteca disposa d'un compte Issuu per a poder-hi publicar les guies de lectura, les guies de novetats, els butlletins d'activitats i qualsevol altre publicació que vulgui difondre.

7. Issuu de Biblioteques de Barcelona ([issuu.com/bibliotequesbcn](http://issuu.com/bibliotequesbcn)).

8. *Més Biblioteques*, butlletí electrònic de Biblioteques de Barcelona ([www.mesbiblioteques.cat](http://www.mesbiblioteques.cat)).

9. A Facebook tenim 9.923 fans i a Twitter 8.469 seguidors (juny 2013).

Aquests continguts s'enllacen a partir de les diferents pàgines web i es difonen a través de les xarxes socials i del blog.

## 2.2. Canals digitals de distribució i comunicació de la informació

Són els canals que utilitzem per difondre els continguts i les activitats de Biblioteques de Barcelona i altres informacions d'interès cultural, vinculats a la ciutat de Barcelona o a l'entorn de biblioteques.

El primer mitjà digital que vam utilitzar va ser el butlletí electrònic *Més Biblioteques*, l'any 2007. Més tard, vam consolidar la nostra presència a les xarxes de Facebook i Twitter, i aquest any hem estrenat la aplicació mòbil *BibliosBCN*.

### Butlletí electrònic *Més Biblioteques*<sup>8</sup>

Tot i ser el mitjà amb més antiguitat és, encara ara, un del més efectius. Amb més de 264.000 subscripcions és el que arriba a més usuaris<sup>9</sup>, i l'efecte que té en crides a participació de concursos o difusió de vídeos és realment molt important.

La informació que es fa arribar és bàsicament sobre activitats i fons, en el cas de cada biblioteca en particular, i avantatges i descomptes del carnet, activitats destacades i els concursos del *Bibarnabloc*, en el cas de Biblioteques de Barcelona.

Tots els butlletins enviats de manera centralitzada els reben tots els usuaris; en canvi, els generats per cada biblioteca s'envien només als usuaris que tenen associats.

Es recomana a les biblioteques que no facin més d'un enviament setmanal i que agrupin informacions, ja que el sistema aglutina fins a cinc notícies en un mateix butlletí. També es recomana que s'envii una setmana abans de l'activitat per donar temps als usuaris a poder-la llegir.

## Facebook<sup>10</sup>

Facebook va ser la primera xarxa social que es va utilitzar a Biblioteques de Barcelona. Va ser creada l'any 2009, i actualment 38 de les 39 biblioteques gestionen la seva pàgina.

Des d'aquestes pàgines es difonen activitats i notícies destacades, les col·leccions (guies de lectura, posts destacats del *Bibarnabloc*, novetats de música...) i serveis destacats (descomptes, inscripcions als cursos TIC...), es difonen els concursos mensuals del *Bibarnabloc*, i s'avisa de les incidències puntuals a les biblioteques i les noves inauguracions.

També s'intenta fer difusió d'esdeveniments més globals, tant a nivell local com dels programes culturals de la ciutat de Barcelona, o de cultura en general (dies internacionals, efemèrides i defuncions...), sempre intentant destacar els fons de les biblioteques amb cerques al catàleg.

Al Facebook s'organitzen, també, concursos puntuals de fidelització amb regals de llibres cedits per editorials, inscripcions gratuïtes a itineraris literaris o tallers de creació.

Es programen entre dos i quatre posts diaris, i a vegades un els caps de setmana.

## Twitter<sup>11</sup>

L'octubre de 2010 es creà el compte de Twitter de Biblioteques de Barcelona. Una de les finalitats és la de difondre

aquella informació rellevant per a l'usuari de Biblioteques de Barcelona, com l'agenda de biblioteques, informació cultural de la ciutat, notícies sobre serveis (descomptes, horaris...), notícies d'actualitat, i altres informacions rellevants, d'una manera més exhaustiva que a Facebook.

Intentem millorar la marca sumant-nos a iniciatives culturals a través de *hashtags*, *trending topics*, sempre amb l'objectiu que siguin missatges amb prou interès per tal de fidelitzar els nostres seguidors i augmentar la nostra influència, com el #DiaBolaño, promogut pel CCCB, la iniciativa de l'@anyespriu pel centenari de Salvador Espriu o el @tricentenari (#tricentenari), per seguir els esdeveniments de la Commemoració del Tricentenari del 1714.

Es programen una mitjana de sis a set tuits de dilluns a divendres i dos els caps de setmana, depenent de les activitats culturals que s'organitzin a les biblioteques. Actualment s'ha començat amb una tasca de difusió de descomptes puntuals que els usuaris poden gaudir amb el carnet de biblioteques.

Així com a Facebook es va creure oportú que cada biblioteca gestionés un compte propi, a Twitter només tenim identitat corporativa de Biblioteques de Barcelona.

## App *BibliosBCN*

A través de l'aplicació mòbil *BibliosBCN*, que es va posar en funcionament l'abril d'aquest any 2013, els usuaris poden consultar al mòbil les activitats que es duen a terme a les biblioteques de Barcelona. L'aplicació permet tenir accés a les activitats destacades de les biblioteques i cada usuari pot configurar la seva selecció personal. Amb l'opció de calendari poden consultar totes les activitats que es fan un dia concret. I amb l'opció de mapa poden saber quina biblioteca tenen més a prop, com s'hi arriba, els horaris i la informació bàsica.

10. Facebook de Biblioteques de Barcelona ([www.facebook.com/bibliotequesbcn](http://www.facebook.com/bibliotequesbcn)).

11. Twitter de Biblioteques de Barcelona ([www.twitter.com/bibliotequesbcn](http://www.twitter.com/bibliotequesbcn)).



©istockphoto/nyul

Tota aquesta informació la poden compartir per SMS, correu electrònic i xarxes socials. L'aplicació és gratuïta i està disponible en iOS<sup>12</sup> i Android<sup>13</sup>, i en el primer mes de funcionament ja l'han descarregada prop de 2.000 usuaris.

### 2.3. Canals digitals d'interacció amb l'usuari: formals i informals

Biblioteques de Barcelona disposa de diferents canals digitals on interacciona amb els usuaris. La interacció és de caràcter formal quan és necessari l'emplenament d'un formulari, i queda un registre en una base de dades de la consulta; i, de forma informal, quan la interacció es realitza a través d'una xarxa social o blog.

Tot el que es recull des d'aquests canals s'analitza i se'n fa seguiment, introduint millores o canvis en els serveis quan els temes exposats pels usuaris així ho requereixin, i sempre que sigui viable, i en la mesura de les nostres possibilitats.

Biblioteques de Barcelona disposa de diferents canals digitals on interacciona amb els usuaris.

En qualsevol cas, no es deixa cap tema sense resposta, malgrat que no sempre es puguin satisfer les necessitats plantejades.

#### Canal formal

##### – Iris

Es tracta d'un servei de l'Ajuntament de Barcelona per atendre les consultes, queixes, suggeriments i agraïments sobre els serveis municipals, a través d'un formulari web. Totes les biblioteques de la xarxa gestionen el seu canal específic a l'Iris, encarregant-se directament de respondre als usuaris sobre els temes adreçats a la seva biblioteca en particular. Des del Departament de Comunicació i Premsa de Biblioteques de Barcelona es realitza la gestió de les respostes que afecten al funcionament general de Biblioteques de Barcelona; a més, se supervisen les respostes elabora-

12. App d'iOS per a Biblioteques de Barcelona ([itunes.apple.com/us/app/bibliosbcn/id638704165?l=ca&ls=1&mt=8](https://itunes.apple.com/us/app/bibliosbcn/id638704165?l=ca&ls=1&mt=8)).

13. App Android per a Biblioteques de Barcelona ([play.google.com/store/apps/details?id=cat.bcn.bibliosbcn&hl=ca](https://play.google.com/store/apps/details?id=cat.bcn.bibliosbcn&hl=ca)).

des per les biblioteques amb l'objectiu de donar coherència de xarxa al tipus de respostes que reben els usuaris, amb independència de la biblioteca que la redacti.

Moltes vegades trobem als canals d'interacció informals queixes i suggeriments que requereixen explicacions més llargues, que generen debat, o missatges de trols. En aquests casos es convida a l'usuari a utilitzar l'Iris per tal de donar un missatge homogeni, fer-ne seguiment i per introduir els canvis i millores en els casos en què el suggeriment és compartit per més usuaris.

## Consultes informals

### – Twitter

El compte de Twitter té com un dels seus objectius posicionar la marca dins d'aquesta xarxa de *microblogging*, millorar-ne la reputació, fer seguiment i monitoritzar els missatges que es fan des d'altres comptes sobre les nostres biblioteques. Es tracta d'una escolta proactiva dels nostres seguidors, establint un *feedback* per rebre els seus comentaris, suggeriments, dubtes i queixes amb l'objectiu de compartir coneixement, fugint de tuits que siguin només titulars.

No defugim tampoc donar resposta a les mencions i missatges directes que incloquin queixes i suggeriments. Intentem donar una resposta clara i ràpida, valorant si és necessari adreçar-los als nostres canals de consultes formals.

### – Facebook

A la *fanpage* de Biblioteques de Barcelona es resolen dubtes i suggeriments sobre els serveis i fons de les biblioteques amb un temps de resposta no superior a 24 hores. També es fa un seguiment de tots els missatges i comentaris que compleixin uns criteris de res-

pecte, tolerància i privacitat, no exercint mai censura, sempre i quan compleixin les condicions d'ús i participació de les xarxes socials<sup>14</sup>.

### – Participació al *Bibarnabloc*

La participació dels usuaris es realitza en forma de comentaris als posts. Sempre que els comentaris necessitin aclariments és responsabilitat de l'autor del post mantenir la conversa amb l'usuari i resoldre els seus dubtes i/o afegir informació de valor complementari per incentivar el préstec dels documents recomanats.

## 3. Dades per a l'anàlisi

Durant l'any 2012 vam participar en un estudi comparatiu entre diverses xarxes de biblioteques que està realitzant Nieves González Fernández-Villavicencio, de l'Àrea de Biblioteconomia y Documentación de la Universidad Pablo de Olavide, amb motiu de la seva tesi doctoral. Es tracta de mesurar la nostra activitat en mitjans socials, el grau de compromís que estem obtenint dels usuaris, el tràfic cap a la nostra pàgina web procedent de mitjans socials i el nivell d'ús dels serveis bibliotecaris o retorn de la inversió.

L'estudi forma part d'un projecte d'investigació sobre la rendibilitat de la biblioteca en la web social, i tot i que encara no ha finalitzat ens ha passat algunes dades que fan referència a la reputació de la marca i que reproduïm a continuació.

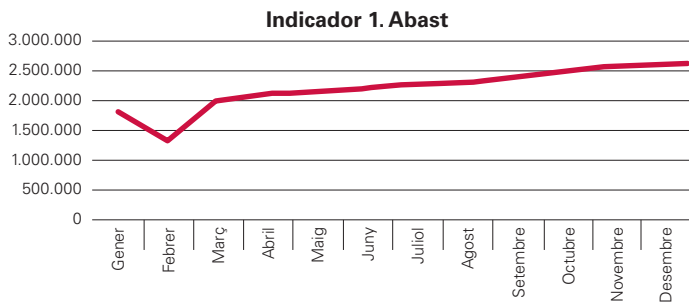
### 3.1 Anàlisi de la reputació de la marca

Impacte de les relacions entre la marca i els seus seguidors en base a quatre variables: **abast**, **influència**, **rellevància** i **fidelització**.

14. Normes d'ús i participació de Biblioteques de Barcelona a Facebook ([www.facebook.com/bibliotecalesRoquetes/app\\_154898304574163](http://www.facebook.com/bibliotecalesRoquetes/app_154898304574163)).

### 3.1.1. Abast

Abast de la marca en els mitjans socials en base al nombre de seguidors de la marca en els seus perfils socials i a la seva evolució.



**Taula 1.** Abast

Amb aquest indicador general d'abast podem veure la suma dels següents indicadors parcials:

*Seguidors a Twitter, Fans a Facebook, Amics dels Fans a Facebook, Suma de l'abast total de cada setmana a Facebook, Subscriptors, Seguidors o Contactes (per cada canal en mitjans socials).*

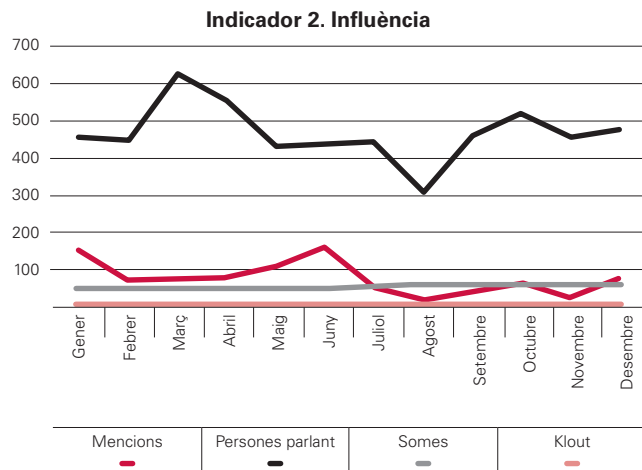
La suma d'aquests ítems mostra una evolució ascendent que confirma una bona i constant presència en els mitjans socials i una audiència que no deixa de créixer i a la qual li arriba el missatge de Biblioteques de Barcelona en aquests mitjans.

El creixement continu significa que els continguts que la biblioteca està aportant a la xarxa són d'interès per a la seva comunitat d'usuaris.

El nombre de seguidors a Twitter ha anat augmentant de forma progressiva durant l'any 2012, començant amb 3.157 seguidors el mes de gener, fins arribar fins al desembre a una xifra de 6.091, que l'autora de l'estudi considera molt elevada. Actualment tenim 8.418 seguidors (19 de juny de 2013).

### 3.1.2 Influència

La influència dels continguts, l'autoritat del contingut de la marca (mencions de la marca a la web 2.0, enllaços als continguts, etc.).



**Taula 2.** Influència

Amb aquest indicador general d'influència podem veure la suma dels següents indicadors parcials: mencions a les xarxes socials, persones que parlen d'això a Facebook, i índexs Klout i Somes. L'índex Klout quantifica la influència social a Internet i ha esdevingut una de les maneres més esteses per mesurar l'impacte a les xarxes socials.

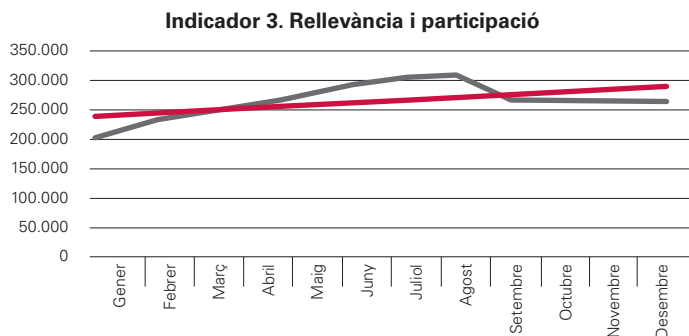
La suma d'aquests indicadors ens mostra un nivell d'interacció constant amb la comunitat a través dels mitjans socials. Els índexs de reputació Klout i Somes es mantenen o pugen al llarg de l'any 2012.

L'índex Klout és molt elevat si el comparem amb altres xarxes de biblioteques públiques catalanes. Amb dades del mes de juny, el nostre índex Klout és de 62, @bibliotequesXBM de 52 i @bibliotequescat de 58.



### 3.1.3. Rellevància

S'entén per rellevància la participació i interacció dels seguidors en els perfils de la marca.



Taula 3. Rellevància i participació

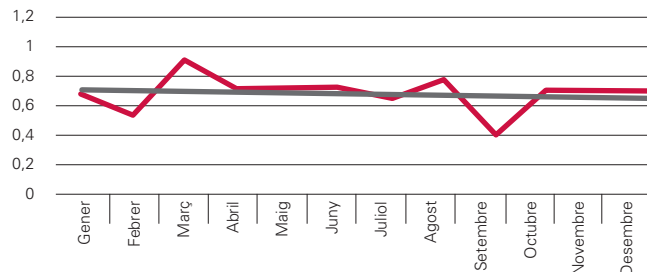
Amb aquest índex general de rellevància podem veure la suma dels següents indicadors: visites al blog, retuits, *M'agrada* a Facebook i arxius mèdia visualitzats, descarregats o compartits (Slideshare, Youtube i Issuu).

Aquest índex també ens indica que la marca és rellevant per als nostres usuaris. La suma d'aquests valors ens mostra una evolució ascendent que demostra un creixement constant en la participació de la comunitat, l'*engagement* i la rellevància de la marca.

### 3.1.4. Fidelització

La fidelització mesura les variables del tràfic generat pel contingut de la marca en els mitjans socials cap a la pàgina web (evolució del tràfic que arriba a la pàgina web des de plataformes socials, usuaris únics, mitjana de temps, etc.).

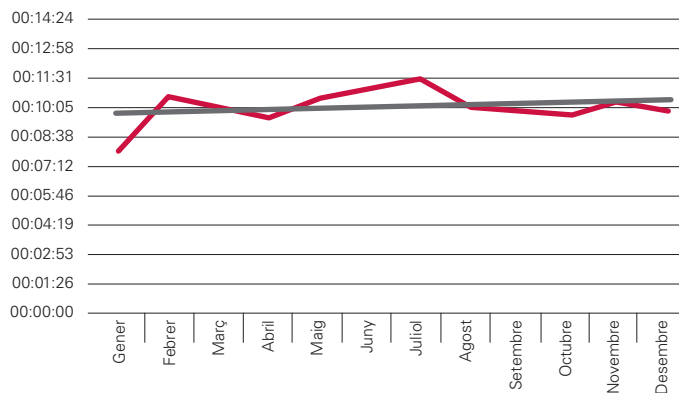
Indicador 4.1. Fidelització  
Percentatge de visites des dels mitjans socials



Taula 4. Percentatge de visites a la web procedents de mitjans socials

El percentatge de visites a la web procedents de la web social és molt poc significatiu, no arriba al 0,8%, però es constata que el percentatge es manté al llarg de l'any.

Indicador 4.2. Fidelització  
Temps mig a la web procedent dels mitjans socials



Taula 5. Temps mig a la web procedent dels mitjans socials

També es manté el temps mig que els usuaris procedents de la web social passen a la web de Biblioteques de Barcelona, amb tendència a créixer.

«la nostra influència se situa per sobre d'altres xarxes de biblioteques públiques del nostre entorn més proper»

#### 4. La gestió de les xarxes socials

S'utilitza el gestor *Hootsuite*<sup>15</sup> per fer el seguiment i la monitorització de les dues xarxes principals de Biblioteques de Barcelona, el perfil de Twitter i la pàgina de Facebook. Es publiquen en diferit totes les activitats i notícies que estiguin penjades prèviament al web de Biblioteques o als de les pròpies biblioteques, deixant uns períodes de temps per tal de poder publicar, en directe, aquelles informacions rellevants del dia.

Amb *Twitterfeed*, s'indiquen a Twitter les publicacions de dos blogs. Amb el *hashtag* #AMPLI, els posts del blog AMPLI. *Associació de musictecaris*<sup>16</sup>; i amb #LlegeixBarcelona, els del blog *Llegeix Barcelona*<sup>17</sup>, que realitzen les biblioteques del districte d'Horta-Guinardó, amb l'Ateneu Barcelonès.

15. Hootsuite ([www.hootsuite.com](http://www.hootsuite.com)).

16. AMPLI. Associació de musictecaris ([musictecaris.blogspot.com](http://musictecaris.blogspot.com)).

17. Llegeix Barcelona ([www.llegeixbarcelona.net](http://www.llegeixbarcelona.net)).



La publicació dels posts del *Bibarnabloc* es realitza en diferit. A primera hora del matí es programa directament amb el Wordpress, el gestor del blog, i amb els hashtags #Bibarnabloc i #recomanem es publiquen els posts que apareixen diàriament. Aquest blog no està sindicat sinó que es publica cada dia, ja que s'inclouen les mencions als autors i a les editorials per aconseguir-ne una major difusió.

#### 5. Conclusions

A partir de l'anàlisi de les dades podem concloure que la nostra estratègia corporativa de presència a Internet i a les xarxes socials apunta en la bona direcció. Tenim uns nivells d'audiència més que acceptables i que no paren de créixer; la nostra influència se situa per sobre d'altres xarxes de biblioteques públiques del nostre entorn més proper i la suma dels indicadors de rellevància ens de-

mostren que la comunitat participa cada cop més i que per als usuaris Biblioteques de Barcelona és una marca rellevant.

També en el sector professional ens han destacat entre els 140 perfils més influents a Twitter en el camp de la Informació i Documentació a Biblogtecarios<sup>18</sup>.

L'aposta per atomitzar la nostra presència a les xarxes socials està donant els seus fruits. No volem que proliferin perfils individuals de cadascuna de les Biblioteques de Barcelona, volem potenciar la marca de xarxa i dedicar tots els recursos a la promoció de la marca mare. La proliferació de perfils i la dispersió de missatges no ajuda gens a la creació d'una identitat digital ni a la construcció de la reputació digital. Volem que els usuaris reconeguin la marca Biblioteques de Barcelona en cadascuna de les biblioteques que formen part de la xarxa.

També constatem que els usuaris valoren la proximitat i una certa informalitat en el llenguatge, i que volen respostes ràpides a les seves consultes. No podem esperar a donar resposta a una pregunta quan l'usuari ja no la necessita, hem de donar-la a temps i de manera pertinent.

El nostre repte actual se centra en el desenvolupament d'aplicacions mòbils i en l'actualització del nostre pla de màrqueting a les xarxes socials a partir dels resultats obtinguts en l'observació i l'anàlisi de les dades.

## Bibliografia

BRÖLL-NADAL, Anna; CABRÉ-SERRA, Anna; GÁNDARA-SANZ, Dolors; FRECHINÉ-PARRA, Pau. «Bibarnabloc, el proyecto líquido de *Biblioteques de Barcelona*» [En línia]. *El profesional de la información*, 2011, noviembre-diciembre, vol. 20, núm. 6, p. 671-676. <<http://dx.doi.org/10.3145/epi.2011.nov.11>>.

BRÖLL-NADAL, Anna; CABRÉ-SERRA, Anna; GÁNDARA-SANZ, Dolors. «Comunicar en la Red. Estrategias de *Biblioteques de Barcelona*» [En línia]. *El profesional de la información*, 2012, julio-agosto, vol. 21, núm. 4, p. 406-412. <<http://dx.doi.org/10.3145/epi.2012.jul.13>>.

CABRÉ SERRA, Anna. «*Biblioteques de Barcelona en las redes sociales*» [En línia]. En: V *Congreso nacional de bibliotecas públicas*. Gijón, 2010. <<http://travesia.mcu.es/portalnb/jspui/handle/10421/4908>>.

CELAYA, Javier. *La empresa en la web 2.0: el impacto de las redes sociales y las nuevas formas de comunicación online en la estrategia empresarial*. Barcelona: Ediciones Gestión 2000, 2011.

DEL PINO. Cristina; CASTELLÓ, Araceli; RAMOS-SOLER, Irene. *La comunicación en cambio constante. Branded content, Community Management, Comunicación 2.0 y Estrategia en medios sociales*. Madrid: Fragua, 2013.

GENERALITAT DE CATALUNYA. *Guia d'usos i estil a les xarxes socials de la Generalitat de Catalunya* [En línia]. Barcelona: Gencat, 2012. <[http://www.gencat.cat/xarxessocials/pdf/v5\\_guia\\_usos\\_xarxa\\_cat.pdf](http://www.gencat.cat/xarxessocials/pdf/v5_guia_usos_xarxa_cat.pdf)>.

GONZÁLEZ FERNÁNDEZ-VILLAVICENCIO, Nieves. «Bibliotecas y marketing en red» [En línia]. *BiD: textos universitaris de biblioteconomia i documentació*, vol. 23 (2009). <<http://www.ub.edu/bid/23/gonzalez2.htm>>.

GONZÁLEZ FERNÁNDEZ-VILLAVICENCIO, Nieves «Demostrar el valor de las bibliotecas en la web social», *Anuario ThinkEPI*, vol. 7, (2013), p. 44-47.

MARQUINA-ARENAS, Julián. *Plan social media y community manager*. Barcelona: UOC, 2012.

18. Biblogtecarios ([www.biblogtecarios.es](http://www.biblogtecarios.es)).

# Preservació digital al núvol

## **Ricard de la Vega**

Cap del servei de Portals i Repositoris. Centre de Serveis Científics i Acadèmics de Catalunya (CESCA)  
*rdelavega@cesca.cat*

### **Resum:**

Ara que ja ningú qüestiona les publicacions digitals, l'accés obert o els repositoris de tipus institucional o temàtic, cal començar a treballar en els temes de preservació. Som responsables d'assegurar l'accés i l'ús futurs dels documents digitals, hagin nascut en format digital o siguin fruit dels projectes de digitalització.

En aquest article es parla sobre la preservació digital i els tipus de continguts, així com els conceptes, tipologies, costos i usos de l'anomenat «núvol». També es mostra un exemple de núvol privat, la preservació de tesis doctorals del TDX i es comenta una experiència de núvol públic, el DuraCloud.

### **Paraules clau:**

preservació digital, núvol, informàtica en núvol, repositoris, DuraCloud, TDX, MetaArchive

## Preservación en la nube

### Resumen:

Ahora que ya nadie cuestiona las publicaciones digitales, el acceso abierto o los repositorios de tipo institucional o temático, deben comenzarse a trabajar los temas de preservación. Somos responsables de asegurar el acceso y el uso futuro de los documentos digitales, hayan nacido en formato digital o sean fruto de diferentes proyectos de digitalización.

En este artículo se trata el tema de la preservación digital y de los diferentes tipos de contenidos, así como los conceptos, tipologías, costes y usos de la denominada «nube». También se muestra un ejemplo de nube privada, la preservación de las tesis doctorales del TDX y se comenta una experiencia de nube pública, el DuraCloud.

### Palabras clave:

preservación digital, nube, repositorios, DuraCloud, TDX, MetaArchive

## Digital preservation on the cloud

### Abstract:

Now that we have reached the point where nobody questions digital publications, open access or institutional or subject repositories, we should deal with preservation issues. We have the responsibility to guarantee future access and use of digital publications, whether they were released in a digital support in their origin or are the result of digitalization projects. In this article we will talk about digital preservation and kinds of content, as well as concepts, typologies, costs and usage of the so called «cloud». We will also show an example of private cloud, the preservation of TDX doctoral theses, and we will make comments on a public cloud experience, DuraCloud.

### Keywords:

digital preservation, cloud, cloud computing, repositories, DuraCloud, TDX, MetaArchive

## 1. Preservació digital

La preservació assegura l'accés i l'ús futur dels documents digitals creats en el present o el passat<sup>1</sup>. Ho fa mitjançant polítiques de conservació, de seguretat informàtica i per tècniques com la migració, l'emulació o la distribució de còpies que permeten el manteniment de les dades i el seu ús a llarg termini.

Canvis en el maquinari, el programari o el format de les dades són els principals motius que fan inestables les dades digitals. A més, els continguts digitals són fràgils, ja que es poden modificar o esborrar de manera conscient, per error, negligència o per una acció externa com, per exemple, que un virus alteri un document.

Una diferència clau entre els continguts digitals i els analògics, com un llibre en paper, és que si els deixes tancats en un dipòsit sense interacció, al cap de molts anys segurament encara es trobaran en un bon estat de conservació. Mentre que si es tanca un contingut digital al mateix dipòsit, per exemple un DVD amb un document escrit amb *Microsoft Word 2003*, al cap d'uns anys probablement no podrem accedir a aquest format i possiblement tampoc no disposarem, almenys de manera habitual, d'un reproductor de DVD per visionar-lo.

En Miquel Térmens, professor de la Universitat de Barcelona, considera que tot i que la preservació digital és la resposta a un problema d'origen tecnològic, no només s'ha d'abordar mitjançant mesures tècniques, sinó també amb mesures organitzatives, financeres i legals, que moltes vegades seran més complicades d'aplicar.

Un aspecte organitzatiu clau és determinar quina informació digital ha de ser preservada i qui és el responsable d'aquesta missió. Això s'aplica directament sobre els

aspectes financers, ja que preservar seguint el model de Diògenes repercuteix directament sobre els costos que se'n deriven. S'ha d'escollir bé quins són els documents a preservar (tasca dels *curators*), ja que actualment el factor econòmic és un fre. Les activitats de preservació són inversions a llarg termini que no generen un rendiment tangible immediat; en paraules d'en Miquel Térmens, «no ajuda a guanyar eleccions i moltes vegades no és capaç de generar beneficis econòmics».

Quant a aspectes legals, no estan resolts tots els dubtes. Existeixen legislacions nacionals i internacionals de protecció de drets d'autor que impedeixen canviar de suport i de format documents sense l'autorització expressa de qui disposi dels drets de propietat intel·lectual. També cal tenir en compte la protecció dels continguts preservats davant d'usos no autoritzats.

Així doncs, la preservació digital ha de ser sostenible a nivell tècnic, organitzatiu, econòmic i legal, i això s'ha de poder demostrar perquè els sistemes de preservació siguin fiables. Aquesta confiança es pot aconseguir mitjançant auditories<sup>2</sup>, mecanismes reconeguts per provar un determinat servei mitjançant un examen professional per part d'experts independents.

En Ciro Lluca, professor de la Universitat de Barcelona i coordinador del projecte PADICAT de la Biblioteca de Catalunya, assenyala cinc regles que poden ajudar a prendre decisions en els processos de preservació digital<sup>3</sup>. Concretament:

1. Dissenyar una estratègia, és a dir, analitzar, pensar, prendre decisions i posar-les per escrit, basant-nos en la tecnologia (avanços tècnics), els recursos (econòmics i humans disponibles) i la pròpia organització (cultura i procediments, responsabilitats i equips).

1. Miquel Térmens. *Preservación Digital*, Barcelona UOC, 2013.

2. En Miquel Térmens, la Núria Casaldàliga i la Cristina Azorín expliquen l'experiència d'aplicació d'una auditoria de l'Esquema Nacional de Seguridad (ENS) sobre el repositori digital de la Universitat Autònoma de Barcelona, el DDD, a «Evaluación con el Esquema Nacional de Seguridad (ENS)» [En línia]. <<http://ddd.uab.cat/record/107685>> [Consulta: 1/07/13].

3. Ciro Lluca, «Preservación digital: 5 reglas para tomar decisiones», XIII Jornadas españolas de documentación (FESABID), 25 de maig de 2013. Toledo.



©istockphoto/rvlsoft

2. Confiar en OAIS, un model teòric creat per la NASA que indica quines funcions han de suportar els sistemes de preservació digital, sense importar quin tipus de dades es custodien ni a quin tipus d'activitat o organització es refereixen. El model es descriu en sis grans blocs: la ingesta, l'emmagatzematge, la gestió de les dades, l'accés, la preservació i els serveis comuns. Les tècniques de preservació que es descriuran després s'inclouen en algun d'aquests sis blocs.
3. Centrar-se en la ingesta, és a dir, en el primer bloc del model OAIS, ja que el 90% de la preservació es realitza en aquest moment del procés, quan l'objecte digital es diposita, junt a les seves metadades. És el moment en què els propietaris de l'objecte autoritzin els aspectes legals pel seu dipòsit, la transformació de format o suport en cas necessari i les condicions de la seva difusió. També és el moment de l'elecció del format, on s'haurien de prioritzar estàndards oberts.
4. Respectar els estàndards; per tal de garantir la interoperabilitat dels sistemes cal usar metadades. N'hi ha de tres tipus: descriptives, tècniques i estructurals. Les primeres descriuen el contingut semànticament (títol, autor...), les segones descriuen el contingut, els drets d'accés i l'històric d'accions de preservació que s'han realitzat (sistema operatiu, resultat antivirus...), i les darreres estableixen relacions entre l'objecte i d'altres continguts al mateix repositori (capítols d'un llibre, versions...). Formats a tenir en compte són Dublin Core (DC), MARC21, METS o PREMIS.
5. Cooperar. Aquesta és una màxima aplicable a molts àmbits, però en el cas de la preservació digital és doblement important disposar de còpies de les dades allunyades geogràficament, i la cooperació entre institucions pot afavorir-ho. També és important en el cas dels projectes de digitalització evitar duplicitats; un bon exemple és la digitalització de les revistes catalanes amb el projecte ARCA<sup>4</sup>, liderat per la Biblioteca de Catalunya.

4. ARCA, Arxiu de revistes catalanes antigues ([www.bnc.cat/digital/arca](http://www.bnc.cat/digital/arca)).

Un cop definit el concepte de preservació digital, mencionats els principals aspectes tècnics, organitzatius, econòmics i legals que fan sostenibles els sistemes, i cinc regles importants per a començar, es detallen a continuació les principals tècniques d'un sistema de preservació. Es treballa amb grans volums de dades que fan inviable un seguiment individualitzat. El personal no pot créixer actualment al mateix ritme que aquests volums de dades, fet que implica la necessitat d'automatitzar al màxim totes les operacions i fer que els formats i altres característiques dels documents siguin el més homogènis possible per tal de simplificar els processos de gestió. Les principals tècniques o mecanismes de preservació són:

1. **Replicació de les dades en llocs geogràficament separats.** D'aquesta manera, si hi ha problemes amb una còpia de les dades sempre es podran restaurar des d'una altra còpia. Un exemple de per què es recomana que les còpies estiguin en llocs separats són els atemptats de l'11 de setembre a les torres bessones de Nova York, on qui disposava d'aquesta redundància a distància va aconseguir la preservació de les dades.
2. **Comprovació de la integritat de les dades.** Cal assegurar que els bytes que formen un fitxer (0s i 1s) no s'alteren amb el pas del temps, ja sigui de manera intencionada o per un error del suport. El mètode més habitual per comprovar la integritat és el càlcul de la suma de verificació (*checksum*), que s'implementa amb funcions de *hash* (MD5, SHA-1, SHA-256). El resultat d'aquesta funció de *hash* sobre el document es guarda com una metadada de l'objecte a preservar per, periòdicament, tornar a calcular la funció i comparar-la amb la metadada. Si els càlculs no coincideixen, s'ha produït una alteració en el contingut del fitxer i serà necessari restaurar-lo d'una còpia correcta.
3. **Refresc de suports.** Els suports tenen una vida útil limitada i amb el temps la informació gravada sofreix alteracions a nivell de bit, fet que es coneix com que els fitxers es corrompen. Els suports poden sofrir alte-

Es treballa amb grans volums de dades que fan inviable un seguiment individualitzat.

racions físiques degudes a factors mediambientals o també per obsolescència tecnològica on, per exemple, els disquets deixen pas als CD, i aquests a les unitats de disc USB.

4. **Migració de formats.** Els formats es poden tornar obsolets amb el pas de temps i cal migrar mentre encara siguin accessibles. Totes les operacions de migració han de quedar registrades (per exemple, en metadades PREMIS) per tal de poder seguir la traçabilitat en el futur, i així poder justificar els canvis de formats i assegurar l'autenticitat del contingut.
5. **Emulació.** Tot i que en molts casos la migració de formats es obligada, també és perillosa pels error que es poden generar, o massa costosa, i moltes vegades s'adopta l'emulació de programaris o maquinari per poder accedir a continguts preservats en formats obsolets.

Després de comentar els principals conceptes de la preservació digital, a continuació es descriuen els del núvol com a base per després relacionar-los en un apartat posterior.



## 2. Núvol

El núvol no és un concepte nou. Vinton Cerf, considerat un dels pares d'Internet, comenta<sup>5</sup> que es tracta d'una extensió natural del *time-sharing* inventat als anys setanta. Simplificant, podríem dir que el pèndol de la història torna a oscil·lar. Es va començar amb els antics *mainframes* (aquells grans ordinadors que ocupaven una habitació), on s'accedia mitjançant terminals lleugers sense gaire capacitat de càlcul o memòria; després es va passar als ordinadors personals, que sí que necessitaven recursos computacionals, per ara tornar al núvol, on es troben aquests recursos i el nou terminal lleuger és el navegador web. La diferència bàsica d'aquesta nova oscil·lació és Internet, que permet que els terminals lleugers no estiguin al mateix edifici dels *mainframes*, sinó repartits per tot el món (una diferència molt considerable).

Com a la natura, existeixen molts tipus de núvols (cúmulus, cumulonimbus...) i molta literatura que en parla. S'ha seleccionat una definició prou completa dels conceptes que es fan servir per a aquesta tecnologia de la Universitat de Califòrnia a Berkeley<sup>6</sup>, diu que la informàtica en núvol (*cloud computing*) es refereix tant al programari lliurat com a servei a través d'Internet, com al maquinari i els sistemes de programari dels centres de dades (*datacenters*) que proporcionen aquests serveis. La paraula clau de la definició és «servei» (*as a service*), a diferència del lliurament de productes que es feia fins ara. Existeixen tres tipus bàsics de serveis al núvol:

1. Els de programari com a servei (*software as a service*, *SaaS*), com les aplicacions de Google, el Gmail o el Calendar, o programaris CRM com Salesforce.com.
2. Els de plataforma (*platform as a service*, *PaaS*), com una imatge de màquina virtual amb un sistema opera-

tiu, un servidor web i una base de dades, o el *Google App Engine* o *Microsoft Azure*.

3. Els d'infraestructura (*infrastructure as a service*, *IaaS*), com els *Amazon Web Services*, l'*EC2* per al còmput i l'*S3* per a l'emmagatzemament.

Tot i que considerem que tots tres formen el núvol, moltes vegades al primer se l'anomena com els serveis del núvol i als altres dos com el propi núvol. Aquests dos darrers s'especialitzen en la infraestructura, i gràcies a grans centres de dades poden reduir costos d'operació i amortitzacions i assolir economies d'escala. Els proveïdors de serveis (*SaaS*) despleguen el programari al núvol, sense necessitat de tenir un centre de dades propi, i es beneficien de l'elasticitat en l'aprovisionament de recursos, és a dir, sense sobredimensionament (més maquinari del necessari) ni infradimensionament (capacitat d'absorbir altes càrregues puntuals). Un altre avantatge és que es poden focalitzar en el seu negoci, sense «distreure's» amb la infraestructura.

Tot aquest núvol hauria de ser transparent per als usuaris finals d'aquests serveis, és a dir, entren al Gmail o el Sales force a través del navegador web com aquell qui pitja l'interruptor de la llum de casa, sense saber-ne res de l'estació transformadora d'electricitat que hi ha al darrera.

Un altre concepte clau és el pagament per ús (*pay-per-use*), que permet pagar només per serveis dels quals se'n fa ús, per exemple, un servidor durant dos dies o cent servidors durant una hora. Per a aquell que necessiti un servidor sempre funcionant segurament serà més barat comprar-lo que fer servir el núvol, però per a aquells que puntualment en necessiten cent és un exemple clar on el núvol guanya la partida, gràcies a l'elasticitat abans mencionada. Quan un núvol es accessible mitjançant aquesta

5. Al pròleg del llibre *Todo va a cambiar*, d'Enrique Dans (Barcelona: Deusto, 2010), en Vinton Cerf també comenta que pels anys setanta ja es parlava també de *computing utility*, terme molt relacionat amb el núvol actual. Es pot consultar en línia a [www.todovaacambiar.com/prologo](http://www.todovaacambiar.com/prologo) [Consulta: 25/06/13].

6. Ambrust, Michael; Fox, Armando; Griffith, Rean; Joseph, Anthony D.; Katz, Randy H.; Konwinski, Andrew; Lee, Gunho; Patterson, David A.; Rabkin, Ariel; Stoica, Ion; Zaharia, Matei. «Above the Clouds: A Berkeley View of Cloud Computing» [En línia], *Technical Report*, No. UCB/EECS-2009-28, febrer de 2009. <<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>> [Consulta: 25/07/13].

fórmula de pagament se l'anomena núvol públic, i el servei que ofereix *utility computing*. S'usa el terme núvol privat pel centre de dades d'una empresa o institució, i encara hi ha un tercer model que és la combinació dels dos models anteriors: el núvol híbrid.

Es pot fer una analogia amb la compartició de cotxe (*car sharing*), on es paga per l'estona que conduïxes el cotxe, i no t'has de preocupar per les revisions, l'assegurança o el canvi de pneumàtics. En comptes de comprar el «producte» cotxe, pagues pel «servei» cotxe<sup>7</sup>.

El darrer concepte important sobre el núvol és la transferència del risc, o els anomenats acords de nivell de servei (*Service Level Agreement, SLA*). Són contractes entre els proveïdors del servei i els seus clients on s'explicita la qualitat del servei com, per exemple, que un servei web estigui sempre funcionant, i quines són les contraprestacions (normalment econòmiques) que el proveïdor haurà de donar al client en cas d'incompliment. És a dir, es transfeix el risc que el servidor estigui en funcionament continu al proveïdor del servei, i aquest no assegura que sempre funcionarà, assegura amb el SLA que, per exemple, si aquest servidor està aturat més de dues hores en un trimestre pagarà una quantitat determinada de diners, fet que l'obliga, d'alguna manera, a vetllar perquè això no passi. En aquest document també s'hi acostumen a incloure clàusules contractuals que recullen garanties de protecció de dades i usos de la informació, la seva seguretat i la seva confidencialitat.

Fins aquí s'han comentat els conceptes bàsics del núvol i també els seus avantatges. Es passa ara a comentar el costat fosc del núvol, la privacitat i localització de les da-



©istockphoto/caracterdesign

des, així com els (de vegades) elevats costos que pot arribar a tenir per segons quins tipus de serveis.

A nivell particular, és habitual penjar fotografies a Flickr o pujar fitxers per compartir-los a Dropbox de manera gratuïta; i és que el concepte de gratuïtat s'ha estès com un estàndard de facto de model de negoci a la xarxa<sup>8</sup>. Però el model de negoci d'aquests serveis al núvol se sustenten, a banda de per la publicitat, també perquè comencen a facturar a partir d'un cert volum de continguts<sup>9</sup>, així que, tot i tenir la sensació de gratuïtat, no hem d'oblidar tenir en compte els costos d'aquests serveis quan s'usen a nivell professional, i més en temes de preservació, on els continguts poden arribar a ser voluminosos i creixents.

7. Un informe del mateix Berkeley apunta que el *car sharing* està guanyant popularitat i que el 2012 s'estimava que dos milions de persones compartien uns 45.000 vehicles a més de 27 països, entre ells, el nostre. Sembla que el pas de productes a serveis no només afecta les TIC. Vegeu «Carsharing» [En línia], Transportation Sustainability Research Center, University of California Berkeley, 2013. <<http://tsrc.berkeley.edu/carsharing>> [Consulta: 25/06/13].
8. Un llibre de culte per entendre els models de negoci gratuïts a la xarxa és Chris Anderson, *Free: The Future of Radical Price*, New York: Hyperion Books, 2009.
9. Dropbox, sistema d'emmagatzematge i compartició de fitxers al núvol, comença a facturar a partir dels 100 GB ([www.dropbox.com/pricing](http://www.dropbox.com/pricing)) i Flickr, el mateix però encarat a la fotografia, als 2 TB ([www.flickr.com/help/limits](http://www.flickr.com/help/limits)).

Respecte a la localització de les dades, es necessari que la legislació vigent aplicable a aquestes permeti que es posin al núvol. En alguns casos pot estar regulat que una determinada informació sensible, per exemple algun tema mèdic, no pugui sortir d'una institució. En aquest cas, s'ha de veure com es tracta això als SLA, si és possible que aquesta informació estigui en un núvol públic o ha de ser en un de privat, i tenir en compte també les legislacions dels proveïdors del núvol<sup>10</sup>. La privacitat de la informació que s'emmagatzema en els núvols públics és un dels temes que més polèmica suscita<sup>11</sup>. De totes maneres, es recorda que molts<sup>12</sup> dels documents que es poden preservar digitalment al núvol, per les biblioteques i els centres de documentació, són precisament continguts en obert<sup>13</sup>, circumstància on la privacitat de la informació no és aplicable.

Per acabar aquest apartat, cal comentar que és important escollir bé el núvol on es volen dipositar continguts, i si convé que sigui un núvol privat o un de públic. En el cas d'escollir un núvol públic, recordem també que n'existeixen molts i amb nivells de serveis i qualitat diferenciats<sup>14</sup>.

### 3. Preservació digital al núvol

En Miquel Térmens ja apuntava<sup>15</sup> el 2009 que la informàtica en núvol tindria en la preservació digital un sector important d'aplicació i de negoci. Però concretament, per a quins mecanismes de preservació és pot fer servir el núvol? DuraSpace, organització sense ànim de lucre de la que parlarem al següent apartat, va plantejar una enquesta<sup>16</sup> el 2010 i va arribar a algunes conclusions: hi ha organitzacions que no inverteixen prou recursos econòmics com per disposar de serveis de qualitat d'aquest estil i un alt percentatge d'institucions probablement començarien al llarg del següent any a fer servir el núvol per temes de preservació.

L'enquesta, amb la meitat de participants d'institucions dels Estats Units, deia que els possibles usuaris del núvol per temes de preservació valoraven positivament, per aquest ordre: l'escalabilitat, l'emmagatzemament remot, la facilitat d'implementació i la flexibilitat. Després, i en segon terme, es creia que seria necessari menys perso-

10. L'associació professional sense ànim de lucre, Business Software Alliance (BSA), ha elaborat un rànquing dels deu millors països per a la informàtica en núvol basant-se en criteris com la privacitat de les dades, la seguretat, la prevenció de la ciberdelinqüència, els drets de protecció intel·lectual, etc. La font és Joe McKendrick, «The 10 Best Countries for Cloud Computing», *Joe McKendrick* [blog], 3 de juliol de 2013, <<http://www.forbes.com/sites/joemckendrick/2013/03/07/the-10-best-countries-for-cloud-computing>> [Consulta: 28/06/13].
11. S'ha publicat que la National Security Agency (NSA) i l'FBI americans, així com els serveis d'intel·ligència britànics, han extret informació (àudio, vídeo, xats, fotografies, documents...) directament de nou proveïdors de núvol, com ara Google, Microsoft, Facebook o Youtube. El programa, batejat amb el nom de PRISM, s'ha convertit en un culebrot amb un protagonista que ha fet la filtració, Edward Snowden, que competeix popularitat amb Julian Assange de Wikileaks (*articles.washingtonpost.com/2013-06-06/news/39784046\_1\_prism-nsa-u-s-servers* [Consulta: 28/06/13]). De totes maneres, l'accés a aquesta informació també es pot obtenir mitjançant requeriments judicials, i aquí és on la localització del núvol és important, com es comenta a l'article de Laia Raventós i David Alandete, «¿Se somete a un juez de California? Sí, acepto» [En línia], *El País*, 13 de gener de 2011. <[http://elpais.com/diario/2011/01/13/sociedad/1294873201\\_850215.html](http://elpais.com/diario/2011/01/13/sociedad/1294873201_850215.html)> [Consulta: 28/06/13].
12. El documental *Google i el cervell mundial* parla sobre el projecte Google Books, on sembla que algunes biblioteques han pujat al núvol públic de Google continguts sense els corresponents drets d'autor i això ha provocat demandes judicials. Vegeu-lo en línia a Televisió de Catalunya (*www.tv3.cat/videos/4614571/Google-i-el-cervell-mundial* [Consulta: 28/06/13]).
13. Un company de feina recelós amb PRISM comentava que ara a la seva tesi, dipositada a Tesis Doctorals en Xarxa (TDX), hi podria accedir l'FBI. De fet, l'FBI pot entrar al repositori i consultar qualsevol de les tesis perquè estan en accés obert. Es va quedar més tranquil amb l'argument d'un altre company que deia que si a la seva tesi no havia posat paraules com *terrorism* o *massive weapon* no havia de patir, que passaria desapercbut pels robots que tracten de manera massiva la informació.
14. És destacable el fet que es va clausurar un proveïdor de serveis al núvol, Megaupload, al·legant infraccions de copyright. Aquesta mesura polèmica va afectar no només els que presumptament cometien delictes compartint pel·lícules sobre les que no es disposaven els corresponents drets d'autor, sinó també a la resta d'usuaris que usaven l'eina per compartir i preservar els seus documents. Vegeu The United States Department of Justice, Justice Department, «Charges Leaders of Megaupload with Widespread Online Copyright Infringement» [En línia], 12 de gener de 2012. <<http://www.justice.gov/opa/pr/2012/January/12-crm-074.html>> [Consulta: 28/06/13].
15. Miquel Térmens, «Preservación digital en 2009», *Anuario ThinkEPI*, 2010, v. 4, p.224-230.
16. Michele Kimpton, «DuraCloud: Open technologies and services for managing durable data in the cloud» [En línia], Repositories and the Cloud Meeting, London, 23 de febrer de 2010. <<http://www.slideshare.net/eduserv/duracloud-open-technologies-and-services-for-managing-durable-data-in-the-cloud>> [Consulta: 28/06/13].



©istockphoto/Nikada

nal per aquests temes a les pròpies institucions, i els elements claus de l'enquesta eren el cost, l'elasticitat i la modalitat de pagament per ús. Aquests dos darrers elements, característiques pròpies del núvol, no tenen gaire aplicació a la preservació, on la demanda d'espai és sempre creixent i acumulativa.

Els principals reptes que l'enquesta troba al núvol són temes de confiança de traspasar actius crítics a tercers, la fiabilitat a llarg termini, la seguretat de les dades, temes de rendiment i amples de banda per moure les dades, la pèrdua de control i complicacions amb SLA, entre d'altres. En preguntar a les institucions com podrien començar a fer servir el núvol, la majoria va respondre que per fer una segona còpia en remot de les dades que es volien preservar, seguit de suport a la preservació, la gestió de col·leccions entre institucions, serveis de computació, i al final de tot, i a molta distància, l'emmagatzemament primari en línia dels continguts.

Després de veure amb l'enquesta necessitats, reptes i usos del núvol, cal plantejar-se en quina de les tipologies encaixa millor la preservació digital.

A l'apartat anterior es detallaven els tipus de núvol. Segons l'informe de Berkeley existeixen tres casuístiques on és millor emprar el núvol públic enlloc d'un núvol privat. Concretament:

1. Quan la demanda fluctua en el temps, ja que equipar un centre de dades amb molts recursos per servir demandes elevades pot quedar infrautilitzat si les demandes decreixen.
2. Quan no es té una previsió de quina serà la demanda i, per tant, no se sap a priori quina ha de ser la magnitud del centre de dades i es necessita elasticitat en l'aprovisionament de recursos.
3. Quan es necessita realitzar anàlisis que poden ser paral·lelitzades en diversos servidors i es vol enllestir ràpidament. El cost d'usar 1.000 servidors d'Amazon Web Services durant una hora és el mateix que el d'un sol servidor durant 1.000 hores.

Cap de les tres aproximacions encaixa amb la naturalesa dels continguts emmagatzemats per a preservació. Com s'ha comentat, la demanda d'espai és sempre creixent (amb major o menor mesura) i acumulativa, i haurien

d'existir estimacions de creixement dins els plans de preservació. A més, si no existissin aquestes estimacions, no cal patir per infrautilitzar infraestructura, només s'haurà d'escollir una arquitectura prou escalable per ampliar-la fàcilment. La tercera casuística tampoc no encaixa amb l'emmagatzemament per preservació, tot i que sí que podria ser útil per a determinats mecanismes de preservació, com per exemple la comprovació d'integritat (*checksums*) de grans volums de fitxers.

Així doncs, la preservació de continguts digitals, tot i ser possible per a qualsevol tipus de núvol, té un millor encaix al núvol privat. Al següent apartat es mostrarà un exemple en aquest sentit: la preservació de tesis doctorals que es realitza a la infraestructura del Centre de Serveis Científics i Acadèmics de Catalunya (CESCA). També es comentarà una experiència de núvol públic que ofereix serveis de preservació, el DuraCloud, de DuraSpace.

## 4. Experiències de preservació digital al núvol

La visió del Centre de Serveis Científics i Acadèmics de Catalunya (CESCA) és ser capdavanter en la gestió i l'ús de les TIC per millorar la qualitat i l'eficiència del sistema universitari i de recerca, aprofitant les economies d'escala mitjançant la cooperació interuniversitària, les bones pràctiques professionals i la compartició de recursos.

Aquesta visió encaixa directament amb el que s'ha anat mencionant a l'article respecte la importància de la cooperació en temes de preservació, i es disposa de diver-

sos serveis que ho exemplifiquen. Es pot considerar com un exemple de núvol privat, tant a nivell d'infraestructura, com de programari com a servei (SaaS).

Des de 2001 s'ha desenvolupat i es realitza el manteniment evolutiu de vuit repositoris digitals<sup>17</sup>. Es preserven tesis, articles de revistes, documents de recerca, imatges o contingut audiovisual de patrimoni cultural i fins i tot llocs web<sup>18</sup>.

A continuació s'aprofundeix en dues experiències. La primera es tracta d'un projecte que ofereix serveis de preservació al núvol, iniciativa que es va avaluar en les seves etapes inicials per veure si era interessant usar un núvol públic de manera complementària pels repositoris digitals gestionats pel Centre. La segona experiència detalla algunes de les tècniques de preservació que s'estan fent servir en un dels repositoris, el TDX, interessant des de la temàtica d'aquest article perquè a la vegada que el repositori es pot considerar situat en un núvol privat per les institucions membres, també s'usa un altre núvol privat extern, la Xarxa MetaArchive.

### 4.1 DuraCloud

DuraCloud és un servei desenvolupat per l'organització sense ànim de lucre DuraSpace, que també gestiona les comunitats de programaris de repositoris digitals DSpace i Fedora. La seva missió és facilitar per a les institucions o usuaris finals serveis de preservació al núvol, i està especialitzat en serveis per a biblioteques acadèmiques, centres de recerca i altres organitzacions amb patrimoni cultural.

17. Conjuntament amb el Consorci de Biblioteques Universitàries de Catalunya (CBUC) les Tesis Doctorals en Xarxa (TDX, [www.tdx.cat](http://www.tdx.cat)), el Dipòsit de la Recerca de Catalunya (RECERCAT, [www.recercat.cat](http://www.recercat.cat)) i Materials Docents en Xarxa (MDX, [www.mdx.cat](http://www.mdx.cat)); amb el CBUC i la Biblioteca de Catalunya (BC), les Revistes Catalanes amb Accés Obert (RACO, [www.raco.cat](http://www.raco.cat)); per a la BC el Patrimoni Digital de Catalunya (PADICAT, [www.padicat.cat](http://www.padicat.cat)); per al Departament de Cultura de la Generalitat i per a la Filmoteca de Catalunya, els seus repositoris digitals, (CALAIX, [calaix.gencat.cat](http://calaix.gencat.cat) i [filmoteca.cesca.cat](http://filmoteca.cesca.cat)), i per a la Fundación Española de Ciencia y Tecnología (FECYT), el repositori RECYT ([recyt.fecyt.es](http://recyt.fecyt.es)).

18. El projecte PADICAT és molt interessant des del punt de vista de la preservació; es pot ampliar informació al següent article: Lluca, Ciro; Cócera-Saló, Daniel; Torres, Natalia; Suades-Méndez, Gerard; Vega-Sivera, Ricard de la, «A ritmo de tweet: archivando elecciones 2.0», *El profesional de la información*, 2011, vol. 20, núm. 3, p. 309-314.

Els objectius inicials el 2010 eren usar el núvol tant per emmagatzemament de contingut com per usar les seves capacitats computacionals per realitzar tasques de mineria de dades (*data mining*). Es va iniciar un pilot amb Biodiversity Heritage Library per pujar al núvol una còpia de 40 TB d'espai d'algunes de les seves col·leccions, i un altre amb la New York Public Library per, a més de pujar continguts, realitzar tasques de transformació d'imatges en TIFF a JPEG2000, format amb més compressió que requereix menys espai en disc.

Un cop enllestits els pilots, DuraCloud comença a oferir serveis el 2011. Actualment, és una interfície per pujar continguts i s'encarrega de redundar-los distribuint-los per diferents núvols públics d'emmagatzematge, com Amazon o RackSpace (*cloud broker*). D'aquesta manera, s'aconsegueix que existeixin còpies dels continguts disperses geogràficament en diferents núvols públics però amb un únic punt de gestió i facturació pel servei.

DuraCloud disposa d'eines i API per pujar continguts de manera cada cop més integrada amb els programaris de repositoris digitals, o restaurar-los en cas de necessitat. També disposa de mecanismes per garantir la integritat de les dades de les diferents còpies dels continguts dels diferents núvols i de sincronització dels continguts, fet que permetria realitzar còpies de seguretat directament en línia.

La iniciativa també disposa de dos serveis complementaris més, una interfície perquè diferents usuaris, o grups de recerca, puguin compartir documents de manera col·laborativa, com un Dropbox o Google Drive; i un sistema per emmagatzemar i reproduir contingut audiovisual que després es pugui reproduir en visors incrustats als webs o repositoris digitals de les institucions, com Youtube o Vimeo.

La possibilitat de contractar en forma de serveis permet a una institució disposar de solucions simples o més robustes. Existeixen, documentats al web de DuraCloud, des

de casos d'ús on es fa servir per preservar una còpia de diverses col·leccions importants que ocupen entre 1 i 2TB d'espai per un cost aproximat de 4.500\$ trimestrals, fins a altres institucions que a banda de disposar de còpies també disposen d'una còpia primària en línia i mecanismes de sincronització entre totes, que amb més de 5 TB d'espai té un cost trimestral d'aproximadament 31.000\$<sup>19</sup>.

Existeixen altres serveis semblants a DuraCloud, o alguns de mixtes que combinen l'emmagatzemament local amb el remot al núvol<sup>20</sup>. Aquests poden ser útils per a determinades institucions, tot i que, com s'ha mencionat a l'apartat anterior, el tema de l'emmagatzemament encaixa de manera més natural per la preservació en els núvols privats. El que pot ser més interessant d'iniciatives com DuraCloud és l'oferiment de serveis no només enfocats a l'emmagatzemament, sinó a la computació, com apunten els objectius del projecte i algun pilot inicial, i que es puguin fer, per exemple, migracions massives de formats.

## 4.2 Preservació de les tesis doctorals del TDX

Tesis Doctorals en Xarxa (TDX) és un repositori cooperatiu que conté, en format digital, tesis doctorals llegides a les universitats de Catalunya i d'altres comunitats autònomes. La consulta de les tesis és d'accés lliure. Els objectius d'aquest repositori, gestionat i coordinat pel Consorci de Biblioteques Universitàries de Catalunya (CBUC) i el Centre de Serveis Científics i Acadèmics de Catalunya (CESCA), i patrocinat per la Generalitat de Catalunya, són difondre els resultats de la recerca universitària, oferir als autors una eina per incrementar l'accés i la visibilitat del seu treball i afavorir l'edició electrònica en obert de la producció científica pròpia.

Les tesis doctorals són un contingut que tradicionalment, per la seva qualitat, s'ha considerat important per ser preservat. A continuació és detallaran, en clau històrica,

19. Existeixen diverses modalitats de subscripció segons els serveis contractats ([duracloud.org/pricing](http://duracloud.org/pricing)) [Consulta: 30/06/13].

20. Com el pilot Kindura, del JISC ([www.jisc.ac.uk/whatwedo/programmes/flexibleservicedelivery/kindura.aspx](http://www.jisc.ac.uk/whatwedo/programmes/flexibleservicedelivery/kindura.aspx)) [Consulta: 30/06/13].

alguns dels mecanismes de preservació que des dels consorcis gestors, que es poden considerar un núvol privat de les institucions participants al TDX, s'han implementat, com ara la migració de suports, la replicació i la comprovació de la integritat dels continguts.

El repositori es va posar en funcionament l'any 2001, i va rebre 8.000 consultes a unes poques tesis que aleshores ocupaven 80 GB d'espai en disc. A finals de 2012 disposava ja de 13.562 tesis en accés obert incorporades per divuit universitats participants. L'espai de les tesis s'ha incrementat de manera moderada fins als 200 GB (es tracta de fitxers en PDF), però els accessos a aquestes tesis han crescut de manera exponencial, fins als 18 milions de consultes<sup>21</sup>.

La infraestructura actual del TDX ha escalat per poder servir la demanda actual d'usuaris que volen consultar les tesis, i això ha implicat haver de fer migracions de suport, tant de programari com de maquinari. La primera instància del repositori es va allotjar en un supercomputador, un HP V2500 amb 16 processadors, 4 GB de memòria i 227 GB de disc. I ja aleshores es realitzava una còpia de seguretat a una cabina de cintes StorageTek TimberWolf 9740, que tenia una capacitat de 2,7 TB.

El 2003 es va migrar el TDX cap a un nou entorn més especialitzat per a aquest tipus de producte. D'una banda, la capacitat de càlcul del repositori va passar a un servidor dedicat HP rp5430, i d'altra banda, l'espai per emmagatzemar les tesis es va separar a una cabina de discs HP EVA v.2 amb 2,8 TB d'espai disponible. El fet de separar l'emmagatzemament a una cabina especialitzada permetia un millor creixement i aprofitament de l'espai, i també més mesures de preservació, més sofisticades, que incorporava la pròpia cabina, més tolerant a fallades de discs que no pas quan aquests estan als propis servidors.

El 2006 es van produir dues migracions més de suports. D'una banda, es va migrar del servidor dedicat cap a un clúster d'alta disponibilitat de 32 nodes HP DL360 amb altres repositoris digitals gestionats pel CESCA. D'altra banda, es va renovar el robot de cintes de *backup*, passant a un Adic Scalar i2000. La primera de les migracions va ser més transparent, ja que les tesis estaven emmagatzemades a la cabina de discs, però la segona va ser més costosa perquè les cintes eren d'una generació tecnològica molt diferent.

El 2009 es va renovar la cabina de discs, passant a un NetApp FAS3170 amb 60 TB d'espai en disc i majors prestacions per augmentar la tolerància a la fallada dels discs<sup>22</sup>. I el 2013 està previst tornar a migrar la capacitat de processament a un nou clúster d'alta disponibilitat.

Com es pot veure, cada pocs anys és normal que es produeixin canvis en la infraestructura i que s'hagin de realitzar migracions d'unes tecnologies de disc o cinta a unes altres per poder seguir preservant les dades. El mateix passa amb el programari. El TDX es va desenvolupar el 2001 amb el programari especialitzat per a tesis doctorals ETDdb, creat a la Virginia Tech University, eina usada aleshores per la xarxa internacional Networked Digital Library of Theses and Dissertations (NDLTD), de la qual el TDX en forma part.

Amb els temps, se li han anat incorporant millores al programari, incloent mòduls per a la publicació pel protocol d'interoperativitat OAI-PMH i d'estadístiques d'ús. Però les necessitats del repositori, més usat cada vegada, no estaven alineades al ritme d'actualització del programari, fet pel qual es va migrar el 2011, any del seu dècim aniversari, cap a la plataforma DSpace. La migració va suposar la reestructuració de la informació i el traspàs de les metadades de l'antic sistema de l'ETDdb a Dublin Core.

21. Es poden consultar estadístiques sobre el TDX a CESCA, *Memòria d'activitats de 2012* [En línia]. <[http://www.cesca.cat/sites/default/files/docs/ma12\\_0.pdf](http://www.cesca.cat/sites/default/files/docs/ma12_0.pdf)> [Consulta: 1/06/13].

22. Com el RAID-DP. Més informació a NetApp, «FAS3100 Series Technical Specifications», <<http://www.netapp.com/us/products/storage-systems/fas3100/fas3100-tech-specs.aspx>> [Consulta: 1/06/13].



©istockphoto/cmcdern1

A banda de la migració de suports, un altre mecanisme de preservació present al TDX és la replicació de les dades. Actualment existeixen onze còpies de les tesis doctorals distribuïdes de la següent manera:

1. La versió en línia del repositori accessible en obert a través de *www.tdx.cat*.
2. Una còpia de seguretat a la llibreria de cintes del CESCA, amb una política de *backup* diària amb retencions de fins a 1 any.
3. Una altra còpia de seguretat en cinta, desada en un armari ignífug dins les instal·lacions del CESCA.
4. Una tercera còpia de seguretat en cinta guardada a uns 150 km del centre de dades, al magatzem cooperatiu de documents de baix ús GEPA.
5. Set còpies fosques a la xarxa MetaArchive, iniciativa de la qual s'amplia informació a continuació.

"... un aspecte bàsic de la preservació és la comprovació de la integritat dels fitxers, és a dir, que la informació dels fitxers no hagi estat alterada i pugui ser accessible."

Tant pel refresc de suports com per disposar de diverses còpies de les tesis, un aspecte bàsic de la preservació és la comprovació de la integritat dels fitxers, és a dir, que la informació dels fitxers no hagi estat alterada i pugui ser accessible. És important aquesta comprovació en cada migració de suport i també que es garanteixi en les còpies distribuïdes. Per a les migracions, les comprovacions són *ad hoc*, però per les còpies es realitza de manera periòdica, mitjançant el propi DSpace per a la versió en línia de les tesis, i amb el programari LOCKSS per a les còpies de la xarxa MetaArchive.

#### 4.2.1 MetaArchive

La cooperativa MetaArchive<sup>23</sup> és una xarxa de preservació digital administrada per l'Educopia Institute, una organització sense ànim de lucre per ajudar les institucions a preservar i fer accessibles els seus actius digitals. La missió de la cooperativa és difondre coneixement sobre mètodes de preservació digital distribuïda i crear una xarxa estable

23. Educopia Institute. «MetaArchive Cooperative Charter» [En línia], march 2007. <[http://www.digitalpreservation.gov/partners/documents/MetaArchiveCharter\\_0707.pdf](http://www.digitalpreservation.gov/partners/documents/MetaArchiveCharter_0707.pdf)> [Consulta: 28/06/13].



i dispersa geogràficament de còpies fosques (*dark archives*) de continguts digitals a la que es recorrerà, en cas necessari, per restaurar col·leccions dels seus membres.

L'experiència va començar el 2002 quan sis biblioteques del sud-est dels Estats Units van treballar plegades per desenvolupar una solució de preservació digital per a les seves col·leccions de continguts especials. Van arribar a la conclusió, també amb la Library of Congress, que cap d'elles podria arribar a tenir individualment una solució segura i sostenible per a la preservació a llarg termini de continguts digitals, però que unides podien crear un sistema distribuït. El fruit d'aquesta col·laboració va ser una iniciativa de i per a biblioteques, arxius i altres organitzacions de memòria digital.

Les organitzacions que comparteixen la missió i els objectius de la iniciativa poden formar-ne part. Hi ha tres tipus de membres, amb uns drets i deures determinats. Concretament:

1. Membre de sosteniment. Institucions que lideren la iniciativa, a través d'un comitè format per un membre de cada institució. Són els responsables principals de provar, desenvolupar i mantenir el maquinari i el programari, la connexió de la xarxa i la transmissió dels estàndards, incloent el disseny i la implementació de noves característiques, tractant amb temes de seguretat, validació de continguts, control de la integritat de les dades i rastreig de metadades. També són responsables de preservar part del contingut digital de la xarxa.

2. Membres de preservació. Institucions responsables de preservar part del contingut digital de la xarxa.
3. Membres col·laboratius. Grup d'institucions que executen un node de preservació centralitzat i comparteixen a través seu continguts digitals. Aquest membre ha de ser una entitat legal i assignar personal tècnic per al desenvolupament i l'adaptació de codi (*plugins*) per tal de preparar les col·leccions per a la seva ingesta a la xarxa, així com la gestió d'aquestes col·leccions (altes i baixes).

Tots els membres han d'instal·lar i mantenir un node de preservació (que anomenen *cache*) amb unes característiques determinades<sup>24</sup>, dedicar un percentatge de recursos humans tècnics i pagar una quota econòmica<sup>25</sup>. A canvi, es beneficien de poder restaurar els continguts preservats en cas de desgràcia i disposen d'informes periòdics de l'estat dels continguts.

Diverses universitats i biblioteques d'arreu del món participen en aquesta xarxa cooperativa i, tot i que la majoria són de l'àmbit nord-americà, també n'hi ha una al Brasil, una a Anglaterra i, des del 2011, hi participa el TDX (vegeu la figura 1). Entre els centres participants destaquen la Georgia Tech, la Virginia Tech i la Library Congress.

La infraestructura de la xarxa es basa en el programari lliure LOCKSS<sup>26</sup>, un sistema de preservació digital dirigit a biblioteques construït sobre el principi que «conservar moltes còpies assegura la preservació del material» (*lots of*

24. Les característiques principals del servidor, actualment, són que ha de disposar de 6 GB de memòria i entre 16 i 24 TB d'espai de disc. Aquest disc ha de ser preferiblement intern del node i no d'un sistema SAN o NAS d'emmagatzemament, ja que un dels requisits dels nodes de la xarxa és que siguin el màxim d'homogenis entre si i quan més aïllats d'altres entorns informàtics millor. Cerquen el concepte de simplicitat més que el de redundància o sofisticació perquè si un node cau, existeixen altres còpies del seu contingut a altres nodes de la xarxa i es pot restaurar un cop el problema s'hagi solucionat. Existeix el compromís entre els membres de renovar, cada tres anys, les característiques del node. La resta de característiques tècniques es poden consultar a Educopia Institute, «Technical Specifications. The MetaArchive Cooperative» [En línia], 23 de gener de 2013. <[http://www.metaarchive.org/public/resources/charter\\_member/2013\\_Technical\\_Specifications.pdf](http://www.metaarchive.org/public/resources/charter_member/2013_Technical_Specifications.pdf)> [Consulta: 28/06/13].

25. Actualment, el cost del node servidor que cada membre ha d'instal·lar és d'aproximadament 4.600 \$ més 1 \$ per cada GB emmagatzemat a l'any i un 2% d'un administrador de sistemes. La quota dels membres de sosteniment és de 5.500 \$/any, 3.000 \$/any els membres de preservació i variable en funció del número d'institucions per als membres col·laboratius ([www.metaarchive.org/costs](http://www.metaarchive.org/costs)) [Consulta: 28/06/13].

26. Programari lliure LOCKSS ([www.lockss.org](http://www.lockss.org)) [Consulta: 28/06/13].

El seu funcionament es basa en la distribució de còpies fosques de les tesis en diferents servidors de les institucions participants, assignats arbitràriament i per criteris de llunyania geogràfica.

*còpies keep stuff safe*)<sup>27</sup>. El seu funcionament es basa en la distribució de còpies fosques de les tesis en diferents servidors de les institucions participants, assignats arbitràriament i per criteris de llunyania geogràfica. Per assegurar la preservació dels documents es fan comprovacions periòdiques bit a bit de les dades per tal de detectar-ne qualsevol corrupció i restablir, en cas que sigui necessari, la integritat dels fitxers.

La xarxa MetaArchive no intenta substituir altres mecanismes de preservació locals, sinó que es tracta d'un núvol privat complementari en el qual actualment s'estan preservant tesis doctorals de les institucions participants. Usant el mateix programari LOCKSS també es podrien fer altres xarxes per preservar de la mateixa manera diferents continguts especials. Això sí, les tesis doctorals ocupen poc espai, però s'han de tenir en compte els costos de disposar d'una solució d'aquestes característiques per a



**Figura 1.** Localització dels membres de la cooperativa MetaArchive

grans volums de dades. Si el que es vol preservar a la xarxa són 10 TB s'ha de posar a disposició de la xarxa entre 70 i 80 TB de disc<sup>28</sup>, fet que limita aquesta solució només a determinats continguts.

## 5. Conclusions

La preservació digital assegura l'accés als continguts digitals a llarg termini. Es treballa amb grans volums de dades on s'automatitzen al màxim totes les operacions per simplificar els processos de gestió. Existeixen mecanismes de preservació que estan prou estandarditzats i documentats com perquè es puguin considerar una *comodity*, és a dir, que s'implementen de la mateixa manera a diferents organitzacions.

27. Al web de MetaArchive modifiquen un parell de cites de Ben Franklin per il·lustrar la missió de la cooperativa, com «A place for everything, everything in its place, and seven secure copies» o «Nothing is certain but death, taxes, and file corruption» ([www.metaarchive.org](http://www.metaarchive.org)) [Consulta: 28/06/13].

28. Castillo, José Manuel; Jorba, Ferran. «Emmagatzematge distribuït i preservació digital: una panoràmica d'alternatives» [En línia]. *BID: textos universitaris de biblioteconomia i documentació*, juny, núm. 20. <<http://www.ub.edu/bid/20casti1.htm>> [Consulta: 30/06/13].

Una opció és oferir aquests mecanismes com a serveis mitjançant el núvol públic sota la modalitat de pagament per ús, concepte que destaca des de fa uns anys a totes les travesses de tendències tecnològiques de les consultores<sup>29</sup>. De totes maneres, els continguts típicament de preservació, on la demanda d'espai és sempre acumulativa i s'hauria de disposar d'estimacions de creixement dins els plans de preservació, no hi acaben d'encaixar. Els costos d'operació són actualment més elevats<sup>30</sup> en serveis al núvol públic que preservar a la infraestructura de les pròpies institucions (núvols privats).

Amb els actuals models de negoci del núvol públic, creiem que només poden fer ombra al núvol privat en temes de preservació quan facin falta recursos computacionals elevats de manera temporal com, per exemple, per realitzar migracions massives de formats. Aquest és un camp que es troba encara a les beceroles del que pot arribar a ser.

S'ha de tenir en compte que cada contingut a preservar té uns requisits determinats que s'han de garantir a l'hora d'escollir mesures de preservació, i que no existeixen solucions globals (sostenibles) per a tot. El futur és híbrid, on diferents mesures o solucions s'hauran de combinar per tal de ser eficients des del punt de vista tècnic i sobretot econòmic. EITDX és un exemple on es complementen actualment la preservació dins la infraestructura tecnològica del CESCO i la xarxa MetaArchive.

La preservació digital no és prou «agraïda»; es tracta d'una disciplina enfocada a llarg termini i on només és notícia si falla, com l'electricitat, en la qual només hi penses quan se'n va la llum. Això la fa poc visible i fa que no existeixi prou pastís encara perquè empreses del núvol apostin fort i ofereixin solucions rendibles a les institucions.

El fet que les institucions apliquin mesures de manera individual també té costos elevats que poden compartir-se

col·laborant entre elles, per exemple, a través de consorcis com el CESCO o el CBUC, o en iniciatives com la xarxa MetaArchive. Generalment, no existeix un valor diferencial entre institucions a l'hora de preservar les seves col·leccions («simplement» volen poder-hi accedir a llarg termini). La col·laboració és clau per a la seva sostenibilitat.

La preservació digital no es tracta només d'un repte tecnològic, sinó que s'ha d'abordar també amb mesures organitzatives, financeres i legals que facin sostenibles els sistemes de preservació.

Aspectes organitzatius com determinar quina informació digital ha de ser preservada, en quin tipus de tecnologia i amb quins costos, són bàsics durant el disseny dels sistemes de preservació, ja que existeixen diferents opcions a implementar depenent d'aquests factors. També és necessari definir clarament dins de les institucions els diferents rols (selecció, *curator*, responsable de seguretat...) que requereix la preservació. Avui en dia és un problema que aquests perfils no estiguin clarament establerts dins les institucions.

Un altre problema actual és que es volen preservar grans quantitats de documents que han estat digitalitzats gràcies a ajudes o subvencions puntuals que no tenien en compte els costos de la preservació. Aquests conceptes, digitalització i preservació, haurien d'anar relacionats d'origen.

Actualment, el factor econòmic és un fre per a les activitats de preservació digital, que no generen un rendiment tangible immediat sinó a llarg termini. Si es vol assegurar l'accés i l'ús futur dels documents digitals cal una bona tria de les eines de preservació, una bona selecció dels documents a preservar i que el sistema sigui sostenible i assumit institucionalment.

29. Pragsis. Tecnología e innovación, «Gartner identifica el Top 10 para 2013 en tendencias de estrategias en tecnología» [En línia]. <[http://bigdata-hadoop.pragsis.com/blog/2012/11/23/gartner\\_identifica\\_el\\_top\\_10\\_para\\_2013\\_en\\_tendencias\\_de\\_estrategias\\_en\\_tecnologia](http://bigdata-hadoop.pragsis.com/blog/2012/11/23/gartner_identifica_el_top_10_para_2013_en_tendencias_de_estrategias_en_tecnologia)> [Consulta: 1/06/13].

30. Rosenthal, David S. H.; Vargas, Daniel L. «Distributed Digital Preservation in the Cloud». [En línia], *International Journal of Digital Curation*. 2013, p. 107-119. <<http://ijdc.net/index.php/ijdc/article/view/248>> [Consulta: 1/08/13].

## Bibliografia

AMBRUST, Michael; FOX, Armando; GRIFFITH, Rean; JOSEPH, Anthony D.; KATZ, Randy H.; KONWINSKI, Andrew; LEE, Gunho; PATTERSON, David A.; RABKIN, Ariel; STOICA, Ion; ZAHARIA, Matei. «Above the Clouds: A Berkeley View of Cloud Computing» [En línia]. *Technical Report*, No. UCB/EECS-2009-28, febrer de 2009. <<http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf>> [Consulta: 25/07/13].

CASTILLO, José Manuel; JORBA, Ferran. «Emmagatzematge distribuït i preservació digital: una panoràmica d'alternatives» [En línia]. *BiD: textos universitaris de biblioteconomia i documentació*, juny, núm. 20. <<http://www.ub.edu/bid/20casti1.htm>> [Consulta: 30/06/13].

Educopia Institute. «MetaArchive Cooperative Charter» [En línia], march 2007. <[http://www.digitalpreservation.gov/partners/documents/MetaArchiveCharter\\_0707.pdf](http://www.digitalpreservation.gov/partners/documents/MetaArchiveCharter_0707.pdf)> [Consulta: 28/06/13].

ELVIRA SILLERAS, María. «Preservación de los documentos digitales: Guía para comenzar. ISO/TC 46/SC 11. Versión española, octubre 2010», *Revista Española de Documentación Científica*, 34, julio-septiembre 2011, p. 461-477.

GALLART MARSILLAS, Núria; KEEFER, Alice. «Preservació de dipòsits digitals» [En línia]. Consorci de Biblioteques Universitàries de Catalunya, 2007. <<http://hdl.handle.net/2072/4078>> [Consulta: 01/08/13].

GALLART MARSILLAS, Núria; KEEFER, Alice. *La preservación digital: el reto para bibliotecas en el siglo XXI*. Barcelona: Editorial UOC, 2007. ISBN: 978-84-9788-567-6.

JORBA, Ferran. «I si deixem de parlar de preservació digital? Variacions sobre un tema de David Rosenthal, amb un poutpourri final amb la CDLib Orchestra» [En línia]. Consorci de Biblioteques Universitàries de Catalunya, 2010. <<http://hdl.handle.net/2072/67597>> [Consulta: 01/08/13].

LAVOIE, Brian F. «The Open Archival Information System Reference Model: Introductory Guide» [En línia]. <[http://www.dpconline.org/docs/lavoie\\_OAIS.pdf](http://www.dpconline.org/docs/lavoie_OAIS.pdf)> [Consulta: 01/08/13].

LLUECA, Ciro. «Preservación digital: 5 reglas para tomar decisiones». XIII Jornadas españolas de documentación (FESABID). 25 de mayo de 2013. Toledo.

ROSENTHAL, David S. H.; VARGAS, Daniel L. «Distributed Digital Preservation in the Cloud» [En línia], *International Journal of Digital Curation*. 2013, p. 107-119. <<http://ijdc.net/index.php/ijdc/article/view/248>> [Consulta: 1/08/13].

TÈRMENS, Miquel. *Preservación Digital*. Editorial UOC - El Profesional de la Información, 2013.

VEGA, Ricard de la. «Repositorios en la nube» [En línia]. Jornadas OS-Repositorios (4as : 2010 : Barcelona), 2010. <<http://hdl.handle.net/10609/1705>> [Consulta: 01/08/13].