DIGITAL.CSIC

OPEN ACCESS
TO DIGITAL DOCUMENTS

Perceptions and participation in the Open Access movement at CSIC: Report of Digital.CSIC survey to researchers

A summary

Isabel Bernal
Digital.CSIC Technical Office
October 2010

Section 1: Objectives of the survey

Digital.CSIC was launched in January 2008 with the aim to facilitate seamless access to research made in CSIC 122 centers and institutes and to organise, archive and preserve it in a centralised digital platform. Backed with more than 70 year history, CSIC is a fundamental producer of science in Spain and the main scientific state agency nation-wide. Digital.CSIC seeks to become its memory of current, past and future research.

Digital.CSIC is a direct consequence of the signing by CSIC Presidency of the Berlin Declaration in 2006 and is a project developed by CSIC Libraries Coordination Unit. In less than 3 years of existence, Digital.CSIC has managed to leave behind its initial phase successfully and has established a distributed work model which rests on the following communities: its Technical Office that undertakes all technological innovations in the platform; carries out a large part of the deposits and organises a wide array of promotional, training and awareness raising activities; CSIC network of 78 libraries that take part in the repository by uploading research made in their corresponding centers and institutes, and last but not least CSIC scientific community who self-archive their own publications. To date, the annual average in the upload of deposits in Digital.CSIC is distributed as follows: a 45% goes on the libraries network, a 37% corresponds to the repository's Technical Office and the remaining 18% represents individual researchers.

> ***Digital.CSIC by the numbers [October 18, 2010][i]***
>
> -More than 26,000 items show the multidisciplinary character of CSIC research. The growth in contents is constant, and 2009 closed with a 42% increase with respect to 2008
>
> -Digital.CSIC houses post-print and pre-print articles, conference papers, technical papers, reports, working papers, data sets, books and book chapters, presentations, divulgative material, patents, software, audiovisual material, maps, music compositions, etc
>
> -More than 3.500.000 visits to the web since January 2008
>
> -More than 3.560.000 full text downloads since January 2008, with US-based users on the top

In Spring 2010 Digital.CSIC Technical Office conducted surveys addressing CSIC researchers and librarians in order to analyze how they perceive and to what extent they are knowledgeable about the open access movement and to see how they value the institutional repository. Both surveys included a high number of open-ended questions to give respondents the opportunity to express their opinions about Digital.CSIC and suggest ways to improve it.

As far as the survey to CSIC researchers is concerned, a first one was already conducted in 2007, and on that occasion all questions focussed on publications habits and general stances towards open access. The 2010 survey has been far wider in scope, as besides reporting about changes in publications trends amongst CSIC scientific community and their degree of knowledge about open access, an evaluation of Digital.CSIC has been added, a component that could not exist 3 years ago.

The survey was divided into 4 parts: one that shedded light on the respondents' profile, a second one that centered on publications habits, a third section that provided detailed information on how they value and how much they know about the 2 roads towards open access and a closing part that focussed on how they evaluate Digital.CSIC. The survey was sent to a CSIC mailing list comprised of 6,879 people, of whom 832 responded. Compared to the 2007 survey, in which 228 researchers participated out of a total of 2,514, we observe a slight increase in participation (from a 9% to a 12%).

Digital.CSIC being a hybrid project, in which the active participation of researchers is key to guarantee the sustainability and usefulness of the repository, it has been quite interesting to collect their opinions about the current platform as well as their suggestions for future improvements. We have thus incorporated their feedback into our work agenda and have got confirmation that most of their suggestions are issues that are already on top of the repository's priorities. However, it is worth mentioning that at the time this survey was released neither the new Digital.CSIC web nor its new statistics module had been made public yet.

The following is a summary of the main findings from this survey. A complete report is available at "Percepciones y participación en el acceso abierto en el CSIC: Informe sobre la Encuesta de Digital.CSIC para investigadores" (http://digital.csic.es/handle/10261/28543).

Section 2: Main findings

Respondents' profile

Most researchers that took part in the survey were male, with an average age between 36 and 55 years old and part of the permanent scientific staff at CSIC. The multidisciplinarity of the Spanish National Research Council showed in the survey, Biology and Biomedicine being the most representative research area, with a 17.2% of respondents, followed by Humanities and Social Sciences (15.4%), Natural Resources (15.3%), Physics (13.9%) and Agrarian Sciences (12.7%). The remaining respondents were affiliated to centers and institutes devoted to Chemistry (9.1%), Materials Science and Technology (9.1%), Food Science (7.5%) and Computer Sciences and Documentation (0.5%).

Publication trends

The survey highlighted that CSIC scientific community continues to publish in journals of leading subscription-based publishers largely, although a relative opening towards open access publishers and/or publications is observed, mostly amongst researchers in the fields of Physics, Biology, Medicine and related disciplines.

As far as top editorial preferences for their 3 latest publications are concerned, Elsevier ranked first, with a 57.3% of respondents, and in this respect there are no changes as compared with the 2007 survey. Second in this classification stood Springer (35.5%), followed by Wiley/Blackwell (26.7%), CSIC Journals and American Chemical Society both with a share of 11.7% and Oxford University Press (7.1%).

On the other hand, a 19.8% of researchers indicated to have published recently with publishers other than those listed in the survey, and within this percentage 55 respondents reported to have done so in open access journals. Next are the open access initiatives most cited in this respect:

- BioMed Central Journals
  [It is worth mentioning that CSIC Libraries Coordination Unit signed an agreement with BioMed Central whereby publications fees are discounted by half for CSIC researchers who wish to publish in its journals]
- Public Library of Science (PLoS) Journals
- Universidad Complutense de Madrid Publications
- Bentham Open
- Hindawi Publishing Group Journals
- PubMedCentral
- Wellcome Trust Centre for the History of Medicine
- European Geophysical Union Journals
- Pre-prints en arXiv.org

- Universidad Autónoma de Madrid Publications
- Microbial Cell Factories
- Demographic Research
- Geologica Acta
- Information Research
- Beech Tree Publications

Equally, within this group we found researchers that reported to have chosen publishers offering paid open access options or publications that after a relatively short span of time make content freely accessible to all. In this group it is worth citing Publications of National Academy of Science, Magnolia Press, Wiley, and The Company of Biologists (Development).

Copyright issues in a digital environment usually stand high amongst main concerns by researchers vis-à-vis open access. It is interesting to point out that 44% of respondents answered negatively when asked whether they knew the copyright policies of publishers with whom they had recently published.

Perceptions and use of open access content and Digital.CSIC

Degree of knowledge about open access benefits

The most widely known benefits derived from open access were reported to be 1) the enhanced visibility and accessibility of research, with a share of 39.7% of answers; 2) its the value as an alternative model with respect to the subscription-based content (36.8%) and 3) its effectiveness to communicate publicly funded research widely (35.7%). However, other benefits are "little" known, such as open access being a useful means to identify plagiarism (48.6% respondents admitted not to know much about it); being a rapid means to achieve promotion and professional impact (36.1%) and a way to preserve works in a digital format (36%) in the long term.

What are the most used tools to search and identify open access resources?

Broadly speaking, search engines, aggregators, harvesters, repositories and other portals put in place in order to facilitate the discovery open access resources are largely unknown and little used by respondents. On the whole, the most frequently used tools to search scientific information are Google (52.4%), followed by the Web Citation Index (37.9%) and Google Scholar (35.8%). However, it is worth remarking that very popular and widespread tools to discover open access content, such as DRIVER (93.5%), OAISTER (95.3%), OPENDOAR (96.3%) and ROAR (96.7%) are to a much lesser extent familiar, as respondents percentages show. Likewise, initiatives to measure the impact of publications in open access such as CITESEER y CITEBASE are alien to a 92.4% and a 91.8% of surveyed researchers respectively.

This being said, the most largely used tools to locate open access scientific information are PubMed Central (with 26.7% of answers), followed by e-Revistas, CSIC Journals, BioMed Central and PLoS. A smaller group of respondents mentioned the regular use of arXiv, NASA ADS or SPIRES, which makes guess the profile of CSIC scientist who is widening her horizons when searching for resources of interest.

To close this brief overview, a large part of respondents does not use Spanish repositories listed in our survey (RECERCAT, eBuah, RUA, E-prints de la UCM, UPCommons de Barcelona, e-Espacio UNED…) in order to find scholarly resources. DIALNET, a pioneering initiative in Spain when open access started to become widespread is an exception in this regard, as it is frequently used by an 8.8%.

Green road or golden road?

Active participation in the open access movement via self-archiving or publishing in open access journals is also a practice limited to a minority yet. As regards repositories and Digital.CSIC aside, 79.9% of respondents declared that they have never uploaded any work in an open repository. Those that answered positively have mostly done it in arXiv, which collects 42 answers, followed by far by e-LIS, ASTROPH, REPEC, HAL, and some repositories in Spain such as RECERCAT, UPC Commons and UCM e-prints (in these cases probably as researchers belong to one of CSIC centers run in partnership with a University).

Equally, publishing in open access journals is still uncommon, and, as previously noted, PLoS and BioMed Central on the one hand and CSIC Journals on the other hand gathered most mentions. Quality and/or the impact factor of the journal are told to be deciding reasons when selecting the journal. It is interesting to remark that some respondents reported to have paid the open access fees from their pocket in order to give wider availability to their research.

In general, the experience to publish in open access has been satisfying and most of those who did so would publish in an open access publication again due to fast publication, ethical reasons and to gain more visibility and accessibility. The possibility to keep track of visits and downloads was another incentive and was regarded as a more transparent way to analyse the impact of research.

Digital.CSIC two years and a half after its launch

The survey closed with a section about CSIC institutional repository. In spite of the many awareness raising sessions and training targeting the CSIC community already conducted, Digital.CSIC is not yet a project that everybody at CSIC is completely familiar with. In part, this is due to the huge geographical distribution of the research centers and institutes across the whole country and abroad. By way of illustration, 52.1% of respondents ignored that Digital.CSIC is a direct result of the signing of Berlin Declaration by the CSIC Presidency in 2006 and 85.6% did not know the implications of the Open Access policy issued by Comunidad de Madrid concerning CSIC research which is co-funded by the region [it is required to deposit this kind of research on Digital.CSIC not later than 6 months after publication].

33% of surveyed researchers claimed to use Digital.CSIC and usual access points are CSIC official web site or the repository platform. Most out of these (62.9%) regularly visit the repository to search and download works of interest while around half self-archive their research –around 5 works yearly–. This section mirrored to a large extent main findings under the section about open access in general: researchers who self-archive or send their works to their libraries or to the Technical Office to be uploaded are encouraged by gains in visibility and accessibility in the first place, followed by their support to the open access movement and their support to recommendations by CSIC. Just a 15.1% consider preservation issues and a mere 7.2% use the repository to prove their authorship when submitting their works. Those that have never uploaded any research output reported a lack of time and/or incomplete knowledge to do so. Curiously enough, only 14.1% of those that never self-archive commented that copyright concerns were a deciding deterrent.

A repository without services and without rich and quality content cannot guarantee its sustainability in the long run. Thus, we closed our survey by asking researchers their suggestions for improvements. Searching functionalities ranked very high, with many asking for mentions advanced search options, including links to other open access resources topically related alongside links to researchers' CVs. Almost 40% was also supportive of adding digitised material into the collections of the repository. A huge majority proved also very favourable that Digital.CSIC and other CSIC databases containing information about scientific production could get integrated into one system. Other suggestions for services included more training sessions and awareness raising over open access as well as Digital.CSIC assistance over copyright issues.

Section 3: Improvements in Digital.CSIC

The survey results have helped us prioritize our agenda of activities and double check whether there is a synergy between the issues perceived by the repository's Technical Office as most urgent and the upfront needs by its end-users, namely, researchers.

Thus, Digital.CSIC has embarked upon a number of work lines that mark a jump into its consolidation phase. So far the following has been accomplished:

-Migration of Digital.CSIC web site: coinciding with the beginning of the Open Access Week 2010, we have migrated into DSpace 1.6.2, mostly in order to improve functioning of the platform and to benefit from new functionalities such as the SWORD protocol.

-Improvements in the repository's web layout, including new sections and changes in its look and feel: since its launch in 2008 a number of technological innovations have been incorporated by the repository's Technical Office in order to better serve end-users. Amongst these services stand the linking to CSIC Virtual Library through SFX resolver; a tool for news update in Digital.CSIC homepage; the auto-complete functionality when depositing items; "Digital.CSIC on your web" API and so on. Now, we have benefited from the web update in order to add sections dedicated to researchers and librarians, a space for copyright issues etc.

-New statistics module: This set of statistics enriches the existing usage data with general and per CSIC center/institute statistics. Our new statistics module is a home grown development and an innovation of the services provided by DSpace, by offering data with a higher granularity and showing growth trends and impact of available research.

-Promotional and awareness raising campaign: we have launched a content development strategy recently to promote the wider diversification of types of research in the repository –for instance, we are focussing these days on identifying and incorporating more patents, data sets and grey literature- and have produced a number of promotional material such as posters, brochures, comics, new handbooks etc to make Digital.CSIC widely known and encourage researchers' participation further.

[i] *Digital.CSIC Technical Office*, "2009 Annual Report", April 2010
http://digital.csic.es/handle/10261/23383