European Commission

Open Access Status of Journal Articles from ERC Funded Projects

European Research Council

Established by the European Commission

http://erc.europa.eu

Open Access Status of Journal Articles from ERC-Funded Projects

June 2012

This document is a staff working paper of the European Research Council Executive Agency, used for discussion and consultation purposes. It does not purport to represent the views of the European Commission, the ERC Scientific Council or the ERC Executive Agency.

Suggested citation:

European Research Council (2012). Open Access Status of Journal Articles from ERC-Funded Projects, Brussels.

For further information please contact:

Alexis-Michel Mugabushaka
Policy Analyst
ERCEA Unit A1
Alexis-Michel.MUGABUSHAKA@ec.europa.eu

Executive summary

The main objective of this analysis is to estimate the extent to which journal articles from ERC funded projects are available in an open access.

A list of 630 journal articles reported in 88 mid-term project reports formed the basis of this analysis. A web-based search to find openly accessible versions of these articles was undertaken from a computer network with no subscription linked to it. Articles were classified as available in open access if the web-based search yielded a copy of the journal article, an author personal copy or a pre-print.

The results show that 62 % of journal articles from ERC funded projects are available in open access. The share of articles in open access varies across research domains. It is close to 70 % in Life Sciences, 65 % in Physical Sciences and Engineering and nearer 50 % in Social Sciences and Humanities. A comparison with the data on open access status provided by the grant holders in their mid-term reports shows that self-reporting leads to an underestimation of the proportion of open access articles.

The potential policy implications of the results are highlighted in the last section of the report. They include among others, the idea of using soft measures to further encourage open access and seeking the cooperation of the libraries of organisations hosting ERC grantees to ensure that the ERC's Open Access policy is well aligned with institutional policies and practices and does not place an undue burden on researchers.

1. Open access: background

The widespread use of the internet has profoundly changed scientific communication. The internet has given researchers unprecedented opportunities to engage in a dialogue with their colleagues, exchange information on on-going research, discuss interim research results and share research materials. The internet (and advances in information and communication technologies in general) has also decreased tremendously the costs associated with both the production and circulation of research literature.

The open access initiatives were inspired by these changes. They seek to take full advantage of the potential of the internet for wider scholarly communication. The switch to paper-based subscriptions to scientific journals promised both wide dissemination of research results and substantial financial savings for university libraries, but the soaring price of journal subscriptions led many libraries to cancel their journal subscriptions. According to some estimates, between 1975 and 1995 the price of journals increased by 200%-300% above inflation (Dewatripont et al. 2006). Post 1995, the adoption of new digital technologies for production and delivery has not lowered journal prices. On the contrary, they have continued to rise, although less sharply than before. Data from the *Association of Research Libraries*, a group of major American university libraries, shows that between 1995 and 2006 the "serial unit cost" (i.e. the subscription price of scholarly journals) increased by about 15 % in real terms (ARL 2008). A paper commissioned by *Oxford University Press White and Creaser* (2007) analysed the evolution of journal prices in the biomedical and social sciences of 11 major publishers (8 commercial and 3 university presses). Their results show that on average the median price of scholarly journals increased by 70 % in real terms between 2000 and 2006. The price increase varies greatly between publishers, and ranges from 42 % (Oxford Journals) to 104 % (Sage).

Against this background, the pioneers of open access initiatives felt that the potential to maximize the access to research results had not been fully realized. In 2001, the *Open Society Institute* organised a meeting to discuss how open access to "peer-reviewed journal literature" could be optimised. The resulting "Budapest Open Access Initiative" calling for the removal of "the barriers to open access" set the principles and strategies which form the basis of calls for open access today. It was followed two years later by the "Bethesda Statement on Open Access Publishing" and the "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities". These statements catalysed the efforts to promote the dissemination of scientific knowledge for the benefit of the public.

A range of initiatives and projects to promote open access, too numerous to be listed here, mushroomed across Europe:

- *At the policy-makers level*, in its meeting of 22-23 November 2007, the Council of the European Union recognised the strategic importance of open access for European science. The European Commission has led the efforts to promote the wider dissemination of research results by coordinating and harmonising the work of the EU countries. Some Member States have already developed legal frameworks to implement open access. An example is the new *Spanish Law for Science, Technology and Innovation* (2011). It requests publicly funded researchers to publish digital copies of the final version of the contents accepted for publication in scientific journals no later than twelve months after the date of publication.
- *Major universities* have committed to the principles of open access. This is demonstrated by the recently published *Roadmap towards Open Access* by the *League of European Research Universities* (LERU 2011) or the *Recommendations on Open Access* adopted by the Council of the European University Association in March 2008 (EUA 2008). Many universities request that the researchers they employ disseminate as widely as possible the fruits of their research and scholarship. In varying degrees they assist them in this process through the provision of institutional repositories and advice on copyright related issues. Universities libraries play a special role in this context. They often operate the institutional repositories and take care of negotiations with publishing houses on the researchers' behalf.
- The European Heads of Research Councils adopted a statement on open access in 2008 recommending that their members (research funding agencies and large public research organisations) adopt open access policies which seek to shorten the embargo period (EUROHORCS 2008). Currently, a EUROHORCS Working Group, led by the Max Planck Society, is coordinating the exchange of information on best practices and experiences and seeking to develop common or complementary policies on open access for research organisations. A growing number of *funding agencies* also encourage, or expressly request, researchers they fund to make their research results accessible in an open access format.

2. Trends and state of play

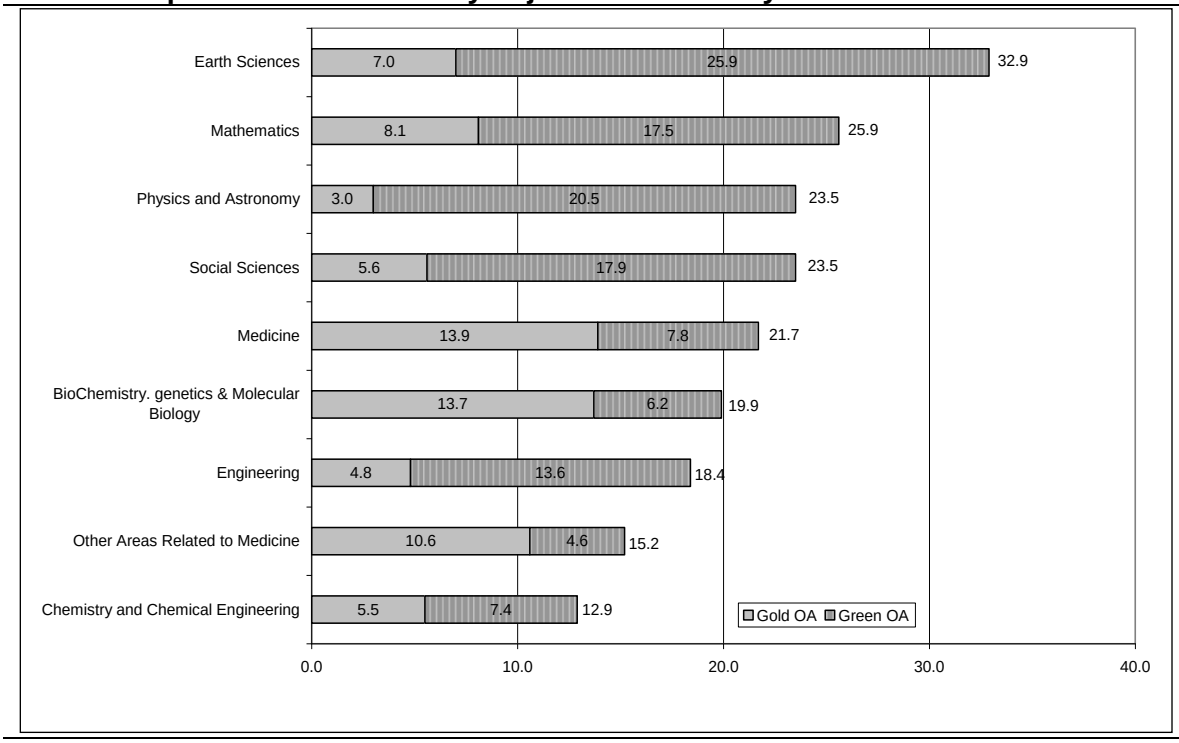
Open access initiatives, both past and present, have had a visible impact on the scholarly communication landscape. This is illustrated by three distinct, but strongly interrelated, trends.

- Open access journals, which were almost unheard of at the beginning of the 1990s, are now a normal phenomenon in the scholarly communication landscape. Lakso et al. (2009, p. 6) estimate that the number of open access journals increased six fold between 2000 and 2009 (from 740 to 4,769) and the number of articles they published increased tenfold. The number of open access journals will continue to grow, as they are increasingly seen as viable business models which are capable of disseminating high quality research. In January 2011 the Nature Publishing Group announced the creation of a new open access journal, *Scientific Reports*, and in June 2011 three main research organisations (Howard Hughes Medical Institute, the Wellcome Trust, and the Max Planck Society) announced the launch of a new open access biomedical and life sciences journal (Wellcome Trust 2011).
- The pre- and post-print directories have established themselves firmly on scholarly communication landscape. ArXiv, an electronic pre-print archive for the physics community which has been active since 1991, has gained wider recognition: it now includes around 700,000 documents and is highly influential in the communities it serves. It is estimated that about 70 % of articles published in 2003 in the *Astrophysical Journal*, had been deposited previously in ArXiv (Kurts et al. 2003). The PubMed Central of the National Institutes of Health (NIH), which archives full text articles (either as post-print or author copies and often after an embargo period), plays a central role in the dissemination of research in the bio-medical sector. The share of articles indexed which are freely accessible rose from about 21% for articles published in 2000 to over 28 % for articles published in 2009. The NIH expects to continue at the same rate and have around 50 % of articles freely accessible and indexed in PubMed in 2012 (*Nature*, 2011).
- Publishers, both for profit and scholarly societies are increasingly offering authors of accepted papers the possibility of having their papers made accessible to all readers by paying an "open access charge". Publishers have embraced this practice to a varying degree. In its open access statements from January 2011, Nature Publishing Group revealed that 80% of its academic journals offer open access options, whilst Oxford University Press lists 90 of the 230 journals it publishes as having open access options

(40 %). The Max Planck Digital Library estimates that 22 % of the 9,500 journals from major publishers offer open access options (quoted by Björk et al. 2009, p. 8).

The study cited above (Björk et al. 2009) tried to reliably assess the open access situation in 2009. The researchers selected a random sample of more than 1,800 articles from peer-reviewed journals across different disciplines published in 2008. They then used web search engines to manually find copies of those articles and record their open access status. Their results show that overall the share of journal articles published in 2008 in an open access format was 20.4 %. This combines the 11.9 % of articles published in non-open access journals for which the researchers could find copies either in repositories or on web sites (Green Open Access) and the 8.5 % which could be found on publishers' sites (Gold Open Access). The study also found that open access of journal articles varied greatly between research disciplines (see Chart 1 below). It was highest in Earth Sciences (33%) and lowest in Chemistry and Chemical Engineering (13 %).

Chart 1 : Open access availability of journal articles by the end of 2009.



Source: Björk et al. 2009, own representation

3. ERC Open Access policy: principles and implementation

The Scientific Council of the ERC has been a strong supporter of open access since its creation. In December 2006, long before any grant was awarded; the Scientific Council

issued its Statement on Open Access in which it committed to making efforts to ensure that research results from projects it funds are available in an open access format (ERC 2006).

The statements refer specifically to the worrying status of access to research articles from scientific journals. The soaring price of some scientific journals led many research organisations to cancel their subscriptions, resulting in a situation in which researchers from public organisations had no access to the results of publicly funded research. For many researchers and librarians, it was obvious that traditional publishing methods were not functioning as well as they should.

The statement was followed a year later by specific guidelines for ERC funded projects (ERC 2007) which required all peer-reviewed publications from ERC funded projects to be made openly accessible shortly after their publication. As an implementing measure, the *Guide for Grant Holders* informs researchers that open access costs can be charged as part of their project costs.

Below we present the results of an analysis which aims to establish baseline data of the open access status of ERC journal articles. The results will be used to monitor the trends and patterns of the open access to journal articles from ERC funded projects. As there are only soft measures to implement the ERC's Open Access policy in place, the results of this analysis can help us to understand the potential impact of any future measures to further encourage open access to the research results of ERC funded projects.

4. Methodology and results of analysis

4.1 Data basis and search methods

The target population of the analysis is the set of journal articles from ERC funded projects. ERC grantees are requested to file an interim project report halfway through their funding period in which they list all publications resulting from their research work during the project. The forms used to collect publication data differentiate between five publications types (journal articles; publication in conference proceedings; books: monographs and edited books; chapters in books and other publication types).

For this analysis, we used **88 mid-term project reports** which were received by the ERC by the end of June 2011. **70** projects were Starting Grants and **18** were Advanced Grants.

Bibliographic data of all listed journal articles was extracted from these project reports (authors, title of article, publication year, journal etc.). After data cleaning (e.g. deleting records erroneously captured as journal articles, eliminating reduplication) a list of **630 journal articles** was compiled and assessed for this analysis.

A two-step approach was used to determine the open access status of a particular journal article. In the first step we used PubMed and ArXiv, the largest discipline specific repositories for Life Sciences and Physics respectively, to search for the articles. For each article PubMed indexes it includes, under its "LinkOut" attributes, the URL which allows you to access the full text. This can be used to check whether PubMed has recorded an open access version of the articles. ArXiV offers a web service interface which can be used to retrieve the full text of the documents it indexes. In the second step all journal articles which were not found in the first iteration were searched for manually using web search engines (mainly Google and Yahoo! Search).

For the purposes of this analysis we define as "open access" the journal articles that an average researcher/reader, without subscription, would easily find on the internet. In the search and assessment of the open access status of the journal articles, the following principles were observed:

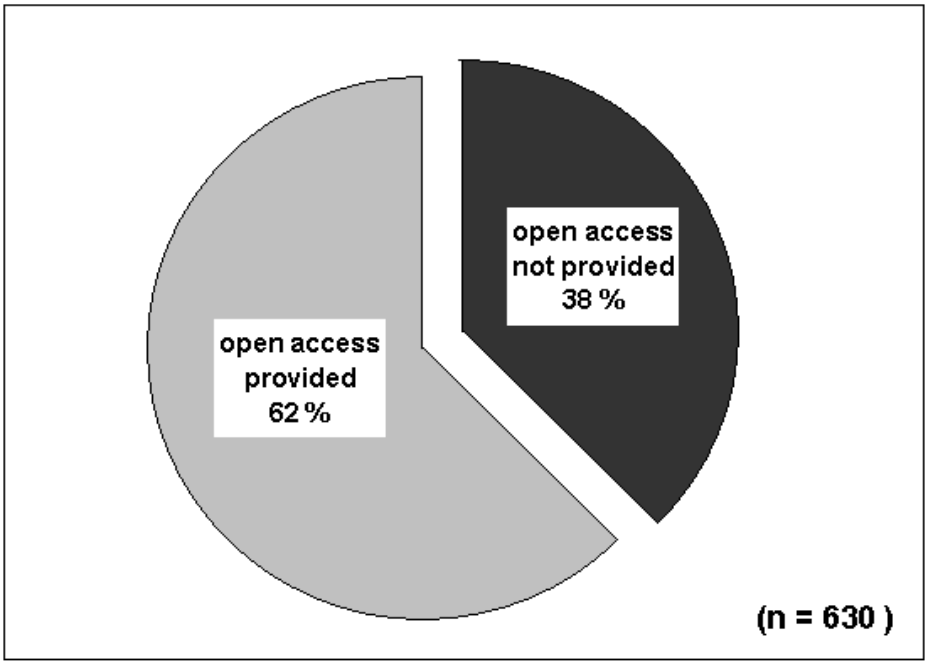
- To avoid mistakenly retrieving articles which are accessible to subscribers through IP-based authentication, all searches were undertaken from a computer network with no journal subscription.
- There were a few cases in which the full text of the articles could be accessed only after registering with the site. In this case, the articles were considered "not open access" (even if the registration was free).
- In some cases, copies of articles were found but were clearly labelled "draft" or "incomplete". In these cases copies of the articles searched were considered not found, and the article as "not open access".
- The search focused on finding any kind of copy of journal articles without distinction being made between author copy of the article and exact copy of the article. In some disciplines researchers publish pre-prints in the form of papers (often without mentioning the journal in which it has been, or will be, published) or in the form of working papers. In such cases, if other bibliographic meta-data (authors and titles) matched the article, it was classified as open access.

The following sections present the results of the analysis.

4.2 Open access in general and by research disciplines

As shown in Chart 2 below, out of the 630 journal articles for which a search was undertaken, for **393 (62 %)** open access was provided and **for 239 (38 %)** no open access versions could be found.

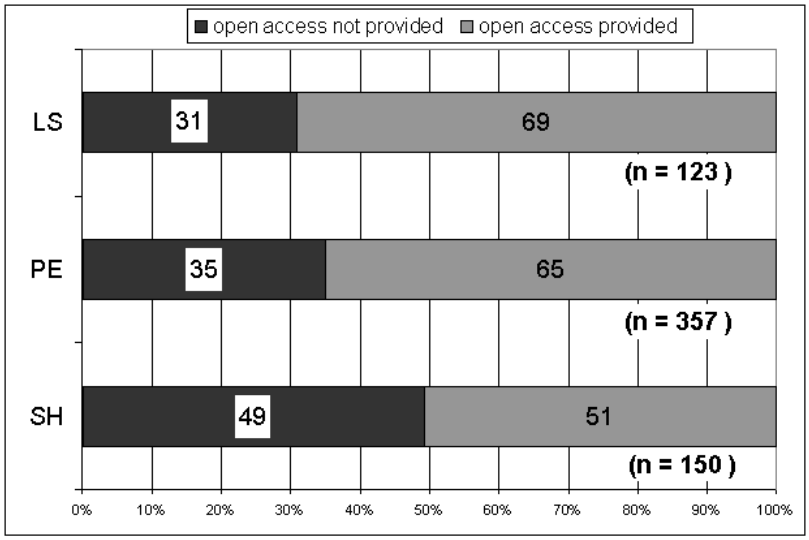
Chart 2 : Open access to journal articles from ERC funded projects



A differentiated picture is provided in Chart 3 below. The share of journal articles for which open access is provided is **69 %** for Life Sciences (LS), **65 %** for Physical Sciences and Engineering (PE) and **51** % in Social Sciences and Humanities (SH).

The relatively high open access availability for Life Sciences and for Physical Sciences and Engineering can be partly explained by the well-developed repositories in these research fields. In fact in the first iteration of the search, 55 out of 123 articles from Life Sciences were found in PubMed **(44 %)** and 101 out of 357 **(28 %)** articles from Physical Sciences and Engineering projects were found in ArXiv.

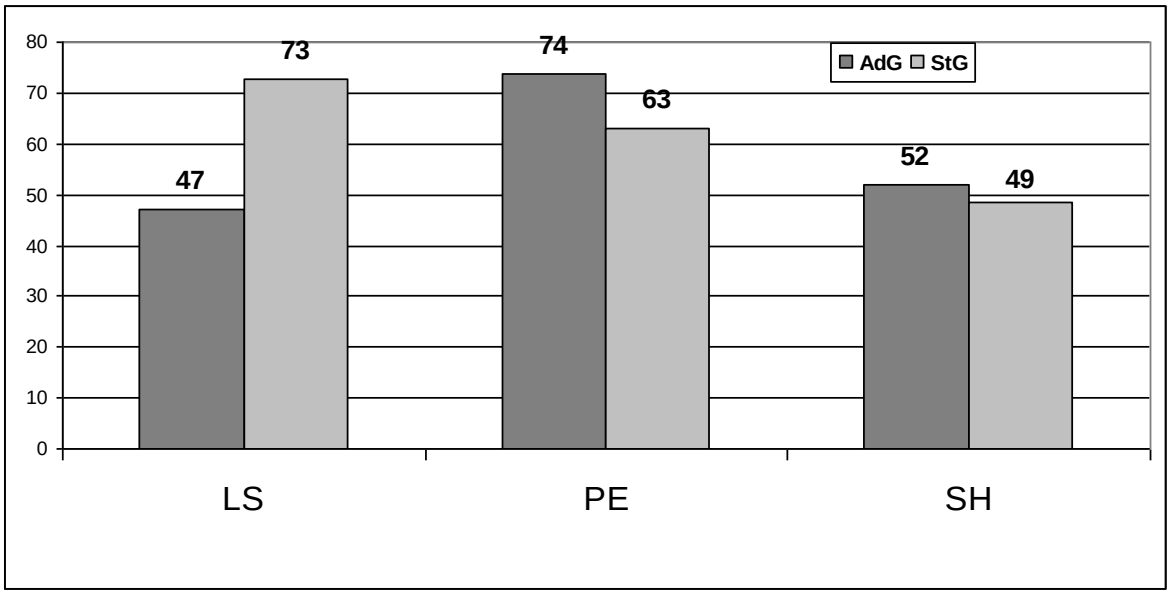
Chart 3 : Open access to journal articles by research domain



When broken down by research domain and funding scheme, the numbers become too small to be meaningfully interpreted (especially for Advanced Grant where only 18 project reports

were analysed). However, as Chart 4 shows, open access availability seems to vary for articles for Starting grantees and from advanced grantees from different research domains.

Chart 4 : Percentage of journal articles available in open access by research domain and funding scheme



For example in Life Sciences **73 %** of journal articles from projects awarded Starting Grants are available in an open access format, whereas the share is **47 %** for Advanced Grants. In Physical Sciences and Engineering the reverse is the case: **74 %** of articles from Advanced grantees are in an open access format while this is the case for only **63 %** of articles from Starting grantees.

4.3 Self-reporting and open access search

In the mid-term reports, the forms used to collect publications from the projects include a question "***is open access (to this publication) provided***?".

For 605 journal articles, the grantees provided this information, indicating that open access is provided for 234 journal articles **(39%)** and not provided for 371 **(61 %).**

When we compare these results, we find that self-reporting grossly underestimates the open access availability for the articles recorded.

This could be explained by the fact that the self-reporting only captures open access status at the time of reporting. This can change: for example when journal publishers have a policy of providing open access after an embargo period has expired. In addition, grantees may not know that open access is provided for the articles reported. This is the case when co-authors (often from other institutions or countries) make journal articles available in their institutional repositories, often in order to comply with open access mandates from their funding bodies.

5. Summary and possible policy implications

The main objective of this analysis was to estimate the extent to which journal articles from ERC funded projects are available in open access. The data basis is a list of 630 journal articles reported in 88 project mid-terms reports. We performed a web search to find openly accessible versions of those articles. The search was undertaken from a computer network which has no subscription linked to it.

The results show an open access availability for 62 % of journal articles from ERC funded projects. The share of articles in open access varies across research domain. It is close to 70 % in Life Sciences, 65 % in Physical Sciences and Engineering and nearer 50 % in Social Sciences and Humanities. A comparison with the data on open access status provided by the Grant Holders in mid-term reports show that self-reporting tends to underestimate the proportion of open access articles.

The limited scope of the analysis did not enable other interesting aspects of open access practices to be evaluated. For example, it would be interesting to know if the articles in open access repositories are predominantly exact copies of journal articles, or are instead pre-prints or personal author copies. Equally interesting would be to see if articles from the journals of commercial publishers for which open access is provided are made available by the publishers (in which case open access fees may have been charged), or are only available through open repositories. Those aspects should be taken into consideration in the future monitoring of the ERC Open Access policy.

Overall, the results show that the availability of articles in an open access format is relatively high for ERC funded projects. Obviously, the ERC Open Access policy, even if it can be said to have played a role, is not the only driving factor behind this. As was made clear in sections 1 and 2 of the analysis, there is a strong trend towards open access and various actors are engaged in initiatives which, taken together, make this possible. Almost all articles in this analysis acknowledge multiple sources of funding. It is possible that they were made available in an open access format to comply with the open access mandate of other funding bodies. In addition, the important role played by institutional repositories (and university libraries in general) should be taken into account.

In light of the results of the analysis, the following potential policy implications should be brought to discussions:

- Although the ERC has not taken strong measures to implement and enforce its Open Access policy, the ERC can maintain and increase open access availability by ***simply further encouraging*** the researchers it funds. Strong measures may not be necessary (and may even be counter-productive if perceived by the grantees to be

too heavy handed). Examples of "soft measures" could be to inform ERC grantees about new developments in the open access activities of the ERC on a regular basis.

- Many universities who host ERC grantees have well developed Open Access policies. It is important that the ERC's implementation of open access is aligned to and compatible with them to minimize the burden on researchers. ***Universities libraries*** play a crucial role in making the publications of the research community they serve available in open access formats. The ERC could start discussions with the libraries of organisations which host several ERC grantees in order to understand how their open access activities, policies and practices could be better coordinated with the ERC's own Open Access policy.
- In some research fields there are ***well-established repositories*** which have been instrumental in the development of open access. PubMed Central for the Life Sciences and ArXiv in the Physical Sciences are particularly notable. A carefully designed cooperation between the ERC and these repositories could increase the share of ERC access publications they contain.
    - For example, PubMed Central works together with funding organisations and publishers to archive on the researchers' behalf the final versions of papers either immediately, or after an embargo period.
    - ArXiv is currently working with selected research organizations to fund and enhance its operations. A possible collaboration area is to index the acknowledgments texts from the documents provided. This would make it easier to track ERC funded research papers.

References

ARL (2008). ARL Statistics 2006–2007. Association of Research Libraries, Washington, DC (Compiled and Edited by Martha Kyrillidou Les Bland)

Björk BC et al. (2010) Open access to the scientific journal literature: situation 2009.PLoS One. 2010 Jun 23;5 (6):e11273.

Dewatripont, M. et al. (2006) Study on the Economic and Technical Evolution of the Scientific Publication Market in Europe. Final report, January 2006. Commissioned by Directorate-General for Research, European Commission. Brussels.

ERC (2006). ERC Scientific Council Statement on Open Access http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf

ERC (2007). ERC Scientific Council Guidelines for Open Access. http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf

EUA (2008). Recommendations from the EUA Working Group on Open Access adopted by the EUA Council on 26th of March 2008 (University of Barcelona, Spain). European University Association, Brussels

Kurtz M. et al. (2005). The Effect of Use and Access on Citations. https://www.cfa.harvard.edu/~kurtz/kurtz-effect.pdf

Laakso M et al. (2011). The development of open access journal publishing from 1993 to 2009.PLoS One. 2011;6 (6):e20961.

LERU (2011). The LERU Roadmap towards Open Access. League of European Research Universities, Brussels.

Nature 2011, Vol 476, p. 11 (see also provided link on the website).

Wellcome Trust 2011. Leading research organisations announce top-tier, open access journal for biomedical and life sciences. Press release 27 June 2011 http://www.wellcome.ac.uk/News/Media-office/Press-releases/2011/WTVM051897.htm

White and Creaser (2007). Trends in Scholarly Journal Prices 2000-2006.