

POSTER PRESENTATION

JOURNALS' POLICIES OF STORAGE AND REUSE OF RAW RESEARCH DATA AND THEIR IMPACT IN FIVE SCIENTIFIC AREAS

Aleixandre-Benavent, R. (1); Ferrer-Sapena, A. (2); Vidal-Infer, A (3); Alonso-Arroyo, A. (3); Sánchez Pérez, E.A. (4), Peset, F. (2).

1. Ingenio (CSIC- Universitat Politècnica de València). UISYS (CSIC – University of Valencia). Spain.
2. DCADHA. Universitat Politècnica de València. Spain.
3. Department of the History of Science and Information Science. School of Medicine and Dentistry. University of Valencia. Spain.
4. Instituto Universitario de Matemática Pura y Aplicada. Spain.

Correspondence:

Aleixandre-Benavent, R.

Instituto de Gestión de la Información y el Conocimiento-Ingenio (CSIC-Universitat Politècnica de València). Unidad de Información e Investigación Social y Sanitaria-UISYS (CSIC- Universitat de València)

Plaza Cisneros, 4

46003-Valencia

Spain

Rafael.Aleixandre@uv.es

BACKGROUND

In order to enhance the fast dissemination of information and exchange of knowledge, it is required to make available for researchers both scientific literature and research raw data. The Open Access movement has allowed accessing the full text of many articles, thereby facilitating universal access to scientific advances. A further step is to open the raw data that support research to be shared and reused (data-sharing). Many scientific manuscripts consist of full text plus supplementary data to support the presented information. These additional data are too extensive to include them in the full text of the article, but some researchers may require these data to ensure that the experiment has been correctly conducted (Murray-Rust, 2008). Moreover, these data can also be useful, if reused, to supplement other work, make comparisons or generate new hypotheses. The purpose of this study is to analyse open-data policies concerning the availability of raw data in journals included in 5 subject categories (SC) of the Journal Citation Reports (JCR) and the relation of these policies with the impact factor and the quartile.

METHODS

We reviewed the websites of 451 journals included in the following JCR subject categories: Food Science and Technology (FS&T) (124 journals), Pediatrics (115 journals), Information Science & Library Science (IS&LS) (85 journals), Substance Abuse (39 journals) and Dentistry (88 journals). For each journal we documented the rules related to public availability of sharing research data as stated in the instructions to authors on the journal's website. Data achieved for each journal included the following variables: a) Statement on deposit of complementary material attached with the manuscript; b) Acceptance of reuse of data included as complementary material; c) Acceptance of storage in thematic or institutional repositories. We also took into account the distribution of journals by impact factor and quartile in the 2014 JCR edition.

RESULTS

Table 1 shows the results obtained when the variable "Statement of complementary material" is accepted. The subject category with the highest percentage of acceptance is Pediatrics (81.7%), followed by FS&T (59%), IS&LS (50.6%) and Substance abuse (41.2%). In Dentistry journals, the percentage of acceptance is 19.3%. Statistical significant differences have been found in this variable through the four quartiles, with a higher acceptance in the upper quartiles (Q1 and Q2) and lower in Q3 and Q4. The highest values for journals ranked in the first quartile correspond to FS&T (24.2%) and Pediatrics (20.9%).

Table 1. Acceptance of the variable "Statement of complementary material"

Quartile in JCR	Dentistry	FS&T	IS&LS	Pediatrics	Substance abuse
1	10 (11.4%)	30 (24.2%)	16 (18.8%)	24 (20.9%)	4 (10.3%)
2	4 (4.5%)	24 (19.5%)	13 (15.3%)	25 (21.7%)	4 (10.3%)
3	3 (3.4%)	16 (12.9%)	7 (8.2%)	26 (23.6%)	4 (10.3%)
4	0	3 (2.4%)	7 (8.2%)	19 (16.5%)	4 (10.3%)
Total*	17 (19,3%)	73 (59%)	43 (50.6%)	94 (81.7%)	16 (41.2%)

*Number of journals and percentage on total number of journals in each subject category

The variable "Reuse" has presented the following results (Table 2): 73.4% of journals included in FS&T, 64.7% of journals in IS&LS, 56.4% of journals in Substance abuse, 30.4% in Pediatrics and only 19.3% in journals of Dentistry. The acceptance is higher in the upper quartiles. The highest values for journals ranked in the first quartile correspond to FS&T (25%), IS&LS (21.2%) and Substance abuse (20.5%).

Table 2. Acceptance of the variable "Reuse"

Quartile in JCR	Dentistry	FS&T	IS&LS	Pediatrics	Substance abuse
1	9 (10.2%)	31 (25%)	18 (21.2%)	6 (5.2%)	8 (20.5%)
2	4 (4.5%)	27 (21.7%)	19 (22.3%)	9 (7.8%)	6 (15.4%)
3	1 (1.1%)	25 (20.2%)	12 (14.1%)	13 (11.3%)	5 (12.8%)

4	3 (3.4%)	8 (6.5%)	6 (7.1%)	7 (6.1%)	3 (7.7%)
Total*	17 (19,3%)	91 (73.4%)	55 (64.7%)	35 (30.4%)	22 (56.4%)

The next variable, "Acceptance of storage in thematic or institutional repositories" (Table 3) presents results near 65% in three subject categories: IS&LS, Pediatrics and FS&T, near 60% for Substance abuse and only 37.5% for Dentistry. Here again, the acceptance is higher in the upper quartiles.

Table 3. Acceptance of "Storage in thematic or institutional repositories"

Quartile in JCR	Dentistry	FS&T	IS&LS	Pediatrics	Substance abuse
1	14 (15.9%)	30 (24.2%)	20 (23.5%)	20 (17.4%)	8 (20.5%)
2	9 (10.2%)	25 (20.2%)	16 (18.8%)	21 (18.38%)	6 (15.4%)
3	5 (5.7%)	20 (16.1%)	12 (14.1%)	22 (19.1%)	6 (15.4%)
4	5 (5.7%)	6 (4.8%)	9 (10.6%)	14 (12.1%)	3 (7.7%)
Total*	33 (37.5%)	81 (65.4%)	57 (67%)	77 (67%)	23 (59%)

CONCLUSIONS

Journals with a better position in the impact factor ranking have an open policy about scientific data. This behaviour has been also observed in previous works analysing policies of open research data in high impact factor journals in several subject areas (Alsheikh-Ali et al, 2011; Aleixandre et al, 2014 and 2016). This study has some limitations that should be considered: first, we analysed journals only in five subject categories included in the JCR; second, after the analysis of these policies we do not know the real frequency of data sharing among scientists. Future research should analyse policies about open research data in journals included in other subject categories of JCR. It will be also noteworthy to investigate the rate of papers in each subject category that really provide data for reuse, as well as the nature of these data.

ACKNOWLEDGEMENTS

Supported by the National R+D+I of the Ministry of Economy and Competitiveness of the Spanish Government (projects: CSO2012-39632-C02-01 and CSO2015-65594-C2-2-R) and 2015-Networks of Excellence call (project CSO2015-71867-REDT).

REFERENCES

1. Aleixandre-Benavent, R. et al. Public availability of published research data in substance abuse journals. *International Journal of Drug Policy*. 2014; 25: 1143–1146.
2. Aleixandre Benavent, R. et al. Correlation between impact factor and public availability of published research data in Information Science and Library Science journals. *Scientometrics* 2016; 107(1):1-13.
3. Alsheikh-Ali, A.A., Qureshi, W., Al-Mallah, M.H. & Ioannidis, J.P.A. (2011) Public Availability of Published Research Data in High-Impact Journals. *PLoS ONE* 6(9), e24357.
4. Murray-Rust, P. (2008). Open data in science. Accessed 23/02/2015. Retrieved from Nature Proceedings: <http://hdl.handle.net/10101/npre.2008.1526.1> (Archived by WebCite® at <http://www.webcitation.org/6dfxgEt8X>).