

Data from “Linking full-text grey literature to underlying research and post-publication data: An Enhanced Publications Project 2011-2012¹”

<https://doi.org/10.17026/dans-zca-t9k5>

Dominic Farace, GreyNet International

ORCID 0000-0003-2561-3631

Project Administration/ Dataset Creator / Draft Text

Jerry Frantzen, GreyNet International

ORCID 0000-0002-3405-7078

Visualization

Plato L. Smith II, University of Florida

ORCID 0000-0003-1814-0151

Review & Editing

Abstract

The data collected is based on the responses of fifty authors/researchers, who have presented full-text papers published in the International Conference Series on Grey Literature. The dataset consists of ten questions, three of which are open-ended. The data contains the number and percentages of responses to questions dealing with the author’s own empirical research data, its current availability, and the author’s willingness to archive the data and make it openly accessible. The data in this study was collected and computer analyzed via Survey Monkey, where it remains stored along with a copy in Excel housed in the DANS Easy Archive. The data from the study was collected in 2011 within the framework of an enhanced publications project². Since then, emphasis on data preservation, open access, and the reuse of data has increased exponentially. This data from six years ago offers a valuable baseline that would allow for a follow-up questionnaire in part or whole. Hereby, one would stand to gain insight into changing attitudes and practices within GreyNet’s research community and renders potential (re)use by other professional research communities.

Keywords: Enhanced Publication; Linking Data; Attitude Change; Data Sharing

Subject Area: Information Science

Methods

● Steps

In producing the intended dataset a questionnaire was drawn-up and implemented using the freeware ‘Survey Monkey’³. The questionnaire contains ten queries set off in pairs each of which is preceded by a subheading. These subheadings are considered relevant in achieving informed responses. The three open-ended questions allowed for specific and detailed responses that were later categorized in order to facilitate further analysis of the data. Responses to the open-ended questions contained the names and email addresses of the respondents. Seen as source data, this allowed for insight into the geographic region of the respondents, their most recent affiliation with the GL-Conference Series, and their gender.

● Sampling strategy

The population of the survey was selected from among the 286 authors and co-authors in the International Conference Series on Grey Literature. It was decided that only first authors would receive the questionnaire, which narrowed the potential population of the survey to 162 authors of which only 95 were actually sent the online questionnaire. The reason the other 67 first authors were not included in the final survey population was due to a number of factors such as no current email address, retired, deceased, etc. The final results are based on the response of 50 of the 95 survey recipients, which amounts to roughly a 53% response rate.

● Quality Control

While there was no specified control on the data, one may assume that if a respondent is willing to provide his/her name and email address some level of guarantee of the data could be expected.

Dataset Description

| | |
|---------------------------------|---------------------------------|
| File name: | Farace et al.xlsx |
| Format name and version: | Excel 2010 |
| Creation dates: | From 2011-06-22 to 2011-09-28 |
| Language: | English |
| License: | CC0 Waiver - no rights reserved |
| Archive name: | DANS EASY Archive |
| Publication date: | 2013-02-14 |

Potential Reuse

The data derived from this questionnaire allows for reuse not only by grey literature communities such as GreyNet but also by other diverse information communities exploring a coherent policy for the collection and sharing of research data. Even six years since the publication of the data and its original analysis, the data still allows for further use, other interpretation, and would serve well in a comparative study. Even though the data is openly accessible and bears a persistent identifier, its potential for reuse would increase with citations and references made possible by way of this data paper. The reuse of the data would further serve to validate its content by demonstrating its adherence to the FAIR-data Principles⁴. One specific reuse of the data will serve as part of a case study in a workshop on Data Papers demonstrating the value of this new document type of grey literature for researchers and data management librarians alike.

Nevertheless, as with long-tail research, limitations of this data can be its size. Fifty respondents do not actually allow for an accepted expression of results in percentages. There is also a potential barrier to the original data, which was compiled and stored via Survey Monkey. This could prove an obstacle, even though a copy of the dataset is openly accessible in the DANS EASY Archive⁵, which provides a Data Seal of Approval, DSA⁶.

References

- ¹ Farace, D., [et al] 2012 Linking full-text grey literature to underlying research and post-publication data: An Enhanced Publications Project 2011-2012. Thirteenth International Conference on Grey Literature. http://www.opengrey.eu/data/70/01/53/GL13_Farace_et_al_2012_Conference_Preprint.pdf
- ² GreyNet's Enhanced Publications Project: Tracking and Backtracking Data, 2012 <http://greyguide.isti.cnr.it/index.php/49-gl14/gl14-slide-share/416-gl14-farace-et-al-2>
- ³ SurveyMonkey, A Survey Platform <https://www.surveymonkey.com/home/>
- ⁴ FAIR-Data Principles <https://www.force11.org/group/fairgroup/fairprinciples>
- ⁵ DANS EASY Archive <https://easy.dans.knaw.nl/ui/home>
- ⁶ DSA, Data Seal of Approval <https://www.datasealofapproval.org/en/>