

Data Papers Project

2011 Enhanced Publications 2017



Attitudes of Authors on Data Sharing



Acquisition Linking and Published Data



Public Awareness Data Citations and Reuse



Data Paper Template

Data Paper TGJ Publication

Data Paper Workshop



Recent Acquisitions in the DANS Data Archive:

Aloia and Naughton

Giannini and Molino

Prost and Schöpfel

Vaska and Vaska

<https://doi.org/10.17026/dans-z28-x27y>

<https://doi.org/10.17026/dans-za8-89yb>

<https://doi.org/10.17026/dans-xwt-w73f>

<https://doi.org/10.17026/dans-zw8-8ksd>

Data from "The GreyNet Report: Understanding the Challenges of Finding Grey Literature"
<https://doi.org/10.17026/dans-z28-x27y>

Danielle Aloia, New York Medical College
 Robin Naughton, New York Academy of Medicine

Abstract
 The data in this case consist of a bibliography of 207 full-text papers analyzed for a literature review. The data was collected using PubMed and the full-text papers were analyzed using the search feature in Adobe PDF to count word hits. All papers were scanned manually for search strategy reporting methods. The full-text papers provide a more robust analysis of search strategies and tools used in systematic review research than the abstract. This data can be helpful to the future when trying to understand the search strategies for finding grey literature for systematic reviews and how to report search results in a paper.

Keywords: Data analysis; Systematic review; Literature reviews; Data Sharing

Subject Area: Information Science

Methods

- Step 1: Published search ("grey literature" OR "grey literature") (Limits: English, review, within 5 years) This search yielded over 1500 results. These results were pared down to over 400. Approximately, 100 were excluded for various reasons.
 - Many were focused on grey literature search strategies for systematic reviews.
 - Twelve didn't report how or where they searched, so we were unable to determine the search methodology.
 - Twelve papers indicated they just searched the traditional databases PubMed, Web of Science to locate grey literature.
 - So we were unable to obtain the full text of the article.
 - Five of the articles only reported on the selection bias of not including grey literature in the systematic review, and
 - Three papers mentioned explicitly that they excluded including grey literature in their searches.
- Sampling strategy: The citation results were analyzed using the full-text article. It was found that the abstract was not robust enough to determine the search strategies reported in the paper.
- Quality Control: Full-text peer-reviewed papers were chosen from one citation database, PubMed.

Data from "Grey Literature Citations in the Age of Digital Repositories and Open Access"
<https://doi.org/10.17026/dans-za8-89yb>

Shiva Giannini, ISTIC CNR Italy
 Anna Molino, ISTIC CNR Italy

Abstract
 The data collected is based on a sample corpus built on all the bibliographical references of articles in four journals over the years 2012-2014 by the proceedings of the International Conference held in 2012 and 2014. The full-text paper, presented at the International Conference on Grey Literature, measures grey literature in the years 2012, 2013 and 2014 and describes the features of GI documents within the areas Computational Linguistics and Computer Science and Engineering. The data from the study was collected and analyzed in 2016. The original information was extracted directly from the primary sources, i.e. the bibliographical references of the articles published in the selected journals and proceedings. The dataset consists of all the analyzed bibliographical references of the abstract journals and proceedings, accompanied by some informative data provided to perform calculations and precise statistical information. It contains the number of bibliographic references and the number and percentage of usage of "grey" references and "grey" document types in the considered time-span. The data derived from this sample are suitable for different types of review and research for anyone interested in citation analysis.

Keywords: Bibliographic references, Citation analysis, Models of scientific communication

Subject Area: Information Science

Methods

- Steps: The dataset has been produced by merging the tables used for data analysis and the tables used for data processing into a single Excel file. It is composed by 11 spreadsheets: the first one (JACM, EIRASIP, CL, EIRAC, EIRL, EIRI) contains the bibliographical references accompanied by the informative context: year, issue number, kind of document - Grey (GI) or Published (P), document type and standardized document type. The remaining spreadsheets contain the tables with measurements of the frequency of GI usage, the frequency of GI use and the intensity of GI use accompanied by graphs (frequency, intensity). The spreadsheet called "GI_label" reports the overall distribution of GI by document type while the spreadsheet called "Journal_GI and Proceedings_GI" reports the distribution of GI documents for each journal and proceeding. The last one includes the graphs measuring the intensity of use of the individual types of document by each resource analyzed (GI, GI/P).
- Sampling strategy: The sample data was selected from journals with an impact factor (IF) over the last three years, indexed by Scopus Citation Database and ISI Web of Science, and from top proceedings of international conferences held in 2012 and 2014. The chosen journals are all indexed under the EIRAC subject category "Computer Science (CS)" except for the *IEEE Journal on Advances in Signal Processing (JASIS)*, which is under the subject category "Engineering, Electrical & Electronic" (EIRL). ACM Transactions on Information Theory (ACM TIT) is under the sub-category "Information Systems", *Computational Linguistics (CL)* is under the sub-category "Artificial Intelligence" and "Interdisciplinary Applications", *Language Resources and Evaluation (LREC)* is under the sub-category "Interdisciplinary Applications". Scopus citation database shows the journals CL and LREC in areas related also to the Humanities and Social Sciences "Language and Linguistics".

Data from "Grey communities: An empirical study on databases and repositories"
<https://doi.org/10.17026/dans-xwt-w73f>

Heather Prost, BECT CARES
 OCEC 0000-0002-7182-2793
 Dataset Creator

Janeen Schöpfel, University of Lille
 OCEC 0000-0002-4000-8074
 Project manager

Abstract
 The study explores grey communities outside the Grey Literature Network Service (GreyNet) and identifies potential members for GreyNet. GreyNet can be compared to a network society specialized in grey literature as a particular field of library and information sciences (LIS). Its relevance is related to its capacity to embrace the terminology and definition of grey literature as LIS research and publication, and its impact and outreach can be assessed through the proportion of reports dealing with grey literature and connected with GreyNet. From five databases (Web of Science, Scopus, LITA, Proquest and Francis) and from open repositories we selected 2,480 reports on grey literature published between 2000 and 2012 by 5,499 authors. Publishing features, preferred journals and the number of publications per author are described for the whole sample. For a subsample of 423 authors strongly committed to grey literature, we present data on geographic origin, place of work, scientific domain and profession. Researchers discuss the characteristics of grey communities in and outside of GreyNet and suggest strategies for the further development of the network.

Keywords: Grey literature, academic community, academic network, social networks

Subject Area: Library & Information Science

Methods

- Steps: The first step was a search for publications on grey literature in selected scientific databases. The search was conducted in March and April 2012 in five databases (ISI, SCOPUS, Web of Science, PASCAL, FRANCIS). We then conducted the same search in different open access directories and search engines (DOI, OpenDOAR, ROAD, EOL, OAJ) but only the results from ISI were analyzed and relevant while the other tools were not specific enough to be reliable to the time period, to search on fields except full-text and/or to published documents. Therefore, we only added the references from the ISI directory. The references from all sources were uploaded to a unique database. The final database contains 2,480 references and allows for the analysis:
 - Study on the publication pattern: the study was conducted in order to know more about these references on grey literature - document types, publication year, preferred journal titles.

Data from "Looking for information that is not easy to find: An inventory of LibGuides in Canadian Post-Secondary Institutions Devoted to Grey Literature"
<https://doi.org/10.17026/dans-zw8-8ksd>

Marcus Vaska, Alberta Health Services, Knowledge Resource Service
 Project Administration/Dataset Creator/Visualization/Design/Text/Review & Editing

Rosetta Vaska, University of Calgary (retired)
 Project Administration/Visualization/Design/Text/Review & Editing

Abstract
 In order to obtain a representative sample of the use of grey literature in LibGuides across Canadian post-secondary institutions, an environmental scan was undertaken, identifying 17 colleges or universities where grey literature resources were directly mentioned and included alongside academic databases. After reviewing the LibGuides within each post-secondary institution, 32 library staff (librarians and information specialists) were identified. A brief online survey was sent to each of the 32 library staff members, to uncover how students and researchers use grey literature, and perhaps, most importantly, to verify from the participant responses whether or not sections of existing LibGuides have been devoted to including the grey literature in information seeking pursuits. None of the 17 institutions polled participated in the survey, yielding a response rate of 52.9%. All respondents confirmed that grey literature was mentioned in the research guides/subject guides/LibGuides used within their institutions. This data set is affiliated with CL 16, the 16th International Library Association Conference, which was held at the New York Academy of Medicine from November 28-29, 2016.

Keywords: LibGuides; grey literature; post-secondary institutions; Canada

Subject Area: Life Sciences, medicine and health care; Health sciences; Humanities; Language and literature studies; Political science

Methods

- Steps: Following the environmental scan of Canadian post-secondary institutions that include and/or dedicate their LibGuides to the study and pursuit of grey literature, the following methodology was undertaken:
 - 32 library staff (librarians and information specialists) from 17 institutions were identified.
 - Tool used: <https://forms.jku.ac.at/>, a free assessment survey tool used to generate a brief online survey to each of the 32 library staff members.
 - The purpose of the survey was to:
 - Uncover how students and researchers use grey literature.
 - To verify from participant responses whether or not sections of existing LibGuides have been devoted to including grey literature in information-seeking pursuits.
- Sampling strategy: None of the 17 institutions polled participated in the survey, yielding a response rate of 52.9%. In such instances, more than one respondent from the same institution replied, however, in such cases, only one respondent per institution was considered. Further, some respondents chose to

In line with the Pisa Declaration on Policy Development for Grey Literature Resources