Citation Analysis and Grey Literature:
Stakeholders in the Grey Circuit

J. Schöpfel and C. Stock, INIST/CNRS, France
D.J. Farace, GreyNet, Netherlands
J. Frantzen, Boer & Croon, Netherlands

Introduction

The goal of this research project is to learn more about the expanding field of information known as 'grey literature'. Citation analysis allows one the possibility to follow the work of the authors in the GL-Conference series, as well as authors whom they have cited both in grey and commercial publishing. A further goal of this paper is to examine the value of the GL- Conference Proceedings for research in the field of grey literature by examining among other things, its impact on the work of contributing authors. This study was sparked by a general interest in citation analysis as an instrument used in various sciences including information science. And, in particular by a paper presented at GL5 in which grey literature as a main channel for publication by an International Marine Scientific Advisory Body was investigated (MacDonald et al., 2004).

The stakeholders in this study are not first and foremost the authors who produce a document type, which falls within the category of grey literature. Instead this study is directed to those authors who focus their research and writing on the topic of grey literature, these are referred to as the meta-authors on grey literature. They have entered the field of information studies and focussed on a particular area, they entered in small numbers and in the past decade they have become contending stakeholders.

In the first five volumes of the GL Conference Proceedings a total of 139 conference papers were published, 30 of which did not contain references. Bibliographic data shows that 152 authors/co-authors, which is a full sample, were responsible for the content of these papers. 108 are 1st authors, 78 of whom are sole authors. 30 of the authors (>20%) have published more than once in the conference series and 6 of whom (>.05%) have published three or more times in this series. It is important that those authors who published within the series are also published outside the series – indicating that their clout extends beyond the grey circuit. In this way, their sphere of influence covers their entire work and reflects well on the conference series as a whole. The direction until now has been grey to commercially published e.g. a conference paper is later published as a journal article; however, with the advent of OAI, Open Archive Initiative, this direction could now even allow for reversal.

Procedure and Method

This research project is divided into three stages:

1. The first stage is the compilation of data from the first 5 volumes in the GL Conference Proceedings (1994-2004) in which the parameters are defined, software chosen, the fields described, a division of labour among the project workers, creation of a database, control of the input, and an informal documentation of the process by way of a project log.
2. The second stage is a literature search on citation analysis previously carried out in the field of grey literature and to some extent beyond.
3. The third stage is the analysis of the primary data from Stage 1 within the framework of the findings from the literature research in Stage 2, as well as a reflection on the problems encountered and the observations made during the first two stages. This final stage also includes the writing of the paper and the creation of a PowerPoint presentation.

Data Collection and Processing

The printed and published volumes of the GL proceedings served as the primary source for our data collection. Since there was no electronic version of the earlier volumes available, the data input had to be done manually. The fact that there was no standard applied to citations from the onset of this series meant that a manual input would have been required even if the e-papers were available. MS Access was the most used database software among the four project workers; therefore this was the most suitable choice. The definition and description of the fields in the database, which would constitute a record was the next task ahead. The final number of fields was 19 and the final number of records was 1374, which brought the total number of possible data entries to 26106. This is also counting the number of fields left blank. For example up to four fields were available for Cited Sources i.e. 1st through 4th author. When there was only one Cited Source the other three fields were left blank. When the comparative analysis was carried out using the index from the Annotated Bibliography on the topic of Grey Literature (4th ed., 2000) the possible number of entries in the Citation Database increased to 31602.

The table below provides an overview of the record fields and a description of each field. The description was important for the data input since four project workers were entering data simultaneously on multiple locations. This however did not detract from the fact that fields would later have to be expanded and data entry corrected and standardized.