Grey communities

An empirical study on databases and repositories

Hélène Prost (corresponding author)
INIST-CNRS
2, Allée du Parc de Brabois
CS 10310
F-54519 Vandœuvre-lès-Nancy
helene.prost@inist.fr
Tel ++33 (0)383 50 47 12

Joachim Schöpfel
Charles de Gaulle University Lille 3
GERiiCO - ANRT
BP 60149
F-59653 Villeneuve d'Ascq Cedex
joachim.schopfel@univ-lille3.fr
Tel ++33 (0)320 41 69 38

Abstract

The paper explores grey communities outside the Grey Literature Network Service (GreyNet) and identifies potential members for GreyNet. GreyNet can be compared to a Learned Society specialised in grey literature as a particular field of library and information sciences (LIS). Its relevance is related to its capacity to enforce the terminology and definition of grey literature in LIS research and publications, and its impact and outreach can be assessed through the proportion of experts dealing with grey literature and connected with GreyNet. From five databases (Web of Science, Scopus, LISTA, Pascal and Francis) and from open repositories we selected 2,440 papers on grey literature published between 2000 and 2012 by 5,490 authors. Publishing features, preferred journals and the number of publications per author are described for the whole sample. For a subsample of 433 authors strongly committed to grey literature, we present data on geographic origins, place of work, scientific domain and profession. We discuss the characteristics of grey communities in and outside of GreyNet and suggest strategies for the further development of the network.

Introduction

In 2012, the Grey Literature Network Service (GreyNet) celebrates its 20th anniversary, with nearly 300 contributors from 30 different countries. In recent years, GreyNet has directed its activities towards open access through the launching of OpenGrey, and the creation of the GreyNet LinkedIn group marked its entry into social networks. The activities of GreyNet, such as workshops, summer schools, curriculum development, discussion list, publications and conferences, contributed to the creation and development of a subject community by means of shared terminology, tools, events, experiences and topics.

GreyNet can be compared to a Learned Society specialised in grey literature as a particular field of library and information sciences (LIS). In fact, it is different from an academic group insofar as it is not confined to LIS but also reaches out to include information professionals and scientists from other disciplines (social sciences, computer sciences, law, economics...). GreyNet is more a kind of special interest group that recruits experts from the field of grey literature in order "to facilitate dialogue, research, and communication between persons and organisations (and to) to identify and distribute information on and about grey literature in networked environments" (GreyNet web site).

The relevance of GreyNet is related to its capacity to enforce the terminology and definition of grey literature in LIS research and publications, and its impact and outreach can be assessed through the proportion of experts dealing with grey literature and connected with GreyNet. In other words, if we want to evaluate the success of GreyNet as a community-creating structure and its potential for future development, we need to know the degree of GreyNet to attain and aggregate all (or at least a significant number of) scientists, academics and information professionals interested in the field of grey literature and contributing to its knowledge.

Our definition of community is pragmatic and follows the sociological approach to science (Kuhn 1962, Latour & Woolgar 1979). We consider a scientific and/or professional community as a social group with interaction and communication, common practice, identity and values, and shared interests, definitions and language (see Callon 1989, Schrecker 2006 or Paganelli 2012).

Our study builds on three other papers on scientific and professional members of GreyNet. In 2005, we analysed the citations of the first five conferences on grey literature (Schöpfel et al., 2005). This first paper defined the "stakeholders" as "those authors who focus their research and writing on the topic of grey literature (...) referred to as the meta-authors on grey literature" and identified 152 authors and co-authors of 139 papers. The citation analysis of 1344 records and 1721 authors or corporate authors revealed that roughly one quarter of these cited authors matched with the "meta-authors on grey literature" of GreyNet, and that another quarter "(...) deals (dealt) with grey literature, but does (did) not explicitly adhere to the term".

Four years later, when the new *OpenGrey* service (formerly *OpenSIGLE*) was launched, Farace et al. (2009) defined the grey community as the "250 authors/researchers in the GL-Conference Series" and described the integrating role of GreyNet and the new open repository. The paper also mentioned two groups of potential interest for further development, i.e. the former EAGLE member institutions and "new stakeholders in Grey Literature" whoever this may be.

More recently, Marzi (2012) highlighted the particular relationship between terminology and community. She compared the usage of specific terms and the syntactic complexity of the proceedings of the GreyNet conference to social networks (Facebook) and subject-based communities (LinkedIn) and focused on "supporting relationships and content sharing". Her

conclusion was that subject-based collections such as the proceedings offer a more "coherent flow of shared and structured knowledge" than social networks but that subject-based communities such as the GreyNet group on LinkedIn can nevertheless contribute to "knowledge building and informative flow", if there is a "strong interaction between medium and content (as in subject-based community exchanges)".

Obviously, sharing common research interests, belonging to a group, attending the same events and speaking the same language are key factors for the definition of a community. GreyNet without a doubt is such a community, with an accepted terminology, reference definitions, social events and vectors of communication.

But if it is relatively easy to determine the core grey community based on membership, publishing behaviour and conference attendance, the frontier between inside and outside the community is hard to find. Surely, often enough we can read on Twitter or elsewhere questions like "what does grey literature really mean?" Often enough we can guess that outside of GreyNet the "inside-terminology" is not really in use and that grey literature remains a rather obscure topic, not quite clear, not well known.

On the other hand, we can also observe that other scientists and professionals from LIS or other domains publish about/on grey literature. Sometimes they apply the concept of grey literature together with the definition of GreyNet, sometimes they don't for instance, when an article about PhD theses does not mention their grey character.

We were interested in the boundary (should we say continuum?) between inside and outside of the GreyNet community. This time we have not tried to define this boundary through citation analysis (= which authors are working together? Who cites whose publications?...) but through the analysis of usage of terminology and choice of topics. Who uses "grey literature" as an object of research and publication? Who works on documents belonging to grey literature without applying the term "grey literature"?

Our expectation is that this double approach may provide evidence of the outreach and impact of GreyNet beyond its community and of its potential for further development.

Methodology

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 (table 1), applying three criteria:

Document type: The search was limited to published papers.

Time period: We considered documents published between 2000 and 2012.

Content: We searched for references that contain "grey literature" or variants in the title, abstract or keyword fields. Subsequently, we added references on PhD theses or Master dissertations.

The exact approach for each database is described in Appendix 1. The results were cleaned and consolidated. Cited publications and documents published by GreyNet (TextRelease) were discarded (Table 1).

Source	Owner	Nb of references	Search date
LISTA	Ebsco	465	4/10 2012

Scopus	Elsevier	1,412	3/31 2012
Web of Science	Thomson	1,206	4/1 2012
Pascal/Francis	INIST	129	3/31 2012

Table 1: Databases with search results

We then conducted the same research in different open access directories and search engines (DOAJ, OpenDOAR, ROAR, E-Lis, OAIster) but only the results from E-Lis were satisfying and relevant while the other tools were not specific enough (no limitation to the time period, to search on fields except full text and/or to published documents). Therefore, we only added the references from the E-Lis directory (Table 2). Again, these references were cleaned and consolidated, and GreyNet publications were discarded.

Source	Owner	Nb of references	Search date
E-Lis	RCLIS/CIEPI	124	4/4 2012

Table 2: Open access directory E-Lis with search results

The references from all sources were uploaded to a unique database. Again, double entries were eliminated, references were cleaned and consolidated. The final database contains 2,440 references and allows for three analyses:
- Study on the publication patterns: the study was conducted in order to know more about these references on grey literature – document types, publication years, preferred journal titles.
- Study on authors: the study was conducted in order to describe this community publishing on grey literature outside of the GreyNet, in particular their institutional affiliation, geographical origin, preferred journals and other vectors of communication.
- Comparison with GreyNet community: The GreyNet community (inside) is defined based on the authors who usually publish in the GreyNet newsletter, in *The Grey Journal*, or in the proceedings of the annual conferences on grey literature partially available in the *OpenGrey* repository. The list was downloaded in April 2012 from the TextRelease web page called "WHOIS in Grey Literature 2012"[1]. Together, this corpus (i.e. the GreyNet community) accounts for 296 members, more or less involved, active and publishing. Our comparison was conducted in order to better understand the specificity of the grey community, its boundaries, outreach and potential.

Results

The corpus of publications on grey literature and/or PhD theses or Master dissertations contains 2,440 references. As mentioned above, these references were retrieved from six different sources. No reference was available in all sources but some are listed in two or three sources;

---

[1]http://www.textrelease.com/whois2012.html

for instance, 762 references (30%) are indexed in both Web of Sciences and SCOPUS databases. All references from GreyNet sources (*The Grey Journal*, GL conferences) were discarded.

50% of the references contain "grey literature" in the title, keywords or abstract while another 49% of the references mention PhD or Master theses (dissertations etc.) in their metadata.

12% of all references can clearly be identified as belonging to scientometrics or bibliometrics, mostly citation analysis.

The results are presented in two stages. First we describe publishing features, preferred journals and the number of publications per author for the whole sample of 2,440 and 5,490 authors. This first part is followed by a more detailed analysis of geographic origins, place of work, domain and profession with a subsample of 433 authors who are strongly committed to grey literature.

Whole sample

Publishing features

82% of the references are from journal articles. Yet, about 18% are from other document types, such as theses or conference proceedings (Table 3).

Document type	Nb of references	in %
Articles	1,990	81.6%
Conference papers	201	8.2%
Theses, dissertations	4	0.2%
Books, sections or reviews	82	3.3%
Others, n/a	163	6.7%
	2,440	100%

Table 3: Document types (n=2,440)

Roughly 8,4% of the papers on grey literature and PhD theses are published and disseminated as grey literature while 82% are "mainstream" article publishing, probably mostly by commercial academic publishing houses.

The scope of our study was limited to papers published between 2000 and 2012. Figure 1 shows the distribution of the year of publication.
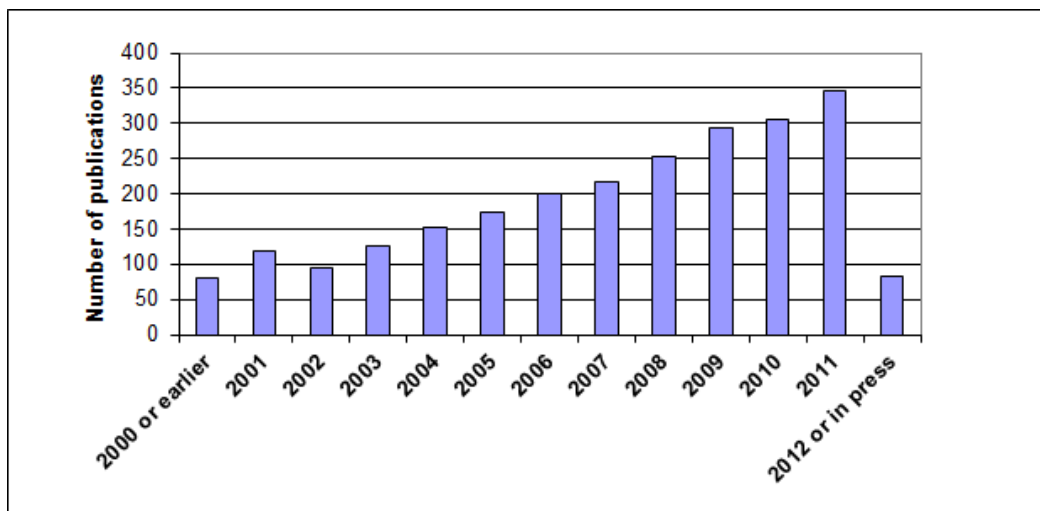
Figure 1: Year of publication (n=2,440)

The median age of publication is four years. 30% of the papers were published during the last two years (2010-2012) or are in press.

More than 36% of all papers were published in medical or public health journals while another 27% appeared in journals from library and information sciences or magazines from the publishing and book trade industry. The other papers were published in different disciplines, for instance in education, computers, biology or psychology.

Preferred journals

Academic journals are still the most important vector for scientific communication. Which titles do scientists and professionals prefer for the submission and publishing of their papers? Our sample provides two different answers.

Articles that mention grey literature in their titles are most often published in LIS journals (Table 7).

Publishing Research Quarterly
Archaeologies
Journal of the Medical Library Association
Interlending and Document Supply
Library Hi Tech News
Collection Building
Journal of Academic Librarianship
INSPEL
Science and Technology Libraries
Documentation et bibliothèques

Table 4: Ten preferred journals (grey literature in article title)

The high ranking of *Publishing Research Quarterly* can be explained by a former agreement with GreyNet that gave permission to PRQ to re-publish the best papers from the international conferences on grey literature.

In 2010 *Archaeologies* published a special issue on grey literature in archaeology that may at least partly explain the second place on the list.

Considering the whole sample and not only the articles with "grey literature" in the title, other titles appear to be as important if not more important than the cited LIS journals (Table 5).

Cochrane Database of Systematic Reviews
Publishing Research Quarterly
Journal of Advanced Nursing
Journal of Academic Librarianship
College & Research Libraries News
Health Policy
Library Hi Tech News
Technical Services Quarterly
Journal of English for Academic Purposes
Chinese Journal of Evidence-Based Medicine

Table 5: Ten preferred journals (all references)

Our data show that journals in medical science and public health (nursing) regularly publish papers that build on grey literature and/or theses and dissertations as valuable sources for literature reviews, scientometric analyses or state of the art contributions. Yet, these journals correspond to about 10% of all articles, and the whole/complete list of journals with articles on grey literature, theses and dissertations is much longer and contains nearly 1,300 different titles. This in other words, means that these publications are more or less scattered in a significant number of journals.

Number of authors

128 references (5%) are anonymous papers and/or without information (n/a) about the author(s). The remaining papers are signed by 5,490 different individual authors. Most of them signed only one paper (88%). The others published two (9%), three to five (3%) or up to fifteen papers (Figure 2).
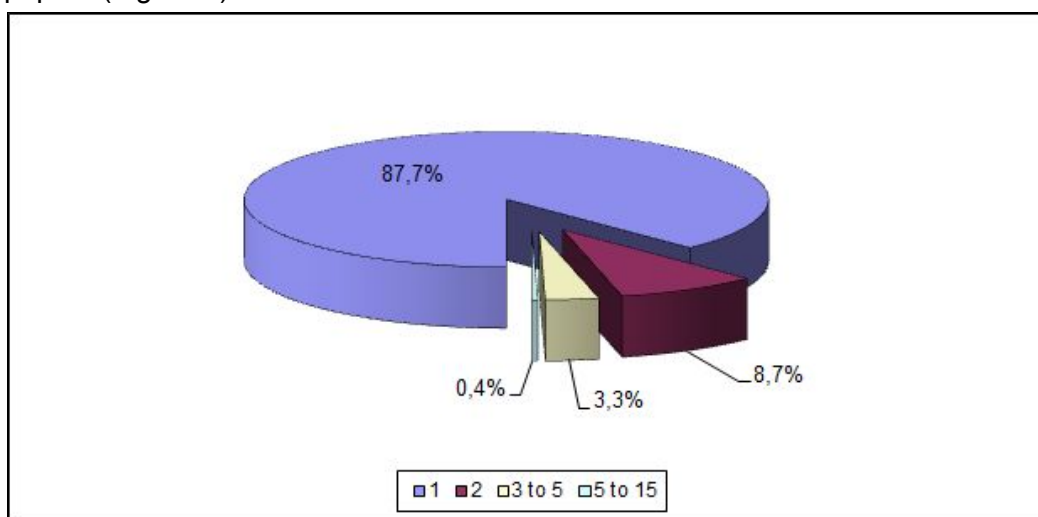
Figure 2: Part/proportion of authors with number of publications (all authors, n=5,490)

All these authors are in some way committed to grey literature or related subjects. Yet, as we were interested in those authors strongly committed to grey literature, we selected a subsample based on two criteria: authors with at least three papers (i.e. 199 authors or 3.6% of the whole sample), and those with at least one paper that mentions "grey literature" in the title. This subsample of authors strongly committed to grey literature contains 432 individuals and one committee of authors (GLISC), that is 8% of the initial sample.

Subsample of authors strongly committed to grey literature

Altogether, these 433 authors have published 550 papers on grey literature and/or PhD theses or Master dissertations.. One quarter of their papers mention "grey literature" in the title. Nearly 13% of the papers present results from scientometric studies.

Geographic origins
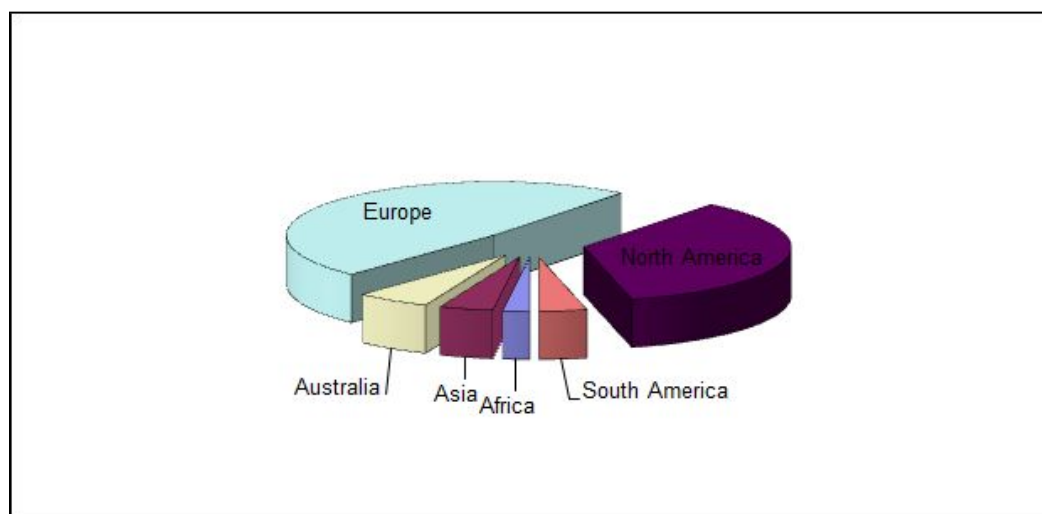Half of these authors are from Europe; another third are from North America (Figure 3).

Figure 3: Geographic origins of authors (n=433)

Even if the authors are concentrated in Europe and North America, all continents are represented. Nevertheless, there is no BRICS effect, at least not in this sample.

Place of work and domain
Where do the authors work, which/what are their professional domains? Unsurprisingly, most of the authors (68%) come from Higher Education, mainly from universities, sometimes from schools. Another 17% are working in research organisations (laboratories, institutes...). The others are working in hospitals, government agencies, non-profit organisations or corporate companies. 4% are from more important and independent libraries, such as INIST, etc. (Figure 4).
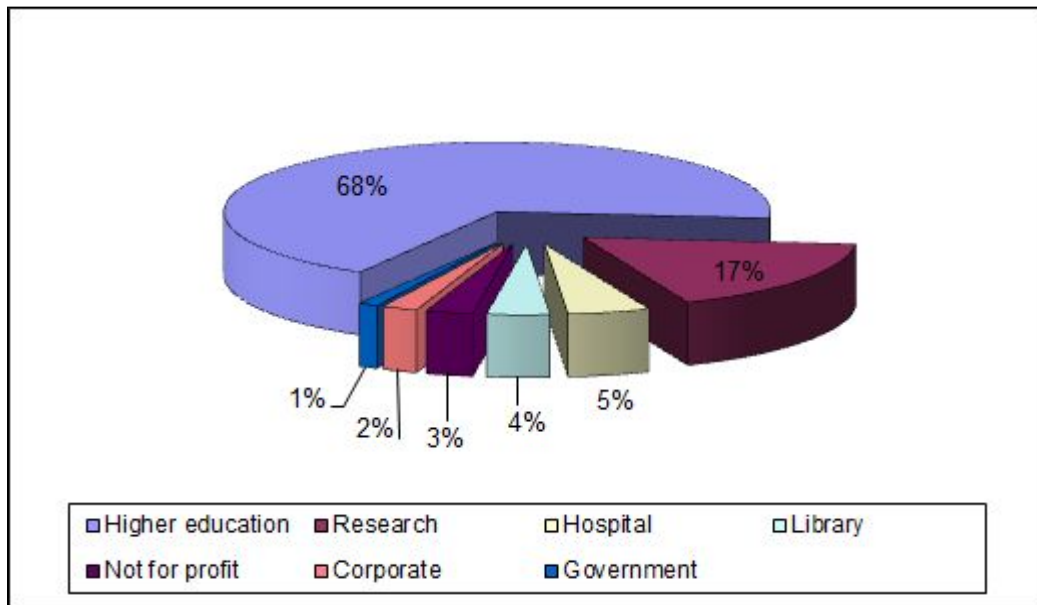
Figure 4: Place of work of authors (n=433)

In which domain are they working? We tried to identify their scientific or professional disciplines from their affiliation or other related information. 40% are working in structures of medical sciences and health, in hospitals or universities, followed by structures in library and information sciences (27%), often in universities (Figure 5).
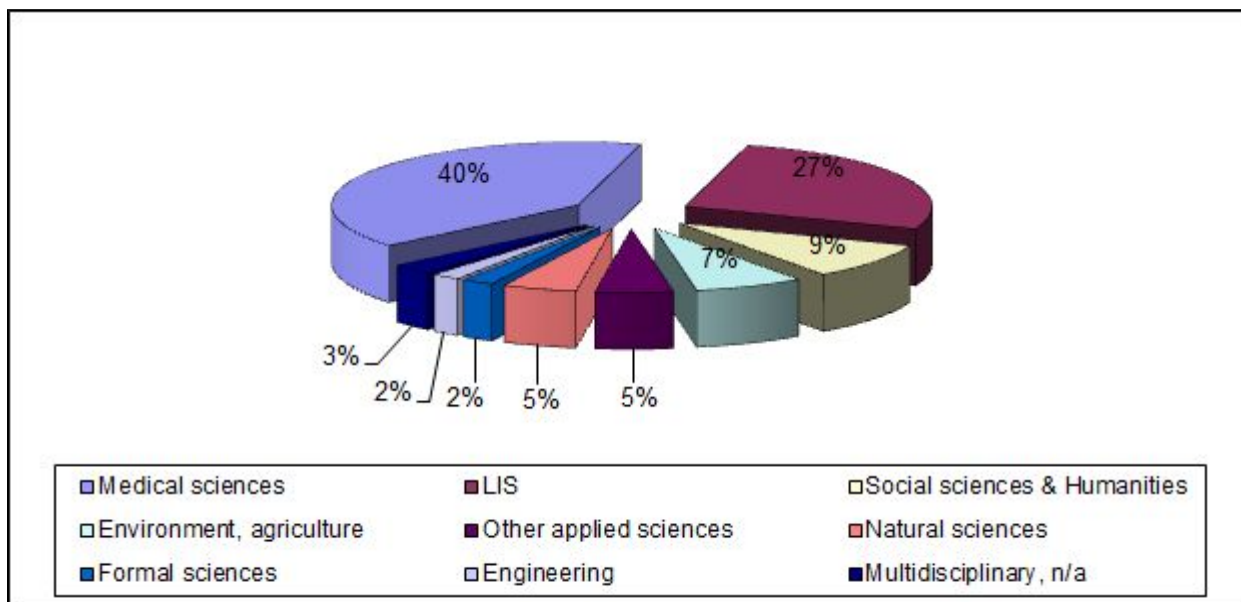

Figure 5: Scientific and professional domains of authors (n=433)

Other significant domains are social sciences and humanities (9%) and environment (7%) while only 5% of the authors are working in disciplines related to natural sciences.

Profession

What are the jobs of the authors publishing about grey literature? More than two thirds are librarians in different functions and settings. Another 25% are scholars teaching Library and Information Sciences at universities or working in research structures (see Figure 6).
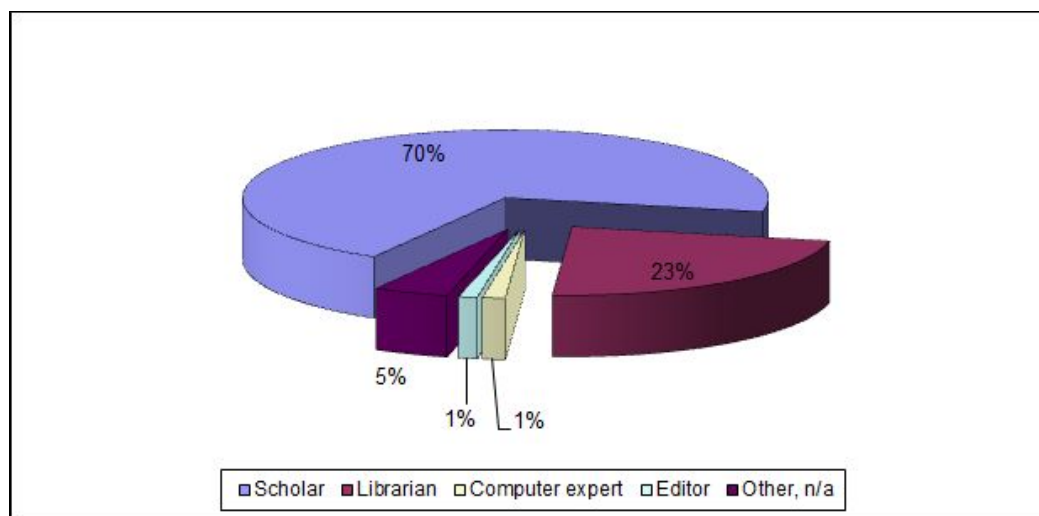

Figure 6: Profession of authors (n=433)

A very small number are computer experts, editors, consultants etc. For others, it was impossible to determine the profession. Most of the academics are teaching Library and Information Sciences while the scientists are often from medical research.

Overlap with GreyNet

Are these authors members of the GreyNet community? Are they listed in the WHOIS, members of the LinkedIn group, authors of papers communicated at a GL conference? We tried to find out by comparing the different samples.

First, we compared the GreyNet community (296 individuals) with our sample of 5,490 authors. 81 people belong to both. This means that 27% of the GreyNet community are represented as authors on grey literature in our initial large sample, but only 1.5% of the authors occasionally writing on grey literature belong to the GL network (sample 1 in figure 7).
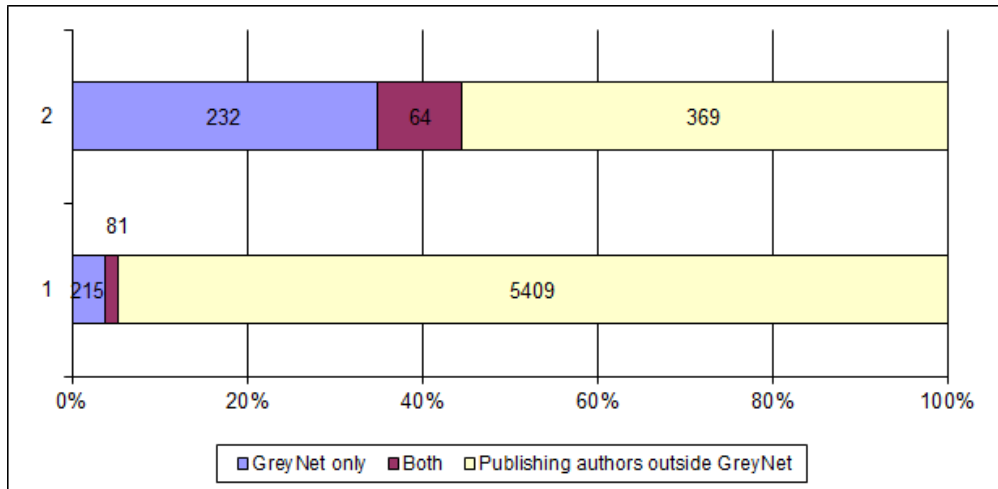
Figure 7: Publishing authors and GreyNet community (sample 1 n=5,490 authors ; sample2 n=433 authors)

Secondly, we limited our comparison to the smaller sample of 433 authors more strongly involved in publishing on grey literature. 64 people are in both communities (sample 2 in Figure 7). Again, this figure means that about 22% of the GreyNet community publish outside of the network and are engaged in grey literature while 15% of these authors are members of the GreyNet community.

The comparison between the two samples 1 and 2 shows that the more an author publishes on grey literature, especially when mentioning the term "grey literature" in the title, the more likely it is for him/her to be a member of the GreyNet community.

But these figures also mean that 70-80% of the members of GreyNet are not identified in databases and repositories as occasional or regular publishers on grey literature or related issues.

Outside of GreyNet - same or different?

85% of the authors strongly committed to grey literature (n=369 of 433) do not belong to the GreyNet community. Are they different?

The comparison of the two groups in our sample of 433 authors shows that the GreyNet authors are not represented in Africa, South America and Australia. Authors from outside of GreyNet represent 36 countries (Figure 7) while GreyNet authors are from only 12 countries.
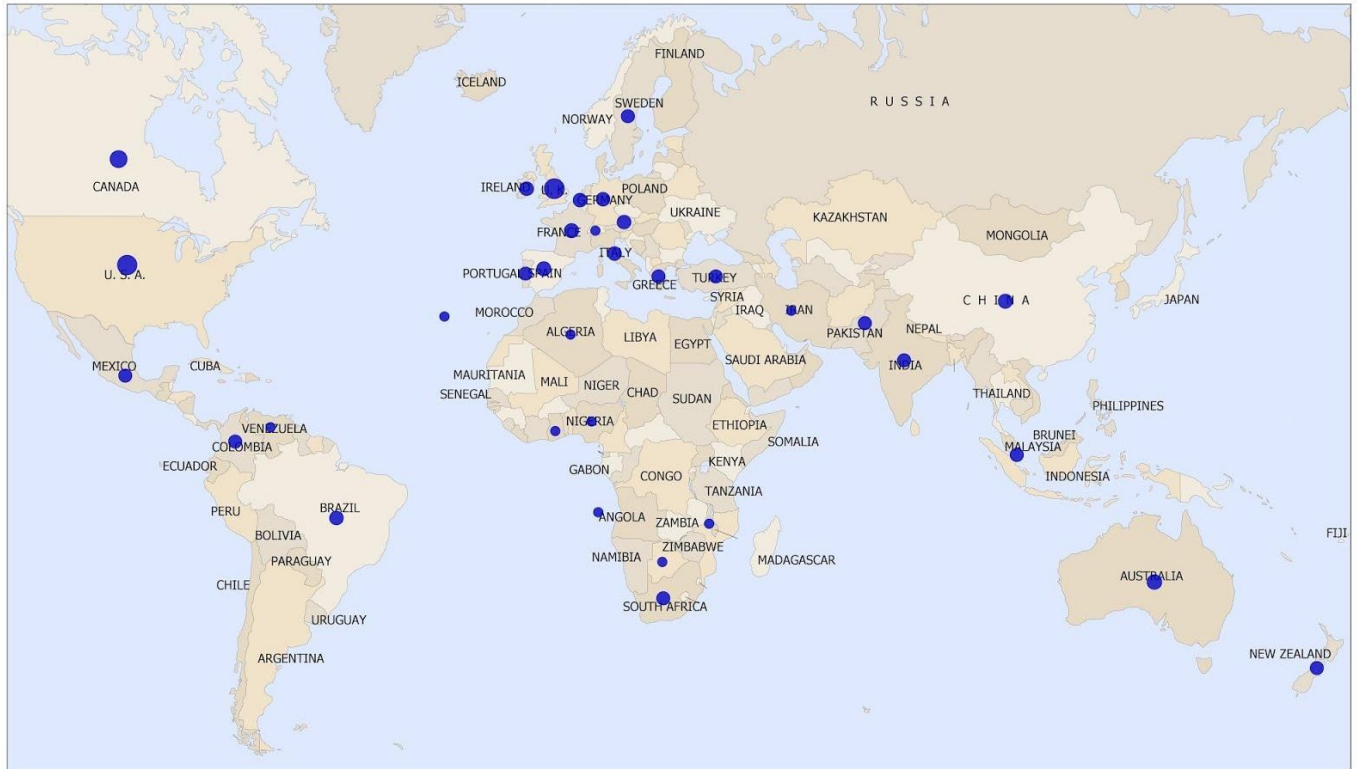
Figure 8: Geographical origin of authors outside of GreyNet (n=369)

Columbia, Mexico, Portugal, Greece, Austria, New Zealand, Pakistan and Turkey are countries so far beyond the reach of the GreyNet. At least this is so in this sample of authors strongly committed to grey literature and publishing more than others.

With regards to the other aspects, we can identify some other differences between GreyNet members and the other authors. For instance:

- GreyNet is less academic than the author sample.
- Research structures, non-profit organisations and libraries are better represented in GreyNet than outside.
- In GreyNet, LIS and natural sciences are better represented, while medical sciences, ecology, applied sciences and social sciences and humanities are underrepresented.
- The relationship between scholars and librarians is completely inverted. Publishing GreyNet members in our study are mostly information professionals while the majority of the publishing community outside of GreyNet is working in Higher Education or research organisations.

These results show two communities which, despite shared interests, are rather different, with a more consistent profile for the GreyNet community.

Discussion

Methodological shortfalls

Our choice to evaluate terminology (grey literature) and choice of topic (theses) as indicators for belongingness to a community is not exhaustive. Also, this kind of scientometric analysis depends partly on the quality and consistency of abstracting and indexing services of databases and repositories and also on the quality and relevance of their search tools and results. Missing search facilities for instance reduced the number of open access directories and most probably the number of identified publications and authors. Also, we had to apply different search strategies (see Appendix) that probably increased inconsistency in the corpus and eliminated relevant items.

Another problem shared with other scientometric studies is the author identification because of misspelling, changes or variants of the first and/or last name and so on. Because of missing person or author identifiers, the consolidation and validation were done manually, a procedure that decreases the reliability of the results.

The same remark applies to the analysis of geographical origins, affiliations and domains of activity. We tried to control this source of error through double-checking and, in some cases, by eliminating the reference or author from the sample. This may reduce the error rate but at the same time decrease the relevance of the overall results.

Grey community or grey communities?

A subject community, as defined above, shares terminology, tools, events, experiences and topics. The "grey community" should therefore share the basic terminology of grey literature, should attend the same events (conferences, workshops, other meetings), make use of the same tools of communication (journals, listservs...) and research (methodology), and discuss the same topics. But does it?

Drawing on our empirical evidence, we can distinguish different groups (Table 6).

Name	Number	Comments
GreyNet	296	Coordination and management by GreyNet Amsterdam and TextRelease.
GreyNet publishing outside	81	The "visible" part of GreyNet outside of the GreyNet events and communication vectors. More information professionals.
Occasionally publishing on GL	5057	One or two papers in ten years, no usage of grey literature terminology. All disciplines.
In 433 authors sample, Regularly publishing on GL with GL terminology	149	Three or more papers. SS&H, LIS. More scholars.
In 433 authors sample, Regularly publishing on GL without GL terminology	31	Three or more papers. Medical sciences. All disciplines. More scholars.

Table 6: Concentric and overlapping groups

These groups are not exhaustive, and we could add others such as:
- GreyNet members attending events,
- GreyNet authors publishing occasionally,
- Groups regularly publishing on grey literature with scientometric approach,

- Groups occasionally publishing on theses,
- Subgroups according to disciplines, organisation or profession, etc.

Some of these groups are defined mainly by practice (publishing features, attendance to events etc.), others by usage of terminology and methodology (citation analysis, grey literature definition etc.) or are mixed.

Probably, members of some groups would not have any problem to identify themselves as part of a "grey literature community" while others would probably prefer to define their affiliation by means of a scientific discipline or a profession.

In and outside of the community

One does not need to use Eskimo words to describe different kinds of snow. As Molière's *Bourgeois Gentilhomme* discovered with delight, one can speak prose without knowing it. And one can publish about grey literature without using the term, or even without awareness of it. We included in our sample papers on theses as a significant and central part of grey literature. We obtained two subsamples, one with references that explicitly mention "grey literature" in the title or in other header information; the others are papers on theses that may or may not mention "grey literature" in the body of the text.

In the whole/complete sample of references, the GreyNet authors represent 15%. But when we consider only the references with "grey literature" in the header information, this part/proportion increases up to more than 40%. On the other hand, when we consider the subsample on theses, the percentage of GreyNet authors decreases to 8%. This means that GreyNet authors when publishing outside of the GreyNet use the grey literature terminology more often but they do not do it in a consistent way. And this means also, that this terminology is known outside of GreyNet but not generally accepted or explicitly used when studying grey literature such as theses.

Comparison between the two samples of authors (the 433 authors and the others) against the two communities (GreyNet and outside) reveals another significant difference. The statistical distribution of the usage and non-usage of the grey literature terminology across these groups is significantly different from the expected values (Table 7).

	GreyNet Terminology	GreyNet No terminology	Outside Terminology	Outside No terminology	Total
433 authors	**64**	0	**336**	*33*	433
Others	*14*	3	3549	**1491**	5057
Total	78	3	3885	1524	5490

Table 7: Usage of grey literature terminology in header across the author samples

The figures in bold are higher than expected, and the figures in italics are lower than expected. In other words: Authors from GreyNet but also from outside who are highly committed to grey literature significantly use the term "grey literature" more often the others when describing their work. Compared to less committed authors, we can speak of a terminology-based community.

One reason for the differences may be that while GreyNet authors (mainly librarians from all disciplines and LIS academics) often choose grey literature as their object of research (Library and Information Sciences), authors outside of the GreyNet (mainly academics from different disciplines) include specific types of grey items in citation analysis or reviews (state of the art), in particular in medical or life sciences, without having a global concept of grey literature.

Our empirical data tend to confirm the reality of two grey communities, with relatively closed frontiers between inside and outside of Greynet - only 3.5% of all papers are co-authored with somebody "from the other side". Obviously in the whole/complete sample we cannot speak of some kind of freedom for the movement of ideas, projects and publications across this border. In so far as both are working with the same material, is this a problem? Perhaps not, as long as mutual understanding and exchange remain possible. Yet, the concept of grey literature draws a framework for research and practice and allows for a better understanding of a specific kind of scientific and technical communication. Therefore, the use and promotion of "grey literature" terminology is all but trivial, not only inside the GreyNet community but also and above all outside of the community.

Conclusion

As we stated at the beginning, a community builds on shared terminology, definitions (concepts) and practice, such as common methodology and events. Our empirical data indicate the existence of different shades of grey communities, with regards to GreyNet membership, publishing features and usage of terminology. Nine out of ten authors appear to be "bouncers" during the observed period 2000 to 2012, occasionally speaking of grey literature or using grey material. Only one out of ten authors can be considered as a kind of "returnee", with a kind of loyalty and respectful behaviour regarding grey literature.

The analysis of 2,440 papers published by 5,490 authors confirms that the concept of grey literature remains more or less a professional affair applied by librarians and LIS academics. The papers can be mapped to/in three concentric circles: at the core, some studies on grey literature as an object of research, followed by a group of papers with a conscious and direct use of the concept. The third and largest circle contains papers that make usage of grey items, with or without awareness of the concept of grey literature.

Thus, the data show a potential for the development of the GreyNet community beyond the actual and often tight frontiers. We conclude with two possible strategies for the development of GreyNet, one based on proximity, the other on exploration.

Proximity

The first strategy is centred on authors outside of GreyNet with a profile adjacent to that of GreyNet members. This calls for contact as a priority:

- Authors who mention grey literature in the header information of their paper.
- Librarians and other information professionals, scholars from library and information sciences.
- Authors of papers published since 2008.

Contact could be established in order to invite publications for *The Grey Journal* or a monograph, to suggest communications for the GL conferences, and to invite to join the GreyNet listserv etc.

Exploration

The second strategy focuses on specific groups of authors that are not necessarily very near to GreyNet but nevertheless in (some kind of) grey literature. For instance:
- The Networked Digital Library of Theses and Dissertations (NDLTD), an international organization dedicated to promoting the adoption, creation, use, dissemination, and preservation of electronic theses and dissertations.
- Other library sections dedicated to special collections, acquisition policy and institutional repositories.
- Identified authors from Latin America, Sub-Saharan Africa, Australia or the BRICS countries.
- Identified authors of papers on grey literature (citation analysis, case studies etc.) in medical sciences.

Here, contact could be made in order to suggest joint publications or events, keynote addresses to GL conferences, or invitation to join the GreyNet listserv etc.

Beyond the community

These strategies could be implemented with the help of the GreyNet group on LinkedIn. This group already radiates beyond the traditional frontiers of GreyNet. Presently (October 2012) it has 271 members, and only around twenty of them usually publish in the GreyNet newsletter, in *The Grey Journal*, or in the proceedings of the annual conferences. Here, the GreyNet community clearly has the potential to expand and to enhance its impact.

However, radiating beyond the community can also mean publishing on grey literature in other media and products, editing special issues related to grey literature not only in LIS journals but also in selected journals from other disciplines that are usual "consumers" of grey literature (medical sciences, social sciences and humanities...), and fostering joint research and publishing projects between professionals and academics.

Integrating authors near to GreyNet, exploring other communities and reaching beyond the GreyNet community could be three different but complementary ways to promote and foster sustainable development of this network.

Bibliography

M. Callon (1989). *La Science et ses réseaux :genèse et circulation des faits scientifiques*. Paris, La Découverte.

Farace, D. (2011) What the Future holds in Store for GreyNet International, Business Report Amsterdam, TextRelease. http://www.greynet.org/images/Business_Report_2011.pdf

Farace, D., Frantzen, J., Stock, C. et al (2009) OpenSIGLE, Home to GreyNet's Research Community and its Grey Literature Collections: Initial Results and a Project Proposal, in *GL10*

Tenth International Conference on Grey Literature: Designing the Grey Grid for Information Society. Amsterdam, 8-9 December 2008. http://archivesic.ccsd.cnrs.fr/sic_00379643

T. S. Kuhn (1962). *The Structure of Scientific Revolutions*.University of Chicago Press, Chicago.

B. Latour& S. Woolgar (1979). *Laboratory life : the social construction of scientific facts*. Sage Publications, Beverly Hills.

Marzi C. (2012). `Knowledge Communities in Grey'.*The Grey Journal***8**(1):27-33.

Paganelli C. (2012). `Recherche en SI. Analyse des discours sur la notion d'« usage » dans deux revues en sciences de l'information : Doc-SI et BBF'. *Documentaliste*49(2):64-71.

Schöpfel, J., Stock, C., Farace, D. (2005) Citation Analysis and Grey Literature: Stakeholders in the grey circuit, in *GL6 Sixth International Conference on Grey Literature "Work on Grey in Progress"* New York Academy of Medicine Conference Center New York City, 6-7 December 2004. http://archivesic.ccsd.cnrs.fr/sic_00001534

C. Schrecker (2006). *La communauté : histoire critique d'un concept dans la sociologie anglo-saxonne*. L'Harmattan, Paris.

Appendix

For each data source, we preferred if available the advanced search interface, eliminated double entries and references from the GreyNet (The Grey Journal, GL conference series etc.), and limited the results to the time period mentioned above (2000-2012).

OAIster was accessed through the OCLC WorldCat but the results were not specific enough. We could not search in the LISA database because we had no access. We did not consider the results from ROAR and OpenDOAR because of missing search functionalities. The search results were much too large and without interest for our study.

The table shows the search strategies. After some exploratory tests, we decided to limit the search in the following way/manner:

- Grey (or gray) literature in the title of the publication.
- Grey (or gray) literature in the abstract.
- Grey (or gray) literature in the keywords.

In order to produce some information about authors dealing with grey literature but not using the concept of grey (or gray) literature, we also searched for papers on theses and dissertations, excluding biographical papers.

Corpus	Extraction
LISTA	all txt contains "grey literature" or "gray literature" or dissertation or "master thesis" or "doctoral thesis"
Scopus Grey	Title abstr keyword contains "grey literature" or "gray literature"
Scopus Thesis	KW contains thesis or Title contains doctoral dissertation
WoS Grey	Topic contains "grey literature" or "gray literature"
WoS Thesis	Title contains dissertation or "master thesis" or "doctoral thesis"

PASCAL and FRANCIS	All text contain "grey literature" or "gray literature"
E-LIS	All metadata & Full Text contain "grey literature" or "gray literature", or subject = "H. Information sources, supports, channels > HB. Gray literature" or "master thesis" or "doctoral thesis"

All references were cleaned and consolidated, and GreyNet publications were discarded.