

OpenGrey = new OpenSIGLE

Petra PEJŠOVÁ

National technical Library, Prague
petra.pejsova@techlib.cz

Christiane STOCK

INIST-CNRS, Nancy, France
Christiane.STOCK@inist.fr

INFORUM 2012: 18. Conference on Professional Informational Resources
Prague, 22. - 24. 5. 2012

Abstract

OpenGrey is a unique repository providing open access to European grey literature references, the result of 25 years of cooperation. OpenGrey is based on the OpenSIGLE/SIGLE database which contains almost 700 thousand records of grey literature. As a multidisciplinary database it covers Science, Technology, Biomedical Science, Economics, Social Science and Humanities. This paper presents new search functionality, design and vision. The cooperation with GreyNet on GL conference preprints will also be mentioned. OpenGrey usage data and feedback information support our advocacy for more grey literature in repositories. Cooperation example will be shown in the National Repository of Grey Literature in the Czech Republic.

OpenGrey is European database of grey literature, based on the OpenSIGLE/SIGLE database. When we look into more than thirty-year history it started as a unique European database of bibliographic records in grey literature SIGLE (System for Information on Grey Literature in Europe). SIGLE was created and maintained between 1980 and 2005 by initially seven and in the end fifteen members of the European Union, represented by major libraries and research organizations. (Stock, 2011) These representatives formed for this purpose non-profit network called EAGLE (European Association for Grey Literature Exploitation).

SIGLE contents covered all scientific disciplines (pure and applied science and technology, economics, social sciences and humanities). As a commercial product, SIGLE was distributed through subscription to hosts, e.g. STN International, and available on a CD-ROM produced by Silverplatter/Ovid.

SIGLE records had the following mandatory information, essential for it's usage: an English title or keywords allowed search across all countries, the SIGLE

classification code for a search by subject and an availability statement (holdings), meant to facilitate the order of a paper copy. (Stock, 2011)

In 2005 the EAGLE discontinued the network because its organisational structure and business model were unable to cope with Internet technology and the Google generation, e.g. SIGLE offered no solution for online cataloguing, metadata harvesting, links to full-text and other resources (Schöpfel et al. 2007).

The French INIST-CNRS (Institute for Scientific and Technical Information of the French National Center for Scientific Research) decided to transfer the results of 25 years of work onto an open access platform. As a result OpenSIGLE went live in December 2007 with almost 700 000 bibliographic records. It was based on DSpace technology and its qualified Dublin Core metadata format. Minimal developments were made for the user interface, which provided search and browse facilities and offered separate information on the partners, the subject classification, and on document supply. (Stock, 2011) No new records were uploaded with the exception of materials from the series of the International Conference on Grey Literature organized by GreyNet¹.

Since summer 2008 OpenSIGLE was indexed in Google and Google Scholar. In the same year it was also integrated in the WorldWideScience.org portal and a year later in OpenDOAR (Directory of Open Access Repositories). OpenSIGLE contents found an ever growing audience and new visitors from all over the world. The database was reached through the Google search engine by an increasing number of users, and visits via Google Scholar amounted to 30% per month. Between 2009 and 2010 the number of visits almost tripled, and the number of page views more than doubled.

Technical limits and limited options for user interface led INIST-CNRS to think of changing the system. Technical limits included problems with uploading and indexing new records, long search response times, needs for referencing old records and exporting. Many visitors reached OpenSIGLE after using a search engine, spent a short time looking for relevant information and left. An important number of users enquired about how to obtain the document. Although this service was one of the principal goals of SIGLE and OpenSIGLE, the information was hidden on a separate page. However, few users took the time to look it up.

¹ <http://www.greynet.org/>



Figure 1: OpenSIGLE homepage

With the goal to improve the access to information on document supply through the general layout, improved record display and additional links and re-open OpenSIGLE for “new” input and close the gap between SIGLE and today INIST-CNRS decided to move OpenSIGLE to a new platform and new technology.



OpenGrey went live at www.opengrey.eu on June 17th, 2011. The complex change included not only software, but also logo, design and the user interface. For search is used the engine Exalead, for the database and for the user interface was chosen in-house developments using php and MySql software. This solution offers main improvements both for the administration and for the user interface, namely re-indexing, faceted search, possibility to export records in XML and Zotero, export search results as RSS file, information how to obtain the document is placed more prominent, etc.

The unique identification of records is very important. Each record in OpenGrey is identified by a unique identifier, *the handle*. These *handles* had to be transferred from the old system to the new one. Redirections were added for bookmarked pages (collections, individual records).

The new solution provides requested improvements for building grey literature content. The most important features are tools for adding records from 2005 onwards from current partners, opening of the database to new European partners and links to digital documents for existing records.



Figure 2: OpenGrey homepage

Directly at OpenGrey homepage is the search field, besides it are three sections for short information about *OpenGrey*, *news* about grey literature in the world and *focus* on the figures on content. The bottom part offers detailed help on search facilities and document delivery, legal mentions, and further information on OpenGrey, grey literature and GreyNet.

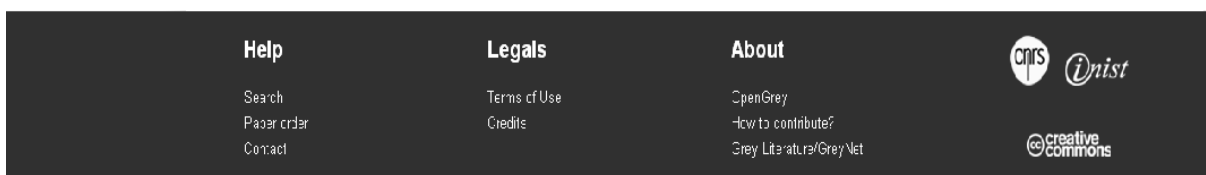


Figure 32: OpenGrey bottom



Figure 32: OpenGrey search

“Basic search in OpenGrey is done in a Google-like field. Exalead offers the possibility to refine the results through faceted search. The new criteria are added to the query field and the search strategy can be copied and saved. The indexes chosen for refinements (such as author, subject, date, language...) are similar those of the Czech repository NUSL/NRGL².” (Stock, 2011)

List of SIGLE Subject Heading is visible in following tab. 22 subject groups are expandable to the next level and every subject heading is also query to the database, just one click.

² <http://www.nusl.cz/>



Figure 32: OpenGrey subject

List of partners contains general information about the organization, information on how to obtain a copy from old SIGLE records and links to grey literature repositories or projects in the country.



An icon permits to start a search on the records from this partner. The last tab provides information about export opportunities from OpenGrey.

Since the launch of OpenGrey in June 2011 to March 2012 the statistics show more than 33 000 visits per month from 200 countries (a total of 70 000 visits from UK, 36 000 visits from US, 36 000 visits from DE). There is also noted increasing access through mobile devices 1,3 %, even though at present OpenGrey doesn't provide any apps for these devices.

In the past, NTK (National Technical Library) was responsible for delivery of grey literature data from the Czech Republic to SIGLE, produced by EAGLE. As a result of the end of the coordinated evidence of grey literature based on SIGLE principles was also national evidence terminated in the Czech Republic. In 2005, NTK organized several meetings with representatives of universities, central libraries and other institutions, concerning potential collaboration in collecting grey literature at the national level. On the basis of these meetings, NTK

gradually started to prepare NRGL (National Repository of Grey Literature). Building NRGL was initiated as a project which was supported by the Ministry of Culture and its full title was The Digital Library for Grey Literature – Functional Model and Pilot Implementation in 2008. The four-year project involved [the University of Economics in Prague](#). The result is that NRGL provides central access to information on [grey literature](#) produced in the Czech Republic in the fields of science, research and education. (National Repository of Grey, 2012) Search interface is available at www.nusl.cz. NRGL fosters cooperation with international grey literature network.

OpenGrey opens possibility for uploading new records about grey literature. INIST-CNRS prefers cooperation through national partners. For data transfer is recommended OAI-PMH protocol. Metadata formats for OpenGrey are DublinCore or MARC XML, in both cases is necessary to prepare conversion table. Every metadata record has to include mandatory fields, SIGLE subject headings, information about document type and link to original repository, where digital document available in open access, or a contact to a service that delivers a copy of the document.

For Czech repositories is cooperation with OpenGrey set through NRGL. This cooperation is free of charge. NTK prefers OAI-PMH set for data transfer but transfer in batches is also possible. Because NRGL uses its own metadata format, there are no requirements for metadata format and the conversion table is done in any case. Mandatory fields are required, namely Author, Title, Title in English, Publishing date, Language, Document type, conditionally more (eg link to document, data about conference, etc.). In NRGL are added other fields which are necessary for cooperation with OpenGrey, these are Producer name, PSH headings, NRGL document type, PID and contact to delivery service in case that link to digital document is missing. The availability of digital document is required information in every record. Highly recommended field are keywords in English. SIGLE Subject Headings, which is a mandatory field, are used for OpenGrey. Because SIGLE Subject Headings are not updated, NTK uses Polythematic Structured Headings (PSH) in NRGL. When the record comes to NRGL at first it is indexed by PSH and in the second step PSH is mapped to SIGLE Subject Headings for OAI-PMH set.

For a complete picture, it is important to mention other fields that are added into OAI-PMH set. The field Country with content "CZ" is added to all records. The field Relation.uri is added only with records which have link to digital document. The content of the field Language has to be transferred from two characters to three characters code.

Connecting with OpenGrey increases³ visibility and usability of Czech research outputs at the international level. Be a part of OpenGrey, contact NTK which is national partner for the Czech Republic⁴.

REFERENCES:

National Repository of Grey Literature [online]. Praha : Národní technická knihovna, [2008] [cit. 2012-03-30]. English version. Available from World Wide Web: <<http://nrgl.techlib.cz/>>.

SCHÖPFEL, J. - STOCK, C. - HENROT, N., From SIGLE to OpenSIGLE and Beyond: An In-depth Look at Resource Migration in the European Context. [online]. *The Grey Journal*, 2007, vol. 3, no. 1, p. 45-51. Available from www: http://arcm.vesic.ccsd.cnrs.fr/sic_00181592/en/

STOCK, Ch., OpenGrey: How it works and how it is used. In. Seminar on Providing Access to Grey Literature 2011: 4th year of the seminar focused on storage and making the grey literature accessible, 25. 10. 2011 [online]. Prague: National Technical Library, 2011. Available at www: http://nrgl.techlib.cz/index.php/4rd_year_of_the_seminar ISSN 1803-6015.

³ „Connecting the NRGL to international databases increases search ability of Czech grey literature. An attendance of NRGL has increased 10 times since July to October.” at http://nusl.techlib.cz/images/Pejsova_text_EN.pdf

⁴ <http://nusl.techlib.cz/index.php/Kontakty>