



HAL
open science

Data from "OpenGrey, System for Information on Grey Literature in Europe"

Joachim Schöpfel, Dominic J. Farace, Jerry Frantzen

► **To cite this version:**

Joachim Schöpfel, Dominic J. Farace, Jerry Frantzen. Data from "OpenGrey, System for Information on Grey Literature in Europe". The Grey Journal, inPress, 10.17026/dans-xtf-47w5 . hal-03160726

HAL Id: hal-03160726

<https://hal.science/hal-03160726v1>

Submitted on 5 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Data from “OpenGrey, System for Information on Grey Literature in Europe”

<https://doi.org/10.17026/dans-xtf-47w5>

URN: urn:nbn:nl:ui:13-i3-x6uc

Joachim Schöpfel, University of Lille, France

<https://orcid.org/0000-0002-4000-807X>

Project Leader / Database Reformat / Final Text

Dominic Farace, GreyNet International, Netherlands

<https://orcid.org/0000-0003-2561-3631>

Project Administration / Draft Text

Jerry Frantzen, GreyNet International, Netherlands

<https://orcid.org/0000-0002-3405-7078>

Technical Support

Abstract

OpenGrey, System for Information on Grey Literature in Europe, is your open access to 1 014 872 bibliographic references of grey literature produced in Europe and allows you to export records and locate the documents. Examples of grey (gray) literature include technical and research reports, doctoral dissertations, conference papers, official publications, and other types of grey literature. In 1980, this information resource first began as a bibliographic database under the name SIGLE. In 2006, the records migrated to OpenSIGLE - an open access database. In 2010 its name was changed to the OpenGrey Repository given that full-texts documents began to be ingested and preserved. And in 2020, Inist-CNRS (service provider) gave notice that OpenGrey will be discontinued. In order to guarantee open access to this resource, it is now archived as a database in DANS EASY – a data archive certified with the Nestor and CoreTrustSeal.

Keywords: Data Paper; SIGLE; OpenSIGLE; OpenGrey; EAGLE; Grey Literature; Gray Literature

Subject Area: Science, Technology, Biomedical Science, Economics, Social Science, Humanities

Methods Applied

● Steps

Each participating European organization used a standard record template (form) with the aid of manuals (e.g., Subject Category List and associated key terms as well as a Corporate Register with preferred Names/spellings). Each participating organization made a selection of their acquisitions for submission to the SIGLE database. If a document title was not in English, then an English translation was provided. As such, each SIGLE record contained informative titles in English and/or in the original language, the author’s name, academic degree, and the research organization or educational institution, the document’s date of publication and type, number of pages, report numbers, and language, as well as subject classifications. Many records include keywords and abstracts. Each national structure submitted records in their own language. A search through the entire database was made possible by providing an English translation of the title or English keywords. One of these fields was mandatory. Moreover, each record contained a clear mention of availability.

● Sampling strategy

The selection of records entered in the SIGLE database was made by several institutions representing the EAGLE member countries. The following list is not exhaustive.

| Country | Institution |
|----------------|----------------------------------|
| Belgium | Université Catholique de Louvain |
| Czech Republic | National Library of Technology |

| | |
|--------------------|---|
| Germany | Fachinformationszentrum Karlsruhe |
| Germany | German National Library of Science and Technology (TIB) |
| Spain | Centro de Informaciòn y Documentaciòn Científica |
| France | Institut de l'Information Scientifique et Technique-CNRS |
| United Kingdom | British Library Document Supply Service |
| Italy | Consiglio Nazionale delle Ricerche |
| Luxembourg | Bibliothèque Nationale de Luxembourg |
| Latvia | Latvian Academic Library |
| Portugal | Fundação para a Ciência e a Tecnologia |
| Russian Federation | VNTIC Scientific & Technical Information Centre of Russia |
| Slovakia | Slovak Centre of Scientific and Technical Information |

The sampling strategy of the former SIGLE records was decided by the EAGLE association, coordinated by the SIGLE technical committee, and carried out by the member institutions following agreed and explicit rules. Later input in the OpenSIGLE/OpenGrey repository that originated from the International Conference Series on Grey Literature was provided by GreyNet International, the Grey Literature Network Service. Regarding the property of the records and the database, all usage rights of the SIGLE database lapsed upon the complete liquidation of the association while the copyright on input remained with each member organization that supplied the records. EAGLE's last General Assembly asked the operating agent, FIZ Karlsruhe, for the interim preservation of the SIGLE records in XML format beyond the liquidation of SIGLE. This was for the purpose of archiving and integration into a new European non-profit project i.e., OpenSIGLE hosted and managed by INIST-CNRS. Nearly all of the former EAGLE members signed a declaration of intention regarding the future use of their existing input in the SIGLE database. The complete liquidation of EAGLE was formally published by the Luxemburg Register of Commerce and Societies on August 23, 2006 (Schöpfel et al., 2006).

- Quality Control

The SIGLE descriptive cataloguing rules were based on those of the International Nuclear Information System (INIS), and the subject classification scheme was a modified version of that endorsed by the Committee on Scientific and Technical Information (COSATI) of the US Federal Council for Science and Technology. Each member institution was responsible for the quality of their national records. The operating agent (technical processing center) played a key role in the quality control, rejecting non-compliant input. The SIGLE technical committee supervised the compliance with the agreed rules. However, "in operating SIGLE, no attempt has been made to be rigid in formulating rules of membership (...) considerable flexibility has enabled a number of countries to participate to a greater or lesser extent depending on their capabilities" (Wood & Smith, 1993, p.21). Furthermore, the last operating SIGLE agent and the host of the OpenSIGLE/OpenGrey repository, FIZ Karlsruhe and INIST-CNRS, executed complementary quality controls at different moments during the transfer, launch, and update of the metadata records (Schöpfel et al., 2006).

Dataset Description

| | |
|--------------------------|---------------------------------|
| File name: | OpenGrey |
| Format: | MySQL |
| Size: | 360,358,816 bytes |
| Path: | mysql.sql.gz |
| Creation dates: | from 1980-01-01 to 2018-07-27 |
| Language: | English |
| License: | CC0 Waiver - no rights reserved |
| Archive name: | DANS EASY Archive |
| Publication date: | 2021-02-06 |
| DOI: | 10.17026/dans-xtf-47w5 |
| URN: | urn:nbn:nl:ui:13-i3-x6uc |

Potential Reuse of the Data

In 1980, this information resource first began as a bibliographic database under the name SIGLE. In 2006, the records migrated to OpenSIGLE - an open access database. In 2010 its name was changed to the OpenGrey Repository given that full-texts documents began to be ingested and preserved. And in 2020, Inist-CNRS (service provider) gave notice that OpenGrey will be discontinued. In order to guarantee open access to this resource, it is now archived as a database in DANS EASY – a data archive certified with the Nestor and CoreTrustSeal. The database file can be downloaded in XML and CSV format. It is accompanied by a ReadMe text file and a document on MySQL. The ReadMe file contains information about the eight tables of the database. The bibliographic metadata data can be useful to enrich scientific bibliographic databases or other academic discovery tools. The data can also be useful for scientometric studies on academic output in the European research area between 1980 and 2005. Also, because of the records' harmonized multidisciplinary and multilingual characteristics, the data can be used for research and development in the field of scientific terminology. Furthermore, the data are published on the DANS EASY platform under a Creative Commons public domain license.

Linked References:

Farace D., Frantzen J., Stock C., Henrot N., & Schöpfel J, (2009). OpenSIGLE, Home to GreyNet's Research Community and its Grey Literature Collections: Initial Results and a Project Proposal. In: Tenth International Conference on Grey Literature - Designing the Grey Grid for Information Society, Science Park Amsterdam, NL. December 8-9, 2008 vol. 10 (1), pp. 118-122. <https://doi.org/10.26069/greynet-2020-000.239-gg>

Farace D., Frantzen J., Stock C., Henrot N., & Schöpfel J, (2010). OpenSIGLE - Crossroads for Libraries, Research and Educational Institutions in the field of Grey Literature. In: The Grey Journal: An International Journal on Grey Literature, Library of Congress, Washington D.C., USA 14-15 December 2009 vol. 11 (1), pp. 122-128. <https://doi.org/10.26069/greynet-2020-000.212-gg>

Schöpfel, J. (2005). MetaGrey Europe, a Proposal in the aftermath of EAGLE-SIGLE. *GL7 Seventh International Conference on Grey Literature*, 5-6 December 2005, Nancy. http://archivesic.ccsd.cnrs.fr/sic_00001722/en/

Schöpfel, J., Stock, C., & Henrot, N. (2006). From SIGLE to OpenSIGLE and beyond: An in-depth look at Resource migration in European context. *GL8 Eighth International Conference on Grey Literature*, 4-5 December 2006, New Orleans. https://archivesic.ccsd.cnrs.fr/sic_00181592/

Wessels R. H, (1998). An Eagle's eye-view of grey literature research. In: Third International Conference on Grey Literature: Perspectives on the Design and Transfer of Scientific and Technical Information, Luxembourg 13-14 November 1997 vol. 3 (1), pp. 212-213. <https://doi.org/10.26069/greynet-2021-000.398-gg>

Wood, D. N., & Smith, A. W. (1993). SIGLE: A Model for International Co-operation. *Interlending & Document Supply*, 21(1), 18–22. <https://doi.org/10.1108/02641619310154331>

History

The European database SIGLE (System for Information on Grey Literature) was created in 1980 in order to collect and to make available grey literature produced in the countries of the European Community by the intermediary of an online database. The main purpose was to provide access to documents, and incidentally to improve bibliographic coverage. SIGLE was launched with the support of the European Community. The base was produced from 1985 onwards by EAGLE (European Association for Grey Literature Exploitation), with an operating agent (KNAW, The British Library, FIZ Karlsruhe) and with input from several Member States of the European Community (Wood & Smith, 1993; Wessels, 1998). Each country was represented by one or more documentation centers or libraries of national importance, and each national center held the national grey collection or at least, guaranteed the document supply of distributed holdings. Until 2005, the SIGLE database was distributed by BLAISE, STN, EI and Ovid (Silverplatter) and in the early nineties by SUNIST in France. Records were also sold to organizations like NERAC for specific use (Schöpfel et al., 2006).

In 2005, the members of EAGLE decided to dissolve the association and to stop the input of new documents. The database SIGLE was taken from all hosts proposing the database online. In the context of the Open Access movement and with the consent and approval of the former EAGLE members, EAGLE's last operating agent (FIZ Karlsruhe) transferred the SIGLE records to the French INIST-CNRS (Institut de l'Information Scientifique et Technique) which then launched the OpenSIGLE project, i.e., all former SIGLE records, in a simplified XML format, were made freely available on a DSpace open repository (Schöpfel, 2005; Schöpfel et al., 2006; Farace et al., 2009, 2010).

In 2009-2010, GreyNet International joined OpenSIGLE, adding the preprints of the GL conference series to the repository. By providing full text access to the conference papers, it offered open access to international research on grey (gray) literature.

In 2011 OpenSIGLE changed its platform and its name and was relaunched as the open repository OpenGrey. OpenGrey provided new features and new content. Thus, all former SIGLE bibliographical records were searchable again, free of charge, with OpenGrey.

Contents

OpenGrey¹ is a multidisciplinary European database, covering science, technology, biomedical science, economics, social science and humanities. The records are in English. They include bibliographical data, the SIGLE classification codes with their wording, and possibly an English abstract (starting in 1997). Each record has either an English title and/or English keywords. Document types include technical or research reports, doctoral dissertations, conference papers, official publications, and other types of grey literature.

The OpenDOAR directory entry on the OpenGrey Repository² indicates 1,014,842 metadata, giving access to the full text of 18,158 items (data provided by the Jisc CORE service).

Its content in March 2021 shown in figures:

| Scientific domain | Records |
|-------------------------------|---------|
| Sciences | 833,500 |
| Social Science and Humanities | 325,000 |
| Biomedical Science | 116,000 |
| Science | 170,000 |
| Technology | 222,500 |

Table 1. Scientific domain of the OpenGrey records (multiple entries)

¹ Former OpenGrey website: <http://www.opengrey.eu/>

² OpenDOAR <https://v2.sherpa.ac.uk/id/repository/2181>

| Language | Records |
|----------|---------|
| English | 568,494 |
| French | 263,215 |
| German | 157,833 |
| Italian | 15,966 |
| Spanish | 9,170 |
| Czech | 4,576 |
| Latvian | 3,238 |
| Dutch | 2,080 |

Table 2. Most important languages of the OpenGrey records

| Document type | Records |
|--------------------------|---------|
| Theses and dissertations | 534,648 |
| Reports | 189,646 |
| Conference papers | 20,521 |
| Miscellaneous | 244,531 |

Table 3. Most important document types of the OpenGrey records

| Country | Records |
|----------------|---------|
| United Kingdom | 451,896 |
| France | 283,689 |
| Germany | 199,322 |
| Italy | 31,350 |
| Russia | 19,536 |
| Spain | 11,286 |
| Czech Republic | 5,927 |
| Belgium | 5,252 |
| Latvia | 3,962 |
| Portugal | 1,782 |
| Slovakia | 851 |

Table 4. Geographical origin: most important countries of the OpenGrey records

| Table | Rows | Count |
|--------------|------|-----------|
| handle2url | 2 | 30 |
| item | 8 | 942,052 |
| item2subject | 2 | 1,501,299 |
| modif | 5 | 0 |
| news | 4 | 0 |
| partner | 5 | 15 |
| subject | 2 | 274 |
| translation | 3 | 85 |

Table 5. Information about the 8 database tables

Sample SIGLE Record

| |
|--|
| Accession Number |
| 98:4462DE SIGLE |
| Title |
| Anaerobic co-treatment of biological waste materials and sewage sludge. Final report. Gemeinsame anaerobe Behandlung von Bioabfall und Klaufschlamm. Abschlussbericht. |
| Author |
| Risse, H. |
| Corporate Source |
| Technische Hochschule Aachen (DE). Lehrstuhl und Inst.fuer Siedlungswasserwirtschaft |
| Funding Organization |
| Oswald-Schulze-Stiftung, Gladbeck (DE) |
| Source |
| Jan 1998. 83 p. |
| Availability |
| Available from TIB Hannover: F98B1076+a. |
| Document Type |
| Numerical Data |
| Country |
| Germany, Federal Republic of |
| Language |
| German |
| Abstract |
| The main goals of co-fermentation are sewage sludge stabilisation, production of a cyclable product and minimisation of process emissions. This research project focused on anaerobic wet fermentation. Criteria of evaluation were the quantity and quality of the resulting fermented sludge, biogas, and process water. In view of the scarcity of data on optimum process conditions, the investigations were to provide information, e.g.: Optimum ratio of biological waste to sewage sludge, influence of process conditions (temperature, charge volume, flow rate) on biogas production and process water burden. (orig.) |
| Classification Code |
| 10U Biomass energy |
| Terms |
| BIOLOGICAL WASTES; SEWAGE SLUDGE; ANAEROBIC DIGESTION; METHANE; CHEMICAL OXYGEN DEMAND; AMMONIA; WASTE WATER; PHOSPHORUS; METALS; TEMPERATURE DEPENDENCE; EVALUATION; EXPERIMENTAL DATA. |

Figure 1. Sample of a former SIGLE record (source: Wikipedia)³

³ Wikipedia https://en.wikipedia.org/wiki/System_for_Information_on_Grey_Literature_in_Europe