

Grey Literature and the Role of the Library of the AS CR, v. v. i.

The 3rd Seminar on Providing Access to Grey Literature

December 8th, 2010 - NTL, Prague, Czech Republic

Jana Doleželová, Library of the AS CR, v. v. i.

dolezelova@iach.cz

Zdeňka Chmelařová, Library of the AS CR, v. v. i.

chmelarova@lib.cas.cz

Kateřina Durecová, Library of the AS CR, v. v. i.

durecova@lib.cas.cz

The Library of the Academy of Sciences (AS CR Library) has been a coordinator for data collecting of publications and other results of basic research in AS CR since 1993. The main purpose is to provide evidence and enable access to data from the database ASEP (Register of Publication Activity of the AS CR) and to provide actual statistical and analytical output for AS CR management. The basic structure of data has been extended and modified from 1993 according to demands of the AS CR institutes and also according to requirements of the collaborating systems.

It has already been decided to export the selected ASEP bibliographic records to the database SIGLE (System for Information on Grey Literature in Europe) in 2000. It was necessary to add some new types of documents. Off-line data collecting was substituted by on-line collecting in 2005 and bibliographic records of the database ASEP were made accessible on-line through the system ARL (Advanced Rapid Library). The Library was ready to enter quality data into the database SIGLE at this time, but the SIGLE discontinued its activity.

The bibliographic records have been gathered in the database ASEP since 1993. According to the interior AS CR instruction there is a duty for each institute to enter the selected types of results into the database. Bibliographies of our fifty four institutes are complete – as concerns fundamental scientific outputs. The database contains more than 185000 records and the annual increase for the last 5 years has been 11500 records per year in average. The evaluation of AS CR institutions in 2010 was based on these data. The records are displayed and searchable in an on-line catalogue and there is also an analytical overlay - webpages ASEP Analytics - with graphic, statistical and bibliographic outputs in different levels - whole institute, authors or the departments of institutes. Selected records are regularly imported to the Information Register of R & D results for evaluation. The grey

literature records metadata are harvested to the National Repository of Grey Literature in National Technical Library.

The database contains records of results important for evaluating institutes: journal article, monograph, monograph chapter, specialized book review, specialized translation, patent document, results of applied research – pilot plan, verified technology, breed, utility model, specialized map, norm and directive, certified methodology, software, prototype and functional sample, grey literature records – research reports, conference papers, conference proceedings, internal reports, theses, presentations and posters.

The database also contains records of conference arrangements, exhibitions and workshops, records of abstracts and newspaper articles and other results.

To enable harvesting metadata and full-texts from ASEP to NRGL it was necessary to take four steps:

1. System ARL has to support protocol OAI-PMH (The Open Archives Initiative Protocol for Metadata Harvesting) for the purpose of providing metadata and enabling links to full-texts in the AS CR Library repository for grey literature. Technical support is provided by Cosmotron Company, the developer of ARL who is also our system administrator.
2. To select document types of grey literature, add new document types and extend the data structure to correspond with the data structure in NRGL.

The records of grey literature in the database ASEP and their number in years 2000-2010, published in the Czech Republic, are displayed in the following table:

Research reports:	2686
Conference proceedings:	630
Conference papers:	10863
Internal reports:	420
Theses:	410
Presentations, Posters:	810

It used to be possible to store only research reports in the database. We added new types of reports – annual reports, technical reports, interim reports of the project, final reports of the project, grant reports, survey reports and statistical reports. We added more types of conference material as conference programs.

The system ARL is based on the UNIMARC format and it follows the simplified version of Anglo-American Cataloguing Rules AACR2. That is why the data structure in ASEP corresponded with the requirements of NRGL – most of the mandatory fields in our system have already been defined. It was necessary to add only three new fields – type of document NRGL, copyright and access limitation.

Types of documents NRGL – types of documents in ASEP are marked by letter codes. The data structure was extended to help us exactly determine the type of document in NRGL.

Copyright – it is specified how to handle the document – whether it is possible to make a copy, make it public, translate it, use it for educational purposes etc. The field is repeatable and it is possible to choose a combination of these options:

Document is protected according to copyright No.121/2000 in the collection of laws:

Licence Creative Commons Author's name 3.0 The Czech Republic

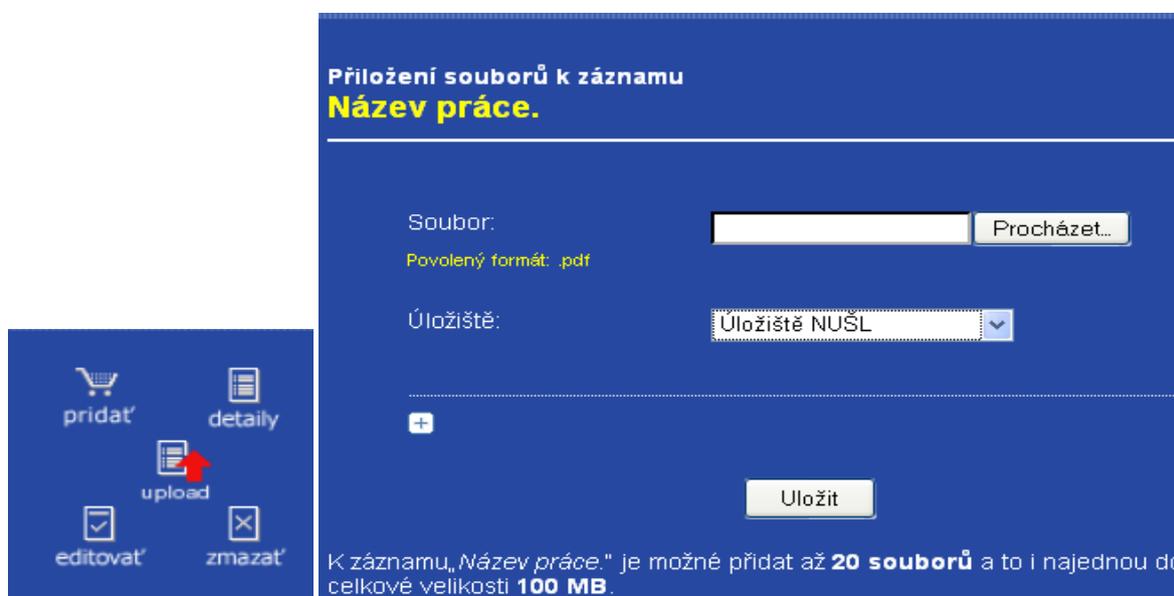
Licence Creative Commons Author's name - retain licence 3.0 The Czech Republic

Licence Creative Commons Author's name - not interfere with the work/publication 3.0 The Czech Republic

Licence Creative Commons Author's name - do not use the work/publication for commercial purposes - retain licence 3.0 The Czech Republic

Restriction of access – it is possible to choose options – a document can be accessible to public, accessible when having permitted IP address, protocol shibboleth or protocol eduroam, or a document is not accessible at all.

3. To add new function to ARL, that enables attaching full-text to a bibliographic record. Data are entered into the system ARL through a web-form. There is a new press-button, which enables uploading of a full-text to a record. A data administrator fills in this simple form:



Přiložení souborů k záznamu
Název práce.

Soubor: Procházet...

Povolený formát: .pdf

Úložiště: Úložiště NUSL

Uložit

K záznamu „Název práce.“ je možné přidat až **20 souborů** a to i najednou do celkové velikosti **100 MB**.

It is possible to add maximum number of 20 files, in total size of 100 MB. Prescribed format is pdf format. An administrator can see a list of attached files and information about their size. The files are saved to repository of grey literature in the Library of the AS CR. Metadata and full-text are prepared to be harvested to the archive NRGL.

4. To prepare a training system for administrators of grey literature data in ASEP. Some institutes of AS CR are interested in storing records with full-text in the digital repository of NRGL.

Fourteen institutes subscribed a cooperation contract with NLT about providing records and full-text to NRGL:



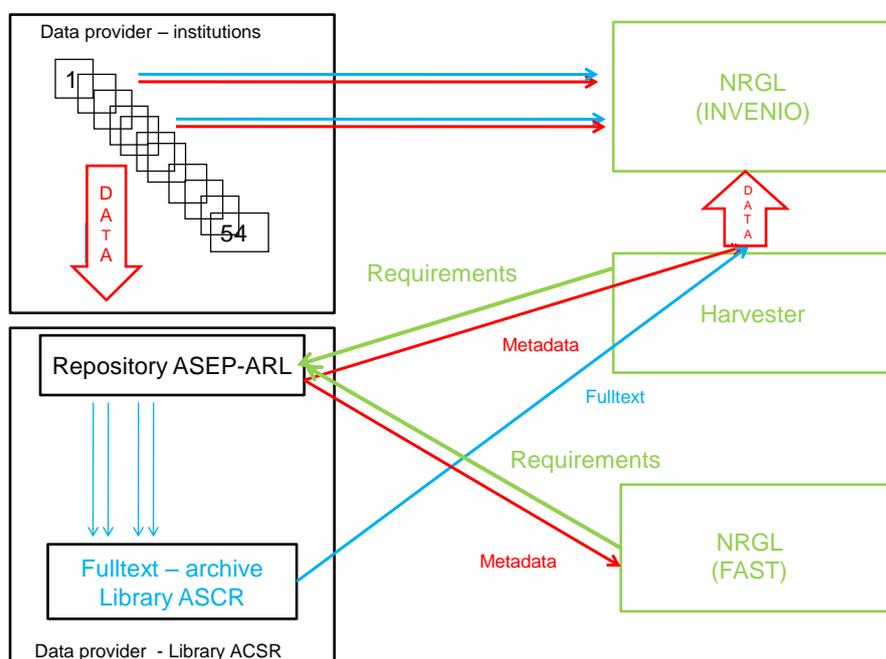
Administrators of institutes can decide if they firstly save data to the database ASEP (data will be prepared for automatic harvesting to NRGL) or if they enter data directly to NRGL.

AS CR library staff members are prepared to train these administrators of grey literature. it will be necessary to distribute competences for ASEP administrator and NRGL administrator (in case it is not the same person) to avoid duplication of the records.

We created the web pages of ASEP-NRGL where administrators can find information about the whole process of preparing data for NRGL. There will be definitions of document types in NRGL, instructions how to create a bibliographic record and a possibility to register for training. We will start the routine administration of grey literature data of ASEP in the beginning of the year 2011.

There are 6057 records of grey literature harvested from ASEP database in the grey literature repository.

A possible method of providing grey literature data from the provider AS CR to NRGL:



NRGL harvester will send the requests to ARL-ASEP repository, ARL system provides metadata with links to the document full-texts in the Library AS CR repository, the NTL system deposits metadata and full-texts to the NRGL repository.

The AS CR institutes do not depend on each other. They can set up their own way of collecting data, which is the most suitable to them and also is easy to change it in case of changing requirements or staff. Grey literature originates concurrently with other significant valuable results of the AS CR institutes. We think that collection should be provided by libraries of institutes. The system of data collecting was set to easy providing of grey literature and automatically harvesting of metadata and full-texts to the NRGL repository. The number of records depends on each institute. We wish that the people who will work in libraries will be aware of the importance of grey literature and will archive it. There is no difference between collecting data in the

database ASEP or straight in the repository NRGL. The Library of AS CR will support and enable data harvesting of the AS CR institutes – that is its role.

Websites for the administrators of grey literature of ASEP:

<http://www.iach.cz/KNAV/nusl.html>

On-line catalogue ASEP:

<http://www.library.sk/i2/i2.entry.cls?ictx=cav&language=3>

NRGL:

http://nrgl.techlib.cz/index.php/Main_Page

ASEP Analytics:

<http://www.lib.cas.cz/arl/?lang=en>

Information Register of R & D results:

<http://www.isvav.cz/prepareResultForm.do>