KNIHOVNA
—
AKADEMIE VĚD ČR

DATA DEPOSIT INTO THE ASEP REPOSITORY

Zdeňka Chmelařová, Jana Doleželová

Academy of Sciences Library

Data repository

- Sharing and long-term preservation of research data
- Internal survey – a need for data repository
- Inspiration - good practice (DataShare, Zenodo, Lindat/Clarin)
- Academy of Sciences Library (KNAV) provides administration of the ASEP database in ARL system (Advanced Rapid Library)
- Data repository CAS is ASEP superstructure

ASEP

Bibliographic database, repository

ASEP – the evidence of the scientific research results of the Czech Academy of Sciences institutes (since 1993)

Full texts repository (from 2012 allows full text storage)

Data repository (since 2017 allows data storage)

Number of bibliographic records	268 152
Number of results types	30
Average annual increase	about 12 000
Number of full texts	12 952

Czech Academy of Sciences (53 institute, 3 scientific disciplines)

Data repository ASEP
institutional multidisciplinary repository

- Various data files (multidisciplinary repository)

- Data files (datasets) with metadata (description)

- Dataset deposit

  1. dataset is stored in **ASEP**

  2. data record stored in ASEP with link to **another repository**

- Dataset access (open access, with embargo or on request)

- Linking between bibliographic and data records

- Preparation of data files

Data repository
datasets storage workflow

KNIHOVNA

AKADEMIE VĚD ČR

1. Preparation of scientific data and documentation
The authors must decide which data they want to store, prepare datasets and documentation.

2. Registration/Log into myASEP account
The depositors (authors, administrators of scientific teams and other authorized persons) have to create myASEP account, log in, complete a data form (create metadata) and save datasets.

3. Transfer of data records to control and to publish
The depositor can edit data records and datasets until he confirms the license agreement. After that he can no longer edit or delete the records or datasets and they are transfered to the account of the institutional data administrator. Thereafter he will control and publish them.

4. Control of data records and datasets
The institutional data administrator will check if all fields in the data record are filled in correctly as well as if the dataset is attached in the right way. In case that he has some comments, he will contact the depositor.

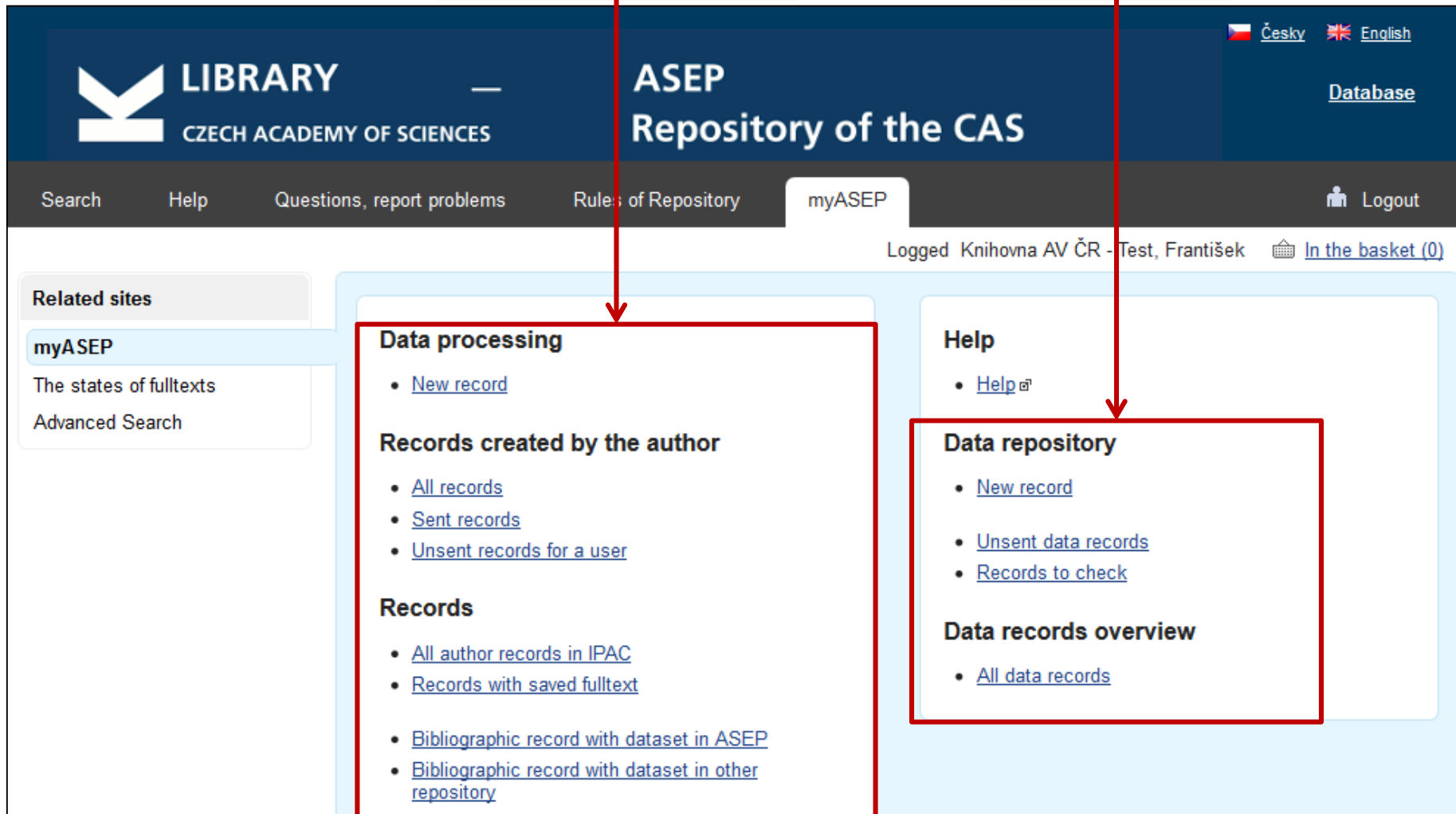**5. Handle identificator and publishing in on-line catalogue**
The handle identificator is assigned to verified data and those are published in the on-line catalogue. The institutional data administrators or the depositors can no longer edit them.

Data repository

MyASEP account - author

Data repository

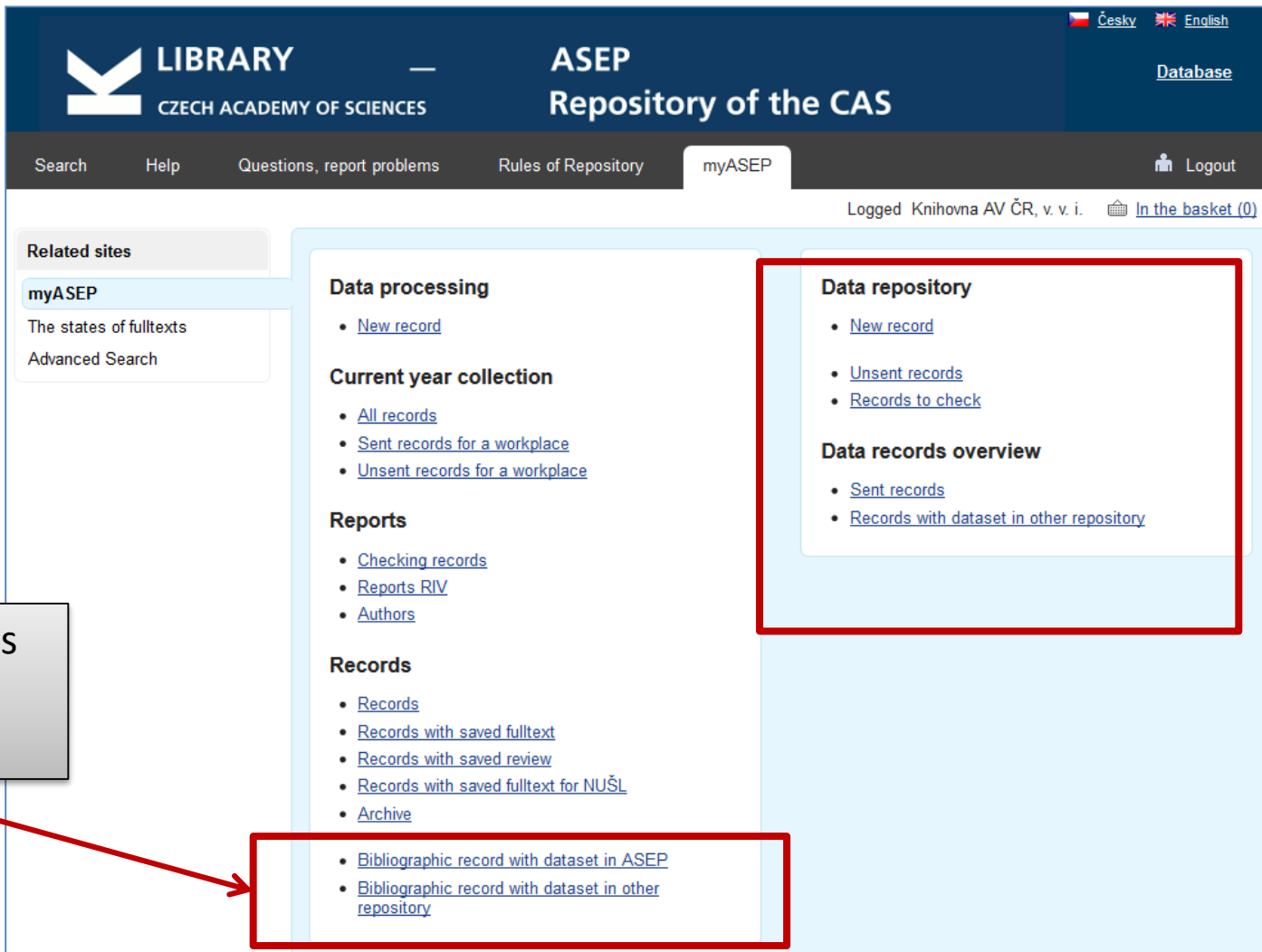## MyASEP account - system administrator

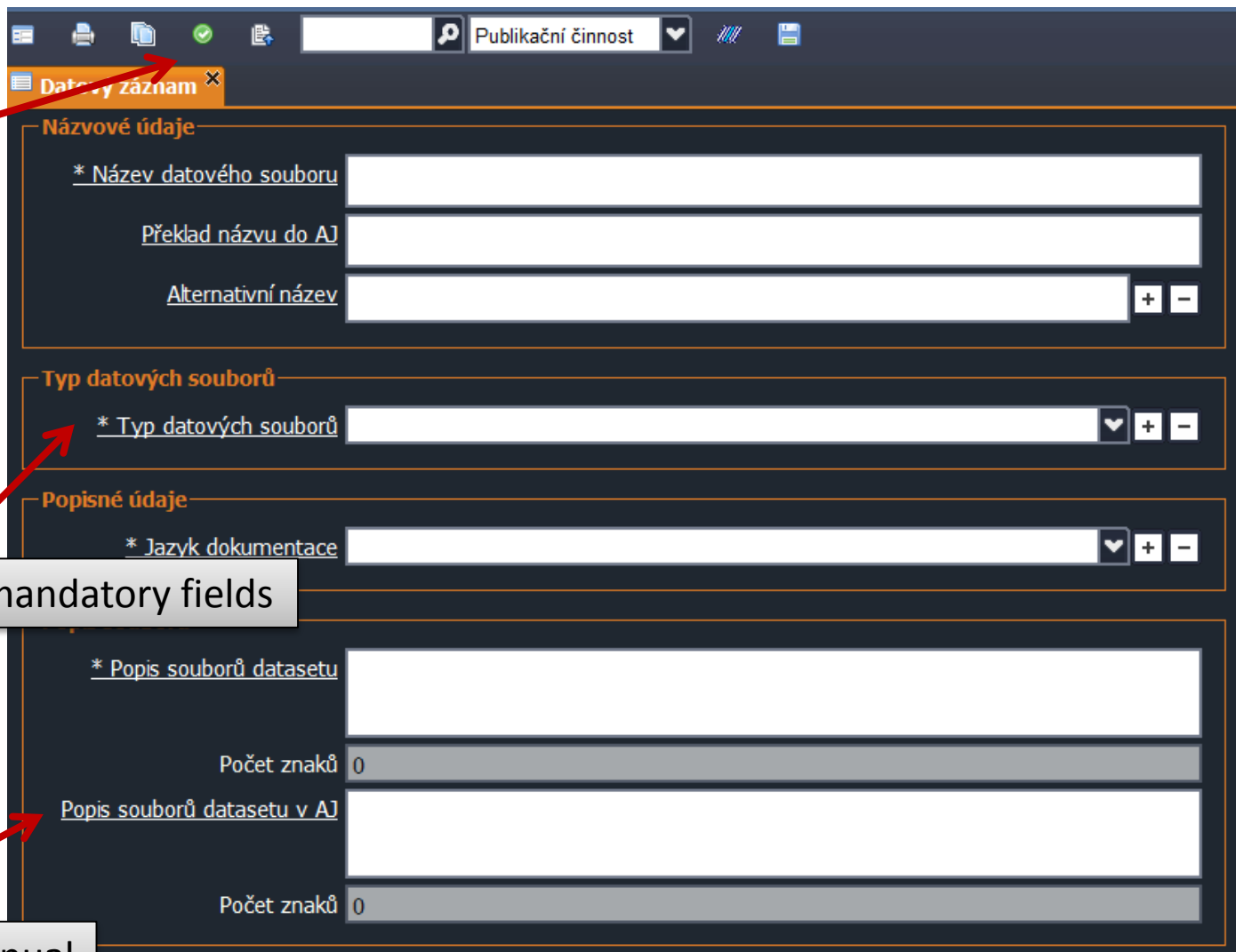Bibliographic records with linking to datasets

Data repository
Metadata
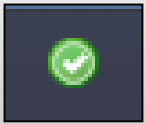
Metadata fields selection - the pursuit of maximum data integrity, but with respect to scientists/authors who will create metadata

- Data structure based on **international standards** *(UNIMARC, Unicode (UTF-8), ISO country and languages codes)*

- **Authority records of authors** – enables unique identification of the author, his affiliation and outputs *(ID authors from WOS, Scopus)*

- KNAV's **own authorities** *(projects, conferences, proceedings)*

- Metadata are filled in the webforms

- **Mandatory fields** are marked *(for instance author, dataset name and translate to eng, datasets description, dataset type, documentation language, keywords, user license setting and access to datasets)* and **conditionally mandatory** *(project, link to ASEP publication…)*

Data repository - the webform

Icone

runs record check

The asterisk indicates mandatory fields

Field name - link to manual

Data repository
Transfer of metadata and dataset to control, consent to storage agreement in the ASEP repository

Storage agreement
1. the author must have required rights to store the data (permission from co-authors)
2. may not store/public sensitive data (identification numbers, names, telephone numbers, etc.)
3. must enter a license for reuse datasets

Předání metadat a datasetu ke kontrole

Upozornění: Souhlas vyplňte, teprve až budete chtít předat metadata a dataset ke kontrole zpracovateli. Metadata a dataset si pečlivě zkontrolujte, protože po potvrzení souhlasu a následném uložení záznamu již nebudete moci metadata ani přílohu datasetu měnit!
Souhlasím s dohodou o zveřejnění datasetu v repozitáři ASEP **odkaz na text dohody**
a předávám datový záznam a datasety ke kontrole.

Souhlasím ☑

Data repository
dataset

ASEP **dataset** - a set of files, that may contain **research data**, **documentation** with important user information or license if the Creative Commons license is not sufficient

- Current **size limit** is 2 GB per file and 20 GB per a single dataset. Upon agreement with the repository administrator it will be possible to save larger files
- When selecting suitable format we recommend standard formats that guarantee a long-term protection

Data repository

The example of a data record with datasets stored in ASEP

0478893 - KNAV-K (2017) **Scientific data**
Novák, František
Phylogeny of Miliusa. Magnoliales: Annonaceae: Malmeoideae: Miliuseae.
[Phylogeny of Miliusa.]

Materials and Methods Molecular phylogenetic analysis. All accessions belong to the subfamily Malmeoideae (Appendix). Twenty-seven accessions of Miliusa covering the entire morphological variation known comprise the ingroup. The outgroups consist of 14 accessions, 11 of which were included as representatives of related genera in the tribe Miliuseae. Seven plastid markers, .e. rbcL exon, trnL intron, trnL-F spacer, matK exon, ndhF exon, psbA-trnH spacer, and ycf1 exon, were amplified (see Table 1 for the number of included, variable, and parsimony informative characters). In total, 7033 characters, including six separately coded indels were included in the analyses
[Materials and Methods Molecular phylogenetic analysis. All accessions belong to the subfamily Malmeoideae (Appendix). Twenty-seven accessions of Miliusa covering the entire morphological variation known comprise the ingroup. The outgroups consist of 14 accessions, 11 of which were included as representatives of related genera in the tribe Miliuseae. Seven plastid markers, .e. rbcL exon, trnL intron, trnL-F spacer, matK exon, ndhF exon, psbA-trnH spacer, and ycf1 exon, were amplified (see Table 1 for the number of included, variable, and parsimony informative characters). In total, 7033 characters, including six separately coded indels were included in the analyses]
Keywords : phylogenetic analysis * phylogenetic analysis
R&D Projects: GA ČR(CZ) GAP503/10/1270
Permanent Link: http://hdl.handle.net/11104/0274948

ASEP publications:
Effect of Chalara longipes on decomposition of humis acids from Picea abies needle litter

Dataset :
    License: BY-SA - Uveďte původ + Zachovejte licenci

File		Download	Size	Commentary
🔓	readme.txt	0	7 B	
🔓	MaterialsandMethods_Molecular_phylogenetic_analysis.zip	ℹ 0	109.8 KB	

```
Archive:  81/064594311.zip
  Length      Date    Time    Name
--------- ---------- -----    ----
        0  11-01-2013 11:41   arl_pos1/
     4710  09-23-1999 18:52   arl_pos1/chang.ico
    31341  11-01-2013 11:38   arl_pos1/client_cav.
     1108  03-29-2011 11:35   arl_pos1/client_jib.
    17825  03-29-2011 11:35   arl_pos1/client_lc_u
    20565  03-16-2012 13:49   arl_pos1/client_nkp_
     1085  04-24-2012 14:25   arl_pos1/client_uk.i
      102  06-12-2012 14:26   arl_pos1/CONFIG.INI
   650240  07-13-2009 10:51   arl_pos1/cs_lib.dll
        0  11-01-2013 09:10   arl_pos1/help/
     1850  07-20-2005 11:58   arl_pos1/help/zclien
    12921  05-09-2006 17:05   arl_pos1/help/ZCLIEN
  6818027  07-20-2005 12:35   arl_pos1/help/ZCLIEN
   888832  05-03-2006 20:57   arl_pos1/iconv.dll
  1545245  04-10-2009 12:16   arl_pos1/libeay32.dl
   961536  07-21-2006 15:16   arl_pos1/libxml2.dll
   161280  07-21-2006 16:26   arl_pos1/libxslt.dll
```

Data repository

The example of a data record with dataset stored in another repository

1.
0478893 - KNAV-K (2017) **Scientific data**
Novák, František
Phylogeny of Miliusa. Magnoliales: Annonaceae: Malmeoideae: Miliuseae.
[Phylogeny of Miliusa.]

Materials and Methods Molecular phylogenetic analysis. All accessions belong to the subfamily Malmeoideae (Appendix). Twenty-seven accessions of Miliusa covering the entire morphological variation known comprise the ingroup. The outgroups consist of 14 accessions, 11 of which were included as representatives of related genera in the tribe Miliuseae. Seven plastid markers, .e. rbcL exon, trnL intron, trnL-F spacer, matK exon, ndhF exon, psbA-trnH spacer, and ycf1 exon, were amplified (see Table 1 for the number of included, variable, and parsimony informative characters). In total, 7033 characters, including six separately coded indels were included in the analyses
[Materials and Methods Molecular phylogenetic analysis. All accessions belong to the subfamily Malmeoideae (Appendix). Twenty-seven accessions of Miliusa covering the entire morphological variation known comprise the ingroup. The outgroups consist of 14 accessions, 11 of which were included as representatives of related genera in the tribe Miliuseae. Seven plastid markers, .e. rbcL exon, trnL intron, trnL-F spacer, matK exon, ndhF exon, psbA-trnH spacer, and ycf1 exon, were amplified (see Table 1 for the number of included, variable, and parsimony informative characters). In total, 7033 characters, including six separately coded indels were included in the analyses]
Keywords : phylogenetic analysis * phylogenetic analysis
R&D Projects: GA ČR(CZ) GAP503/10/1270
Permanent Link: http://hdl.handle.net/11104/0274948

Data in other repository: Zenodo

Publication ASEP:
🔍 Effect of Chalara longipes on decomposition of humis acids from Picea abies needle litter

Dataset :
  License: BY-SA ⧉ - Uveďte původ + Zachovejte licenci

To the basket
Edit
Citation/Review
Upload dataset
Publication ASEP

Data repository

The example of link from data record to bibliographic record

0395512 - KNAV-K 2014 RIV AT eng C - Conference Paper (international conference)
Doleželová, Jana - Chmelařová, Zdeňka
Asep Analytics. A source for evaluation at the Academy of Sciences of the CR.
Proceedings of ISSI. Volume 2. Vienna: Austrian Institute of Technology, 2013, s. 1874-1876. ISBN 978-3-200-03135-7. ISSN 2175-1935.
[International Society of Scientometrics and Informetrics Conference /14./. Vienna (AT), 15.07.2013-19.07.2013]
Institutional support: RVO:67985971
Keywords : evaluation * Academy of Sciences Library
Subject RIV: IN - Informatics, Computer Science
http://www.issi2013.org/Images/ISSI_Proceedings_V
Permanent Link: http://hdl.handle.net/11104/0223530
Dataset in ASEP:
  Analytika ASEP. Zdroj pro evaluaci v Akademii věd Č
  Datový pokusný individiuální

File	Downloa
Asep_Analytics_ISSI2013.pdf	19

- To the basket
- Edit
- Citation/Review
- Administration of
- Upload attachmen
- RIV
- WOS
- SCOPUS

0474236 - KNAV-K (2017) **Scientific data**
Doleželová, Jana - Chmelařová, Zdeňka
Analytika ASEP. Zdroj pro evaluaci v Akademii věd ČR - TEST. Poster ASEP - ISSI 2013.
[ASEP Analytics. A source for evaluation at the Academy of Sciences of the CR - TEST.]

Dataset obsahuje poster a obrázky ve formátu jpg. Poster, který byl prezentován na mezinárodní konferenci 14. International Society of Scientometrics and Informetrics Conference ve Vídni.
[Dataset - consists of jpg files, poster and pictures, presented on conference the 14th International Society of Scientometrics and Informetrics Conference in Vienna (AT).]

Keywords : ASEP
Institutional support: RVO:67985971
Permanent Link: http://hdl.handle.net/11104/0272012

ASEP publications:
  Asep Analytics. A source for evaluation at the Academy of Sciences of the CR
  ISSI 2013

Dataset :
  License: BY-NC-ND

File		Download	Size	Commentary
obrazky.zip	i	3	1.2 MB	
poster-aa.pdf		5	10.5 MB	

Data repository
Summary

KNAV offers to the CAS institutes:
- methodology and space for storing data
- long-term protection and datasets archiving
- user support

The system is open and will be expanded and modified based on user requirements.

In the **future**, we would like to:
- store **large data files** and provide long-term preservation
- join the Register of scientific data repositories **Re3d**
- meet the conditions for the export of data records in the **Data Citation Index** and **OpenAIRE** database
- obtain a **certification** of digital repositories - a repository certification is not only a formal proof that the repository meets the necessary criteria but also a tool for self-checking the correct functionality of the repository

Thank you for your attention

Contact: arl@lib.cas.cz