

Datový repozitář ASEP

Rudolf Sýkora, Knihovna AV ČR, v. v. i.

23.10. 2023

Open Access Week 2023

Osnova

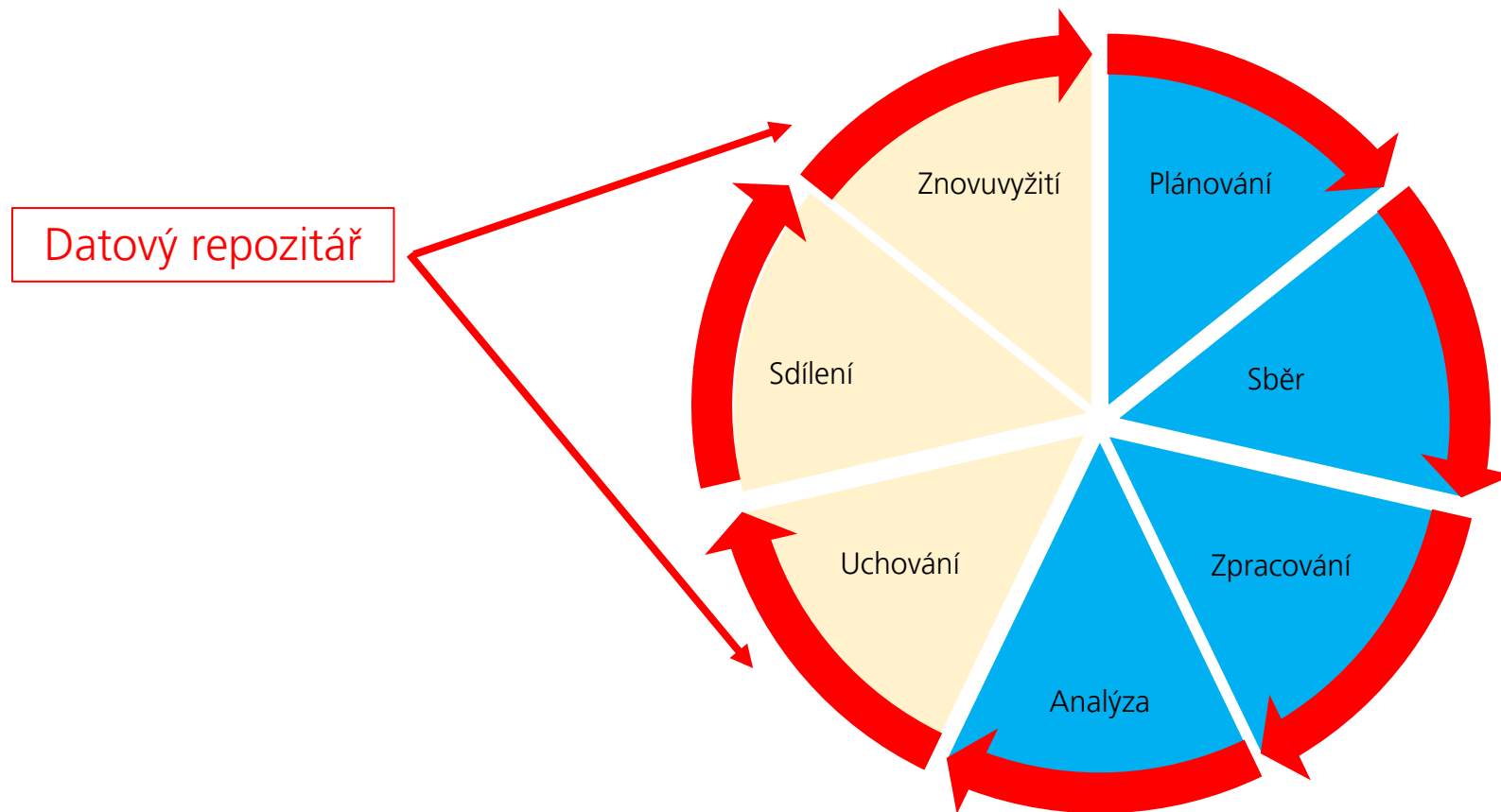
- Datové repozitáře
- Výběr repozitáře
- Charakteristika DR ASEP
- Formáty a licence
- Vytvoření záznamu v ASEP
- Další rozvoj repozitáře

Datové repozitáře

Datový repozitář – online databáze pro ukládání vědeckých dat.

- Datové repozitáře nabízejí nejvhodnější a nejbezpečnější způsob, jak dlouhodobě **uchovávat, sdílet a uspořádat výzkumná data**.
- Řada poskytovatelů financí (např. Evropská komise v projektech [Horizon Europe](#)) vyžaduje uložení dat do důvěryhodného repozitáře.
- Ukládání vědeckých dat vyžadují i někteří čeští poskytovatelé financí ([TA ČR v projektech Kappa](#), Projekty z [OP JAK](#)).

Životní cyklus výzkumných dat



Výběr datového repozitáře

Při výběru vhodného repozitáře pro uložení dat je možné se řídit např. těmito doporučeními od [OpenAIRE – How to find a trustworthy repository for your data](#):

1. Pokud existuje vhodný oborový repozitář, je nejlepší využít ten,
2. další na řadě je institucionální repozitář,
3. můžete využít univerzální “catch-all” repozitář Zenodo, provozovaný CERNem,
4. nebo vyhledejte vhodný repozitář v některém z mezinárodních registrů repozitářů: [Re3data.org](#), [OpenDOAR](#) nebo [FAIRsharing](#)

Výběr datového repozitáře – Re3data

Je možné vyjít z **registru repozitářů** re3data.org, kde je u každého repozitáře piktogram, který může pomoci při výběru.

ASEP Repository

Repository of Czech Academy of Sciences

Subject(s)

Humanities and Social Sciences Life Sciences Natural Sciences Engineering Sciences

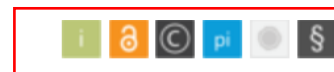
Content type(s)

Standard office documents Images Audiovisual data Archived data Standard office documents
Scientific and statistical data formats Raw data

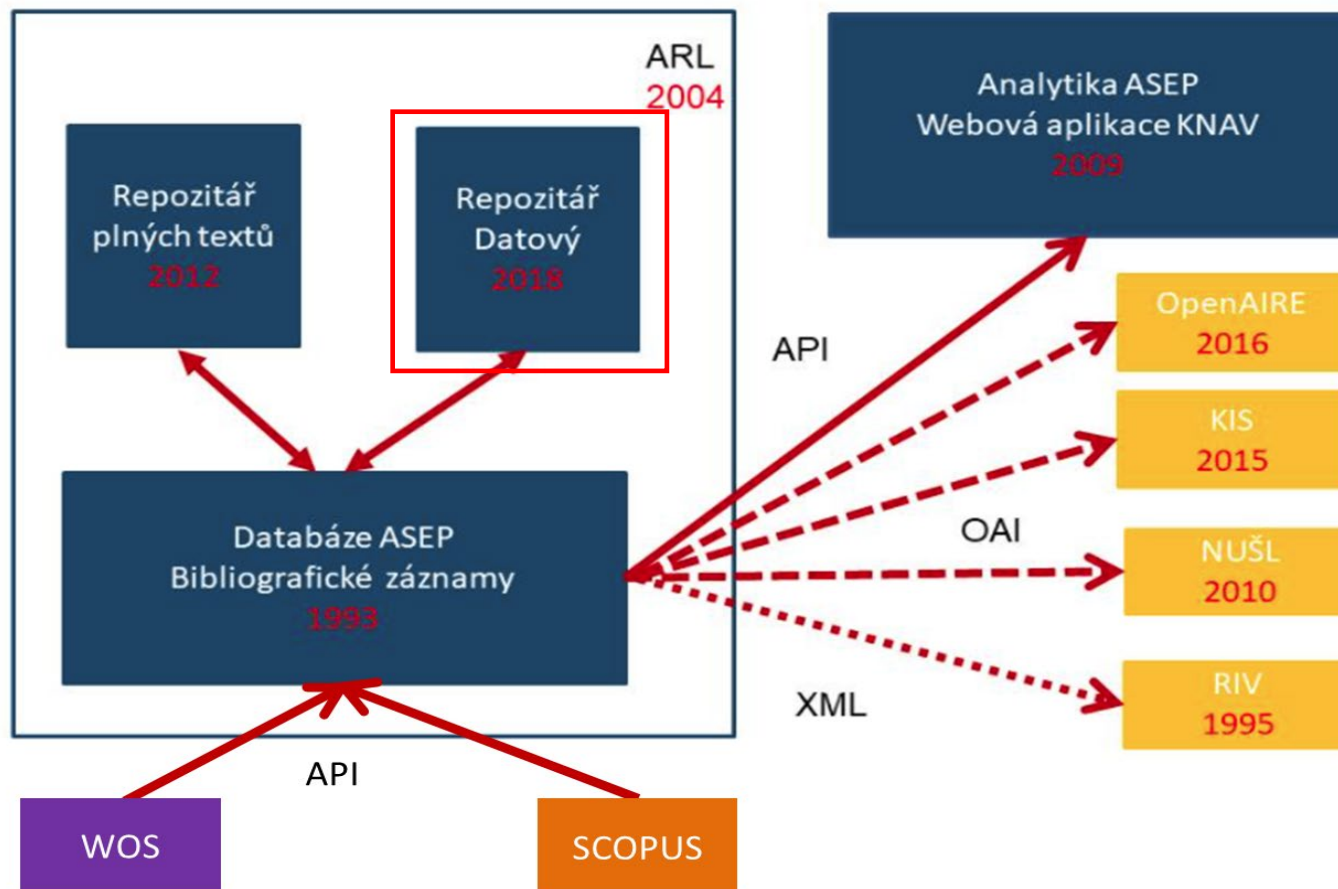
Country

Czech Republic

The ASEP Data Repository is an institutional multidisciplinary on-line repository that stores scientific outputs - bibliographic records, full texts and datasets of the institutional authors from The Czech Academy of Sciences. The repository is hosted by Library of the Czech Academy of Sciences. Data that are stored in database are accessible in the on-line catalogue. Each dataset has its own description and metadata according to international standards.



System ASEP



ASEP v kontextu Open Science

- Zveřejnění v online katalogu (plné texty a datasey v otevřeném přístupu/na vyžádání).
- **Institucionální** = Poskytovateli financí je považován za **důvěryhodný**.
- Spolehlivá **dlouhodobá archivace**.
- Počítáme s připojením do EOSC.CZ – Národní datová infrastruktura.
- Data v ASEP splňují **FAIR principy** (**F**indable), přístupná (**A**ccessible), čitelná a strojově zpracovatelná (**I**nteroperable) a znovupoužitelná bez změny své povahy (**R**eusable).



Licence a přístup k datovým souborům

U každého datového souboru musí být uvedeno, jakým způsobem může uživatel s daty pracovat a jak budou data přístupná.

Licence

Bud' si autor vybere z licencí **Creative Commons** (CC) nebo přiloží **licenční smlouvu**.

Přístup k datovým souborům

- **Veřejně přístupný** – soubory jsou ihned přístupné.
- **Přístupný pro ústav** – soubory jsou přístupné pro ústav, pro další uživatele na vyžádání.
- **Přístupný s embargem** – soubory jsou přístupné až po určité době.
- **Veřejně nepřístupný** – soubory jsou na vyžádání.



Velikost a formáty souborů ukládaných do ASEP

- Maximální **velikost jednoho** ukládaného souboru je **10 GB**.
- Celková maximální velikost uložených souborů u jednoho datasetu je **50 GB**.
- Aby byl zajištěn dlouhodobý přístup a využitelnost dat v repozitáři, je vhodné používat **standardní formáty**, které zaručují dlouhodobou ochranu.
- Důležitý je **podrobný metadatový popis** uložených souborů a jejich formátů včetně informace o programech pro otevření.



Datový záznam - struktura

Povinné	Podmíněně povinné	Nepovinné	Automaticky přiřazené
Název datového souboru	Překlad názvu do angličtiny	Alternativní název souboru	Kontaktní osoba
Popis datasetu	Popis datasetu v angličtině	Odkaz	Handle
Autor	Odkaz na soubor v jiném repozitáři	Poznámka	Vlastník záznamu
Licence	Datum ukončení embarga	Lokalizace	Datum vytvoření
Typ datových souborů	Návaznost na projekty	Časové období	Datum zveřejnění
Jazyk dokumentace	Odkaz na publikace v ASEP		DOI
Přístup k souboru	Autorské identifikátory – Wos ID, ORCID		
Klíčová slova, FORD kategorie			
Souhlas se zveřejněním v katalogu			



Export metadatového záznamu

- Z ASEP je možný export XML v Datacite formátu.
- Připravujeme i export v JSON formátu.

Vybrat způsob exportu

Znaková sada:

UTF-8

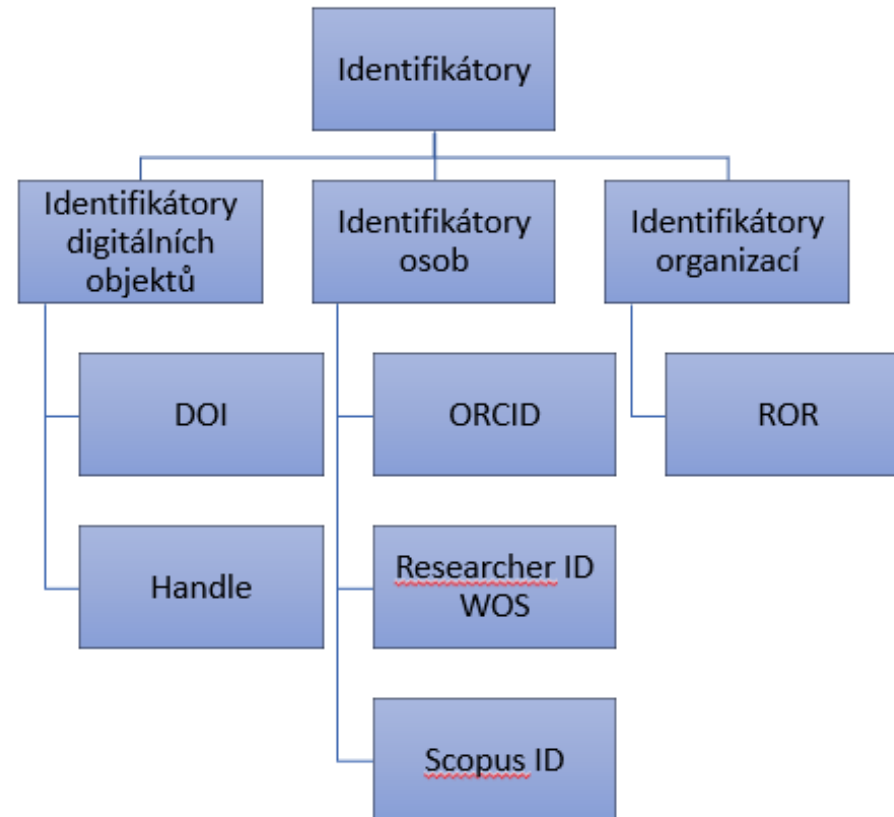
Formát

- EXCEL - csv
- EXCEL - csv (s duplicitními záznamy spolupracujících pracovišť)
- EXCEL - csv - datové záznamy
- XML - Popis polí ASEP [↗](#)
- RIS - exportní formát RIS kompletní
- RIS - ASCII
- XML Datacite - exportní formát XML do Datacite
- Řádkový MARC - [MARC](#) [↗](#)

Exportovat



Persistentní identifikátory v DR ASEP



Uložení datasetu, provázání bibliografických a datových záznamů

Možnosti uložení datasetu

1. Datový záznam i dataset je uloženy v ASEP.
2. Datový záznam je uložen v ASEP a dataset je uložen na jiném úložišti. (pokud úložiště nemá možnost popisu uložených datasetů)

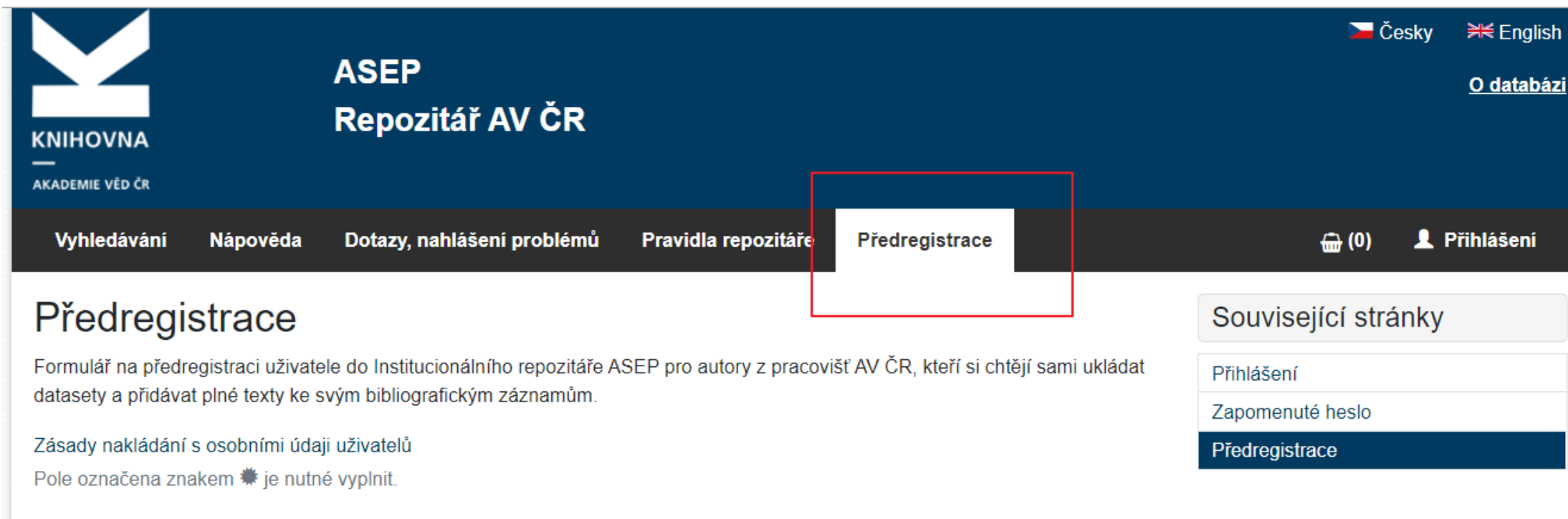
V **bibliografických záznamech** lze vytvořit odkaz na výzkumná data uložená na různých úložištích. Je tu možnost provázat publikace a data uložená v ASEP.



Kdo může ukládat data do DR ASEP?

- **Ústav** AV ČR musí mít s **KNAV** podepsánu **Rámcovou smlouvu o využívání repozitáře ASEP** pro ukládání datových souborů.
- Do repozitáře ASEP mohou ukládat datové soubory opatřené metadaty všichni **autoři**, kteří působí na pracovištích **AV ČR**, případně další **pověřené osoby – zpracovatelé, individuální zpracovatelé**. **Vkladatel** ukládá metadatové záznamy a datové soubory do repozitáře přes **individuální účet**.
- Individuální účet se vytvoří vyplněním jednoduché tabulky [předregistrace](#) na hlavní stránce on-line katalogu.
- **Zpracovatel pracoviště** (supevizor pracoviště) **registraci zkontroluje** a následně žadateli přijde **e-mail o aktivaci individuálního účtu**.

Vytvoření uživatelského účtu



The screenshot shows the header of the ASEP website. On the left is the logo of the National Library of the Czech Academy of Sciences (KNIHOVNA AKADEMIE VĚD ČR). In the center, the text reads 'ASEP Repozitář AV ČR'. On the right, there are language options for 'Česky' and 'English', and a link 'O databázi'. Below the header is a dark navigation bar with several menu items: 'Vyhledávání', 'Nápověda', 'Dotazy, nahlášení problémů', 'Pravidla repozitáře', 'Předregistrace', and 'Přihlášení'. The 'Předregistrace' item is highlighted with a red rectangular box. To the right of the navigation bar, there is a shopping cart icon with '(0)' and a user icon.

KNIHOVNA
AKADEMIE VĚD ČR

ASEP
Repozitář AV ČR

Česky English
[O databázi](#)

Vyhledávání Nápověda Dotazy, nahlášení problémů Pravidla repozitáře **Předregistrace** (0) Přihlášení

Předregistrace

Formulář na předregistraci uživatele do Institucionálního repozitáře ASEP pro autory z pracovišť AV ČR, kteří si chtějí sami ukládat datasey a přidávat plné texty ke svým bibliografickým záznamům.

Zásady nakládání s osobními údaji uživatelů
Pole označena znakem * je nutné vyplnit.

Související stránky

- Přihlášení
- Zapomenuté heslo
- Předregistrace**

Přihlášení MyASEP

Bibliografické záznamy/publikační činnost

[Nový bibliografický záznam](#)

Aktuální rok sběru

[Všechny záznamy zpracovatele](#)

[Záznamy akt. roku sběru \(odeslané\) zpracovatele](#)

[Rozepsané záznamy \(neodeslané\) zpracovatele](#)

Výstupy

[Kontroly dat](#)

[Výstupy RIV](#)

[Autority autorů a projektů](#)

Datový repozitář

[Nový datový záznam \(metadata\)](#)

[Rozpracované datové záznamy zpracovatele](#)

Datové záznamy autorů ke kontrole

Přehledy datových záznamů

[Odeslané datové záznamy v online katalogu](#)



Metadatový formulář

ASEP

Publikace, datasety

Datový záznam

Manuál 1. Záznam vyplňte, uložte. 2. Uložte dataset nebo odkaz na úložiště. 3. Záznam zkontrolujte (zelená ikonka nahoře), opravte chyby. 4. Pro předání záznamu zpracovateli, zaškrtněte checkbox, do té doby lze záznam editovat, přidávat datasety.

Identifikační číslo záznamu

[Názvové údaje](#) [Popisné údaje](#) [Data v jiném repozitáři](#) [Práva/licence](#) [Klasifikace](#) [Návaznost](#) [Publikace v ASEP](#) [Předání ke kontrole](#)

Autorské údaje 1.

* Příjmení

* Jméno

Pracoviště

Oddělení

Země

Názvové údaje

* Název datového souboru

Překlad názvu do AJ

Alternativní název


Typ datových souborů

* Typ datových souborů



Datový záznam v online katalogu

0539495 - ÚBO DATA Vědecká data

 Do košíku

Bryja, Josef

Genomic ddRAD datasets of Ethiopian Lophuromys - raw data.

The files contain pair-end reads that were submitted to ipyrad (Eaton & Overcast, 2020) for denovo assembly of ddRAD loci. The reads were demultiplexed and barcodes removed in Skewer (Jiang, Lei, Ding & Zhu, 2014) and pre-processing of read ends in PEAR (Zhang, Kobert, Flouri & Stamatakis, 2014). Every pair of reads was merged (with minimum required overlap of 10 nucleotides) and cut in two halves. If they were not successfully merged, we trimmed off the last 10 nucleotides from the reverse read. The PEAR-treatment was necessary to avoid poor quality mergers in a substantial proportion of loci. The data are used and analysed in Komarova et al., submitted to Molecular Ecology.

Klíčová slova: ddRAD * Lophuromys * Ethiopia

Obor OECD: Zoology

DOI: <https://dx.doi.org/10.57680/asep.0539495>

Handle: <http://hdl.handle.net/11104/0317292>

Datum publikování: 15.2.2021

Autor

Název


Popis

Obor, klíčová slova

Identifikátory

Soubory, licence Financování

Licence

Licence: BY-NC  - Uveďte původ + Neužívejte komerčně

	Název souboru	Staženo	Velikost	Komentář	Přístup
	Lophuromys_melanonyx.zip	10	1.2 GB	 Dělená	povolen

Soubory

Přístup



Data na jiném úložišti popsaná v ASEP


5.

0506600 - Úl **DATA** Vědecká data

Jeřábek, M. - Šrotýř, M. - Tvrdý, K. - Brabec, Marek - Kasanický, Ivan - Konár, Ondřej - Malý, Marek - Pelikán, Emil - Mlynářová, T. - Zaoral, O. - Termer, J. - Rak, J. - Nosek, V. - Vajda, P. - Sklenář, P.

PrePark.

[PrePark.]

 Do košíku

PrePark je softwarový statistický nástroj, který na základě dat z mýtného systému predikuje obsazenost jednotlivých parkovacích míst na dálnicích v ČR.

Klíčová slova: toll system * prediction * availability of parking spaces * highway

Grant CEP: GA TA ČR TA02031411

Institucionální podpora: RVO:67985807


Obor OECD: Transport engineering

DOI:

Handle: <http://hdl.handle.net/11104/0297820>

Datum publikování: 16.10.2019

Publikace ASEP:

 PrePark

Licence:

Jiná licence. Cena licence je stanovena na 100.000 Kč. Tvůrci mohou užívat výsledek k nekomerčním účelům, tj. k prezentačním a výukovým účelům, další na základě zvláštních licenčních ujednání. Smlouva o využití výsledku projektu "Zvýšení využití parkovací kapacity na dálnicích za pomoci predikčních modelů" uzavřená dne 19.12.2014.

Úložiště

Přístup

Komentář

[Úložiště Úl](#)

[vyžádat](#)

Propojení datového a bibliografického záznamu

0534139 - ÚBO DATA Vědecká data

[Bryja, Josef](#)

Anchored phylogenomics data used for phylogenomic analysis of Arvicanthini rodents.

The files represent datasets used in the paper: Mikula et al.: Nuclear phylogenomics, but not mitogenomics, resolves the most successful Late Miocene radiation of African mammals (Rodentia: Muridae: Arvicanthini). Molecular Phylogenetics and Evolution, under review.

Arvicanthini_40sp_377loci.nex = alignments of 377 nuclear loci obtained by anchored phylogenomic approach in 40 species of Arvicanthini rodents (partitioned nexus format)

Arvicanthini_40sp_377loci_genetrees.nwk = Bayesian gene trees (for 377 genes) used as input for the ASTRAL analysis (newick format)

Klíčová slova: anchored phylogenomics * Arvicanthini rodents * alignments

Obor OECD: Zoology

DOI: <https://dx.doi.org/10.57680/asep.0537176>


Handle: <http://hdl.handle.net/11104/0312390>

Vkladatel: Bryja, Josef

Datum publikování: 9.11.2020

Publikace ASEP:

Q Nuclear phylogenomics, but not mitogenomics, resolves the most successful Late Miocene radiation of African mammals (Rodentia: Muridae: Arvicanthini)

-  Do košíku
-  Editovat 
-  Změnit rok sběru 
-  Změnit druh dok. 
-  Nová verze dat 
-  Smazat záznam
-  Smazat dataset
-  Nahrát dataset
-  Publikace ASEP



Verzování

0571890 - ÚFP DATA Vědecká data

Zorek, M. - Škvára, V. - Šmídl, V. - Pevný, T. - Seidl, Jakub - Grover, Ondřej

Plasma state identification in the COMPASS tokamak.

The dataset contains selected diagnostic measurements from the COMPASS tokamak and manually created labels of plasma state in the form of csv files. The dataset defines 4 plasma states: L-mode, H-mode, ELM leading edge and ELM trailing edge. The data is intended to be used for training and testing deep neural networks in the task of plasma state identification. The dataset is divided into two parts, supervised (labeled) and unsupervised (unlabeled), each containing 20 COMPASS discharges. Within the unsupervised dataset there are 11 labeled discharges that can be used as a test set.

Klíčová slova: COMPASS * tokamak * h-mode * l-mode

Obor OECD: Fluids and plasma physics (including surface physics)

DOI: <https://dx.doi.org/10.57680/asep.0571890>

Handle: <https://hdl.handle.net/11104/0342782>

Vkladatel: admin

Datum publikování: 18.5.2023

Publikace ASEP:

Q Semi-supervised deep networks for plasma state identification


[Soubory, licence](#) [Financování](#)

 Do košíku

 Editovat 

 Změnit rok sběru 

 Změnit druh dok. 

 Nová verze dat 

 Smazat záznam

 Smazat dataset



 Nahrát dataset

 Publikace ASEP

Nová verze dat

Novou verzi záznamu vytvořte pouze v případě, že budete měnit uložené soubory, neboť bude mít přiděleno nové DOI.

Pokud nebudete měnit soubory, ale pouze metadatový popis, editujte současnou verzi záznamu.

 Vytvoření nové verze 


Zavřít

Data v OpenAire

Research Data . 2022

Highly luminescent scintillating hetero-ligand MOF nanocrystals with engineered Stokes shift for photonic applications.

ENGLISH

Authors: Perego, J.; Bezuidenhout, C.X.;  Villa, Irene; Cova, F.; Crapanzano, R.; Frank, I.; Pagano, F.; Kratochwill, N.; Auffray, E.; Bracco, S.; +6 more

DOI: [10.57680/asep.0565195](https://doi.org/10.57680/asep.0565195) 

Publisher: ASEP repository

The archive contains 6 text files (including this annotation.txt) and one pdf file (ESI of the paper): 1. Annotation.txt (this file) 2. 'Nature_Comm_figures 3_Photophysical properties of hetero-ligand MOF nanocrystals.TXT' 3. 'Nature_Comm_figures 4_Luminescence properties of MOF-based nanocomposites under X-ray exposure.TXT' 4. 'Nature_Comm_figures 5_Scintillation...

 Average popularity

 Average influence

+ ADD TO ORCID 



Shrnutí

- Repozitář splňuje FAIR principy, uložení dat je rychlé a bezpečné.
- Autoři, kteří chtějí ukládat data do DR ASEP, by se měli obrátit na zpracovatele pracovišť, kteří je proškolí, případně zajistí vše co je potřebné.
- Sledujeme vývoj v oblasti ukládání vědeckých dat a Open Science a systém ASEP upravujeme dle aktuálních potřeb a požadavků.
- Zapojení do projektu EOSC.CZ a integrace do **Národní datové infrastruktury**.



Další informace

- [Knihovna AV ČR](#)
- [System ASEP](#)
- [ASEP – Online katalog](#)
- [Informace o datovém repozitáři ASEP](#)

Kontakt: arl@lib.cas.cz

