

Vědecká data v ASEP

Seminář pro knihovny a SVI ústavů
AV ČR

24.5.2022

Rudolf Sýkora
Knihovna AV ČR, v.v.i.

Osnova

- Proč ukládat a popisovat data, FAIR principy
- Výběr repozitáře
- Charakteristika datového repozitáře ASEP
- Výsledky dotazníkového výzkumu
- Podpora – stránky o Otevřené vědě a správě výzkumných dat v KNAV

Proč ukládat vědecká data

- Efektivní využití zdrojů – stejná data se nemusí vytvářet opakovaně.
- Sdílení dat může urychlit výzkumný proces.
- Ukládání dat je v podmínkách poskytovatelů finančních prostředků (Horizon Europe některé projekty TA ČR a GA ČR).

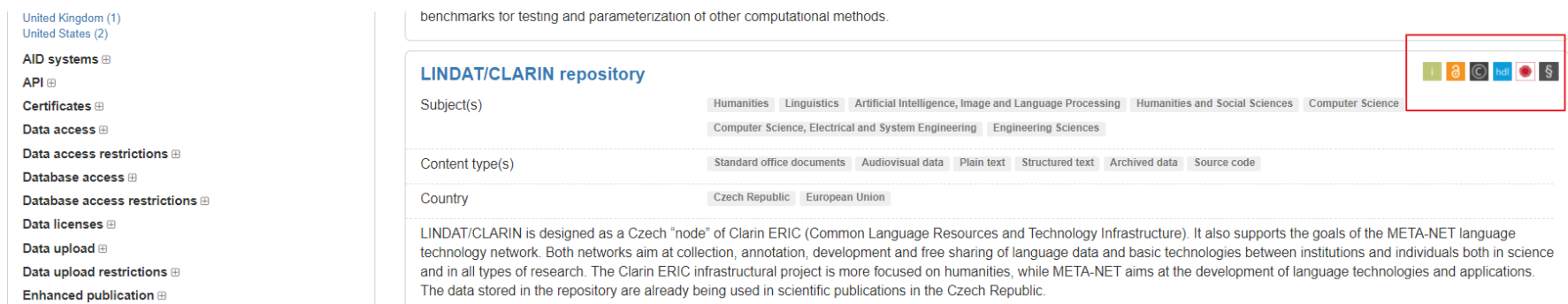
FAIR data

1. **Vyhledatelnost (Findable)** - je třeba zajistit aby lidé a stroje mohli data najít. Pro tento účel je klíčové dataset popsat **strojově čitelnými a strukturovanými metadaty**.
2. **Dostupnost (Accesible)** - K datům by měl být zajištěn **otevřený přístup**, ideálně prostřednictvím repozitáře. Není – li možné zajistit otevřený přístup k vědeckým datům, pak by měla být **volně přístupná alespoň metadata**.
3. **Interoperabilita (Interoperable)** data jsou uložena ve standardizovaných formátech a popsána pomocí řízených slovníků a standardizovaných výrazů.
4. **Reusable (Opětovně využitelná)** – data jsou dostatečně popsaná a sdílená pod otevřenou licencí (např. [Creative Commons](#)) aby uživatelé dat věděli, jak data vznikla, co popisují a jak s nimi mohou nakládat.

[FAIR self assesment tool](#)

Výběr repozitáře

- Při výběru vhodného oborového repozitáře **je dobré se držet oborových zvyklostí.**
- Je možné vyjít z **registru datových repozitářů** re3data.org, kde je u každého repozitáře piktogram, který může pomoci při výběru.



United Kingdom (1)
United States (2)

AID systems

API

Certificates

Data access

Data access restrictions

Database access

Database access restrictions

Data licenses

Data upload

Data upload restrictions

Enhanced publication

benchmarks for testing and parameterization of other computational methods.

LINDAT/CLARIN repository

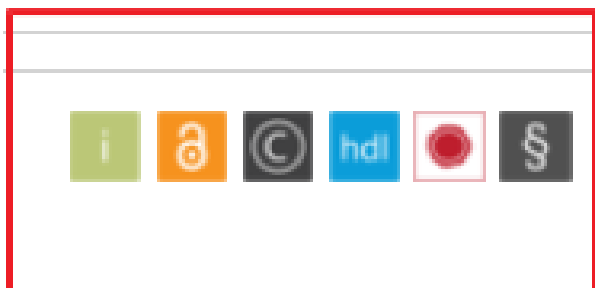
Subject(s) Humanities Linguistics Artificial Intelligence, Image and Language Processing Humanities and Social Sciences Computer Science

Computer Science, Electrical and System Engineering Engineering Sciences

Content type(s) Standard office documents Audiovisual data Plain text Structured text Archived data Source code

Country Czech Republic European Union

LINDAT/CLARIN is designed as a Czech "node" of Clarin ERIC (Common Language Resources and Technology Infrastructure). It also supports the goals of the META-NET language technology network. Both networks aim at collection, annotation, development and free sharing of language data and basic technologies between institutions and individuals both in science and in all types of research. The Clarin ERIC infrastructural project is more focused on humanities, while META-NET aims at the development of language technologies and applications. The data stored in the repository are already being used in scientific publications in the Czech Republic.



- Poskytování **otevřeného přístupu** k uloženým datům.
- Je repozitář **důvěryhodný/certifikovaný**.
- Přidělení **perzistentního identifikátoru**(např. DOI, handle).
- Přidělení **licence** k datům.
- **Metadata, verzování.**

Charakteristika DR ASEP

- Do repozitáře ukládají výzkumná data **autoři**, kteří působí na pracovištích **AV ČR**.
- Obsahuje **konečné** datové soubory opatřené **popisem** a **licencí**.
- Datový záznam má přiděleny **jednoznačné identifikátory** – **handle** (všechny záznamy v ASEP) a **DOI** (pouze datové záznamy uložené v ASEP).
- V ASEP lze **propojit** bibliografické a datové záznamy.
- V ASEP je možné vytvořit **metadatový popis záznamů** uložených **na externích uložiscích**.

Uložení a popis dat v ASEP

- **Uživatel**, který bude ukládat data si musí vytvořit **uživatelský účet myASEP**.
- Vyplní metadatový **formulář** a uloží **datové soubory**.
(Zvolí **licenci** CC nebo přiloží licenci vlastní pro nakládání s daty.)
- Potvrdí, že souhlasí s „**Dohodou o uložení v repozitáři ASEP**“ a předá záznam k **formální kontrole** zpracovateli pracoviště.
- Po kontrole zpracovatel zveřejní záznamy v **online katalogu**.
- Systém přidělí datovému záznamu persistentní identifikátory **handle** a **DOI**.

Formulář pro zápis datového záznamu AJ

Data record ×

Data record

Manual

System number

Title information

Descriptive information

Data in other repository

Rights/license

Classification

Projects/Grants

Publication in

Asep

Handover to review

Author information 1.

* Surname

* Name

Department

Institute

Country

+ -

Title information

* Title of dataset

Title in English

Alternative title

+ -

Data file type

* Data file type

+ -

Dataset description

Perzistentní identifikátory

- **Autorské identifikátory** (ORCID, RID, SAI)
- **Institucionální identifikátory** (ROR)
- **Persistentní identifikátory digitálních objektů** (HANDLE, DOI)
- DOI pro datové záznamy je přidělováno prostřednictvím registrační agentury [Datacite](#).
- KNAV se stala prostřednictvím konsorcia EUDAT členem Datacite.
- DOI bude přiděleno pouze datovým záznamům uloženým přímo v ASEP, ne záznamům uloženým v externím repozitáři.

Vyhledávání datasetů v ASEP

Export	☐ RIV ☐ not RIV
Zařazení WOS/SCOPUS	☐ ID SCOPUS ☐ ID WOS ☐ ID PubMed
Plný text	☐ Plný text u záznamu
Dataset	☐ Datový záznam v ASEP ☐ Bibliografický záznam s odkazem na data
Země vydání ▲	
Jazykové verze ▲	
Datum zápisu (dd.mm.rrrr) ⓘ	Od Do
Spolupráce	☐ zahraniční ☐ domácí
Impakt faktor ⓘ	Od Do

Odznačit limity

Hledat

Související stránky

Jednoduché vyhledávání

Rozšířené vyhledávání

Historie vyhledávání

Slovník

Změna zdroje

Bibliografický záznam s odkazem na data

0555726 - ÚBO 2023 RIV GB eng J - Článek v odborném periodiku

Horák, Kryštof - Bobek, Lukáš - Adámková, Marie - Kauzál, Ondřej - Kauzálová, Tereza - Manialeu, J. P. - Nguelefack, T. B. - Nana, E. D. - Jønsson, K. A. - Munclinger, P. - Hořák, D. - Sedláček, O. - Tomášek, Oldřich - Albrecht, Tomáš

Feather growth and quality across passerines is explained by breeding rather than moulting latitude. *Proceedings of the Royal Society B-Biological Sciences*. Roč. 289, č. 1970 (2022), č. článku 20212404. ISSN 0962-8452. E-ISSN 1471-2954

Grant CEP: GA ČR GA17-24782S; GA ČR(CZ) GA21-17125S

Institucionální podpora: RVO:68081766

Klíčová slova: comparative analysis * fault bars * life-history * long-distance migration * pace-of-life syndromes * ptilochronology

Obor OECD: Ecology

Impakt faktor: 5.349, rok: 2020

Způsob publikování: Omezený přístup

<https://royalsocietypublishing.org/doi/10.1098/rspb.2021.2404>

Trvalý link: <http://hdl.handle.net/11104/0330248>

 Do košíku

 DOI 

 WOS 

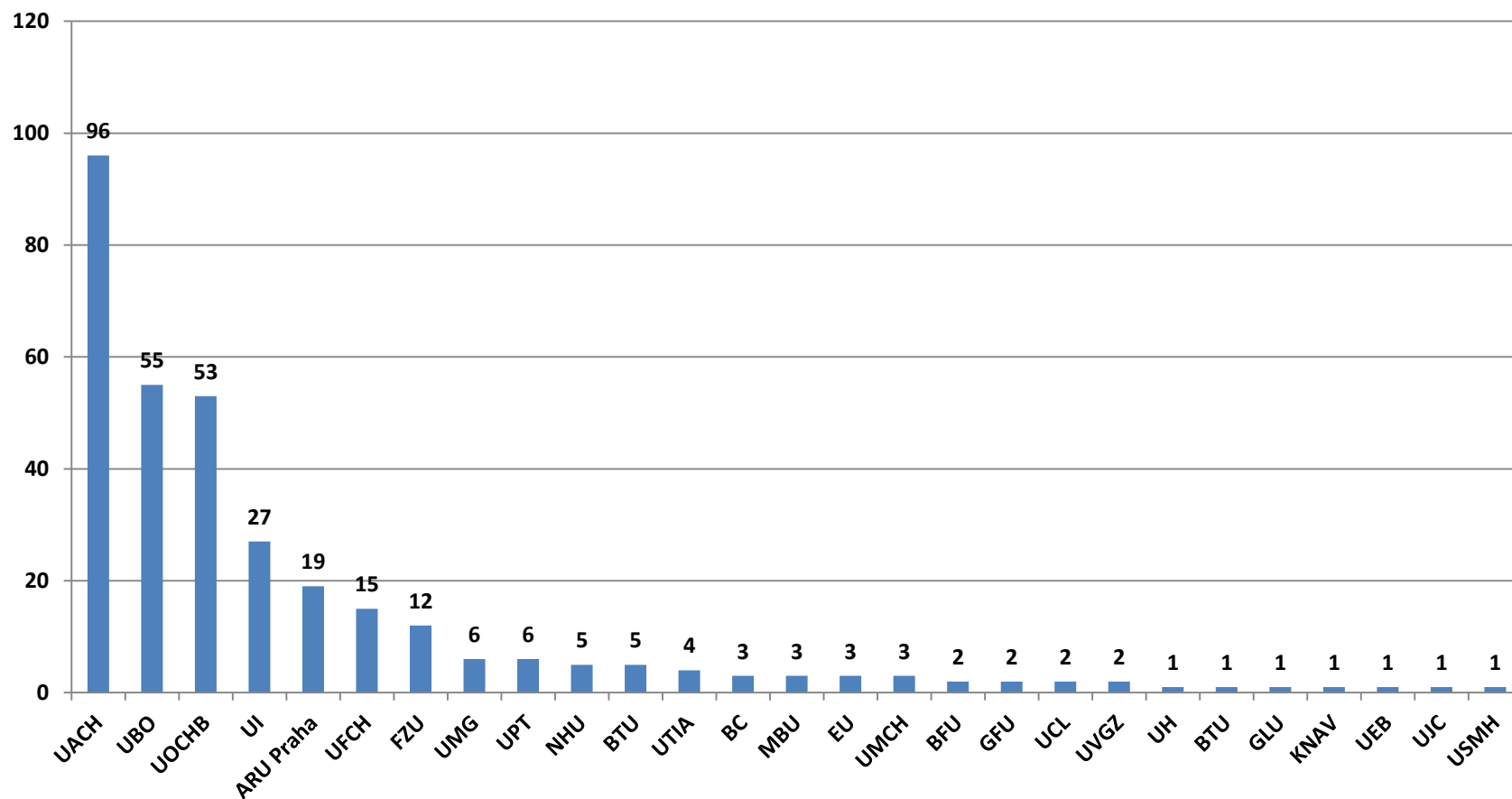
 SCOPUS 

 PUBMED 

 Bookmark

Vědecká data: [Dryad](#)

Počty odkazů na data



Datový záznam v ASEP

1.
0556584 - OÚ DATA Vědecká data
Tříšková, Hana
Mluvte čínsky hezky. Příkladové věty prozodie hovorové čínštiny. Soubor nahrávek příkladových vět, jež byly publikovány coby součást plného textu učebnice TRÍŠKOVÁ, H. Mluvte čínsky hezky. Prozodie hovorové čínštiny. Praha: Academia, 2021. ISBN 978-80-200-3228-7.
[Speak Chinese with ease. Example sentences of colloquial Chinese prosody.]

Dataset obsahuje 81 zvukových nahrávek příkladových vět, které jsou nedílnou součástí učebnice prozodie hovorové čínštiny. Každý výtisk učebnice doprovází USB kartička s plným textem knihy ve formátu .pdf, v němž jsou tyto nahrávky příkladových vět vnořeny. Příkladové věty namluvila rodilá mluvčí paní Tchang Jün-ling Rusková. V učebnici jsou zapsány v nové prozodické transkripci CHIPROT, již zde autorka představuje vůbec poprvé. Učebnice osvětluje, co se děje s čínskými slovy, když vstoupí do reality souvislé řeči. Představuje základní prozodické jevy mluvené hovorové čínštiny, jako jsou přízvuk a jeho souhra s tóny, oslabená výslovnost pomocných slov, frázování, včetně intonace či vlivy informační struktury výpovědi.
[The dataset contains 81 audio recordings of example sentences which are an integral part of the textbook on the prosody of colloquial Chinese. Each copy of the textbook is accompanied by a USB card with the full text of the book in .pdf format, in which the sound files of example sentences are directly immersed. These sentences were recorded by a native Beijing speaker, Mrs. Tang Yunling Rusková. All recorded utterances were transcribed by means of the novel prosodic transcription of the author. It is called CHIPROT (Chinese Prosodic Transcription). The textbook sheds light on what happens to Chinese words as they enter the reality of continuous speech. It presents the basic prosodic phenomena of spoken Chinese (colloquial speech style), such as stress and non-stress, their interaction with four lexical tones, weakened pronunciation of auxiliary words, phrasing, utterance-level intonation, or impact of the information structure on the prosodic structure.]
Technické informace: Zvukové nahrávky ve formátu .wav v zazipované složce.



Metadatový popis

Klíčová slova: Chinese * Mandarin * pronunciation * phonetics * prosody * textbook

Institucionální podpora: RVO:68378009

Obor OECD: Linguistics

Trvalý link: <http://hdl.handle.net/11104/0330758>

Datum publikování: 14.4.2022

Publikace ASEP:

🔍 Mluvte čínsky hezky. Prozodie hovorové čínštiny



PID



Propojení s publikací

Dataset :

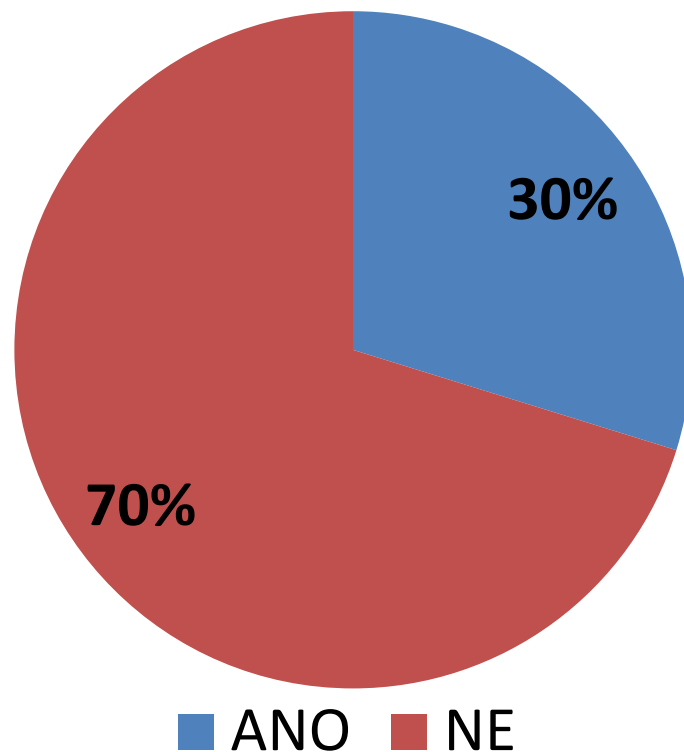
Licence: BY-NC-SA  - Uveďte původ + Neužívejte komerčně + Zachovejte licenci

Název souboru	Staženo	Velikost	Komentář	Přístup
🔒 Mluvte čínsky hezky. Nahrávky příkladových vět.zip	🕒 Přehled souborů	6	26.7 MB	vyžádat

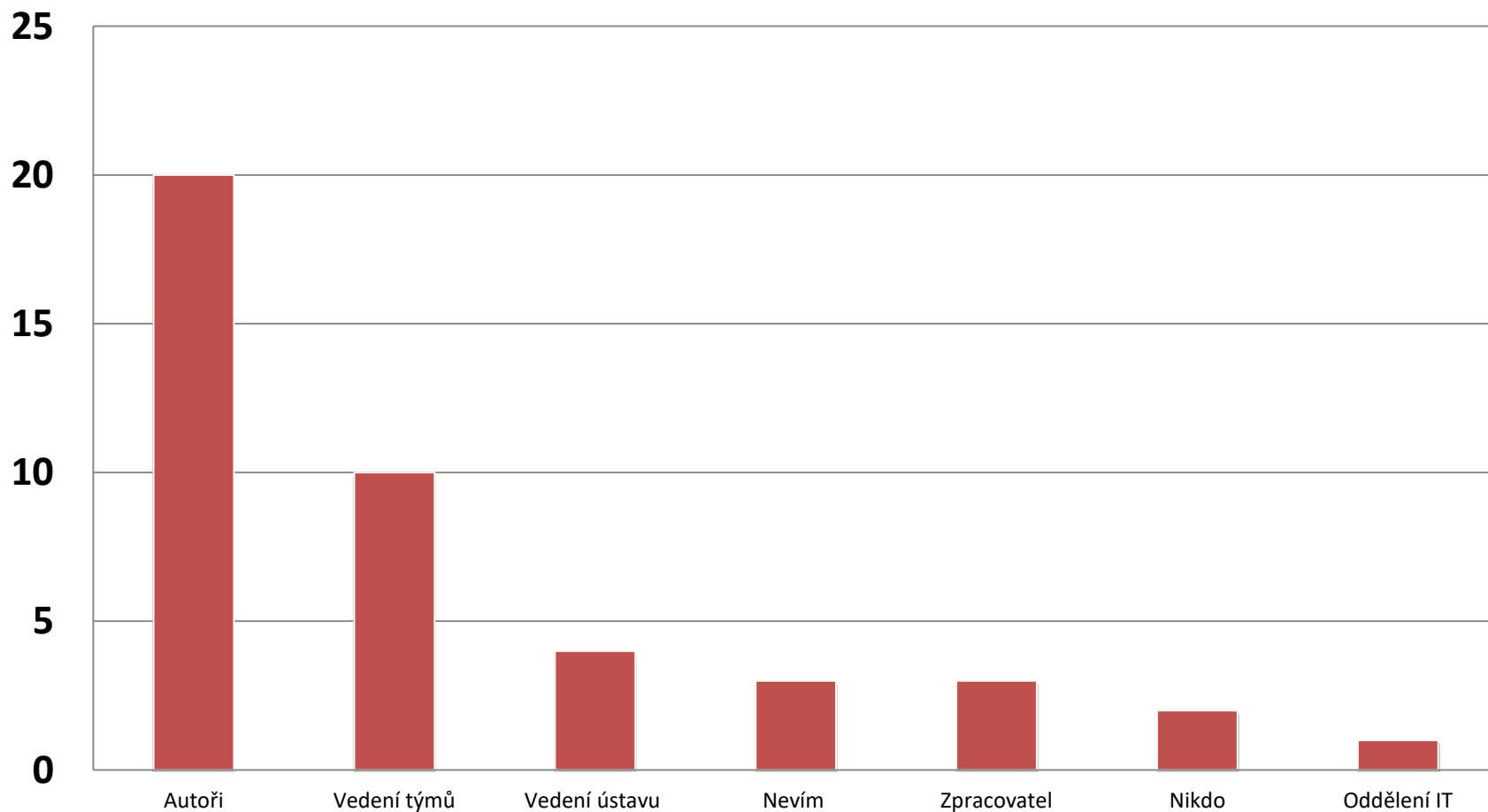


Uložené soubory

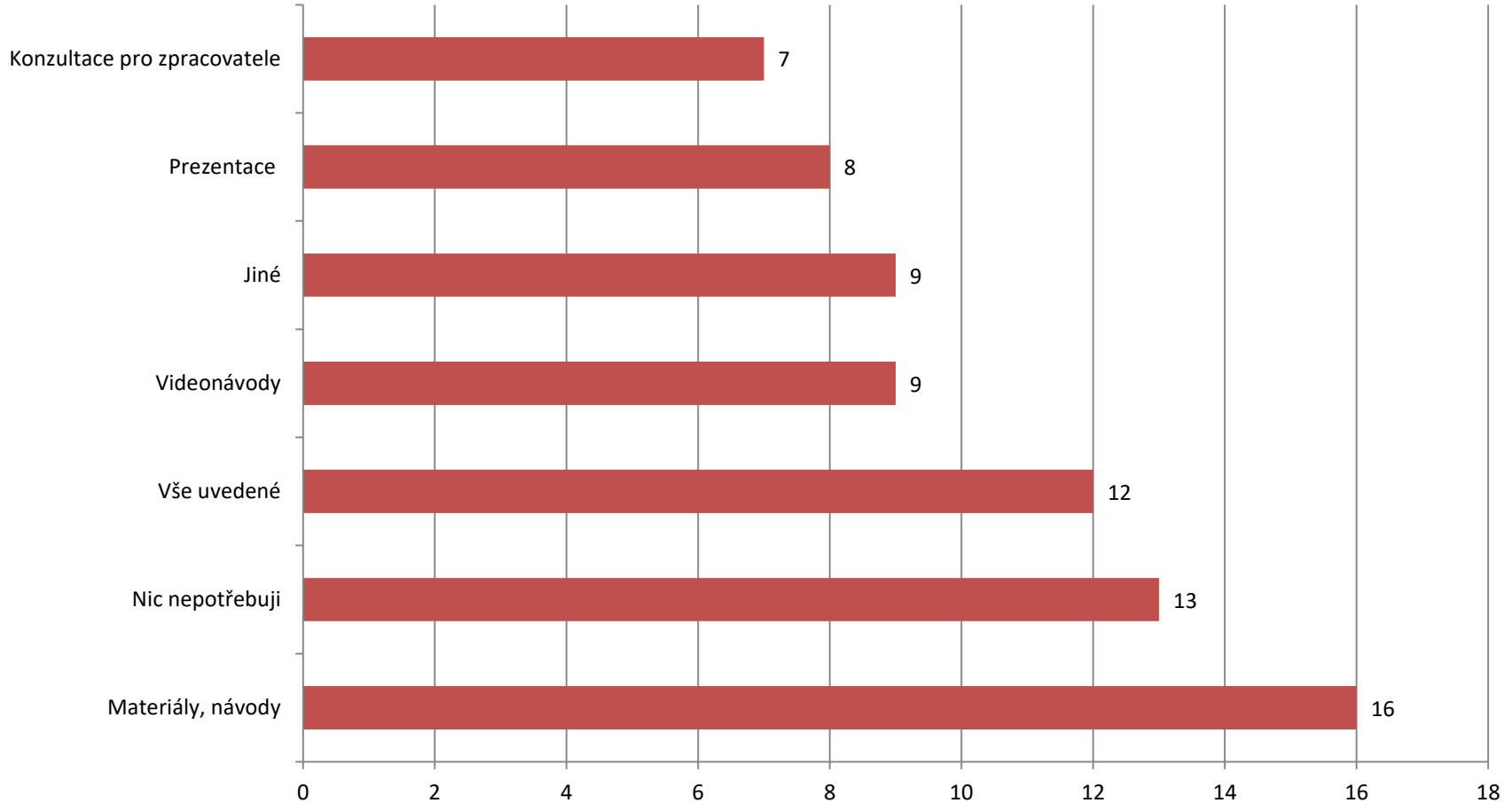
Setkali jste se s požadavkem vědců na ukládání vědeckých dat



Kdo koordinuje ukládání dat



Jakou podporu byste potřebovali od KNAV pro informování vědců?



Výzkumná data

Informace o výzkumných datech

▼ Pro autory

- Uživatelské účty
- Autorské identifikátory
- Zápis publikační činnosti
- Zápis datového záznamu
- Příprava dat v ASEP
- ▼ **Výzkumná data**
- OpenAIRE

[ASEP](#) » [Pro autory](#) » Výzkumná data

Výzkumná data

Otevřený přístup k vědeckým informacím

Publikování a sdílení dat v režimu otevřeného přístupu, který umožňuje okamžitý, bezplatný, trvalý a neomezený přístup k vědeckým výsledkům, je celosvětově stále více podporováno jak vědeckými institucemi, tak poskytovateli finanční podpory.

Otevřený přístup přispívá ke spolupráci, ověřitelnosti i využitelnosti výsledků tím, že usnadňuje přístup k textům, datům, softwaru, a podporuje tak zvyšování kvality i dopad vědeckých výsledků.

V AV ČR jsou principy publikování a zveřejňování výsledků výzkumu zakotveny v [Etickém kodexu výzkumných pracovníků AV ČR](#). Vědecká komunita by měla veřejně sdílet data svého výzkumu, dle toho, jak povaha dat umožňuje.

- **Definice výzkumných dat**
- **FAIR principy**
- **Otevřená data (Open data)**
- **Poskytovatel EU a výzkumná data**
- **Poskytovatelé ČR a výzkumná data**



**REPOZITÁŘ
ASEP**



**ANALYTIKA
ASEP**

Shrnutí

Připravujeme

- Verzování datasetů
- Sklizení datasetů do OpenAire (nyní jsou sklizeny záznamy o publikacích)
- Úpravy systému na základě praxe

Webové stránky ASEP

- [Podrobné informace o datovém repozitáři po uživatele](#) (obecné, [pro autory](#), [pro zpracovatele](#))
- Návod a uživatelské příručky ([vytvoření uživatelského účtu](#), [popis polí](#), struktura metadat)
- [Příprava dat](#) a [DMP](#)