

Otevírání dat v NK ČR v roce 2022

Lenka Maixnerová, Linda Jansová, Petra Šťastná
lenka.maixnerova@nkp.cz, linda.janova@nkp.cz, petra.stastna@nkp.cz
Národní knihovna ČR

Archivy, knihovny, muzea v digitálním světě 2022
Národní archiv, 1. prosince 2022

Co jsou otevřená data

„informace zveřejňované způsobem umožňujícím **dálkový přístup v otevřeném a strojově čitelném formátu**, jejichž **způsob ani účel následného využití není** povinným subjektem, který je zveřejňuje, **omezen** a které jsou evidovány v **národním katalogu otevřených dat**“ <https://www.zakonyprolidi.cz/cs/1999-106#p3a-5>



Zveřejnění prvních datových sad v podobě otevřených dat



Zpřístupnění databází Národní knihovny ČR pod licencí CC0

20. 4. 2022

Praha, Klementinum: Národní knihovna České republiky jako ústřední knihovna České republiky a centrum knihovních služeb v ČR se v souladu se současnými mezinárodními trendy rozhodla podřídit databáze Česká národní bibliografie (ČNB), Databáze národních autorit NK ČR (AUT) a Centrální adresář knihoven a informačních institucí v ČR (ADR) licenčnímu režimu CC0. „Zpřístupnění knihovních metadat pod veřejnou licencí CC0 a jejich propojování v digitálním prostředí je jedním ze strategických cílů Koncepce rozvoje knihoven v České republice na léta 2021–2027 s výhledem do roku 2030“, upřesňuje Lenka Maixnerová, náměstkyně sekce Knihovní fondy a služby v NK ČR.

Veřejná licence CC0 umožňuje nejširší možné užití databází, a to jak pro nekomerční, tak komerční účely, bez nutnosti uvedení autorství. Zveřejnění databází pod veřejnou licencí CC0 bylo provedeno se schválením Ministerstva kultury České republiky, které

https://www.nkp.cz/o-knihovne/soubory/ostatni/tz_db_cc0.pdf

Tři datové sady

 Česká národní bibliografie

 Databáze národních autorit NK ČR

 Centrální adresář knihoven a informačních institucí v ČR

Přidání dílčích datových sad autorit



Personální autority



Korporativní autority



Tematické autority



Geografické autority



Chronologické autority



Formální autority

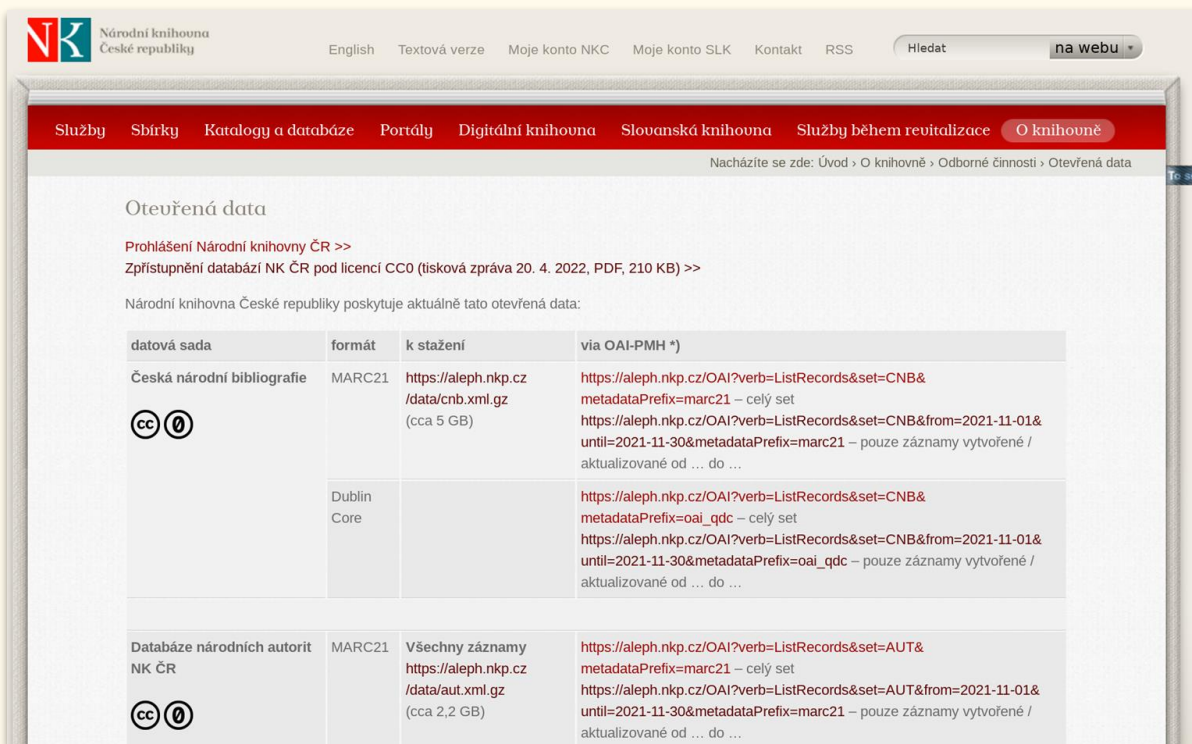


Konspekt



Unifikované názvy

Webová stránka Otevřená data (v češtině)



NK Národní knihovna České republiky

English Textová verze Moje konto NKC Moje konto SLK Kontakt RSS Hledat **na webu**

Služby Sbírky Katalogy a databáze Portály Digitální knihovna Slovanská knihovna Služby během revitalizace **O knihovně**

Nacházíte se zde: Úvod > O knihovně > Odborné činnosti > Otevřená data

Otevřená data

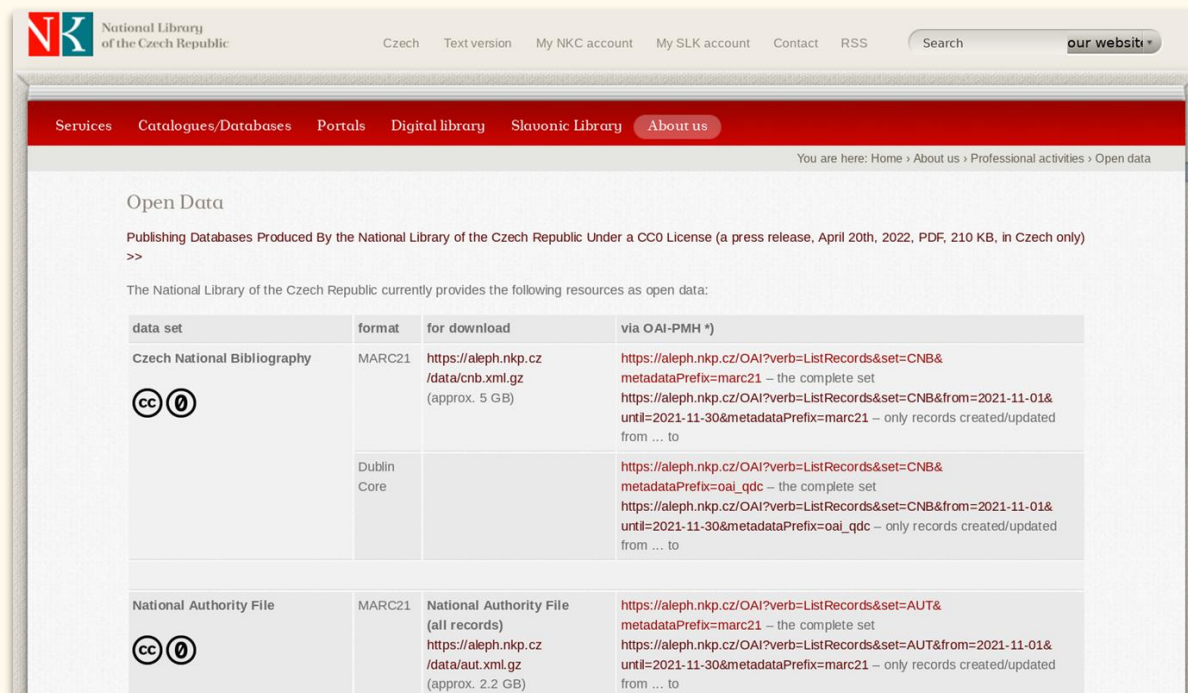
Prohlášení Národní knihovny ČR >>
 Zpřístupnění databází NK ČR pod licencí CC0 (tisková zpráva 20. 4. 2022, PDF, 210 KB) >>

Národní knihovna České republiky poskytuje aktuálně tato otevřená data:

datová sada	formát	k stažení	via OAI-PMH *)
Česká národní bibliografie 	MARC21	https://aleph.nkp.cz/data/cnb.xml.gz (cca 5 GB)	https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&metadataPrefix=marc21 – celý set https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&from=2021-11-01&until=2021-11-30&metadataPrefix=marc21 – pouze záznamy vytvořené / aktualizované od ... do ...
	Dublin Core		https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&metadataPrefix=oai_qdc – celý set https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&from=2021-11-01&until=2021-11-30&metadataPrefix=oai_qdc – pouze záznamy vytvořené / aktualizované od ... do ...
Databáze národních autorit NK ČR 	MARC21	Všechny záznamy https://aleph.nkp.cz/data/aut.xml.gz (cca 2,2 GB)	https://aleph.nkp.cz/OAI?verb=ListRecords&set=AUT&metadataPrefix=marc21 – celý set https://aleph.nkp.cz/OAI?verb=ListRecords&set=AUT&from=2021-11-01&until=2021-11-30&metadataPrefix=marc21 – pouze záznamy vytvořené / aktualizované od ... do ...

<https://www.nkp.cz/o-knihovne/odborne-cinnosti/otevrena-data>

Webová stránka Open Data (v angličtině)



The screenshot shows the 'Open Data' page of the National Library of the Czech Republic. The page header includes the NK logo, navigation links (Czech, Text version, My NKC account, My SLK account, Contact, RSS), a search bar, and a 'our website' button. A red navigation bar contains links for Services, Catalogues/Databases, Portals, Digital library, Slavonic Library, and About us. Below this, a breadcrumb trail reads: 'You are here: Home > About us > Professional activities > Open data'.

The main content area is titled 'Open Data' and contains the following text:

Publishing Databases Produced By the National Library of the Czech Republic Under a CC0 License (a press release, April 20th, 2022, PDF, 210 KB, in Czech only)
>>

The National Library of the Czech Republic currently provides the following resources as open data:

data set	format	for download	via OAI-PMH *)
 Czech National Bibliography	MARC21	https://aleph.nkp.cz/data/cnb.xml.gz (approx. 5 GB)	https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&metadataPrefix=marc21 – the complete set https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&from=2021-11-01&until=2021-11-30&metadataPrefix=marc21 – only records created/updated from ... to
	Dublin Core		https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&metadataPrefix=oai_qdc – the complete set https://aleph.nkp.cz/OAI?verb=ListRecords&set=CNB&from=2021-11-01&until=2021-11-30&metadataPrefix=oai_qdc – only records created/updated from ... to
 National Authority File	MARC21	National Authority File (all records) https://aleph.nkp.cz/data/aut.xml.gz (approx. 2.2 GB)	https://aleph.nkp.cz/OAI?verb=ListRecords&set=AUT&metadataPrefix=marc21 – the complete set https://aleph.nkp.cz/OAI?verb=ListRecords&set=AUT&from=2021-11-01&until=2021-11-30&metadataPrefix=marc21 – only records created/updated from ... to

<https://www.en.nkp.cz/about-us/professional-activities/open-data>

Národní katalog otevřených dat

OTEVŘENÁ DATA

Vzdělávání Články Datová kvalita

Datové sady

Poskytovatelé Klíčová slova

Další 

Poskytovatelé (1)

Národní knihovna České republiky (9) 

temata (1)

Vzdělávání, kultura a sport (9)

Klíčová slova (16)

autoritní záznamy (7)

soubor národních autorit (6)

věcné autority (4)

Vyhledat

Zobrazit pokročilé filtry

Smaž filtry

Seznam datových sad 

Nalezeno 9 datových sad

Název vzestupně 

Národní knihovna České republiky

Centrální adresář knihoven a informačních institucí v ČR

Datová sada obsahuje údaje o téměř 8000 institucích, které spolupracují nebo spolupracovaly na budování souborných katalogů CEZL/CASLIN, účastní se meziknihovních služeb, jsou v kontaktu s Knihovnickým institutem NK ČR nebo jsou...

XML

Česká národní bibliografie

Datová sada obsahuje přes 1 180 000 bibliografických záznamů dokumentů vydaných na území České republiky. Obsahuje i bibliografické záznamy publikací vydaných v zahraničí od českých autorů nebo psané v českém jazyce nebo pojednávající o České republice.

XML



Česká národní bibliografie

Národní knihovna České republiky

Datová sada obsahuje přes 1 180 000 bibliografických záznamů dokumentů vydaných na území České republiky. Obsahuje i bibliografické záznamy publikací vydaných v zahraničí od českých autorů nebo psané v českém jazyce nebo pojednávající o České republice.

[národní bibliografie](#)[česká národní bibliografie](#)

Téma

[Vzdělávání, kultura a sport](#)

Související geografické území

[Česká republika](#) 

Dokumentace

[Zobrazit dokumentaci](#)

Kontaktní bod

[Odbor zpracování fondů](#)

Specifikace

[Zobrazit specifikaci](#)[Zobrazit specifikaci](#)[Zobrazit specifikaci](#)

Periodicita aktualizace

[týdenní](#) 

Distribuce datové sady

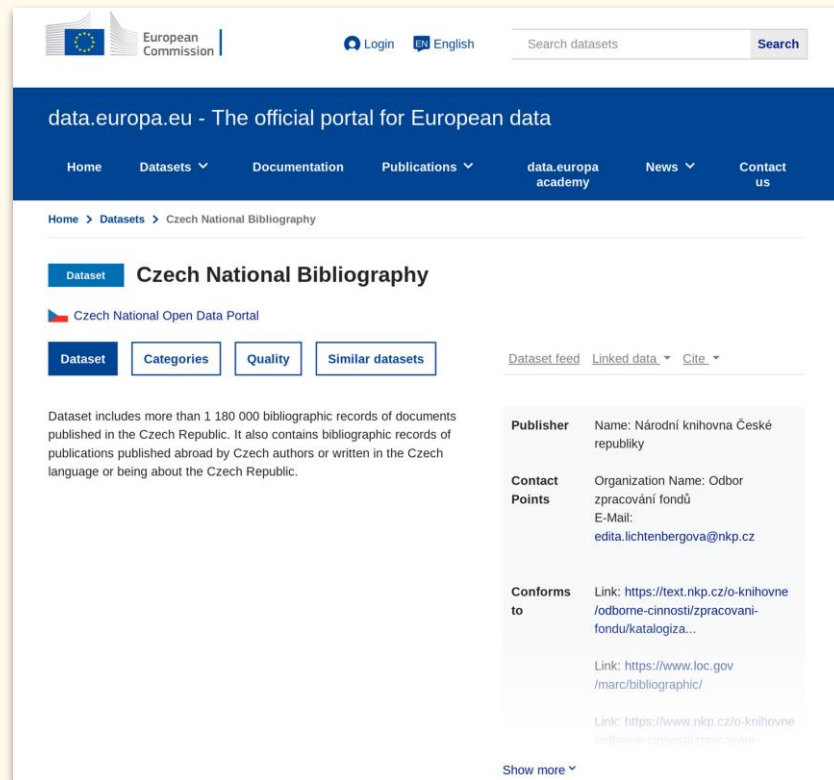
[XML](#)

<https://data.gov.cz/datov%C3%A11-sada?iri=https%3A%2F%2Fdata.gov.cz%2Fzdroj%2Fdatov%C3%A9-sady%2F00023221%2F1042474980>

Záznam České národní bibliografie

Data Europa EU

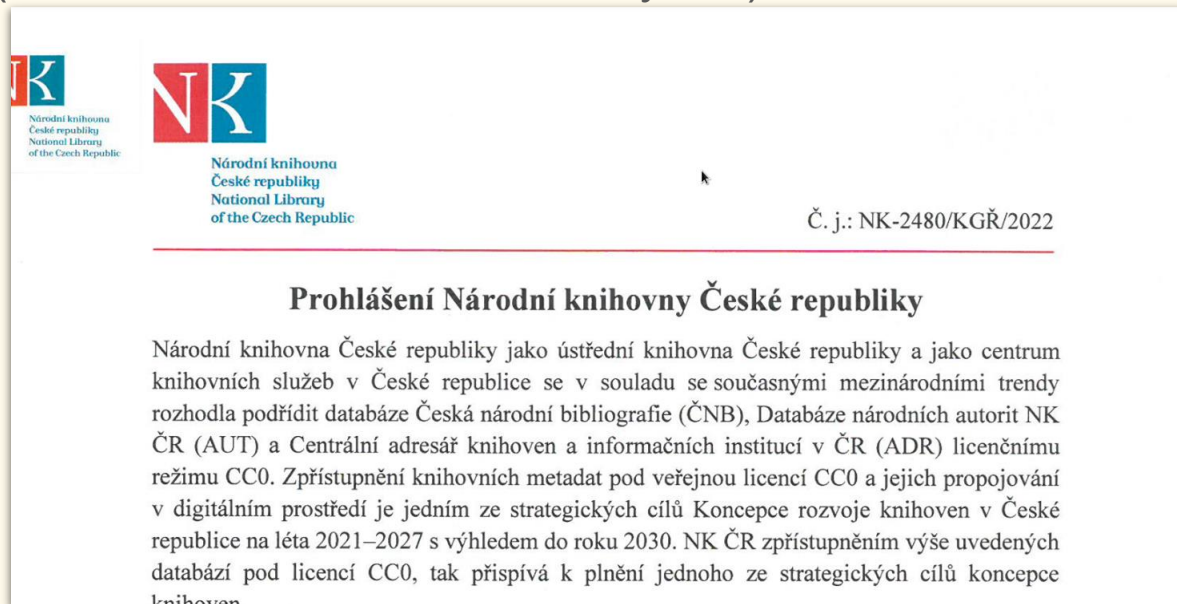
- záznamy se na portál Data Europa EU dostávají automaticky z Národního katalogu otevřených dat
- údaje jsou k dispozici v češtině a angličtině
- příklad záznamu České národní bibliografie



The screenshot shows the Data Europa EU portal interface. At the top, there's a header with the European Commission logo, a search bar, and navigation links for Login and English. Below the header, a blue banner reads "data.europa.eu - The official portal for European data". A navigation menu includes Home, Datasets, Documentation, Publications, data.europa academy, News, and Contact us. The main content area displays the "Czech National Bibliography" dataset page. It features a "Dataset" tab, a "Czech National Open Data Portal" logo, and buttons for Dataset, Categories, Quality, and Similar datasets. The dataset description states it includes more than 1 180 000 bibliographic records. On the right, there's a sidebar with "Dataset feed", "Linked data", and "Cite" options. The sidebar also contains contact information for the National Library of the Czech Republic, including organization name, email, and links to their website and research data portal.

Postup při zveřejňování 1/2

první tři datové sady: žádost schválená Ministerstvem kultury
(zřizovatelem Národní knihovny ČR)



https://www.nkp.cz/o-knihovne/soubory/ostatni/prohlaseni_nk_opendata.pdf

Postup při zveřejňování 2/2

další datové sady již bez žádosti:

využití **novelizace zákona o svobodném přístupu k informacím účinné od 1. září 2022:**

lze se vzdát práv pořizovatele databáze – za účelem vytěžování a zužitkování poskytované informace třetím osobám bez omezení způsobu a účelu následného užití

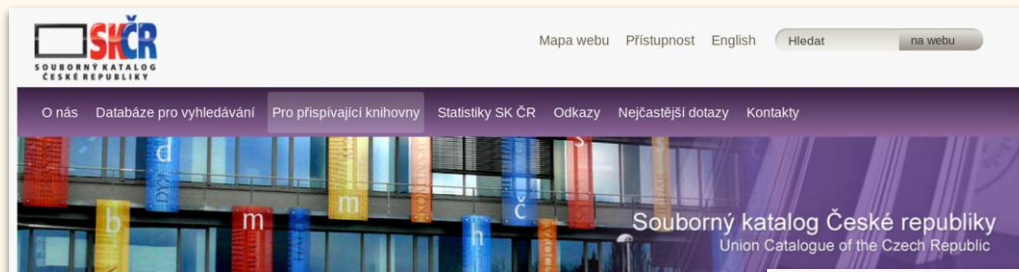
§ 14c

Zvláštní právo pořizovatele databáze

Má-li být poskytnuta informace, která je předmětem zvláštního práva pořizovatele databáze podle autorského zákona^{2b)}, a majetková práva k tomuto právu vykonává povinný subjekt uvedený v § 2 odst. 1 nebo v § 2a, tento povinný subjekt se při poskytování informace těchto práv vzdá, případně poskytne informaci na základě takové licence, která umožní vytěžování a zužitkování poskytované informace třetím osobám bez omezení způsobu a účelu následného užití informace.

<https://www.zakonyprolidi.cz/cs/1999-106#p14c>, § 14c zákona č. 106/1999 Sb.

Jak je důležité mít dokumentaci



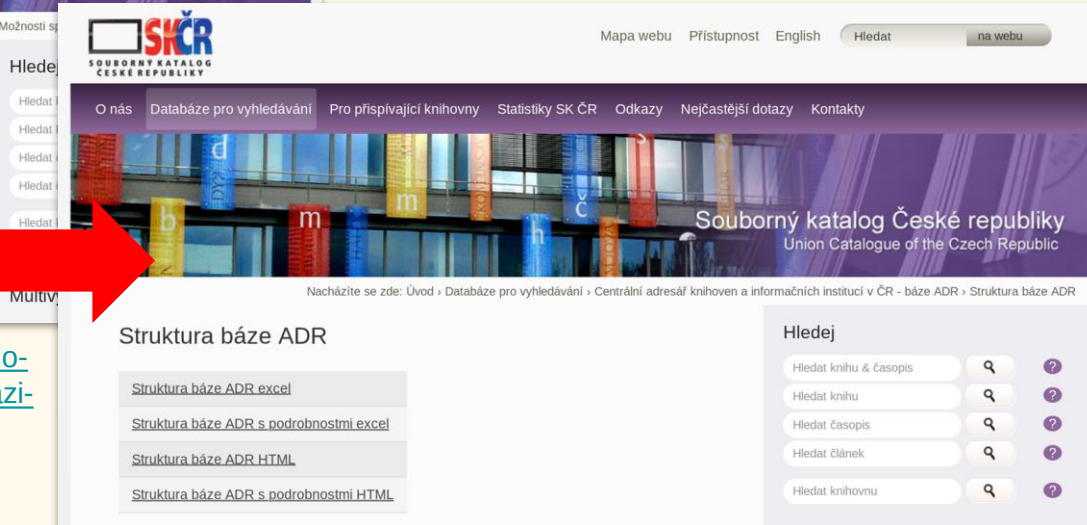
On-line aktualizace dat v bázi ADR

Formulář pro aktualizaci údajů v bázi ADR umožňuje knihovníkům "on-line" zadávat změny údajů, které jsou v bázi ADR o jejich knihovnách (institucích) shromážděny.

Obsah

- [Základní informace](#)
- [Práce s formulářem](#)
- [Aktualizace údajů](#)
- [Potvrzení správnosti údajů](#)
- [Zpráva pro správce báze](#)

<https://www.caslin.cz/caslin/spoluprace/jak-prispivat-do-sk-cr/spoluprace-se-sk-cr/on-line-aktualizace-dat-v-bazi-adr>



<https://www.caslin.cz/caslin/database-pro-vyhledavani/adresar/struktura-baze-adr>

Získávání údajů o využívání dat

česko-anglický formulář

https://docs.google.com/forms/d/e/1FAIpQLSfAwDAUa9y5NzuMuhTmfBt3ZPYguEa9vXKDD_dXvmnueZQukQ/viewform

Využívání otevřených dat Národní knihovny ČR / Using Open Data Published by the National Library of the Czech Republic

Využili jste některé z [otevřených datových sad zpřístupněných Národní knihovnou ČR](#)?
Pokud ano, uvítáme, dáte-li nám vědět, jak a kde jste tato data použili, a poskytnete-li nám
zpětnou vazbu, abychom mohli naši datovou nabídku stále vylepšovat.

Děkujeme za Váš čas!

Lenka Maixnerová
Národní knihovna ČR, náměstkyně pro knihovní fondy a služby
lenka.maixnerova@nkp.cz

Have you used some of the [datasets made available by the National Library of the Czech Republic](#)? If so, we will appreciate very much if you let us know how and where you have used the data and provided us with your feedback so that we can continuously improve our data offer. 14

Údaje o Národní knihovně jako zdroji

vzhledem k publikování dat s doložkou CC0

(<https://creativecommons.org/publicdomain/zero/1.0/deed.cs>)

je uvádění zdroje **nepovinné**

na webu jsou uvedeny věty:

„Přestože jsou datové sady k dispozici pod licencí CC0, můžete ve svém projektu Národní knihovnu ČR zmínit jako zdroj dat.”

„Although the datasets are available under a CC0 license, you can consider mentioning the National Library of the Czech Republic as a source of data for your project.”

Přehled prezentací/publikací

v češtině a angličtině

Publikace/prezentace o otevírání dat v Národní knihovně ČR

Publications/Slides on Opening Data at the National Library of the Czech Republic

2022

- [Na Databázi národních autorit spouštíme automatickou aktualizaci, svatý grál Wikidat](#)
příspěvek na blogu Wikimedia ČR, v češtině, listopad 2022
a post on the Wikimedia Czech Republic Blog, in Czech, November 2022
- **Otevírání dat v NK ČR v roce 2022**
prezentace na konferenci Archivy, knihovny, muzea v digitálním světě 2022, v češtině, v přípravě
slides for the Archives, Libraries, Museums in the Digital World 2022 conference, in Czech, forthcoming

<https://docs.google.com/document/d/1n80p9WnKOG-xOfJt5Xh3ZHV5L5OXHpTw57M7UWIJbmDc/edit?usp=sharing>

Praktické příklady práce s daty



NKLINK

NKlink je nástroj určený pro provozovatele databází, v jejichž záznamech jsou identifikátory **Národních autorit** (NKČR ID) – tedy pro různé knihovny, zejména české instituce. Pokud tento identifikátor ve své databázi pomoci nĚj stahovat různé další užitečné údaje z projektu **Wikidata** na články na Wikipedii a na související záznamy v některých dalších databázích.

NKLINK
CO NÁSTROJ UMÍ
JAK NAHLÁSIT CHYBU
PODOBNÉ NÁSTROJE
ODKAZY, DOKUMENTACE

<https://nklink.toolforge.org/>



NTK

50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Library of Technology



VŠCHT PRAHA



**ÚOCHB AV ČR
IOC PRAGUE**

Souhrnné vyhledávání v katalogích NTK, VŠCHT a ÚOCHB

Vše

Q H

[Hlavní strana](#) / Publikace dané osoby

Raab, Miroslav, 1938-

Biografické údaje: Narozen 1938. Fyzik, vědecký pracovník Ústavu makromolekulární chemie Akademie věd, věnuje se materiálům, popularizátor vědy.

Zdroj: Materiály a člověk
NKC
Vesmír 3/2009, str. 186

Další informace:









<https://vufind.techlib.cz/Authority/Record?id=jn20000620300>

Spolupráce se spolkem Wikimedia ČR

Jak Národní knihovna ČR spolupracuje s
Wikidaty

*How the National Library of the Czech Republic
Collaborates With Wikidata*

Přehledy

Overview

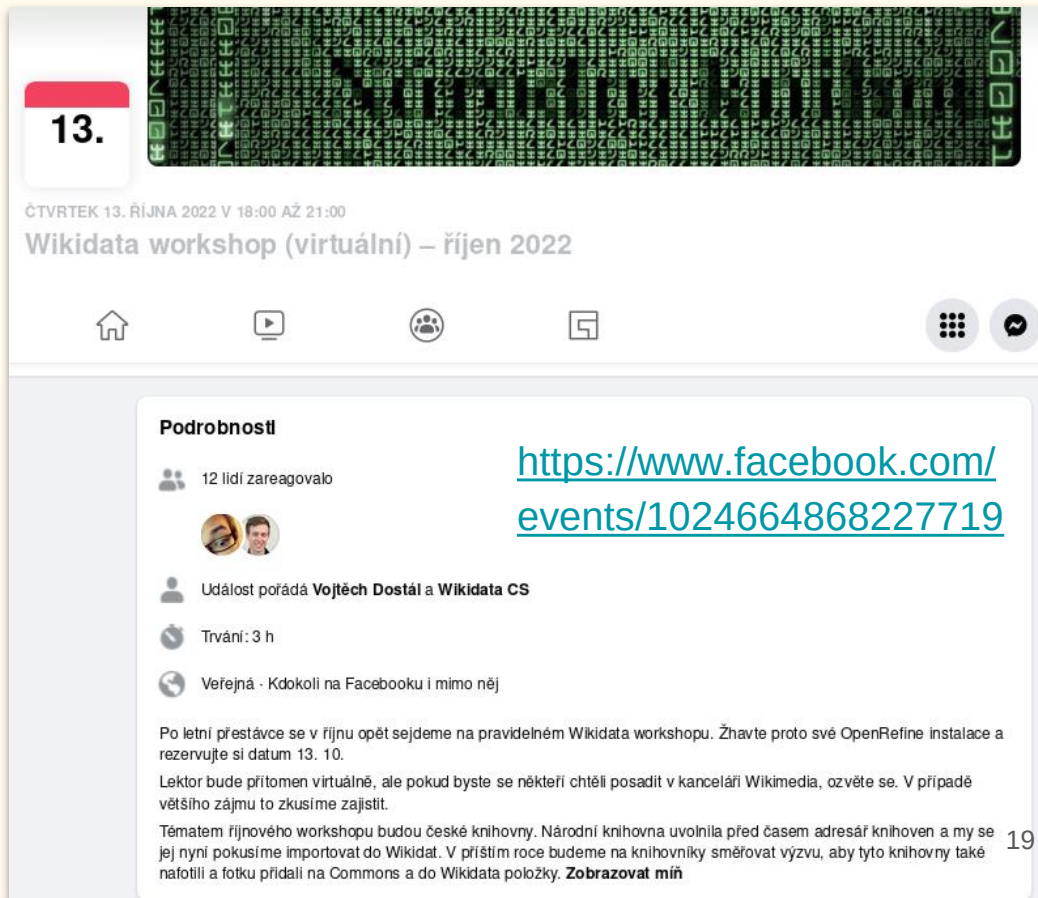
- [wiki stránka o spolupráci Národní knihovny ČR a Wikidat](#)
a wiki page focused on National Library's collaboration with Wikidata
 - především sekce [Validace dat](#) (různé monitorovací a analytické dotazy apod.)
especially a section entitled Validace dat [Data Validation] (various monitoring and analytical queries etc.)

+ mj. účast na
pracovních schůzkách
týkajících se spolupráce
s paměťovými
institucemi

<https://docs.google.com/document/d/1GUf6kGpwVopqmiRDtNAmiBed3S6hCTZGwEV6VB-h7hM/edit?usp=sharing>

Wikidata workshop (virtuální) – říjen 2022

praktické využití dat z adresáře knihoven a informačních institucí



13.

ČTVRTEK 13. ŘÍJNA 2022 V 18:00 AŽ 21:00

Wikidata workshop (virtuální) – říjen 2022

<https://www.facebook.com/events/1024664868227719>

Podrobnosti

12 lidí zareagovalo

Událost pořádá **Vojtěch Dostál a Wikidata CS**

Trvání: 3 h

Veřejná · Kdokoli na Facebooku i mimo něj

Po letní přestávce se v říjnu opět sejdem na pravidelném Wikidata workshopu. Žhavte proto své OpenRefine instalace a rezervujte si datum 13. 10.

Lektor bude přítomen virtuálně, ale pokud byste se někteří chtěli posadit v kanceláři Wikimedia, ozvěte se. V případě většího zájmu to zkusíme zajistit.

Tématem říjnového workshopu budou české knihovny. Národní knihovna uvolnila před časem adresář knihoven a my se jej nyní pokusíme importovat do Wikidat. V příštím roce budeme na knihovnický směřovat výzvu, aby tyto knihovny také nafotili a fotky přidali na Commons a do Wikidata položky. **Zobrazovat miň**

Wikidata workshop (virtuální) – listopad 2022

praktické využití autoritních dat

29.



ÚTERÝ 29. LISTOPADU 2022 V 18:00 AŽ 20:30

Wikidata workshop (virtuální) – listopad 2022

 Wikidata workshop (virtuální) – listopad 2022

Zúčastním se

Pozvat

18 lidí zareagovalo



Událost pořádá Vojtěch Dostál a Wikidata CS

Trvání: 2 h 30 min

Veřejná · Kdokoli na Facebooku i mimo něj

<https://www.facebook.com/events/422789256717614/>

Listopadový workshop se bude věnovat méně programování a více ruční práci na Wikidatech. Ukážeme si, jak jednoduše vytvářet položky lidem, jejichž záznamy byly právě vytvořeny v Bázi národních autorit NK, jak je propojit s dalšími databázemi a vědeckými články.

Naučíme se tedy pracovat s těmito nástroji:

- tabulky nových záznamů - https://www.wikidata.org/.../Wikidata.../New_authorities
- Author Disambiguator <https://author-disambiguator.toolforge.org/>
- MixNMatch funkce doplňování ID dle jména <https://mix-n-match.toolforge.org/?#/search/undefined> **Zobrazovat méně**

20

Wikidata CS



Diskuze

Lidé

Multimédia

Soubory

Reels



Vojtěch Dostál je s uživatelem **Jirka Sedláček**.

3. listopadu v 10:17 · 🌐

Spustili jsme **pravidelnou synchronizaci identifikátorů z národních autorit do Wikidat**. Je to jeden z prvních automatizovaných procesů na "českých Wikidatech" a je to součást dlouhodobé spolupráce s Národní knihovnou.

Jak to funguje: Frettiebot se podívá do polí 024 v národních autoritách, kde jsou uloženy identifikátory Wikidata QID, ISNI a ORCID. Když je schopen napojit se na existující položku, přidá do ní všechny chybějící identifikátory. Předtím zkontroluje, že identifikátor je ve správném tvaru. Pokud je daná hodnota na Wikidatech zavržena (deprecated), k editaci nedojde.

Příklady editací:

<https://www.wikidata.org/w/index.php?title=Q115019783...>

<https://www.wikidata.org/w/index.php?title=Q114562953...>

Kód:

https://github.com/wmcz/nkcr_catmandu

Kompletní výpis editovaných položek:

<https://quarry.wmcloud.org/query/68097>

Spuštění pravidelné synchronizace identifikátorů z národních autorit do Wikidat 1/2

Spuštění pravidelné synchronizace identifikátorů z národních autorit do Wikidat 2/2



CS / UDÁLOSTI

0

Na Databázi národních autorit spouštíme automatickou aktualizaci, svatý grál Wikidat

BY VOJTĚCH DOSTÁL · PUBLISHED 14 LISTOPADU, 2022 · UPDATED 14 LISTOPADU, 2022

Wikidata jsou pravděpodobně největším agregátorem českých databází vůbec. Databáze se ale neustále vyvíjí, mění a aktualizují. Jak to řešit? Nezbyvá než jednou za čas stáhnout aktuální obraz databáze a provést synchronizaci. Jenže je to pracné, nudné a zbytečné. Daleko lepší je v pravidelných intervalech provést import automaticky. Právě to se nyní podařilo zařídit u Databáze národních autorit Národní knihovny ČR.

Automatické aktualizace jsou výzvou

<https://blog.wikimedia.cz/2022/11/14/na-databazi-narodnich-autorit-spoustime-automatickou-aktualizaci-svaty-gral->

Hlavní letošní akce 1/5

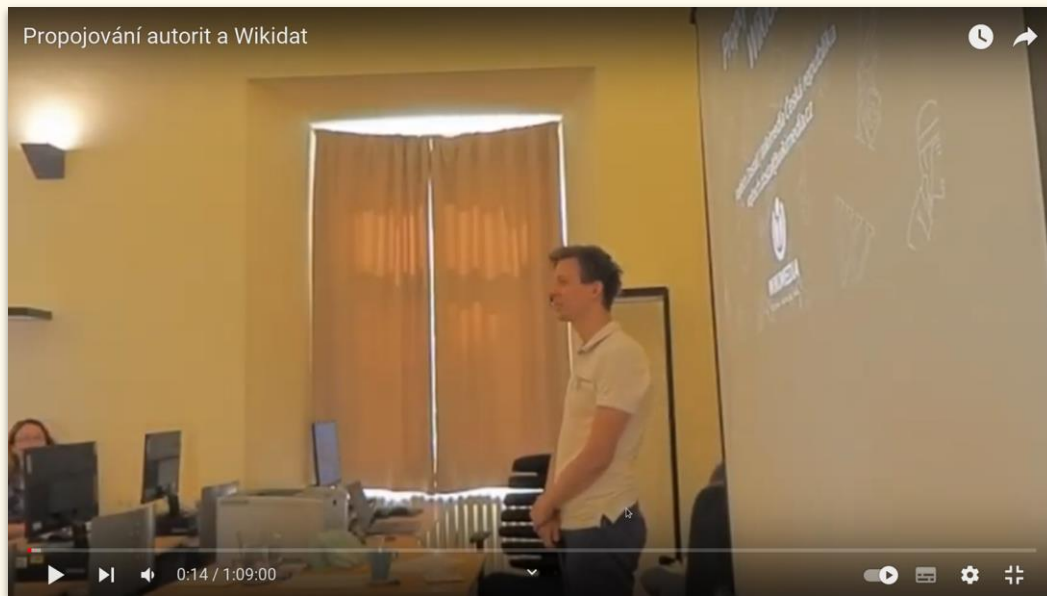
Panelová diskuse **Svobodné licence a volná díla: Archivy, muzea a knihovny otevřené digitálnímu světu** (leden 2022)



<https://youtu.be/SjXVWb8MOmY>

Hlavní letošní akce 2/5

Workshop **Propojování autorit a Wikidat** (květen 2022)



<https://youtu.be/SjXVWb8MOmY>

<https://www.skipcr.cz/knihovnicke-akce/propojovani-autorit-wikidat>

Hlavní letošní akce 3/5

Setkání koordinátorů projektu GLAM s českými knihovníky (září 2022)



Část výsledků zářijového setkání koordinátorů GLAM v Praze ve svém příspěvku připomene **Eva Vele**.

<https://youtu.be/SjXVWb8MOmY>

<https://www.skipcr.cz/knihovnicke-akce/setkani-koordinatoru-projektu-glam-s-ceskymi-knihovniky>

Hlavní letošní akce 4/5

Open Data Expo 2022 (listopad 2022)



<https://osf.cz/open-data-expo-2022/>



Hlavní letošní akce 5/5

A konečně

... DNES zde v Národním archivu :-)!

Čas na oslavu

Wikidata v říjnu 2022 oslavila své **desáté narozeniny!!!**



Wikidata is turning 10! **#WikidataBirthday** Distributed birthday events in October 2022

https://www.wikidata.org/wiki/Wikidata:Tenth_Birthday/Celebration_video

Doporučené zdroje

Portál otevřených dat (včetně Národního katalogu otevřených dat) provozovaný Ministerstvem vnitra:

<https://data.gov.cz/>

Wikidata:WikiProject Czech Republic/Pro instituce (návod pro instituce chystající se na zveřejnění otevřených dat zpracovaný spolkem Wikimedia ČR):

<https://database.wikimedia.cz/>

Plány do budoucna

zpřístupnění dalších našich databází

doplnění dokumentace k databázím

příprava výběrů dat ve formátu CSV (např. výpis děl vybraných autorů z České národní bibliografie)

pokračování v aktivní spolupráci se spolkem Wikimedia ČR a ve spolupráci s dalšími externími uživateli našich otevřených dat

...

Otevřeným datům zdar!

Děkujeme za pozornost!