

Implementace iniciativy European Open Science Cloud (EOSC) v ČR v oblastech společenských a humanitních věd

Jan Hajič

LINDAT/CLARIAH-CZ

Ústav formální a aplikované lingvistiky

Informatická sekce

Matematicko-fyzikální fakulta

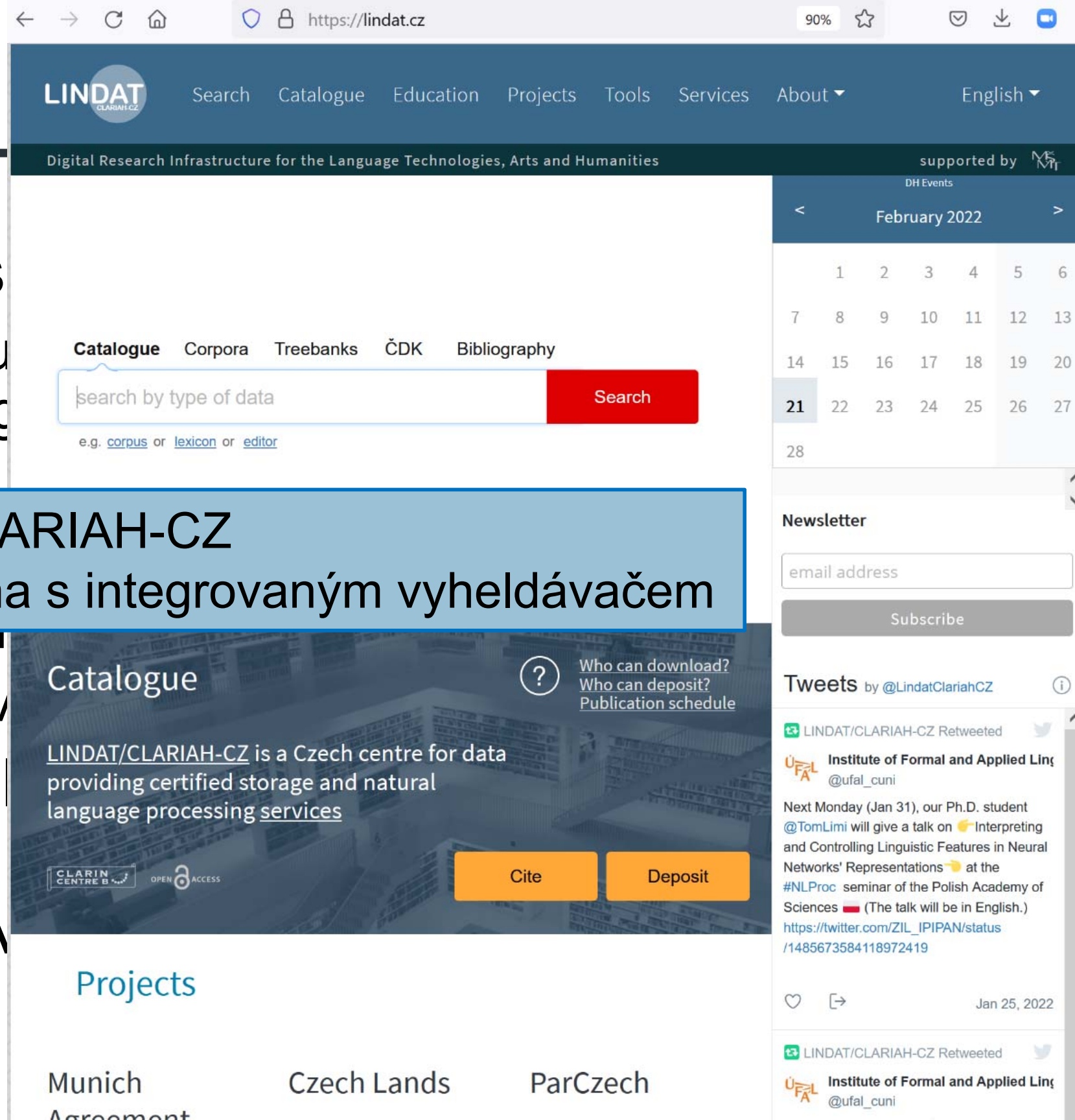
Univerzita Karlova, Praha

- Reprezentuje většinu komunity humanitních věd v ČR
 - Jazykověda a příbuzné obory (kognitivní vědy, psycholinguistika, sociolinguistika, ...)
 - Historie a příbuzné obory
 - Knihovny a knihovní vědy
 - Umění – vizuální, hudební, dramatické, muzikologie
 - ... ale také: jazykové technologie
- Spolupracuje s dalšími třemi VI v této oblasti
 - ČNK, ČLB, AIS
- Je součástí sítí CLARIN ERIC a DARIAH ERIC
- <https://lindat.cz>

- Reprezentuje větš
 - Jazykověda a příbu
 - lingvistika, socioling
 - Historie a příbuzné
 - Knihov
 - Umění – vizuální, m
 - ... ale také: jazykov
- Spolupracuje s da
 - ČNK, ČLB, AIS
- Je součástí sítě CLA
- <https://lindat.cz>

LINDAT/CLARIAH-CZ

Titulní strana s integrovaným vyhledávačem



The screenshot shows the homepage of the LINDAT/CLARIAH-CZ website. The header includes the LINDAT logo, navigation links (Search, Catalogue, Education, Projects, Tools, Services, About), and a language selector (English). Below the header, there's a banner for Digital Research Infrastructure for the Language Technologies, Arts and Humanities, supported by DH Events. The main content area features a search bar with the text "search by type of data" and a red "Search" button. Below the search bar, there are links for "e.g. corpus or lexicon or editor". The right sidebar contains a calendar for February 2022, a newsletter subscription form, and a section for tweets by @LindatClariahCZ. The footer includes a "Catalogue" section with a description of the center's mission and two orange buttons labeled "Cite" and "Deposit".

- VI jako podpora výzkumu v oblasti humanitních věd
 - Data
 - Knihovní katalogy, bibliografie, přístup k hist. dokumentům
 - Jazyková data (korpora), anotovaná data pro jazykový výzkum
 - Text a mluvená řeč (audio)
 - Obrazová data (umění), také ve spojení s textem
 - Video nahrávky (film, ale také tzv. orální historie)
 - Softwarové nástroje
 - Repozitář
 - Analýza jazyka (psaného, mluveného, tj. vč. audia) a obrazu
 - Softwarové a webové služby a aplikace
 - Technické, přístup k jazykovému i vizuálnímu obsahu
 - Jazyková příručka ÚJČ (i pro veřejnost)

- VI jako platforma
 - Data
 - Knihovny
 - Jazyky
 - Texty
 - Obrazy
 - Videy
 - Software
 - Repozitáře
 - Analýzy

Česká digitální knihovna

← → ↺ 🏠 🔒 https://cdk.lib.cas.cz/search?q=Peroutka&page=search ☆ 📧 ⬇️ 📺 ☰

← | **ČDK** Česká digitální knihovna | 🔍 Peroutka | ✕

Results: 1 - 60 / 3689

Advanced Search

Ordering: Relevance ▾

Active filters ✕

🔍 Peroutka

Accessibility ▾

Public 493

Private 3196

All 3689

Document Type ▾

Book 3041

Newspaper 841

Map 2

Graphic 0

0

0

0

0

0

0

Keyword ▾

Politika a vláda 7

Novináři 5

Lidové noviny

👤 Weiner, Richard

📅 22.1.1925

1614 occurrences

Newspaper

Ferdinand Peroutka, muž přítomnosti

👤 Otáhal, Milan

📅 1992

Book

Novinář Peroutka o T.G.M

👤 Houška, Vítězslav

📅 2010

Book

Ferdinand Peroutka: život v novinách

👤 Kosatík, Pavel

📅 2003

Book

Česká literatura: časopis pro literární vědu

👤 Holý, Jiří

125 occurrences

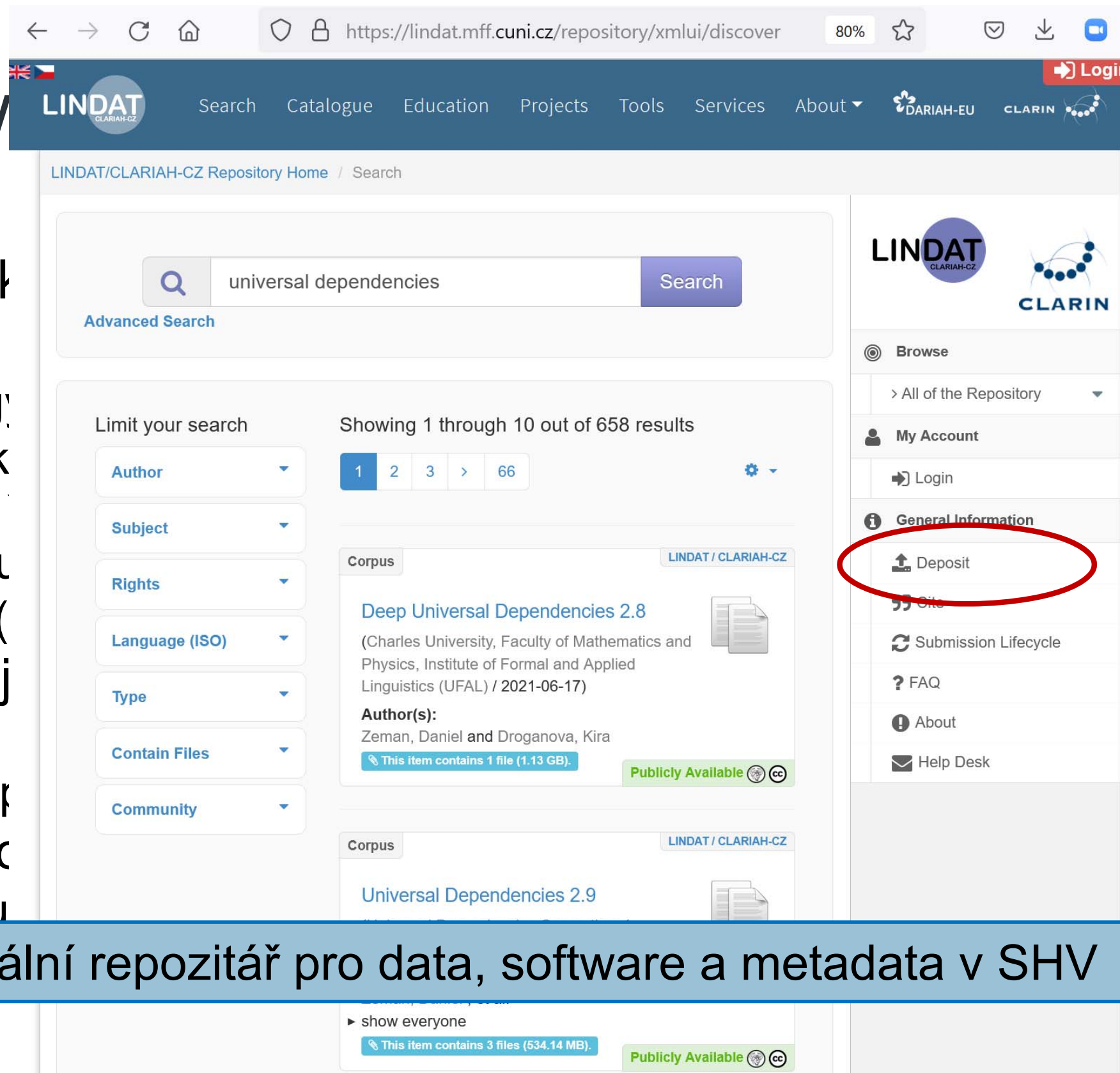
Newspaper

Mluví k vám Ferdinand Peroutka. Díl 1.

📅 2003

Book

- VI jako podpora výzkumu
 - Data
 - Knihovní katalogy
 - Jazyková data (korpora)
 - Text a mluvená slova
 - Obrazová data (obrázky)
 - Video nahrávky (audio)
 - Softwarové nástroje
 - Repozitář
 - Analýza jazyka (překlady)
 - Softwarové a webové služby
 - Technické příslušenství
 - Jazyková



LINDAT/CLARIAH-CZ Repository Home / Search

Search universal dependencies Search

Advanced Search

Limit your search

Author Subject Rights Language (ISO) Type Contain Files Community

Showing 1 through 10 out of 658 results

Corpus LINDAT / CLARIAH-CZ

Deep Universal Dependencies 2.8

(Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics (UFAL) / 2021-06-17)

Author(s): Zeman, Daniel and Drohanova, Kira

This item contains 1 file (1.13 GB).

Publicly Available

Corpus LINDAT / CLARIAH-CZ

Universal Dependencies 2.9

This item contains 3 files (534.14 MB).

Publicly Available

General Information

Deposit

Submission Lifecycle

FAQ

About

Help Desk

Centrální repozitář pro data, software a metadata v SHV

- VI jako podpora
 - Data
 - Knihovní ka
 - Jazyková d

Internetová jazyková příručka pro veřejnost

- Video nahrá
- Softwarové ná
 - Repozitář
 - Analýza jazy
- Softwarové a
 - Technické, p
 - Jazyková př

[Hlavní stránka](#)

[O příručce](#)

[Nápověda](#)

[Mobilní verze](#)

[Návštěvnost](#)

Související odkazy:

[Jazyková poradna](#)

[ČSN 01 6910](#)

[Zajímavé dotazy](#)

[Databáze dotazů](#)

Hledání konkrétního slova nebo tvaru slova.

infrastruktura

dělení: in-fra-struk-tu-ra¹

rod: ž.

	jednotné číslo	množné číslo
	infrastruktura	infrastruktury
2. pád	infrastruktury	infrastruktur
3. pád	infrastrukturu	infrastrukturám
4. pád	infrastrukturou	infrastrukturami
5. pád	infrastrukturou	infrastrukturami
6. pád	infrastrukturou	infrastrukturami
7. pád	infrastrukturou	infrastrukturami

Heslové slovo bylo nalezeno také v následujících slovnících: [ASCS](#), [SSČ](#), [SSJČ](#)
Další slovní charakteristiky a příklady: [ČNK](#)

Odkazy k výkladové části Internetové jazykové příručky:

¹[Dělení slov na konci řádku](#)

- VI jako
- Data

- k
- u
- (

Strojový překlad

- Software a modely
- Web aplikace (i pro veřejnost)
- Dálková webová služba (API)

- Soft

- 7
- u

LINDAT Translation

Translate

Docs

The translation service is available for *personal and non-commercial use* (see [terms of use](#) for more details).

Source

Target

Czech

English

☐ advanced

Input sentences

Translation

Translate

Choose file

Credits: The service runs systems trained by:

Martin Popel
CUBBITT models, en<->cs, en<->fr, en<->pl as described in
Popel, M., Tomkova, M., Tomek, J. et al. *Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals*. Nat Commun 11, 4381 (2020). <https://doi.org/10.1038>

VI jako

Dat

Sof

Sof

LINDAT Translation

Translate

Docs

The translation service is available for *personal and non-commercial use* (see [terms of use](#) for more details).

Source

Target

Czech

English

☐ advanced

Input sentences

Translation

Hlavní budova bývalého železničního nádraží Praha – Těšnov se mohla v časech pravidelného provozu na daném nádraží chlubit neoficiálním titulem nejkrásnější nemovitosti tohoto druhu v Praze. Někdy bývalo nádraží klasifikováno i jako nejkrásnější ve střední Evropě a z úst těch nejodvážnějších mu po stránce estetické nebylo rovno v celé Evropě. Otevřeno bylo v roce 1875, v té době ho dnešním pohledem označujeme názvem Praha – severozápadní nádraží. Uvedené světové strany v označení nádraží nás nesmí zmást, neměly nic společného s jeho prostorovým situováním v Praze, ale vycházely z názvu společnosti, na

Translate

Choose file

Credits: The service runs systems trained by:

Martin Popel
 CUBBITT models, en<->cs, en<->fr, en<->pl as described in
 Popel, M., Tomkova, M., Tomek, J. et al. *Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals*. Nat Commun 11, 4381 (2020). <https://doi.org/10.1038>

VI jako

• Data

- |
- \

• Soft

- |
- \

• Soft

- |
- \

LINDAT Translation

Translate

Docs

The translation service is available for **personal and non-commercial use** (see [terms of use](#) for more details).

Source

Czech

Target

English

☐ advanced

Input sentences

Hlavní budova bývalého železničního nádraží Praha – Těšnov se mohla v časech pravidelného provozu na daném nádraží chlubit neoficiálním titulem nejkrásnější nemovitosti tohoto druhu v Praze. Někdy bývalo nádraží klasifikováno i jako nejkrásnější ve střední Evropě a z úst těch nejodvážnějších mu po stránce estetické nebylo rovno v celé Evropě. Otevřeno bylo v roce 1875, v té době ho dnešním pohledem označujeme názvem Praha – severozápadní nádraží. Uvedené světové strany v označení nádraží nás nesmí zmást, neměly nic společného s jeho prostorovým situováním v Praze, ale vycházely z názvu společnosti, na

Translation

The main building of the former railway station Prague – Těšnov could boast the unofficial title of the most beautiful property of this kind in Prague in times of regular operation at the station. Sometimes the station was also classified as the most beautiful in Central Europe and from the mouths of the most adventurous aesthetically it was not equal in the whole of Europe. It was opened in 1875, at that time we call it Prague – Northwest Railway Station. The mentioned world sides in the name of the station must not confuse us, they had nothing to do with its spatial location in Prague, but they

Translate

Choose file

Credits: The service runs systems trained by:

Martin Popel

CUBBITT models, en<->cs, en<->fr, en<->pl as described in

Popel, M., Tomkova, M., Tomek, J. et al. *Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals*. Nat Commun 11, 4381 (2020). <https://doi.org/10.1038>


- Projekt SSHOC (EC, INFRAEOSC) <https://sshopencloud.eu>
 - *Social Sciences and Humanities Open Cloud*
 - 2018-2022, konference 6.-7.4.22
 - Koordinuje CESSDA ERIC
- Hlavní cíle
 - Vytvořit SHV oblast v EOSC
 - Data management: Open Science, FAIR
 - Spojovat ERICy v oblasti SHV
 - Koordinace s ostatními disciplínami při zachování specifik SHV



- Součást EOSC pro SHV – finální verze hotova
 - <https://sshopencloud.eu/ssh-open-marketplace>
- Obsahuje
 - Softwarové nástroje a webové služby
 - Datasets, databáze, korpusy
 - Materiály pro výuku
 - Publikace, příklady workflows, ...
 - Případové studie, recenze, ...
- Datové katalogy – metakatalogy (např. CLARIN VLO)

- Součást EOSC pro SHV – f
 - <https://sshopencloud.eu/ssh>
- Obsahuje
 - Softwarové nástroje a web
 - Datasety, databáze, korpus
 - Materiály pro výuku
 - Publikace, příklady workflow
 - Případové studie, recenze,
- Datové katalogy – metakatalogy (např. CLARIAH-VLEO)

The screenshot shows a web browser window with the URL <https://marketplace.sshopencloud.eu/search?q=udpipe>. The search bar contains 'udpipe' and a 'Search' button. Below the search bar, the page shows 'Search results (2)'. On the left, there are filters for 'CATEGORIES' (Tools & Services: 2) and 'ACTIVITIES' (Analyzing: 1, Annotating: 1, Natural Language Processing: 1, Parsing: 1, Structural Analysis: 1). The main content area displays two results for 'UDPipe'. The first result includes a description: 'UDPipe is a trainable pipeline for tokenization, tagging, lemmatization and dependency parsing of CoNLL-U files. UDPipe is language-agnostic and can be trained given annotated data in CoNLL-U format. Trained models are provided for nearly all UD treebanks. UDPipe is available as...'. The second result is a snippet: 'UDPipe: tool for lemmatization, morphological analysis, POS tagging and dependency parsing in multiple languages'.

- Autorské právo a intelektuální vlastnictví
 - Autorské právo je třeba respektovat...
 - ... ale současná úprava (vč. TDM v CD EU 2019) omezuje
 - Naopak, IPR výsledků nedostatečně chráněno (v praxi)
- Tradiční přístup ke tvorbě a správě dat
 - „Data jsou moje“ (nejsou...!)
 - ... ale tvorba dat v SHV má nedostatečný kredit (vs. publikace)
- Sdílení výsledků
 - Nutnost skloubit „otevření“ výsledků a kredit pro autory
 - Neustálená nebo žádná správná citační praxe
 - Vykazování využití otevřených dat a služeb: prakticky nemožné

- Práce s komunitou (i mimo LINDAT/CLARIAH-CZ a VI)
 - „Otevírání“ dat
 - Zahrnuje technické, právní, organizační zajištění
 - Tvorba metadat (katalogů)
 - Standardy, konverze, tvorba, workflow
 - „Best practices“
- Technické zajištění
 - Centrální vyhledávání v metadatech napříč všemi repozitáři
 - Standardy a certifikace repozitářů
 - Záložní úložiště („default“ pro instituce bez repozitáře)
- Vzdělávání – semináře, tutoriály, workshopy, hands-on, ...
- Národní EOSC projekt pro humanitní vědy

... děkuji za pozornost!

Jan Hajič, LINDAT/CLARIAH-CZ

<https://lindat.cz> Twitter: [@LindatClariahCZ](https://twitter.com/LindatClariahCZ) hajic@ufal.mff.cuni.cz

<https://ufal.mff.cuni.cz/jan-hajic>

<https://www.linkedin.com/in/jan-hajic-9a09a21/>