

LINDAT/CLARIAH-CZ

OBOROVÝ REPOZITÁŘ NEJEN PRO JAZYKOVÁ DATA

PAVEL STRAŇÁK

LINDAT/CLARIAH-CZ

- Národní projekt “Velké výzkumné infrastruktury”
- Integrace do **CLARIN** ERIC a **DARIAH** ERIC
- Operační fáze pro jazykovou infrastrukturu od 2014, pro další humanitní obory od 2020 fáze konstrukční (= pracujeme na tom)
- Jazykové technologie: pro jazykový výzkum i pro jiné obory
- Technologie: datový repozitář, Webové aplikace a služby (API)
- Znalosti, podpora, školení

UŽIVATELÉ

- Všichni
 - komunikují a pracují v přirozených jazycích!
- ... nejbližší užití infrastruktury:
 - Výzkum v oboru jazykových technologií / NLP
 - Potřeba velkých, dostupných a jasně licencovaných dat
 - Uživatelé “obsahu” (jazyk jen jako prostředek):
 - Historici, učitelé, literáti, sociologové, ...
 - Potřeba jednoznačně identifikovaných, popsaných, prohledatelných a licencovaných dat

Datový Repozitář

BEZPEČNÉ ULOŽENÍ A SDÍLENÍ JAZYKOVÝCH DAT A NÁSTROJŮ

[Advanced Search](#)

| Author | Subject | Language (ISO) |
|-------------------------------|-------------------------------|-------------------------------|
| Hajič, Jan (47) | Germanistik (47) | English (222) |
| Žabokrtský, Zdeněk (32) | machine translation (39) | Czech (192) |
| Straka, Milan (29) | corpus (34) | German (159) |
| Zeman, Daniel (29) | treebank (30) | Dutch (92) |
| Bojar, Ondřej (28) | morphology (26) | Spanish (83) |
| ... View More | ... View More | ... View More |

What's New

ToolService

LINDAT / CLARIN

CorpusExplorer

Author(s):
Rüdiger, Jan Oliver

Description:

Software for corpus linguists and text/data mining enthusiasts. The CorpusExplorer comprises over 45 interactive



? What can you do?

DEPOSIT




CITE

© Browse

> All of the Repository

DATOVÝ REPOZITÁŘ

- **OPEN  ACCESS** (tak moc, jak to jde – [Public License Selector](#))
- > 900 registrovaných uživatelů
 - Nahrání dat a podpis některých licencí (ne vše může být OA)
- 400+ datasetů
 - > 1100 metadatových záznamů
 - >100 jazyků
- 100 TB+ Data in Repository (+ 1PB of UCS Shoah Foundation Archive)

DATOVÝ REPOZITÁŘ

- > 900 registrovaných uživatelů
 - Nahrání dat a podpis některých licencí (ne vše může být OA)
- 400+ datasetů
 - > 1100 metadatových záznamů
 - >100 jazyků
- 100 TB+ Data in Repository (+ 1PB of UCS Shoah Foundation Archive)

The screenshot displays the LINDAT/CLARIN Repository search interface. At the top, navigation links include 'LINDAT', 'Repository', 'TreeQuery', 'TreeX', 'More Apps', and 'All'. Below the navigation is a search bar with a magnifying glass icon and a 'Search' button. A link for 'Advanced Search' is located below the search bar.

The main content area is divided into two columns. The left column, titled 'Limit your search', contains several dropdown menus for filtering results: 'Author', 'Subject', 'Rights', 'Language (ISO)', 'Type', 'Contain Files', and 'Community'. The right column, titled 'Showing 1 through 10 out of 1038 results', features a pagination control with buttons for '1', '2', '3', '>', and '104'. Below the pagination, three search results are displayed, each with a category label, title, source information, author(s), and file size. The first result is 'AKCES 2 ver. 2' (Corpus) by Charles University in Prague, ÚČJTK, dated 2013-12-18, with 1 file (3.85 MB). The second result is 'A Gold Standard Word Alignment for English-Swedish (2015-10-12)' (LexicalConceptualResource) by Linköping University, dated 2015-10-12, with 1 file (590 KB). The third result is 'MorphoDiTa: Morphological Dictionary and Tagger' (ToolService) by Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics (UFAL), dated 2014-02-14, with no files. Each result includes a 'Publicly Available' badge and a Creative Commons license icon.

DATOVÝ REPOZITÁŘ

- Bezpečné uložení
(nahrát a nestarat se)
- Najít a použít data
- Přímé citace dat
(počítají se i v Google Scholar)
- Licencování
(Open Access, ale i jinak)
- Verzování
- Data i software
- Celosvětově, jednoduše

The screenshot shows the LINDAT/CLARIN Repository search results page. At the top, there are navigation links for LINDAT, Repository, TreeQuery, Treex, More Apps, and All. Below the navigation is a search bar with a magnifying glass icon and a 'Search' button. Underneath the search bar is an 'Advanced Search' link. The main content area is divided into two columns. The left column, titled 'Limit your search', contains several dropdown menus for filtering results: Author, Subject, Rights, Language (ISO), Type, Contain Files, and Community. The right column, titled 'Showing 1 through 10 out of 1038 results', displays a list of search results. Each result is shown in a card format with a title, a brief description, the author(s), and a 'Publicly Available' badge. The first result is 'AKCES 2 ver. 2' by Charles University in Prague, ÚČJTK / 2013-12-18, with author(s) Šebesta, Karel ; Goňáková, Hana. The second result is 'A Gold Standard Word Alignment for English-Swedish (2015-10-12)' by Linköping University / 2015-10-12, with author(s) Ahrenberg, Lars ; Holmqvist, Maria. The third result is 'MorphoDITA: Morphological Dictionary and Tagger' by Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics (UFAL) / 2014-02-14, with author(s) Straka, Milan ; Straková, Jana. The page also features a pagination control showing '1 2 3 > 104'.

NAHRÁT A NESTARAT SE

How to Deposit

Only authenticated users can deposit items. If you cannot find your home organisation in the Login dialog list of organisations then register at clarin.eu and authenticate using "clarin.eu website account". In case you cannot use any authentication method above or if you encounter a problem, do not hesitate to contact our [Help Desk](#) and we can create a local account for you.

Step 1: Login

To start a new submission you have to login first. Click Login under My Account: in the right menu panel.

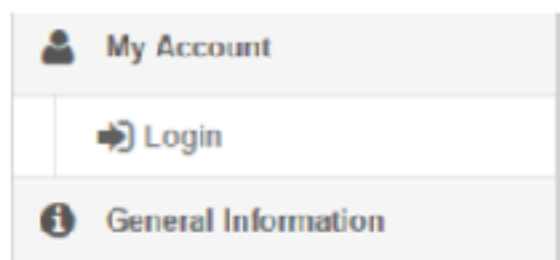
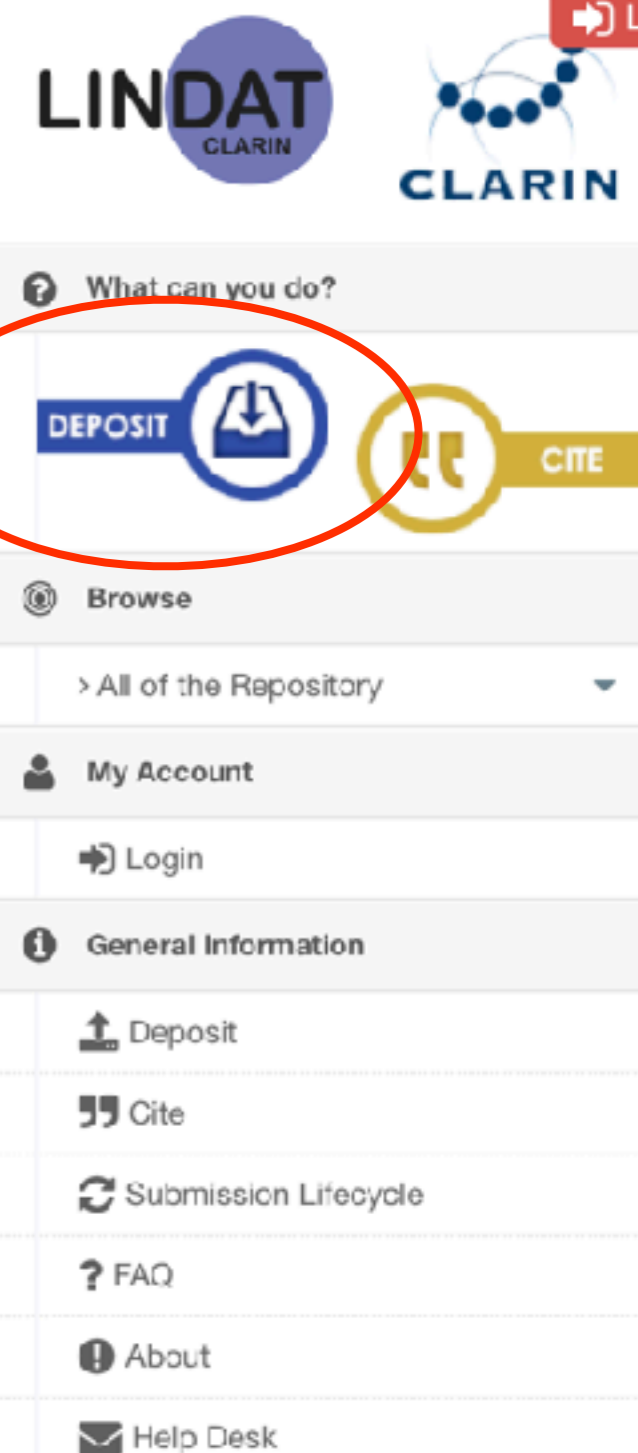
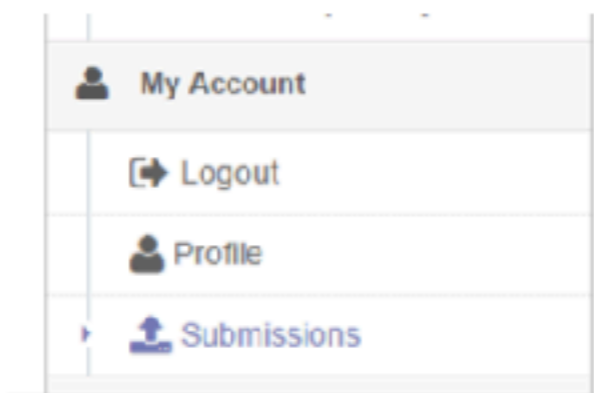


Fig1. Menu Login

Step 2: Starting a new submission

Now you have a new menu item 'Submissions' under My Account. Click on Submissions to go to the Submissions screen.



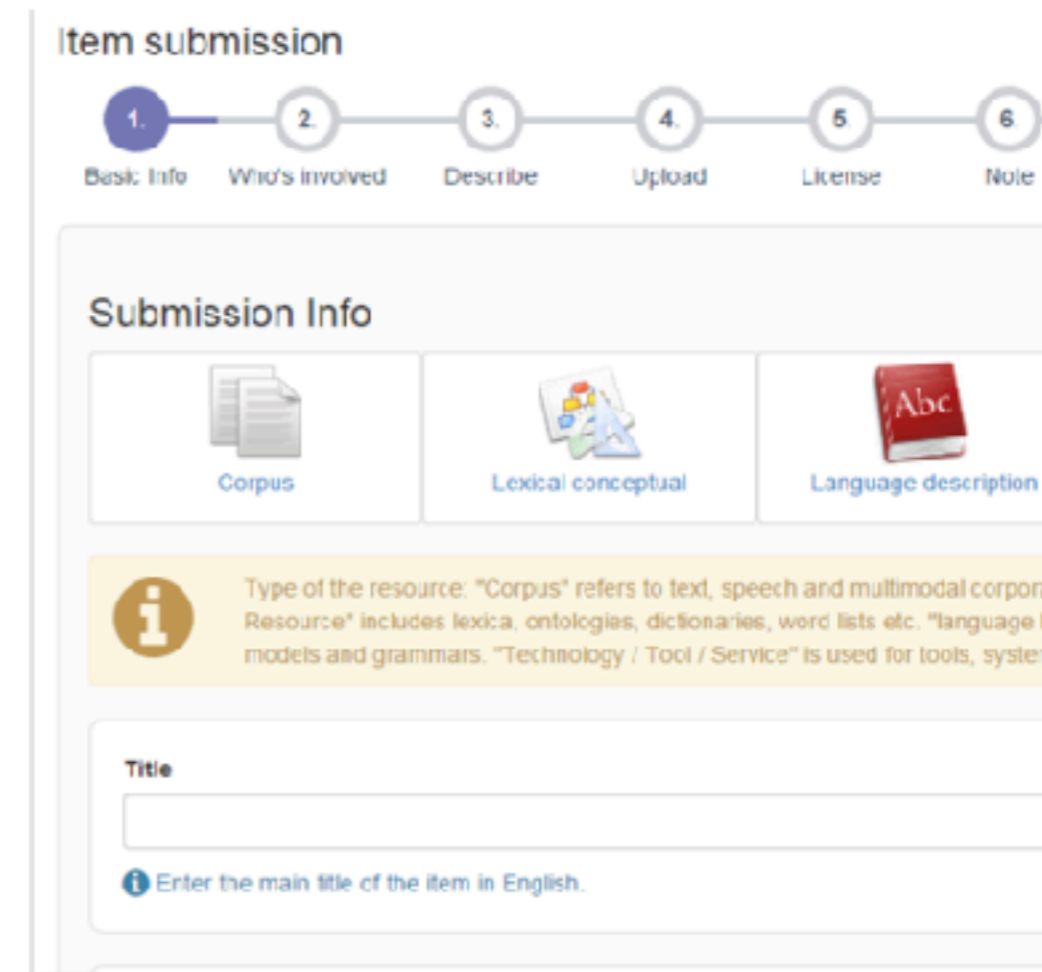
POSTUP ULOŽENÍ

Průvodce krok za krokem

1. login (vlastní institucí)
2. popisná metadata
3. drag&drop nahrát data
4. zvolit licenci
5. zkontrolovat a odeslat

Step 3: Select type of your submission

You have initiated a new workflow item. In the next few steps you will provide the details, select the type of the resource you are about to submit.



Item submission

1. Basic Info 2. Who's Involved 3. Describe 4. Upload 5. License 6. Note

Submission Info

Corpus Lexical conceptual Language description

Type of the resource: "Corpus" refers to text, speech and multimodal corpora. "Resource" includes lexica, ontologies, dictionaries, word lists etc. "Language description" includes models and grammars. "Technology / Tool / Service" is used for tools, systems etc.

Title

Enter the main title of the item in English.

Fig4. Submission info

Click on one of the type buttons e.g. Corpus. Proceed with filling the basic information in the following step.

Step 4: Describe your item

In the following two steps you will provide more details for your item. First describe the item.



Item submission

Basic Info Who's Involved Describe Upload License Note

LOGIN PRO ULOŽENÍ


- vlastní instituce
(člen EduID-cz, EduGAIN)
- CLARIN účet pro
"homeless researchers"
- ukládáme minimum
informací

Sign in to **LINDAT/CLARIN Repository**
Login via Your home institution (e.g. university)

- Univerzita Karlova v Praze
Czech Republic 6 km
- Institute of Biotechnology CAS, v.v.i.
Czech Republic 4 km
- Czech University of Life Sciences Prague
Czech Republic 4 km
- Institute of Art History of the Academy of Sciences of the Czech Republic
Czech Republic 4 km
- College of Polytechnics Jihlava
Czech Republic 4 km
- Global Change Research Institute CAS
Czech Republic 4 km
- Czech Language Institute of the Czech Academy of Sciences
Czech Republic 4 km
- Institute of Chemical Process Fundamentals of the AS CR
Czech Republic 4 km
- Institute of Theoretical and Applied Mechanics AS CR

Q: or search for a provider, such as Example University

▶ **Please help, I cannot find my provider**

 Locate me and show nearby providers

Show providers in **Czech Republic** show all countries

DiscoJuice © UNINETT

V REPOZITÁŘI (S FILTRY)

Advanced Search Search

Limit your search

- Author
- Subject
- Rights
- Language (ISO)
- Type
- Contain Files
- Community

Showing 1 through 10 out of 1038 results

1 2 3 > 104

Corpus LINDAT / CLARIN

AKCES 2 ver. 2

(Charles University in Prague, ÚČJTK / 2013-12-18)

Author(s):
Šebesta, Karel ; Goláňová, Hana

This item contains 1 file (3.85 MB).

Publicly Available

LexicalConceptualResource LRT + Open Submissions

A Gold Standard Word Alignment for English-Swedish (2015-10-12)

(Linköping University / 2015-10-12)

Author(s):
Ahrenberg, Lars ; Holmqvist, Maria

This item contains 1 file (590 KB).

What can you do?

DEPOSIT **CITE**

Browse

> All of the Repository

My Account

Login

General Information

Deposit

Cite

Submission Lifecycle

FAQ

About

Help Desk

V REPOZITÁŘI (S FILTRY)

The screenshot displays the LINDAT/CLARIN repository search interface. At the top, there is a search bar with a magnifying glass icon and a 'Search' button. Below the search bar, the text 'Advanced Search' is visible. On the left side, a 'Limit your search' section is highlighted with a red border, containing several filter categories: Author, Subject, Rights, Language (ISO), Type, Contain Files, and Community. The main content area shows search results, starting with 'Showing 1 through 10 out of 1038 results'. The first result is 'AKCES 2 ver. 2' from Charles University in Prague, dated 2013-12-18, by Šebesta, Karel and Goláňová, Hana. It is a Corpus item, publicly available, and contains 1 file (3.85 MB). The second result is 'A Gold Standard Word Alignment for English-Swedish (2015-10-12)' from Linköping University, dated 2015-10-12, by Ahrenberg, Lars and Holmqvist, Maria. It is a LexicalConceptualResource item, publicly available, and contains 1 file (590 KB). On the right side, there is a navigation menu with options: 'What can you do?', 'DEPOSIT', 'CITE', 'Browse', 'All of the Repository', 'My Account', 'Login', 'General Information', 'Deposit', 'Cite', 'Submission Lifecycle', 'FAQ', 'About', and 'Help Desk'. The LINDAT CLARIN logo is also present in the top right corner.

HLEDÁNÍ DAT A NÁSTROJŮ

GOOGLE



prague dependency treebank 3.0



Vše

Obrázky

Nákupy

Mapy

Videa

Více

Nastavení

Nástroje

Přibližný počet výsledků: 13 900 (0,44 s)

Vědecké články o prague dependency treebank 3.0

Prague dependency treebank 3.0 - **Bejček** - Počet citací tohoto článku: 47

Prague Dependency Treebank - **Hajič** - Počet citací tohoto článku: 385

The Prague dependency treebank - **Böhmová** - Počet citací tohoto článku: 423

Prague Dependency Treebank 3.0 | ÚFAL

<https://ufal.mff.cuni.cz/pdt3.0> ▾ Přeložit tuto stránku

Introduction. The Prague Dependency Treebank 3.0 (PDT 3.0) annotates the same texts as the PDT 2.0 (Hajič et al. 2006), PDT 2.5 (Bejček et al. 2011), and the Prague Discourse Treebank 1.0 (PDiT 1.0, Poláková et al. 2012). The annotation on the four layers was further fixed and improved in various aspects. Moreover ...

The Prague Dependency Treebank 2.0.

<https://ufal.mff.cuni.cz/pdt2.0/> ▾ Přeložit tuto stránku

The Prague Dependency Treebank 2.0 (PDT 2.0) contains a large amount of Czech texts with complex and interlinked morphological (2 million words), syntactic (1.5 MW) and complex semantic annotation ... Please note that new versions of this corpus have been published: PDT 3.0 (2013), PDiT 1.0 (2012), PDT 2.5 (2012).

Prague Dependency Treebank 3.0 (PDT 3.0)

https://lindat.mff.cuni.cz/repository/xmlui/bitstream/.../PDT30_index_lindat.html?...

Prague Dependency Treebank 3.0 (PDT 3.0). Overview. The Prague Dependency Treebank 3.0 (PDT 3.0) annotates the same texts as the PDT 2.0 (Hajič et al. 2006), PDT 2.5 (Bejček et al. 2011), and the Prague Discourse Treebank 1.0 (PDiT 1.0, Poláková et al. 2012). The annotation on the four layers was further fixed ...

HLEDÁNÍ DAT A NÁSTROJŮ

GOOGLE



prague dependency treebank 3.0



Vše

Obrázky

Nákupy

Mapy

Videa

Více

Nastavení

Nástroje

Přibližný počet výsledků: 13 900 (0,44 s)

Vědecké články o prague dependency treebank 3.0

Prague dependency treebank 3.0 - Bejček - Počet citací tohoto článku: 47

Prague Dependency Treebank - Hajič - Počet citací tohoto článku: 385

The Prague dependency treebank - Böhmová - Počet citací tohoto článku: 423

Scholar

Prague Dependency Treebank 3.0 | ÚFAL

<https://ufal.mff.cuni.cz/pdt3.0> ▾ Přeložit tuto stránku

Introduction. The Prague Dependency Treebank 3.0 (PDT 3.0) annotates the same texts as the PDT 2.0 (Hajič et al. 2006), PDT 2.5 (Bejček et al. 2011), and the Prague Discourse Treebank 1.0 (PDiT 1.0, Poláková et al. 2012). The annotation on the four layers was further fixed and improved in various aspects. Moreover ...

The Prague Dependency Treebank 2.0.

<https://ufal.mff.cuni.cz/pdt2.0/> ▾ Přeložit tuto stránku

The Prague Dependency Treebank 2.0 (PDT 2.0) contains a large amount of Czech texts with complex and interlinked morphological (2 million words), syntactic (1.5 MW) and complex semantic annotation ... Please note that new versions of this corpus have been published: PDT 3.0 (2013), PDiT 1.0 (2012), PDT 2.5 (2012).

Prague Dependency Treebank 3.0 (PDT 3.0)

https://lindat.mff.cuni.cz/repository/xmlui/bitstream/.../PDT30_index_lindat.html?...

Prague Dependency Treebank 3.0 (PDT 3.0). Overview. The Prague Dependency Treebank 3.0 (PDT 3.0) annotates the same texts as the PDT 2.0 (Hajič et al. 2006), PDT 2.5 (Bejček et al. 2011), and the Prague Discourse Treebank 1.0 (PDiT 1.0, Poláková et al. 2012). The annotation on the four layers was further fixed ...

GOOGLE DATASET SEARCH



Last updated

Download format

Usage rights

Topic

Free

Saved datasets

33 datasets found

Prague Dependency Treebank
3.5
lindat.mff.cuni.cz
Updated 19. 2. 2018

Data from: Prague Czech-
English Dependency Treebank...
lindat.mff.cuni.cz
Updated Mar 31, 2013

Prague Dependency Treebank -
Consolidated 1.0 (PDT-C 1.0)
lindat.mff.cuni.cz
Updated 12. 1. 2021

Prague Dependency Treebank
of Spoken Czech 2.0 (PDTSC...)

Prague Dependency Treebank - Consolidated 1.0 (PDT-C 1.0)

[Explore at lindat.mff.cuni.cz](https://lindat.mff.cuni.cz)

Dataset updated 12. 1. 2021

Authors

Hajlč, Jan; Bejček, Eduard; Bémová, Alevtina; Buráňová, Eva; Fučíková, Eva; Hajlčová, Eva; Havelka, Jiří; Hlaváčová, Jaroslava; Homola, Petr; Irčing, Pavel; Kárník, Jiří; Kettnerová, Václava; Klyueva, Natalia; Kolářová, Veronika; Kučová, Lucie; Lopatková, Markéta; Mareček, David; Mikulová, Marie; Mirovský, Jiří; Nedoluzhko, Anna; Novák, Michal; Pajas, Petr; Panevová, Jarmila; Peterek, Nino; Poláková, Lucie; Popel, Martin; Popelka, Jan; Romportl, Jan; Rysová, Magdaléna; Semecký, Jiří; Sgall, Petr; Spoustová, Johanka; Straka, Milan; Straňák, Pavel; Synková, Pavlína; Ševčíková, Magda; Šindlerová, Jana; Štěpánek, Jan; Štěpánková, Barbora; Toman, Josef; Uřešová, Zdeňka; Vidová Hladká, Barbora; Zeman, Daniel; Zikánová, Šárka; Žabokrtský, Zdeněk

License

[Attribution-NonCommercial-ShareAlike 4.0 \(CC BY-NC-SA 4.0\)](#)

License information was derived automatically

Description

A richly annotated and genre-diversified language resource, The Prague Dependency Treebank – Consolidated 1.0 (PDT-C 1.0, or PDT-C in short in the sequel) is a consolidated release of the existing PDT-corpora of Czech data, uniformly annotated using the standard PDT scheme. PDT-corpora included in PDT-C: Prague Dependency Treebank (the original PDT contents, written newspaper and journal texts from three genres); Czech part of Prague Czech-English Dependency Treebank (translated financial texts, from English), Prague Dependency Treebank of Spoken Czech (spoken data, including audio and transcripts

enTenTen

“ Please use the following text to cite this item or export to a predefined format: ”

BIBTEX CMDI

Masaryk University, NLP Centre, 2011, *enTenTen*, LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University. <http://hdl.handle.net/11858/00-097C-0000-0001-CCDF-8>.



This resource is also integrated in following services:

Share:

KonText

| | |
|-----------------|---|
| Item identifier | http://hdl.handle.net/11858/00-097C-0000-0001-CCDF-8 |
| Date issued | 2011-12-16 |
| Type | corpus |
| Language(s) | English |
| Description | Very large English web corpus enTenTen, comprising 3,268,798,627 tokens. |
| Publisher | Masaryk University, NLP Centre |
| Acknowledgement | Lexical Computing Ltd. |
| Subject(s) | English large corpus |
| Collection(s) | LINDAT / CLARIN Data & Tools |

Show full item record

LINDAT
CLARIN

CLARIN

Pavel Straňák | Logout

What can you do?

DEPOSIT



CITE



Browse

> All of the Repository

My Account

Logout

Profile

Submissions

Context

> Edit this item

> Export Item

> Export Metadata

Administrative

Control Panel

Access Control

TAK OTEVŘENÉ, JAK TO JDE

Choose a License

Answer the questions or use the search to find the license you want

↻ Start again



What do you want to deposit?

Software

Data

Search for a license...

Public Domain Mark (PD)

The work identified as being free of known restrictions under copyright law, including all related and neighboring rights.

Publicly Available



Public Domain Dedication (CC Zero)

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

Publicly Available



OPEN DATA

JEN TAK OTEVŘENÉ, JAK TO JDE

 **Publisher** Faculty of Arts, Institute of the Czech National Corpus, Charles University in Prague

 **Acknowledgement** Ministerstvo školství, mládeže a tělovýchovy

Project code: LM2011023

Project name: Český národní korpus

 **Subject(s)** representative corpus written language

 **Collection(s)** LINDAT / CLARIN Data & Tools

[Show full item record](#)

 **Files in this item**



Download instructions for command line

This item is Academic Use and licensed under:
Czech National Corpus (Shuffled Corpus Data)



| | |
|--------------------|----------------------------------|
| Name | syn2015.gz |
| Size | 1.48 GB |
| Format | application/x-gzip |
| Description | corpus |
| MD5 | e0242cc77e999794af6cfaf57f843c12 |



[Download file](#)

JASNÁ PRAVIDLA
VLASTNÍ LICENCE
PODPISY LICENCÍ

ZVÝRAZNIT NEJNOVĚJŠÍ, POSKYTNOUT VŠECHNY

Project name: Internet jako jazykový korpus

Ministerstvo školství, mládeže a tělovýchovy České republiky

Project code: LN00A063

Project name: Centrum počítační lingvistiky

Ministerstvo školství, mládeže a tělovýchovy České republiky

Project code: MSM 0021620838

Project name: Moderní metody, struktury a systémy informatiky

Subject(s)

MorphoDiTa

Czech

morphological analysis

morphological generation

PoS tagging

Collection(s)

LINDAT / CLARIN Data & Tools



This item is replaced by a newer submission:

<http://hdl.handle.net/11234/1-1836>

Please refer to the submission above for the latest available data. If you nevertheless need the original data, please click [here](#).

List all versions ▼

ZVÝRAZNIT NEJNOVĚJŠÍ, POSKYTNOUT VŠECHNY

Collection(s)

LINDAT / CLARIN Data & Tools

Other versions

List all versions ▾

- ▶ Czech Models (Morfflex CZ 161115 + PDT 3.0) for MorphoDiTa 161115
- Czech Models (Morfflex CZ 160310 + PDT 3.0) for MorphoDiTa 160310
- Czech Models (Morfflex CZ + PDT) for MorphoDiTa

[Show full item record](#)

Files in this item



Download instructions for command line

This item is **Publicly Available** and licensed under:

Creative Commons - Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)



| | |
|--------------------|---|
| Name | czech-morfflex-pdt-161115.zip |
| Size | 69.18 MB |
| Format | application/zip |
| Description | Czech Models (Morfflex CZ 161115 + PDT 3.0) for MorphoDiTa 161115 |
| MD5 | adde38cd363219759e19165b06baa4ce |



[Download file](#)

[Preview](#)

Open SDP

Please use the following text to cite this item or export to a predefined format: [BIBTEX](#) [GMD](#)

Flickinger, Dan; Hajič, Jan; Ivanova, Angelina; et al., 2016, Open SDP, LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics, Charles University in Prague, <http://hdl.handle.net/11234/1-1742>.

Share: [f](#) [t](#) [s](#)

LINDAT / CLARIN

| | |
|-------------|---|
| Authors | Flickinger, Dan ; Hajič, Jan ; Ivanova, Angelina ; Kühnmann, Marco ; Mlyao, Yusuka ; Depen, Stephan ; Zeman, Daniel |
| Project URL | http://sdp.delph-in.net/ |
| Date issued | 2016-06-25 |
| Type | corpus |
| Size | 391062932 bytes |
| Language(s) | Czech English |
| Description | <p>The original SDP 2014 and 2015 data collections were made available under task-specific 'evaluation' licenses to registered SemEval participants. In mid-2016, all original data has been bundled with system submissions, supporting software, an additional SDP-style collection of semantic dependency graphs, and additional background material (from which some of the SDP target representations were derived) for release through the Linguistic Data Consortium (with LDC catalogue number LDC2016T10).</p> <p>One of the four English target representations (viz. DM) and the entire Czech data (in the PSD target representation) are not derivative of LDC-licensed annotations and, thus, can be made available for direct download (Open SDP; version 1.1; April 2016) under a more permissive licensing scheme, viz. the Creative Commons Attribution-NonCommercial-ShareAlike license. This package also includes some 'richer' meaning representations from which the English bi-lexical DM graphs derive, viz. scope-underspecified logical forms and more abstract, non-lexicalized 'semantic networks'. The latter of these are formally (if not linguistically) similar to Abstract Meaning Representation (AMR) and are available in a range of serializations, including in AMR-like syntax.</p> |

What can you do?

DEPOSIT [CITE](#)

Browse

> All of the Repository

My Account

Login

Statistics

Statistics **BETA**

General Information

Deposit

Cite

Submission Lifecycle

FAQ

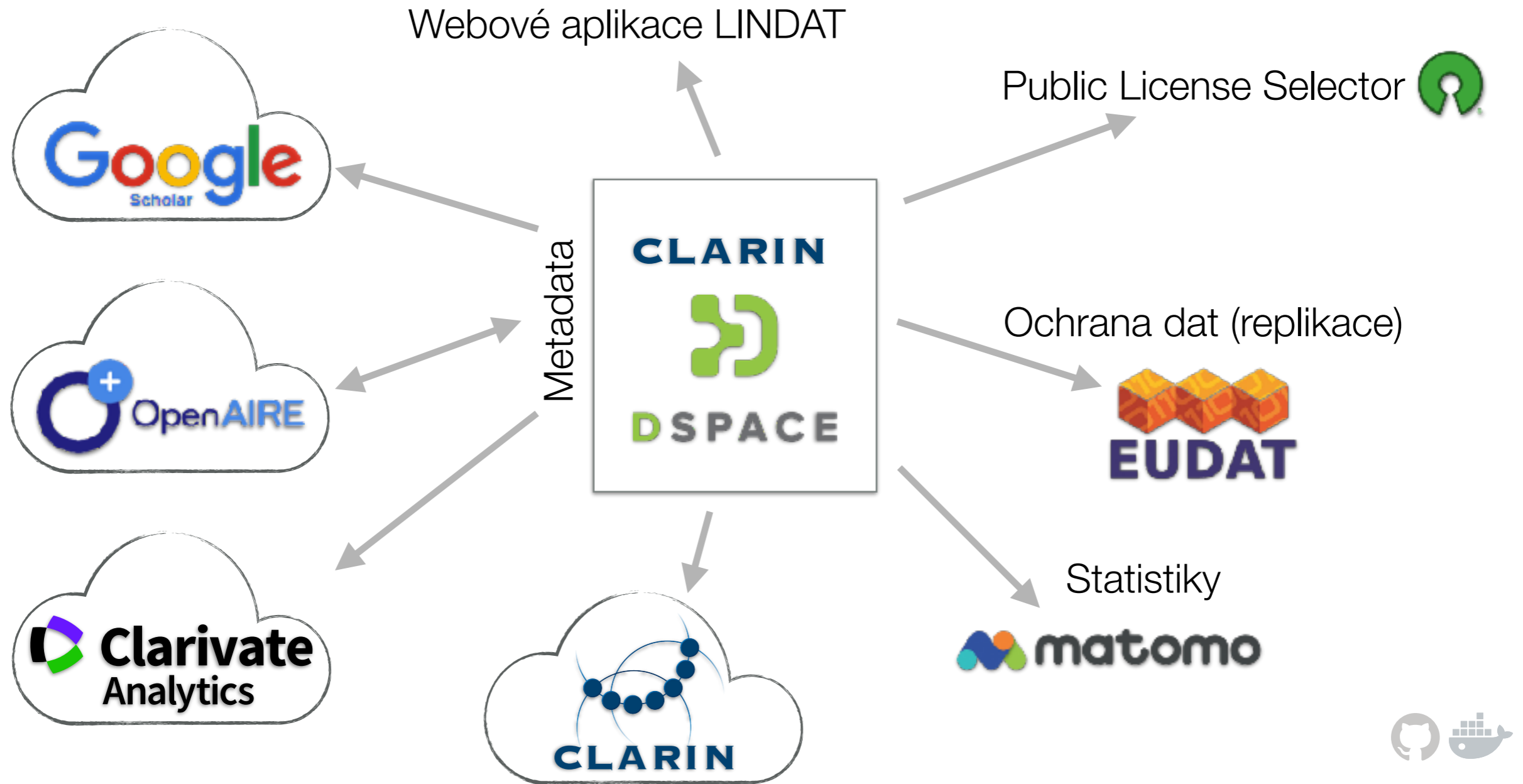
ABOUT

Help Desk

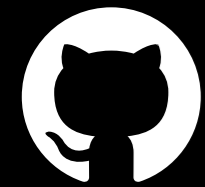
Statistiky datasetu

- Počty návštěv a stažení
- Měsíční reporty mailem (sledování datasetu)

Integrace



PUBLIC LICENSE SELECTOR



<https://github.com/ufal/public-license-selector>

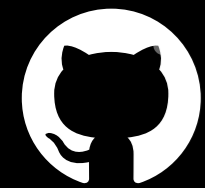
- veřejné licence: bez podpisu, pro veřejné šíření
- data / software
- Průvodce s vysvětlením
- Směřuje k co největší otevřenosti
- Open source / Open data
 - Výběr nejvhodnějších licencí
- Snadné užití a integrace



The screenshot shows the GitHub repository page for 'ufal / public-license-selector'. The repository has 95 commits, 3 branches, and 5 releases. A pull request #9 is open, titled 'Merge pull request #9 from Aldarone/Aldarone-patch-1'. The pull request details show a list of files and their changes:

| File | Change |
|-------------------|--|
| ci | Removed remnants of Clarin version and changed b |
| src | Change software copyleft: question wording to be c |
| .editorconfig | Removed remnants of Clarin version and changed b |
| .gitignore | Removed remnants of Clarin version and changed b |
| .mversionrc | Removed remnants of Clarin version and changed b |
| .travis.yml | Fix year in license file and set node version to stabl |
| LICENSE.md | Fix year in license file and set node version to stabl |
| Makefile | Removed remnants of Clarin version and changed b |
| README.md | Update README.md |
| bower.json | Releasing 0.0.6 |
| index.html | Removed remnants of Clarin version and changed b |
| package.json | Releasing 0.0.6 |
| webpack.config.js | Removed remnants of Clarin version and changed b |

CLARIN-DSPACE



<https://github.com/ufal/clarin-dspace>

- MIT licence (naprostá volnost)
- DSpace upravený pro optimální datový repozitář pro naše data
- + licencování, verzování, snazší správa, náhledy dat, atd.
- dokumentace a otevřený vývoj na githubu
- 15+ nasazení v 10+ zemích
- od roku 2009



The screenshot shows the GitHub repository page for 'ufal / clarin-dspace', which is a fork of 'DSpace/DSpace'. The repository has 8,658 commits and 43 branches. The current branch is 'clarin', which is 1368 commits ahead and 2394 commits behind the 'DSpace:master' branch. A recent merge by 'kosarko' is shown, merging the 'release-2018.01' branch into 'clarin'. Below the merge, a list of sub-projects is displayed, each with a link to its respective issue or pull request.

| Sub-project | Issue/Pull Request |
|-----------------|---------------------------------------|
| dspace-api | Resolves #412 - Access other versions |
| dspace-jspui | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-lni | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-oai | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-rdf | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-rest | Resolves #412 - Access other versions |
| dspace-services | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-solr | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-sword | Resolves #816 - Merge DSpace-5.8 (# |
| dspace-swordv2 | Resolves #816 - Merge DSpace-5.8 (# |

FAIR REPOZITÁŘ PRO (BRZY NEJEN) JAZYKOVÁ DATA

F*indable*: Google, Google Scholar, Data Citation Index, CLARIN VLO, OLAC... I hledání v samotném repozitáři

A*ccessible*: metadata + data (i citlivá), kompletní licencování, Open Access (Public License Selector), login jen tam, kde je nutný

I*nteroperable*: běžné datové formáty, plná dokumentace (rozšířená metadata, uložené soubory s dokumentací vedle souborů samotných dat)

R*eusable*: záznamy s plnými daty, zveřejněná data nelze měnit, licence, verzování, přímé citace dat (PID), maximální OpenAccess

PODPORA

- projekt LINDAT/CLARIAH-CZ (LM2018101; dříve LM2010013, LM2015071) podporovaný Ministerstvem školství, mládeže a tělovýchovy v programu LM "Velké výzkumné infrastruktury"
- Tato prezentace byla podpořena projektem OP VV Rozvoj kapacit pro výzkum a vývoj UK II, CZ.02.2.69/0.0/0.0/18_054/0015222

Děkujeme!
<http://lindat.cz>

Dotazy?

