

# DiDa

## Dokumentace k výsledku A ve formě elektronické databáze



**Nářečí českého jazyka interaktivně.**

**Dokumentace a zpřístupnění mizejícího jazykového dědictví jako nedílné součásti regionálních identit.**

DG20P02OVV029; poskytovatel podpory Ministerstvo kultury, Program NAKI II.



# Obsah

<b>Základní údaje</b>	<b>3</b>
<b>Charakteristiky výsledku</b>	<b>4</b>
Co DiDa obsahuje?	4
K čemu DiDa slouží?	4
Čím je databáze nová a jedinečná?	5
Naplnění cílů programu NAKI II a jeho očekávaných přínosů	6
<b>Databázové schéma</b>	<b>7</b>



# Základní údaje

**Název výsledku:**

DiDa

**Verze:**

1.0

**Druh výsledku:**

A - audiovizuální tvorba, elektronické dokumenty

**Lokalizace výsledku (dump databáze):**

<https://www.ceskanareci.cz/geoportal/dida/dump.sql>

**Webová prezentace výsledku:**

<https://www.ceskanareci.cz/geoportal/diama/>

**Název vlastníka:**

Ústav pro jazyk český AV ČR, v. v. i.

**Technické parametry výsledku:**

SQL databáze PostgreSQL

**Ekonomické parametry výsledku:**

Databáze obsahuje přes 50 000 nářečních ukázek a zhruba 3 200 zvukových exemplifikací, které jsou součástí 154 heslových slov (102 morfologických a 52 hláskoslovných).

**Druh možnosti využití výsledku jiným subjektem:**

A - k využití výsledku jiným subjektem je vždy nutné nabytí licence

**Požadavek na licenční poplatek:**

N - poskytovatel licence na výsledek nepožaduje licenční poplatek

**Typ licence:**

ODbL

Databáze je zpřístupněna pod licencí ODbL 1.0 (Open Data Commons Open Database License), která umožňuje uživatelům databázi svobodně sdílet, upravovat a používat při dodržení zásady uvedení autora databáze, zachování licence a otevřenosti dat.

# Charakteristiky výsledku

**DiDa** (dialektologická databáze elektronických dokumentů) obsahuje různorodý nářeční materiál, spravovaný prostřednictvím elektronických formulářů softwaru ProPONAR, který taktéž vznikl v rámci projektu *Nářečí českého jazyka interaktivně. Dokumentace a zpřístupnění mizejícího jazykového dědictví jako nedílné součásti regionálních identit.* (DG20P02OVV029; poskytovatel podpory Ministerstvo kultury, Program NAKI II; výsledek byl uplatněn 2021)

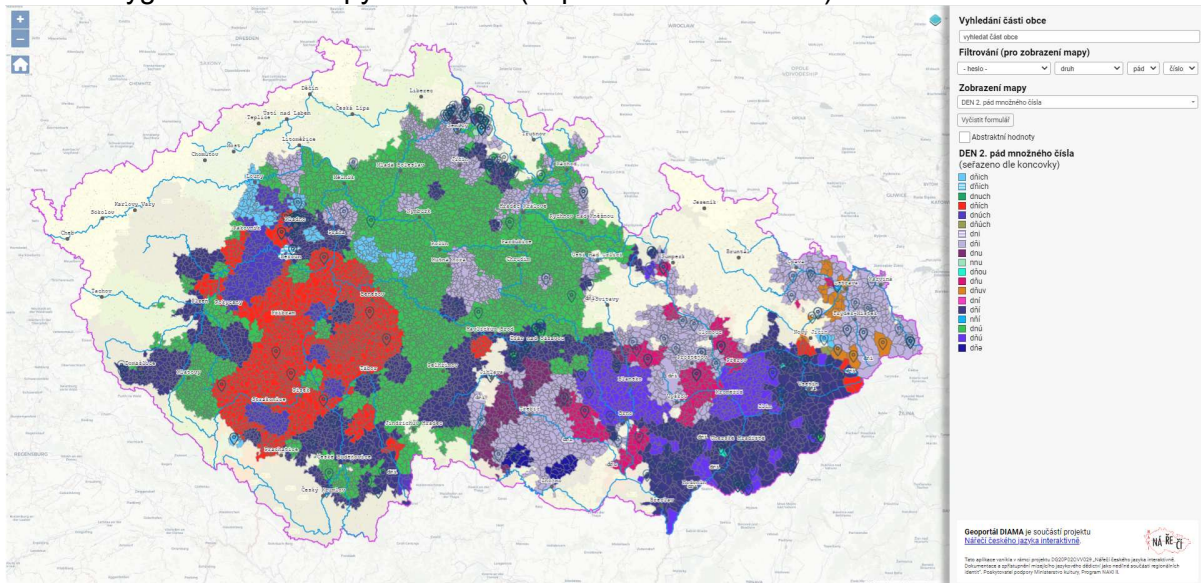
## Co DiDa obsahuje?

Databáze se skládá z textových a zvukových nářečních ukázek, které byly získány z více než 1 000 různých dialektologických zdrojů obsahujících data z let 1847–2022: asi třetinu tvoří nahrávky a excerpta z terénních nářečních výzkumů, uskutečněných dialektologickým oddělením ÚJČ, dále jde o soubor nářečních textů, z nichž mnohé má ÚJČ ve výhradním vlastnictví nebo jde o publikace a texty dnes mimo archiv oddělení velmi obtížně dostupné. Využity byly i exemplifikace z nářečních monografií. Celkem bylo takto získáno přes 50 000 textových exemplifikací a přibližně 3 200 zvukových exemplifikací, a to v rámci 154 hesel (102 morfologických, 52 hláskoslovných). Každý záznam je opatřen informací o roku a lokalitě sběru a zdroji, ze kterého pochází.

## K čemu DiDa slouží?

Data uložená v databázi slouží k sestavení interaktivních analytických map (tedy morfologických a hláskoslovných map s interaktivními body, po jejichž rozkliknutí je možné zobrazit nebo přehrát multimediální obsah lokalizovaný do daného bodu: text, audio), které jsou zpracovány pomocí programu ProMAP a prezentovány prostřednictvím geoportálu DiaMa (<https://www.ceskanareci.cz/geoportal/diama/>). Zmíněné nástroje byly vyvinuty v rámci téhož projektu.

Ukázka vygenerované mapy hesla *den* (2. pád množného čísla):



Ukázka detailu konkrétní lokality se zvukovým záznamem, jeho přepisem a informacemi o významu daného slova, zdroji záznamu a roku sběru:

Ukázka úseku ze seznamu exemplifikací daného slova v databázi: součástí je zvukový soubor, přepis nahrávky a informace o zdroji záznamu a významu slova:

🔊	📄	a druhé (deř, 4 sg.), díš sem chcela, aš si abichme feca mřeli, tak sem musela pjeřve sřest a uš sem se mosela uzeřti po porođe, abch se najeda, protože sem maminku nemřela, žádnú nemřela	Zbraslav BO	N001_01 Zbraslav BO, 1969. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:45:49
🔊	📄	no a vichovala sem je take. A pře řich sem řšachno uřelala. Bilo tomo tedi řtmážd (dní, 2 pl.), uš sem a tēm řla na roli	Zbraslav BO	N001_01 Zbraslav BO, 1969. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:48:12
🔊	📄	no a pak se mřářilo cepama, diki ve žlich se mřářivalo cepama, řtmážd (dní, 2 pl.) sme třeba mřářili	Zbraslav BO	N001_01 Zbraslav BO, 1969. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:50:02
🔊	📄	potom díř tam bila hořne tráva, tak to se muselo nechat (deř, 4 sg.) nebo dva a eše se to muselo vobřáři, no, a pak sme to vřáři	Horní Poříčí BK		časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.10.2022 12:22:06
🔊	📄	no tak na desed (dní, 2 pl.), naz bilo doma sedum, pjeđ deři a voři dvá. To tak... Tak na desed (dní, 2 pl.). A to se speřlo pjetadvacet kil mři	Písařov SU	N093_01 Písařov SU, 1972. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:53:14
🔊	📄	no to bili krásni zábavi, třeba ti (dní, 4 pl.) sme bili u zábavi, po ti (dní, 4 pl.). A do rána	Písařov SU	N093_01 Písařov SU, 1972. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:54:25
🔊	📄	a daž Viktorce balo sedumdesád (dní, 2 pl.)... e.. padesád (dní, 2 pl.), sedum nedel a (deř, 1 sg.), tak mře umřela maminka	Chroměč SU	N150_03 Chroměč SU, 1967. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:55:24
🔊	📄	von tu běl pan kanovřik na penzi, mřival každé (deř, 4 sg.) mři svato	Chroměč SU	N151_01 Chroměč SU, 1967. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:56:51
🔊	📄	a daš Mařence balo bes řtinářř (dní, 2 pl.) rok, tak se narofila Viktorka	Chroměč SU	N150_03 Chroměč SU, 1967. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:57:52
🔊	📄	akorád za rok se narofil prvni řhasep. Z. bez hořina za řtera (dní, 4 pl.) umřel	Chroměč SU	N150_03 Chroměč SU, 1967. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:58:45
🔊	📄	mi sme mřeli prařnicu a ta mřela selata. A voni ti selatika bili starí asi řtmážd (dní, 2 pl.)	Hodějice VY	N149_01 Hodějice VY, 1967. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 10:59:55
🔊	📄	na řřeti (deř, 4 sg.) nřz vzal řž vř, ale řž nřz netahí na takovi mřsta	Otaslavice PV	N091_02 Otaslavice PV, 1969. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 11:02:00
🔊	📄	tož nřz nechal dva (dní, 4 pl.) doma a mřel řzme škol	Otaslavice PV	N091_02 Otaslavice PV, 1969. Archiv zvukových záznamů nářečnických promluv.	časový úsek 24 hodin, zprav. od půlnoci do půlnoci	04.12.2022 11:02:36

## Čím je databáze nová a jedinečná?

Databáze zpřístupňuje doposud nezveřejněný nářeční materiál v podobě textových i audiálních záznamů. Jejří propojení s dalšími nástroji vybudovanými v rámci tohoto projektu navíc umožňuje prezentaci odborných dialektologických dat pomocí přehledných map, které jsou srozumitelné i pro laickou veřejnost. Mohou tak najřt využití nejen u odborníků z řad lingvistů, kartografů, etnologů či historiků, ale též mezi pedagogy a studenty nebo vůbec zájemci o nářečí. V tomto rozsahu a mře podrobnosti dosud dialektologický materiál v českém prostředí zpracován nebyl.

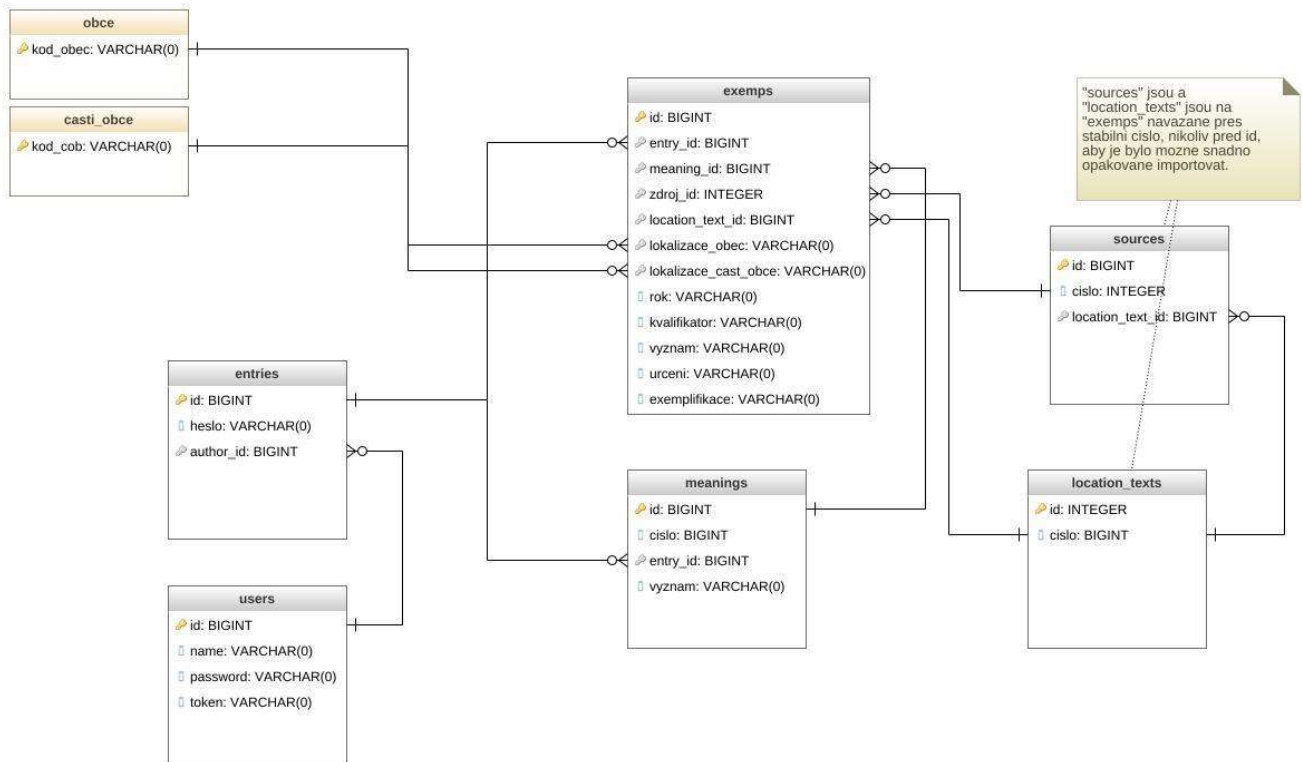
## Naplnění cílů programu NAKI II a jeho očekávaných přínosů

Databáze DiDa naplňuje řadu očekávaných přínosů programu NAKI II:

- Digitalizace nářečního materiálu a jeho shromáždění na jednom místě napomáhá k **uchování nehmotného kulturního dědictví pro další generace** (Očekávané přínosy bod 1).
- Náplň uložených dat je často dokladem o reáliích starých, již mizejících, a jejich uchování tedy přispívá k **záchraně a zlepšování hodnot kulturního dědictví** (Očekávané přínosy bod 2).
- Databáze je podkladem pro vizualizace dat na interaktivních mapách, a umožňuje tak **zpřístupnění kulturního dědictví široké obci zájemců a uživatelů** (Očekávané přínosy bod 3).
- Předpokládá se využití databáze prostřednictvím geoportálu DiaMa ve výuce, čímž budou **využity výsledky výzkumu ve vzdělávání, založeném na poznání národní identity a kultury** (Očekávané přínosy bod 5).

## Databázové schéma

Databáze se skládá z tabulek, které zabezpečují lokalizaci (okresy, obce, části obce...), části pro modelování hesel a exemplifikací (hesla, významy, exemplifikace...) a dalších podpůrných částí (zdroje, uživatelé...).



(<https://github.com/proponar/dida-entry/blob/master/doc/DIDA-base-schema-white.png>)