



MAPOVÁNÍ A KARTOGRAFICKÁ VIZUALIZACE NÁŘEČNÍCH JEVŮ ČESKÉHO JAZYKA

Vít Voženílek, Jakub Koníček

Univerzita Palackého v Olomouci, Katedra geoinformatiky

Martina Ireinová

Ústav pro jazyk český AV ČR, dialektologické oddělení

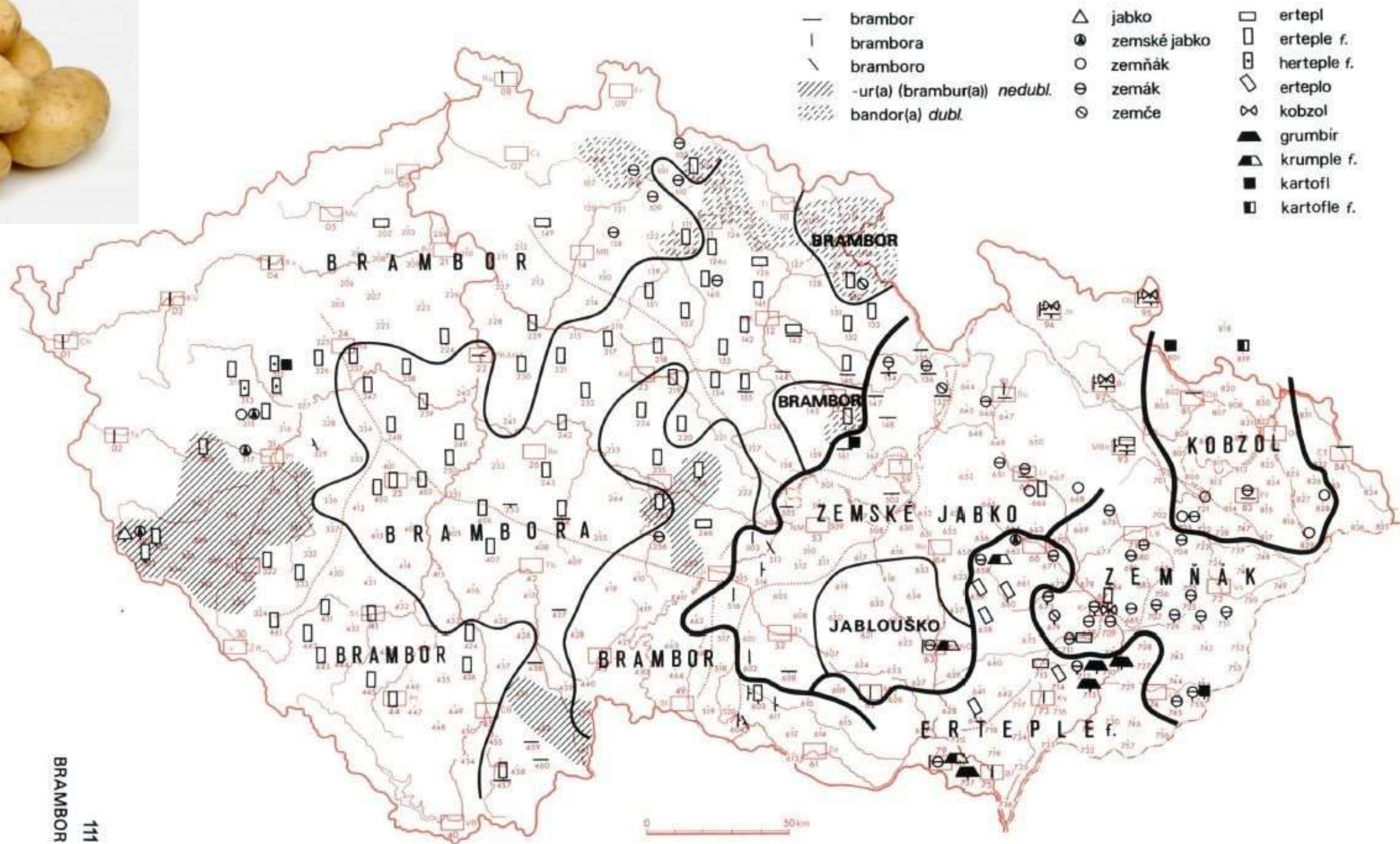


KATEDRA GEOINFORMATIKY
Univerzita Palackého v Olomouci



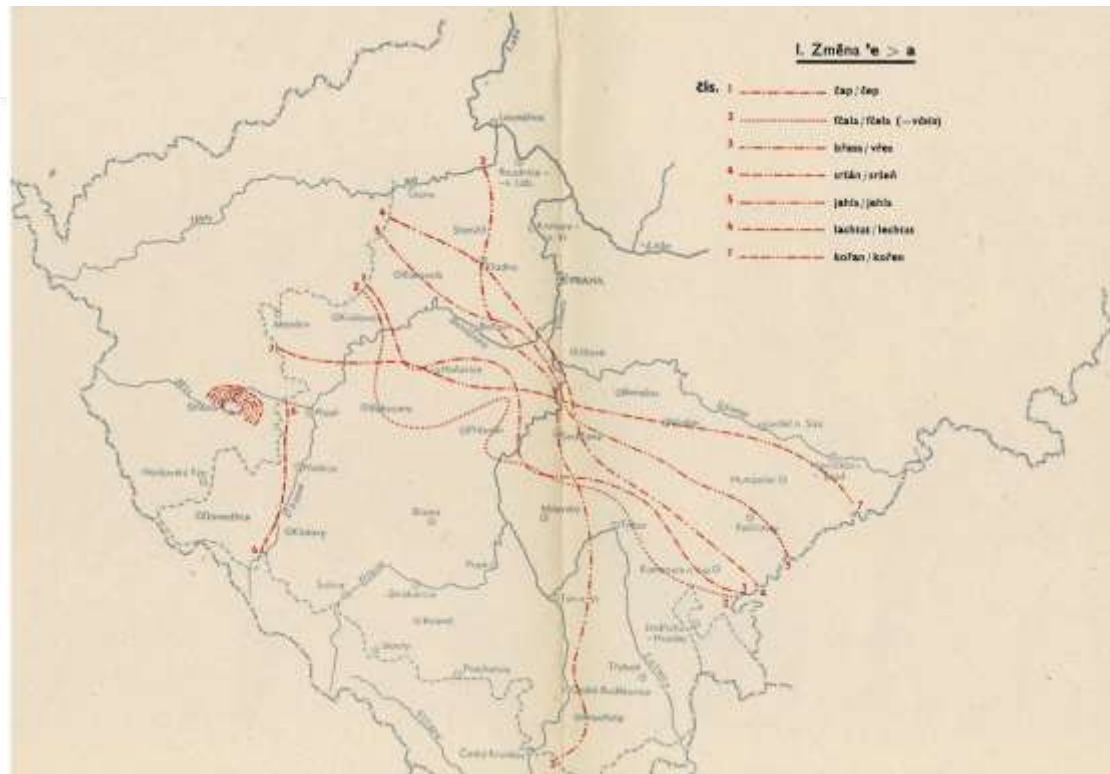
Ústav pro jazyk český
Akademie věd České republiky







NÁŘEČNÍ VÝZKUMY



2. Jak se vyslovují slova:

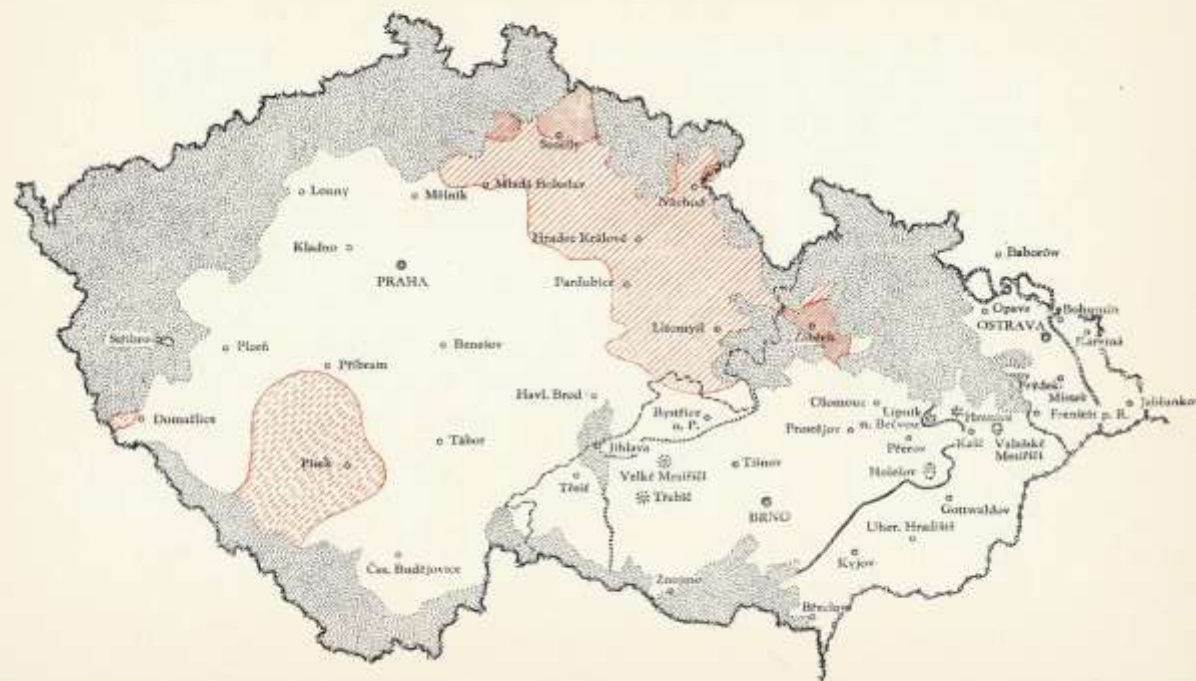
ať, aťsi, teď, dokud, pokud, buď, buďte, plať, plaťte, seď, seďte, raď, raďte ?

/na př. ať, aťsi, dokud, pokud, buď, buďte a pod. Uveďte v nářečí obce! /

odtud - potud
votaj - potaj
dohad - potud
dohad - potud
skoni - potud
ploj - potud
zaplat - potud
sej - potud
sej - potud
poraj - potud
raj - potud

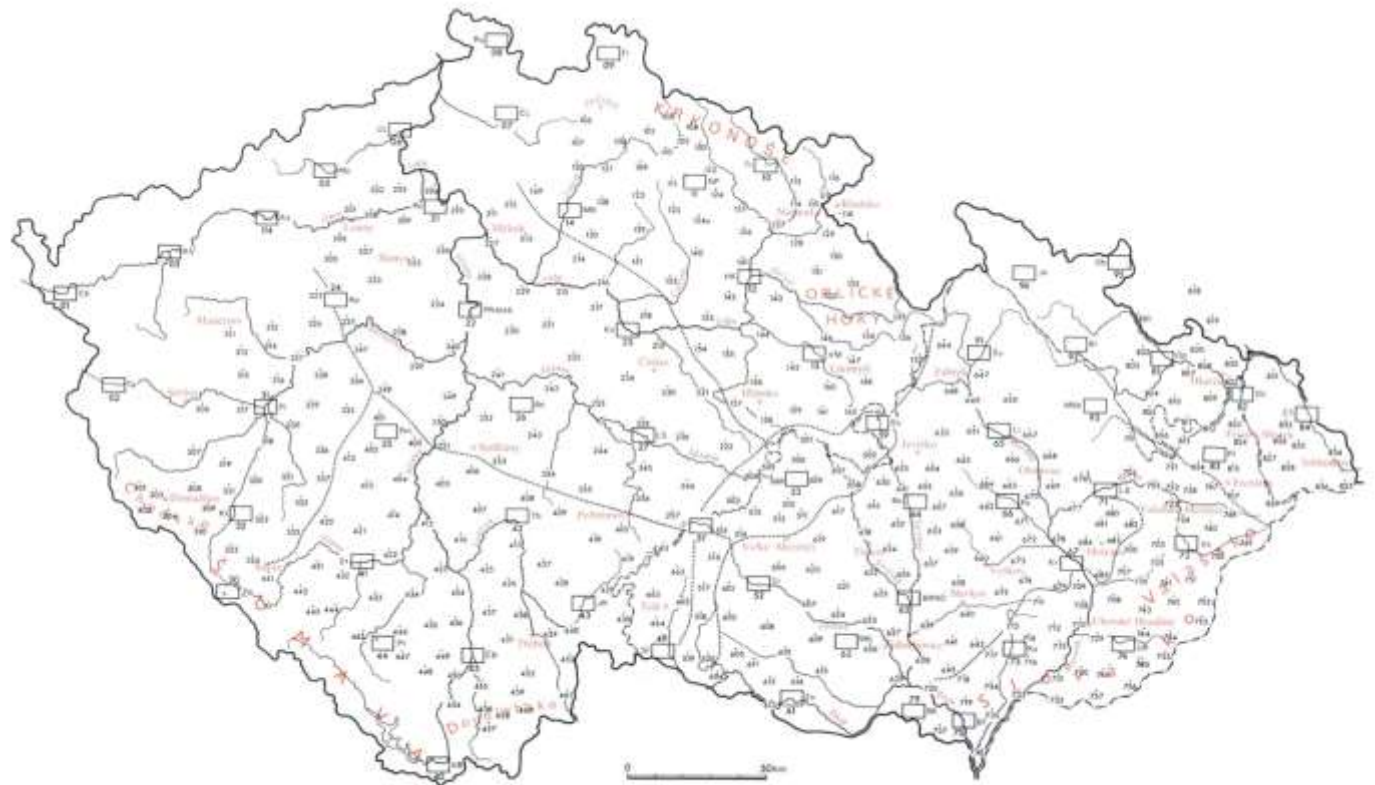


4. Zbytky obouretného w a stopy po něm



Výzkumná síť lokalit

- **420 venkovských obcí** (223 v Čechách, 197 na Moravě a ve Slezsku) + 5 vesnic ležících v dnešním Polsku
- **57 měst**: 37 měst vnitrozemských, 20 pohraničních



IL. POMOCNÁ ORIENTAČNÍ MAPA

Informátoři

- 4 364 informátorů na venkově: 65–75 let
- 1 032 informátorů ve městech: 65–75 let (365), 14–15 let (667)

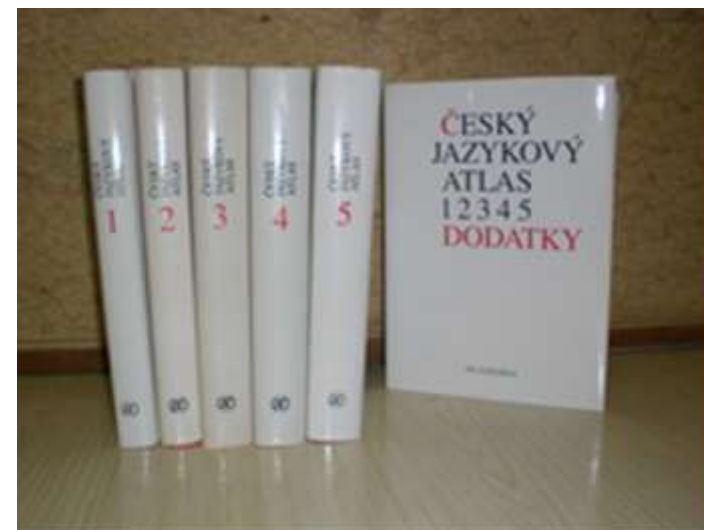


Terénní výzkum



ČESKÝ JAZYKOVÝ ATLAS (1992–2011)

- 1 558 map a 1 578 komentářů
- 1., 2. a 3. díl – nářeční lexikální zásoba
- 4. díl – jevy morfologické
- 5. díl – jevy hláskoslovné a syntaktické, tvoření adverbíí, výzkum městské mluvy, svazky izoglos
- **Dodatky** – 70 nahrávek (2 CD), rejstřík, ...



<https://cja.ujc.cas.cz/>
<https://cja.ujc.cas.cz/e-cja/>

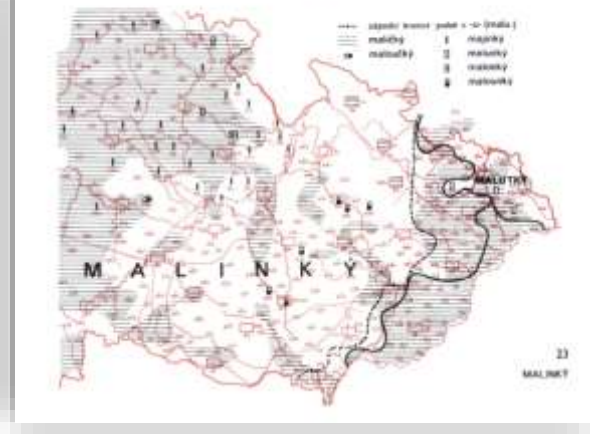
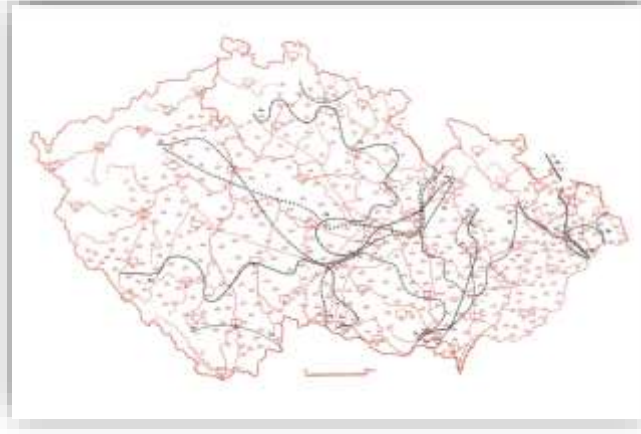
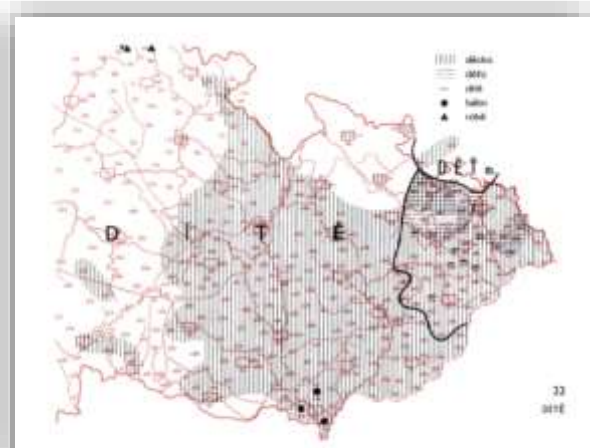
Projekt 2020–2022

- ***Nářečí českého jazyka interaktivně. Dokumentace a zpřístupnění mizejícího jazykového dědictví jako nedílné součásti regionálních identit***
(poskytovatel podpory Ministerstvo kultury, Program NAKI II, č. DG20P02OVV029)
- <http://geoinformatics.upol.cz/app/nareci/>



Český jazykový atlas pohledem kartografa

- **1 600 map**
 - I. Celostránkové základní mapy zkoumaného území
 - II. Půlstránkové mapy celého zkoumaného území
 - III. Půlstránkové výseče s moravskou a slezskou oblastí
 - IV. Svodné mapy izoglosové
- **Jednoduché grafické provedení**
- **Komplikovaný znakový klíč**
- **Využití metod bodových, liniových a plošných znaků**
- **(Pseudo)izolinie – izoglosy**



Atlas nářečí českého jazyka: krácení vokálů

upgrade 1 – digitalizace a kolorizace



Atlas nářečí českého jazyka: krácení vokálů

upgrade 1 – digitalizace a kolorizace

- Systematická vizualizace **prostorové syntézy dialektologických dat** se zaměřením na krácení vokálů
- **16 vybraných slov z 5. dílu** původního Českého jazykového atlasu
- **12 slov vstupovalo do syntézy**
- Vektorizace a vizualizace map
 - **16 analytických map**
 - **2 syntetické mapy – typizace a regionalizace**
- Návrh nového znakového klíče

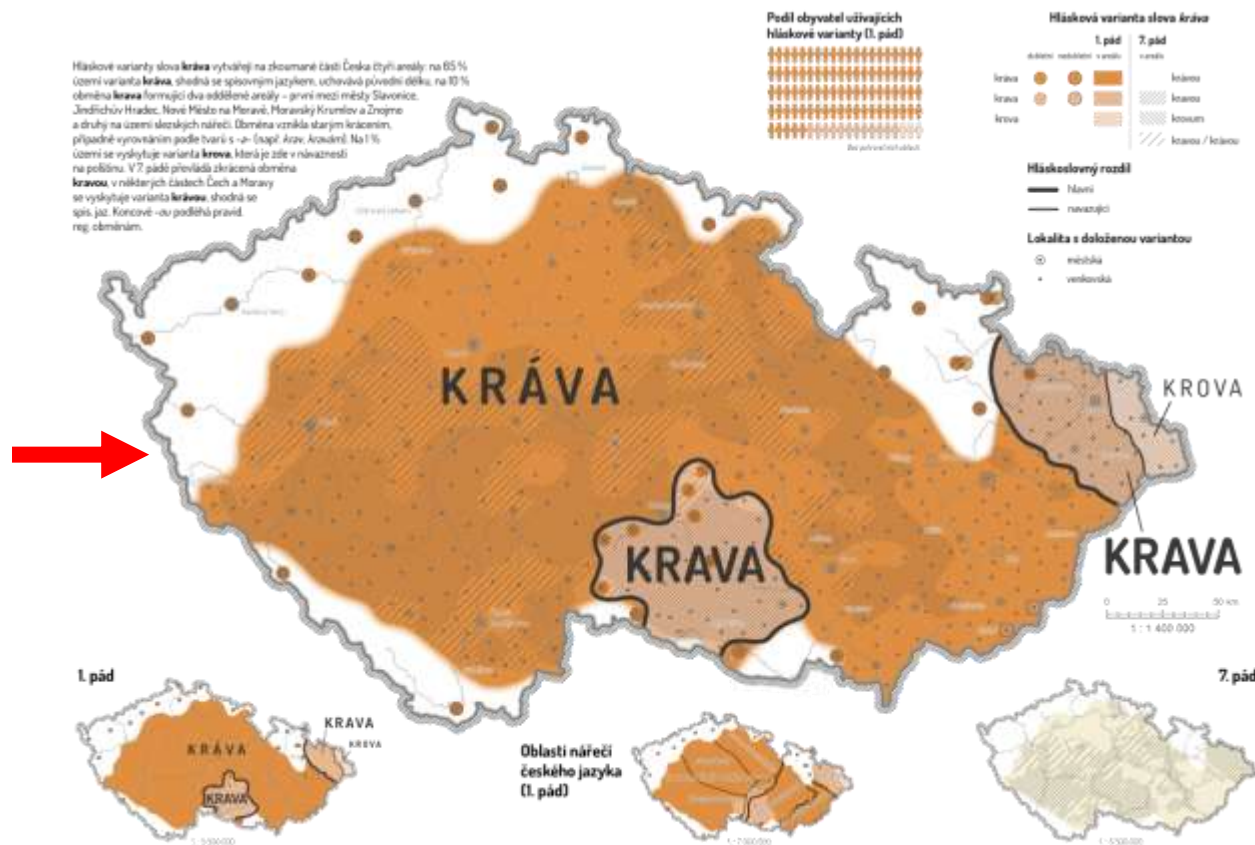


Atlas nářečí českého jazyka: krácení vokálů

upgrade 1 – digitalizace a kolorizace



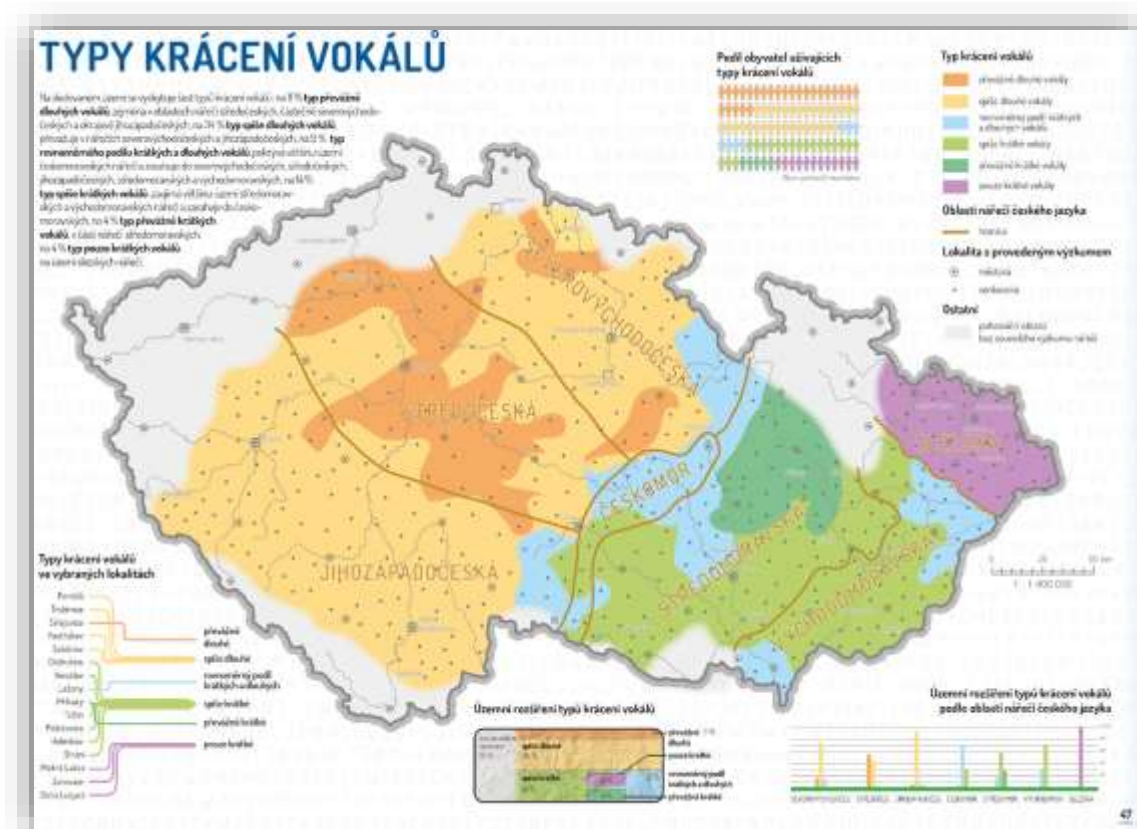
139
a) KRÁVA
b) KRÁVOU
kvantita



Reklasifikace map podle vztahu fonetických variant k množství vokálů

6 kroků:

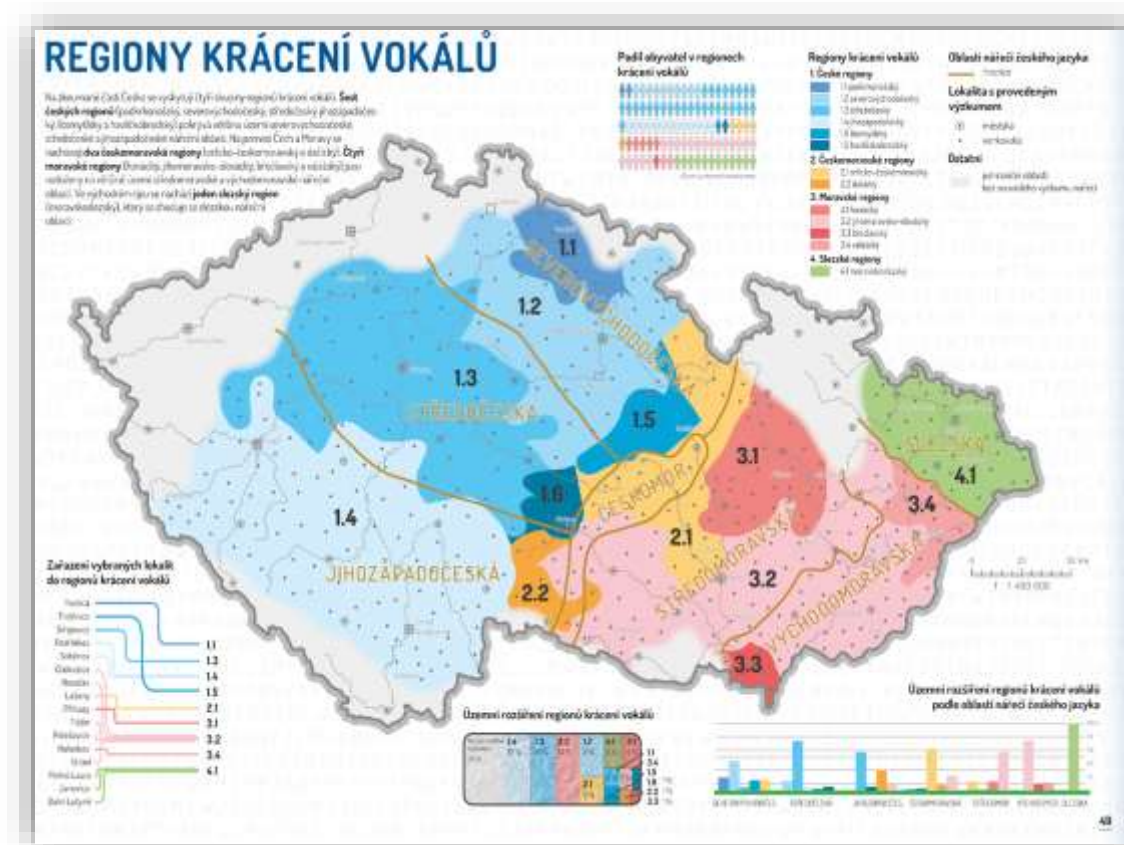
1. Výběr vhodných vstupních map
2. Reklasifikace do dvou tříd dle SHORT atributu
 - +1 dlouhá varianta
 - -1 krátká varianta
3. Nový dataset 541 polygonů
4. Generalizace malých polygonů ($<150 \text{ km}^2$)
5. Kvantifikace krátké a dlouhé varianty
 - od -12 do +10
6. Typizace do šesti intervalů



Atlas nářečí českého jazyka: krácení vokálů

regionalizace

- Kartografická syntéza 12 negeneralizovaných analytických map
- Vizuální analýza založená na nalezení center ve zkoumané oblasti a vizualizaci hranic
- Zohlednění typů, provázanosti a geografické charakteristiky oblasti
- 13 nových regionů
 - 6 českých regionů
 - 2 českomoravské regiony
 - 4 moravské regiony
 - 1 slezský region



Atlas nářečí českého jazyka: instrumentál plurálu

upgrade 2 – digitalizace, automatizace, kartografická správnost

ČESKÝ
JAZYKOVÝ
ATLAS
1 2 3 4 5
DODATKY



Atlas nářečí českého jazyka: instrumentál plurálu

upgrade 2 – digitalizace, automatizace, kartografická správnost

- 39 vybraných slov ze 4. dílu
původního Českého jazykového atlasu
- Vizualizace dubletních lokalit
- Data softwarově generována
- Vektorizace a vizualizace map
 - 39 analytických map
 - syntetické mapy – 4x typizace a 1x regionalizace
- Nový znakový klíč



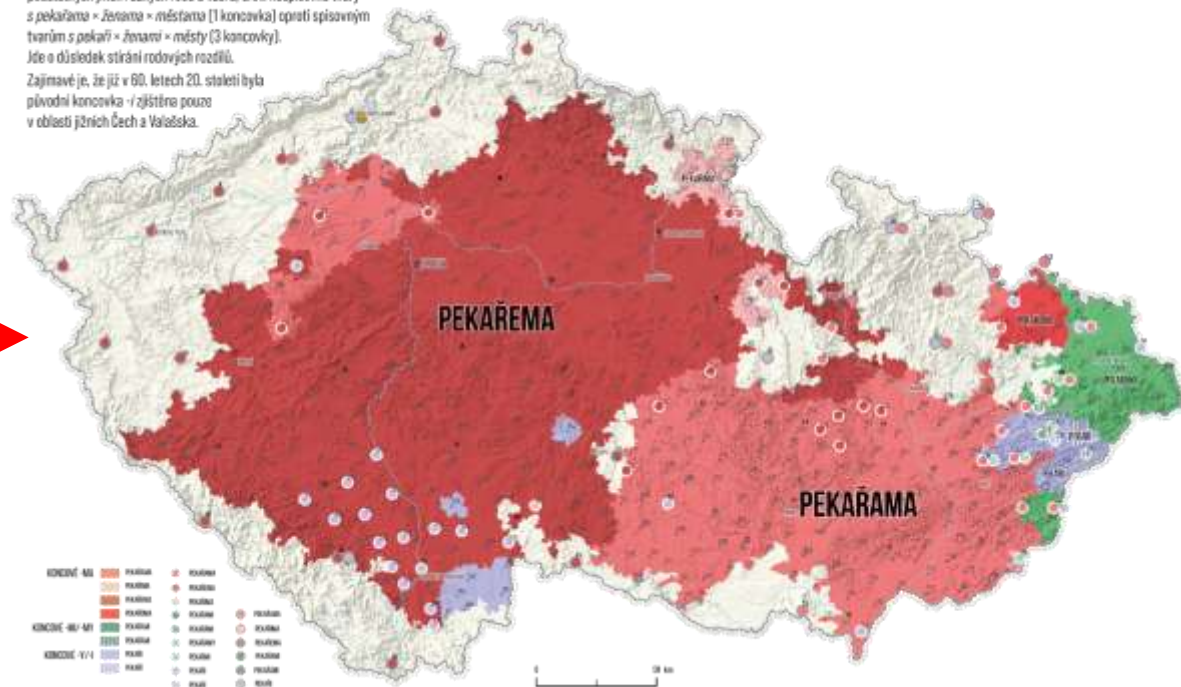
Atlas nářečí českého jazyka: instrumentál plurálu

upgrade 2 – digitalizace, automatizace, kartografická správnost



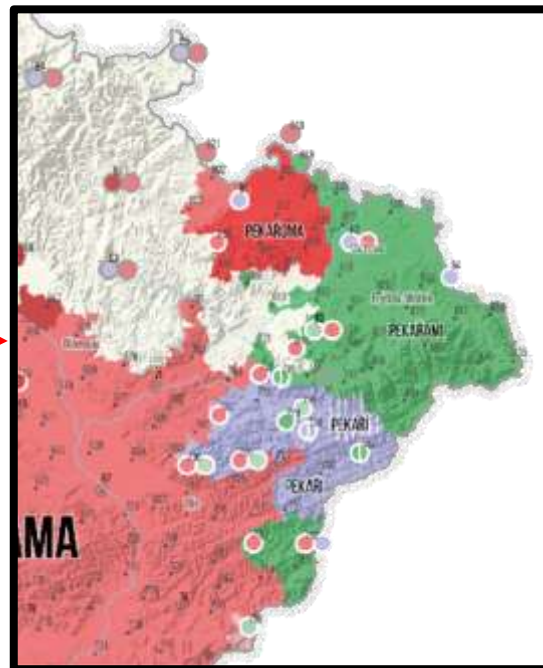
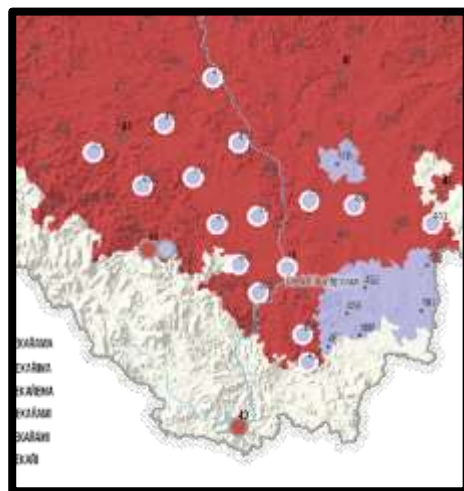
185
PEKAŘI
instr. pl. m.

Původní koncovka -i (s pekaři) je v mluvené češtině vytlačována zakončením -ma (s pekařama/pekařema atd.). Tvary s -ma se užívají i u dalších podstatných jmen různých rodů a vzorů, srov. nespisovné tvary s pekařama - ženama - městama (1 koncovka) oproti spisovným tvarům s pekaři - ženami - městy (3 koncovky).
Jde o důsledek střání rodových rozdílu.
Zajímavé je, že již v 80. letech 20. století byla původní koncovka -i zjištěna pouze v oblasti jižních Čech a Valašska.



Atlas nářečí českého jazyka: instrumentál plurálu

upgrade 2 – digitalizace, automatizace, kartografická správnost



Atlas nářečí českého jazyka: instrumentál plurálu

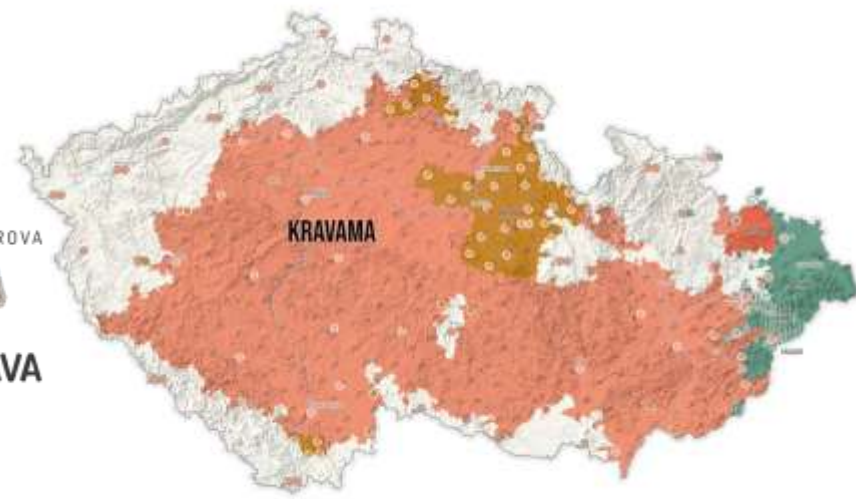
upgrade 2 – digitalizace, automatizace, kartografická správnost



2006



2020



2022

100

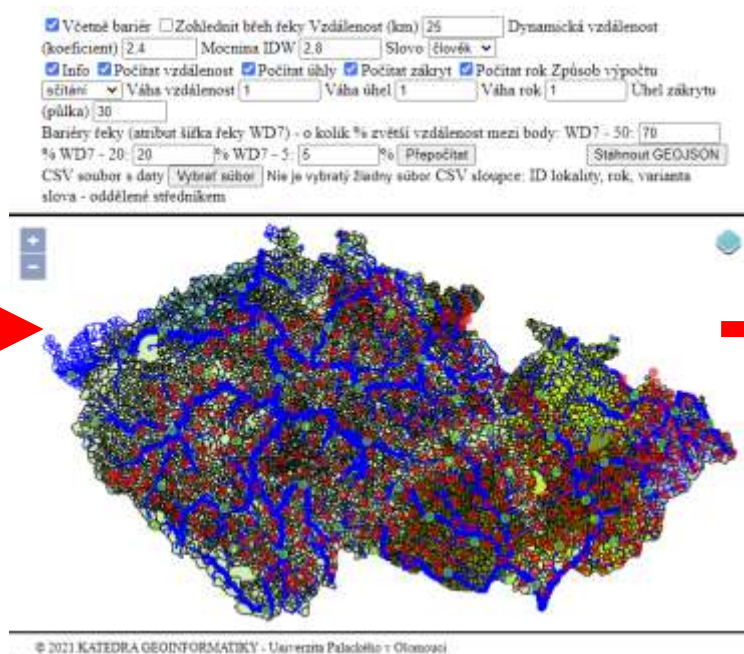


Atlas nářečí českého jazyka: instrumentál plurálu

upgrade 2 – digitalizace, automatizace, kartografická správnost

	A	B	C
1	1		PEKAŘEMA
2	2		PEKAŘEMA
3	3		PEKAŘEMA
4	4		PEKAŘEMA
5	5		PEKAŘEMA
6	5		PEKAŘAMA
7	6		PEKAŘI
8	6		PEKAŘEMA
9	7		PEKAŘEMA
10	8		PEKAŘEMA
11	9		PEKAŘEMA
12	10		PEKAŘEMA
13	11		PEKAŘEMA
14	12		PEKAŘEMA
15	13		PEKAŘEMA
16	14		PEKAŘEMA
17	21		PEKAŘEMA
18	22		PEKAŘEMA
19	23		PEKAŘEMA
20	24		PEKAŘEMA
21	24		PEKAŘI
22	25		PEKAŘEMA
23	26		PEKAŘEMA
24	27		PEKAŘEMA
25	30		PEKAŘEMA
26	31		PEKAŘEMA
27	32		PEKAŘEMA
28	40		PEKAŘEMA
29	41		PEKAŘEMA
30	42		PEKAŘEMA
31	43		PEKAŘEMA

Vstupní data



Generování
prostorových dat

```
{
  "type": "FeatureCollection",
  "features": [
    {
      "type": "Feature",
      "geometry": {
        "type": "MultiPolygon",
        "coordinates": [
          [
            [
              [
                752192.9149577866,
                5533689.438253784
              ],
              [
                752390.1293520274,
                5533600.854802031
              ],
              [
                751945.3909043972,
                5532292.527187389
              ],
              [
                751932.6420538511,
                5532367.0405153265
              ],
              [
                751912.6554693487,
                5532588.953856232
              ],
              [
                751833.1438334134,
                5532681.4481700
              ],
              [
                751791.2559138538,
                5532786.421251542
              ]
            ]
          ]
        ]
      }
    }
  ]
}
```

GeoJSON

	A	B	C	D	E	F
1	ID	DUBLETA	DUBLETA2	NEDUBLETA	NEDUBLETA2	
2	1			PEKAŘEMA		
3	2			PEKAŘEMA		
4	3			PEKAŘEMA		
5	4			PEKAŘEMA		
6	5			PEKAŘEMA	PEKAŘAMA	
7	6			PEKAŘI	PEKAŘEMA	
8	7			PEKAŘEMA		
9	8			PEKAŘEMA		
10	9			PEKAŘEMA		
11	10			PEKAŘEMA		
12	11					
13	12					
14	13					
15	14					
16	21					
17	22					
18	23					
19	24	PEKAŘI				
20	25					
21	26					
22	27					
23	30			PEKAŘEMA		
24	31					
25	32					
26	40			PEKAŘEMA		
27	41					
28	42					
29	43					
30	44			PEKAŘEMA	PEKAŘI	
31	45	PEKAŘI				
32	49			PEKAŘEMA		
33	51			PEKAŘEMA		

Dubletní data

ProPONAR



DIDA		Exemplifikace					?	📄	OOHLÁSIT
Heslo člověk	🔍	✎	+	+					
Exemplifikace	Lokalizace	Zdroj	Kvalifikátor	Význam	Čas				
☉ další si z nás dobrej deň (lidí, 1 pl.), diš šik kolem	Světla pod Ještědem LB	České nářeční texty		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:52:32				
☉ to bil ohromně honnej (člogjek, 1 sg.)	Držkov JN	Držkov JN		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:51:58				
☉ tam bili strašně honní (lidí, 1 pl.), strašně honní (lidí, 1 pl.), vopraudu	Držkov JN	Za života se stane ledacos. Vyprávěčky ze Železnobrodská.		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:52:01				
☉ dřij jich tam bilo na dje sfe padesát, a tejt maj tam pár (lidu, 2 pl.)	Držkov JN	Držkov JN		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:51:58				
☉ přiváželi sem šmelc a to rozváželi potom do Zásadi, do Loužnice, co bili ti faktoři, co dávali zas naučkar (lidem, 3 pl.)	Držkov JN	Držkov JN		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:51:58				
☉ to bil bezvannej film vo Držkovje, vo těch starejch (líděch, 6 pl.)	Držkov JN	Držkov JN		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:51:58				
☉ neska ti (lidí, 1 pl.) sou zalezši doma, nejvedejš nic	Loužnice JN	Za života se stane ledacos. Vyprávěčky ze Železnobrodská.		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:52:01				
☉ řem (lidem, 3 pl.) to znárodník, ti (lidí, 1 pl.) buřto vimfeli, nebo přešli na jinou práci	Loužnice JN	Za života se stane ledacos. Vyprávěčky ze Železnobrodská.		nejvyspělejší živa bytost, lidský jedinec	07.06.2021 21:51:58				

ProMAP



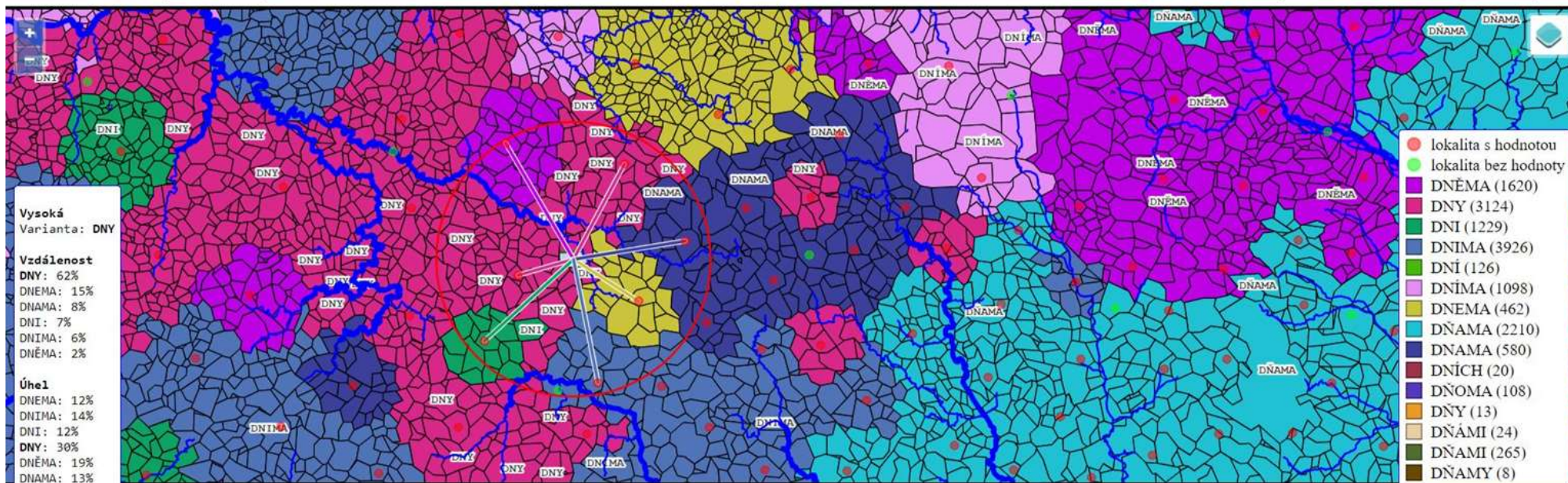
PROMAP | Nářečí českého jazyka interaktivně

☒ Včetně bariér ☒ Zohlednit břeh řeky Vzdálenost (km) Dynamická vzdálenost (koeficient) Mocnina IDW Slovo

☒ Info ☒ Počítat vzdálenost ☒ Počítat úhly ☒ Počítat zákryt ☒ Počítat rok Způsob výpočtu Váha vzdálenost Váha úhel Váha rok Úhel zákrytu (půlka)

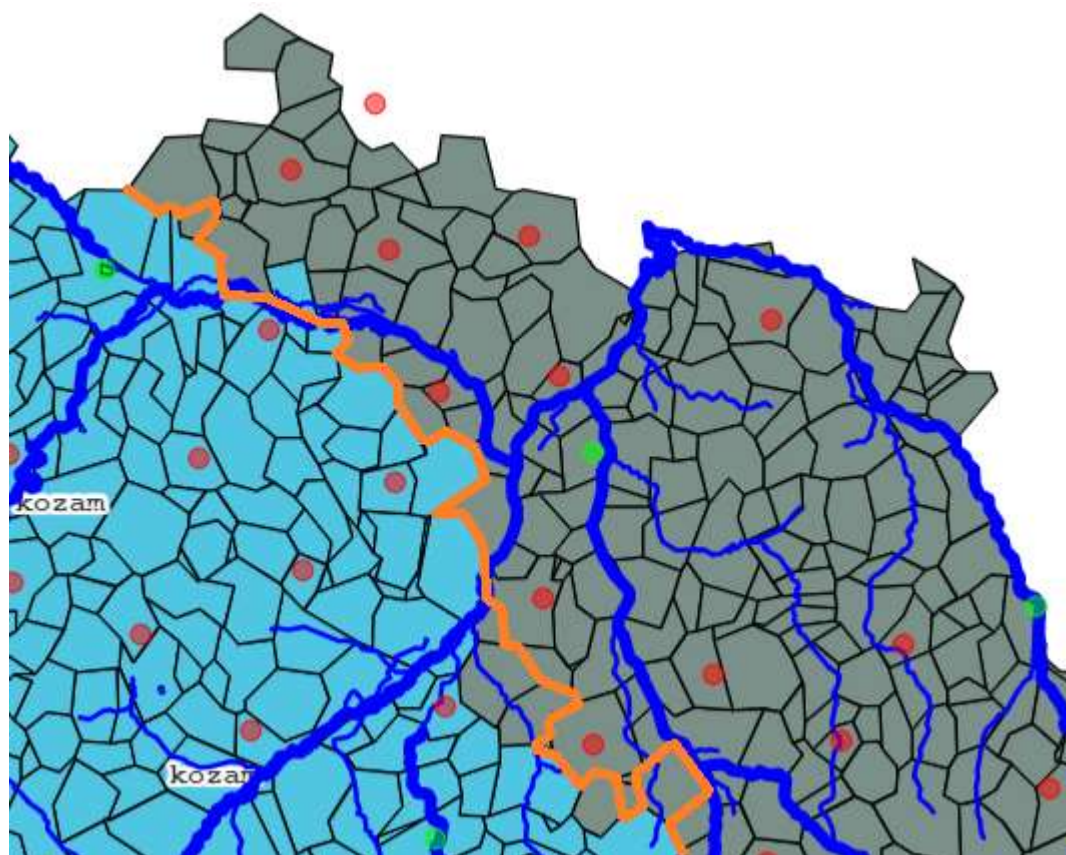
Bariéry řeky (atribut šířka řeky WD7) - o kolik % zvětší vzdálenost mezi body: WD7 - 50: % WD7 - 20: % WD7 - 5: %

CSV soubor s daty Soubor nevybrán CSV sloupce: ID lokality, rok, varianta slova - oddělené středníkem



RA GEOINFORMATIKY - Univerzita Palackého v Olomouci

ProMAP



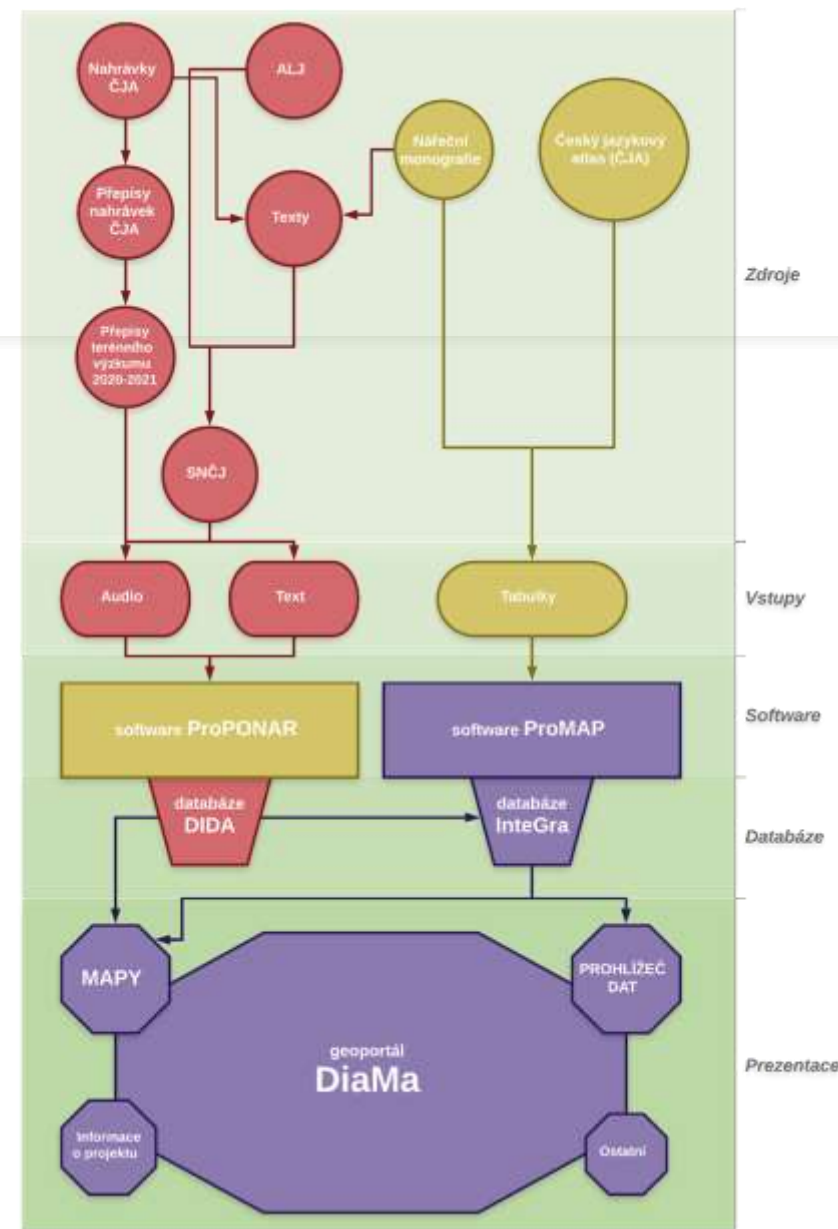
vypočtená izoglosa



reálná hranice

DiaMa

- více než **200** morfologických map
- více než **50** hláskoslovných map
- *volitelný topografický podklad*
- *vyhledávání*
- *několik úrovní přiblížení*
- *srovnání map: překryv, dlaždice*
- *prohlížeč dat: volitelný řez multidimenzionálními daty*



DĚKUJEME ZA POZORNOST

Vít Voženílek, Jakub Koníček

Univerzita Palackého v Olomouci, Katedra geoinformatiky

Martina Ireinová

Ústav pro jazyk český AV ČR, dialektologické oddělení

