



PLIN041

Transkripce mluvené češtiny

Úvodní informace

Transkripce mluvené češtiny je klíčovým krokem v lingvistickém výzkumu a v oblasti zpracování přirozeného jazyka. Budeme zkoumat, jak se mluvená čeština liší od psané formy a jak se s touto specifickou variantou jazyka pracuje.



by **Dana Hlavackova**

S využitím AI.

Struktura kurzu

- **Cílem** předmětu je naučit studenty pracovat s mluveným jazykem s využitím počítačových nástrojů.
- **Vedlejším cílem** je začátek budování nového moravského mluveného korpusu ve Sketch Engine.
- Kurz *PLIN041 Transkripce mluvené češtiny* je rozdělen do několika tematických celků, které pokrývají klíčové aspekty **pořízení nahrávek, transkripce a automatické zpracování** mluveného jazyka.
- Od seznámení se specifickými vlastnostmi mluvené češtiny až po praktické ukázky nástrojů pro automatické rozpoznávání řeči, kurz poskytuje komplexní přehled a praktické znalosti.





Hlavní úkoly v semináři

📄 K zápočtu - prezenze na semináři (možnost jedné neomluvené absence)

Pořízení nahrávek

- odevzdání nahrávek ke kontrole
- schválení
- případná úprava (oříznutí)

Ruční transkripce

- podle stanovených pravidel
- v textovém editoru 5 min záznamu
- odevzdání ke kontrole
- schválení transkripce

Testování nástrojů pro rozpoznání řeči

- vyhledání vhodného nástroje
- využití vlastních nahrávek
- následná editace přepisu
- odevzdání přepisu

Závěrečné zhodnocení

- představení nástroje
- plusy a zápory testovaného nástroje
- porovnání ruční transkripce s automatickou

Osnova

1. Úvodní informace (18. 2.)
2. Instrukce k pořízení nahrávek (4. 3.)
3. Specifika mluveného jazyka, specifika mluvy středomoravské oblasti (11. 3.)
4. Zpracování mluvené češtiny a budování mluvených korpusů (18. 3.)
5. Transkripce nahrávek, pravidla ortografické transkripce (25. 3.)
6. Workshop transkripce (1. 4.)
7. Automatické rozpoznávání češtiny (8. 4.)
8. Testování vhodných nástrojů (15. 4.)
9. Workshop editace automatických přepisů (29. 4.)
10. Závěrečné prezentace (6. 5.)
11. Závěrečné prezentace (13. 5.)
12. Rezerva (20. 5.)



Instrukce

Co budete potřebovat

- **notebook**
- **aplikaci** na pořízení nahrávky v mobilu (běžný formát)
- (případně diktafon)
- kvalitní **sluchátka**

Důležité termíny

- do 11. 3. odevzdání nahrávek
- do 1. 4. odevzdání ručních transkripcí
- do 29. 4. odevzdání editovaných automatických transkripcí
- do konce semestru odevzdat a odprezentovat závěrečné zhodnocení



Dokumenty a soubory

Poskytne student do ISu

- nahrávky (celkem 15-20 min)
- prohlášení nahrávajícího se svým podpisem
- informace o mluvčích
- ruční přepis v textovém editoru
- automatický a editovaný přepis ve formátu .txt
- závěrečná zpráva

Poskytne vyučující do ISu

- pravidla transkripce (pro ruční přepis i editaci)
- evidence údajů o mluvčích
- nabídku automatických nástrojů a seznam sledovaných rysů