

Лекція від
проф. Капранова Яна Васильовича

У межах цієї дисципліни формується усвідомлення аспірантами необхідності дотримання норм публікаційної етики, прийнятої в європейському академічному середовищі, передовсім академічної доброчесності, та розширення загальнокультурного світогляду.

Co-funded by the
Erasmus+ Programme
of the European Union

«Мультилінгвальний корпус та його ресурси для дослідження європеїстики»

Лекційний курс

Дисципліна
“Мультилінгвальний корпус та його ресурси для дослідження європеїстики”
належить до переліку дисциплін професійної підготовки аспіранта.

Відкриття проєкту
«Multilingual Corpus and its Resources for European Studies Research»
за Програмою Erasmus+ напрямку Jean Monnet Activities

УЧАСНИКИ ПРОЄКТУ
Multilingual Corpus and its Resources for European Studies Research (MCRESR)

КООРДИНАТОР проєкту:
Алла КОРОЛЬОВА,
доктор філологічних наук,
професор, проректор з наукової роботи ІМЛІУ
(м. Київ, УКРАЇНА)

КООРДИНАТОР проєкту:
Ян КАПРАНОВ,
доктор філологічних наук,
доцент, професор ІМЛІУ
(м. Київ, УКРАЇНА)

УЧАСНИК проєкту:
Гетяна АНОХІНА,
викладачка проєкту
MCRESR ІМЛІУ
(м. Київ, УКРАЇНА)

УЧАСНИК проєкту:
Ольга АНДРУШЕНКО,
викладачка проєкту
MCRESR ІМЛІУ
(м. Київ, УКРАЇНА)

Робоча програма модуля
“Мультилінгвальний корпус та його ресурси для дослідження європеїстики”

Під'єднайтеся до події +
Microsoft Teams

Corpus	Download	# words	District	Time period	Genre(s)
News on the Web (NOW)	18.6 million	20 countries	2010-yesterday	Web	News
Web: The Intelligent Web-based Corpus	14 million	6 countries	2017	Web	News
Global Web-based English (GloWbE)	1.9 billion	20 countries	2012-13	Web (incl. blog)	
Internet2 Corpus	1.9 billion	Internet2	2014	Webpage	
Commonsense Corpus	1.5 billion	20 countries	Jan 2020-Dec 2022	Web	News
Corpus of Contemporary American English (COCA)	1.0 billion	American	1990-2019	Balanced	
Corpus of Historical American English (COHA)	475 million	American	1820-2019	Balanced	
The TV Corpus	325 million	6 countries	1950-2018	TV shows	
The Movie Corpus	200 million	6 countries	1930-2018	Movies	
Corpus of American Soap Operas	100 million	American	2001-2012	TV shows	
Parliament Corpus	1.6 billion	British	1803-2005	Parliament	
Early English Books Online	755 million	British	1470s-1600s	(Out-of-scope)	
Corpus of US Supreme Court Opinions	130 million	American	1780s-present	Legal opinions	
TIME Magazine Corpus	100 million	American	1923-2006	Magazine	
British National Corpus (BNC)*	100 million	British	1980s-1995	Balanced	
Stanley Corpus (Canada)	50 million	Canadian	1950-2000s	Balanced	
COE Corpus	50 million	6 countries	2014	Web	
From Google Books n-grams (compare)					
American English	155 billion	American	1500s-2000s	(Out-of-scope)	
British English	34 billion	British	1500s-2000	(Out-of-scope)	

ПЛАН

- Європейські студії чи європеїстика?
- Корпусна лінгвістика і корпус текстів
- Типи корпусів
- Корпусний підхід до вивчення європеїстики
- Традиційні корпусні ресурси англійської та української мов
- Мультилінгвальний корпус
- Створення власного корпусу текстів

MULTILINGUAL MULTI-TIERED CORPUS OF TEXTS ON EUROPEAN STUDIES

During the first year the following corpora have been compiled:

Subcorpus of the modern European media(English)	https://app.sketchengine.eu/#share?corpname=user%2Fankhina%2Fsubcorpus_of_modern_european_media
European Sociology Subcorpus(English)	https://app.sketchengine.eu/#share?corpname=user%2Fankhina%2FEuropean_history_4
European Union Literature Subcorpus(English)	https://app.sketchengine.eu/#share?corpname=user%2Fankhina%2FEuropean_history_2

```

graph TD
    1[1. Select and Analyze a Suitable Corpus] --> 2[2. Research Available Annotation Schemes]
    2 --> 3[3. Define a Custom Annotation Scheme]
    3 --> 4[4. Annotate Sample Text]
    4 --> 5[5. Establish Knowledge Representation Model OWL + T-BOX Axioms]
    5 --> 6[6. Define IE Extraction Scheme]
    6 --> 7[7. Extract Information, Concepts and Relations]
    7 --> 8[8. Automatically Populate OWL Ontology with A-BOX Axioms]
    8 --> 9[9. Observe Inferred Relations by Description Logic]
    9 --> 4
  
```

«Мультилінгвальний корпус та його ресурси для дослідження європеїстики»

Co-funded by the Erasmus+ Programme of the European Union



Концептуальні аспекти МОДУЛЯ

ПЕРШЕ

Стандарти освітньо-наукової діяльності європейських ЗВО

набуття компетентностей для інтеграції:
 (а) в систему освіти країн ЄС;
 (б) в інші сфери життя, які становлять об'єкт вивчення дисципліни

Європеїстика

ДРУГЕ

Набуття навичок роботи з корпусними технологіями:

(а) дослідницько-пошукова робота;
 (б) формування умінь аспірантів укладати тематичні ресурси, якими є тексти з Європеїстики

ТРЕТЄ

Мультилінгвальний корпус з ЄВРОПЕЇСТИКИ



Тренінг

29 березня 2024 року

“Ресурси та інструменти Британського національного корпусу для побудови мережевої структури концепту ЄВРОПА”

ТРЕНІНГ
 Ресурси та інструменти Британського національного корпусу для побудови мережевої структури концепту ЄВРОПА / Resources and instruments of the British National Corpus for building the network structure of the EUROPE concept

Алла КОРОЛЬОВА,
д.філол.н., проф., проректор з наукової роботи і міжнародного співробітництва КНЛУ, координатор Проекту

Ольга ВАХОВСЬКА,
к.філол.н., доц., докторантка кафедри германської і фіно-угорської філології КНЛУ, запрошений тренер

Тренінг

29 квітня 2024 року

“Практики застосування мультифункційної платформи корпусних сервісів AntLab”

ТРЕНІНГ
 Практики застосування мультифункційної платформи корпусних сервісів AntLab

Координатор Проекту
Алла КОРОЛЬОВА,
д.філол.н., проф., проректор з наукової роботи і міжнародного співробітництва Київського національного лінгвістичного університету

Запрошений тренер
Вікторія ЖУКОВСЬКА,
к.філол.н., проф., професор кафедри міжкультурної комунікації та прикладної лінгвістики Житомирського державного університету імені Івана Франка

Тренінг

08 травня 2024 року

Індивідуальні практики застосування ресурсів Мультилінгвального корпусу у дисертаціях аспірантів

