

ДЕРЖАВНИЙ ВИЩИЙ НАВЧАЛЬНИЙ ЗАКЛАД  
«УЖГОРОДСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ»  
ФАКУЛЬТЕТ МІЖНАРОДНИХ ЕКОНОМІЧНИХ ВІДНОСИН  
КАФЕДРА ПРИКЛАДНОЇ ЛІНГВІСТИКИ

**«ЗАТВЕРДЖУЮ»**  
Декан факультету міжнародних  
економічних відносин  
Микола ПАЛІНЧАК  
«27» червня 2024 р.



**РОБОЧА ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

**КОМП'ЮТЕРНА ЛІНГВІСТИКА**

Рівень вищої освіти	Другий (магістерський)
Галузь знань	03 Гуманітарні науки
Спеціальність	035 Філологія
Освітня програма	«Прикладна лінгвістика»
Статус дисципліни	обов'язкова
Мова навчання	українська

Ужгород 2024

Робоча програма навчальної дисципліни «Комп'ютерна лінгвістика» для здобувачів вищої освіти галузі знань 03 Гуманітарні науки спеціальності 035 Філологія освітньої програми «Прикладна лінгвістика».

**Розробник:** Кунанець Н.Е., професор, д.н.с.к.

Робочу програму розглянуто та затверджено на засіданні кафедри прикладної лінгвістики

протокол №5 від «27» 06 2024 р.

Завідувач кафедри Лав М. Полюшин

Схвалено науково-методичною комісією факультету міжнародних відносин

протокол №10 від «27» 06 2024р.

Голови науково-методичної комісії



© Кунанець Н.Е., 2024 р.

© ДВНЗ «Ужгородський національний університет», 2024 р.

## 1. ОПИС НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Найменування показників	Розподіл годин за навчальним планом	
	Денна форма навчання	Заочна форма навчання
Кількість кредитів ЄКТС – 6	Рік підготовки:	
Загальна кількість годин – 180	1 - й	1 - й
Кількість модулів – 4	Семестр:	
Тижневих годин для денної форми навчання: аудиторних – 2 самостійної роботи студента – 4	1 – й	1 - й
	Лекції:	
	16	10
	Практичні (семінарські):	
	20	10
Вид підсумкового контролю: іспит	Лабораторні:	
	-	-
Форма підсумкового контролю: усна	Самостійна робота:	
	108	160
	Індивідуальна робота:	
	-	-

## 2. МЕТА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Програма з курсу «Комп'ютерна лінгвістика» призначається для студентів 1 курсу магістерського рівня спеціальності 035 «Філологія» галузі знань 03 «Гуманітарні науки».

Програма курсу передбачає ознайомлення з основними тенденціями у сучасній комп'ютерній лінгвістиці, основними типами сучасних комп'ютерних систем, встановити їхні можливості і з'ясувати методи створення їх лінгвістичного забезпечення на сучасному етапі. Курс викладається для студентів денної і заочної форм навчання.

**Мета вивчення дисципліни.** Мета курсу – ознайомити студентів із основами опрацювання текстів природною мовою з використанням сучасних інформаційних технологій, методів штучного інтелекту та технологій Великих даних, підходами до автоматичної генерації текстових даних, способами аналізу особливості репрезентації мовних одиниць у лінгвістичних базах даних; а також засвоїти практичні навички роботи в середовищах розробки програмних продуктів для опрацювання текстів природної мови, проведення досліджень текстів різних предметних областей.

Основні завдання курсу:

1. Надання базових знань з комп'ютерної лінгвістики та опрацювання текстів природною мовою.

2. Ознайомлення з методами та інструментами опрацювання текстових даних за допомогою технологій Великих даних та систем ШІ.

3. Формування навичок роботи у середовищах розробки програмного забезпечення для NLP.

4. Вивчення сучасних методів машинного навчання для опрацювання текстової інформації, включаючи нейронні мережі та глибинне навчання.

5. Опанування методів оцінки якості опрацювання природної мови та постредагування, згенерованих автоматично текстів, технологій формування електронних корпусів текстів.

У результаті вивчення дисципліни “Комп'ютерна лінгвістика” студент повинен **знати:**

1. основні ідеї провідних українських і зарубіжних вчених у галузі комп'ютерної лінгвістики;

2. методи оцінювання якості опрацювання текстів природною мовою;

3. особливості формування баз даних і баз знань, сховищ даних;

4. основні методи машинного навчання та особливості функціонування систем штучного інтелекту;

5. можливості інформаційних систем та інформаційних технологій для проведення лінгвістичних досліджень.

Студенти повинні **вміти:**

1. опрацьовувати текстову інформацію з використанням методів машинного навчання;

2. використовувати технології Великих даних та системи ШІ для опрацювання текстів природною мовою;

3. здійснювати оцінку якості опрацювання текстів природною мовою;

4. формувати електронні корпуси текстів;
5. здійснювати постредагування автоматично згенерованих текстів;
6. проводити лінгвістичні дослідження з використанням сучасних інформаційних технологій та інформаційних систем лінгвістичного спрямування.

**Міждисциплінарні зв'язки** – наука комп'ютерна лінгвістика має тісні взаємозв'язки з загальною лінгвістикою, інформаційними технологіями, логікою, семіотикою, соціолінгвістикою, психолінгвістикою, лінгводидактикою, теорією та практикою перекладу.

Відповідно до освітньої програми, вивчення дисципліни сприяє формуванню у здобувачів вищої освіти таких **компетентностей**:

**Загальні компетентності (ЗК)**

- Здатність спілкуватися іноземною мовою (ЗК 06).
- Здатність спілкуватися з представниками інших професійних груп різного рівня (з експертами з інших галузей знань/видів економічної діяльності) (ЗК 10).

**Фахові компетентності (ФК)**

- Здатність вільно користуватися спеціальною термінологією в обраній галузі філологічних досліджень (ФК 07).
- Усвідомлення ролі експресивних, емоційних, логічних засобів мови для досягнення запланованого прагматичного результату (ФК 08).
- Здатність редагувати, анотувати, реферувати різножанрові письмові та усні тексти мовами оригіналу (ФК 09).
- Здатність застосовувати методи та прийоми наукового спілкування іноземною мовою (здійснення комунікативно навчальної функції) (ФК 17).

### 3. ПЕРЕДУМОВИ ДЛЯ ВИВЧЕННЯ НАВЧАЛЬНОЇ ДИСЦИПЛІНИ

Передумовою вивчення навчальної дисципліни «Комп'ютерна лінгвістика» є попереднє опанування навчальної дисципліни «Комп'ютерна лексикографія».

### 4. ОЧІКУВАНІ РЕЗУЛЬТАТИ НАВЧАННЯ

Відповідно до освітньої програми «Прикладна лінгвістика», вивчення навчальної дисципліни повинно забезпечити досягнення здобувачами вищої освіти таких програмних результатів навчання (ПРН):

Програмні результати навчання	Шифр ПРН
Знаходити оптимальні шляхи ефективної взаємодії у професійному колективі та з представниками інших професійних груп різного рівня.	ПРН-05
Застосовувати знання про експресивні, емоційні, логічні засоби мови та техніку мовлення для досягнення запланованого прагматичного результату й організації успішної комунікації.	ПРН-06

Очікувані результати навчання, які повинні бути досягнуті здобувачами освіти після опанування навчальної дисципліни «Прикладна лінгвістика»:

<b>Очікувані програмні результати навчання</b>	<b>Шифр ПРН</b>
Вільно послуговуватись іноземною мовою у фаховому середовищі та використовувати набуті навички у професійній діяльності	ПРН 05
Володіти базовими знаннями методології ведення наукової розвідки та культури наукової комунікації.	ПРН 06

## **5. ЗАСОБИ ДІАГНОСТИКИ ТА КРИТЕРІЇ ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ**

### **Засоби оцінювання та методи демонстрування результатів навчання**

Засобами оцінювання та методами демонстрування отриманих знань та компетентностей з навчальної дисципліни є:

1. виступи на практичних заняттях;
2. реферати;
3. презентації;
4. модульні контрольні роботи;
5. екзамен.

### **Форми контролю та критерії оцінювання результатів навчання**

Форми поточного контролю: усне опитування, оцінювання виконання практичних завдань, написання самостійних робіт.

Форма модульного контролю: письмова контрольна робота.

Форма підсумкового семестрового контролю: екзамен.

### **Особливості використання засобів діагностики та контролю за умов дистанційного навчання**

В умовах використання формату онлайн-навчання (дистанційного навчання) із застосуванням корпоративної мережі Google Meet названі засоби, методи і форми визначаються за домовленістю зі студентським колективом і, в залежності від зручного виду взаємодії, застосовуються з допомогою існуючих функцій групових чатів та відео-конференцій.

Для ефективного засвоєння тематики є можливість демонстрації необхідних матеріалів на робочому столі комп'ютерного технічного засобу під час занять.

Зокрема, у разі потреби, під час онлайн-заняття можна надати доступ до свого екрану, щоб показати презентації або іншу тематичну інформацію на робочому столі.

Планування лекційних і практичних (семінарських) занять, модульних контрольних робіт, а також підсумкова перевірка знань у формі екзамену (заліку) здійснюється заздалегідь за допомогою прив'язки до гугл-календаря.

Синхронізація запланованих заходів виконується автоматично на всіх зручних для їх проведення пристроях.

**Розподіл балів, які отримують здобувачі вищої освіти (модуль 1)**

Поточне оцінювання та самостійна робота				Модульна контрольна робота	Сума
T1	T2	T3	T4	70	100
6	6	6	6		

T1,T2... – теми

**Розподіл балів, які отримують здобувачі вищої освіти (модуль 2)**

Поточне оцінювання та самостійна робота				Модульна контрольна робота	Сума
T5	T6	T7	T8	70	100
6	6	6	6		

T5,T6... – теми

**Розподіл балів, які отримують здобувачі вищої освіти (модуль 3)**

Поточне оцінювання та самостійна робота				Модульна контрольна робота	Сума
T9	T10	T11	T12	70	100
6	6	6	6		

T9,T10... – теми

**Розподіл балів, які отримують здобувачі вищої освіти (модуль 4)**

Поточне оцінювання та самостійна робота				Модульна контрольна робота	Сума
T13	T14	T15	T16	70	100
6	6	6	6		

T13,T14... – теми

**Оцінювання окремих видів навчальної роботи з дисципліни**

Вид діяльності здобувача вищої освіти	Модуль 1		Модуль 2	
	Кількість	Максимальна кількість балів (сумарна)	Кількість	Максимальна кількість балів (сумарна)
Практичні заняття	5	20	5	20
Презентація	1	5	1	5
Реферат	1	5	1	5

Модульна контрольна робота	1	70	1	70
<b>Разом</b>		<b>100</b>		<b>100</b>

### **Критерії оцінювання модульної контрольної роботи**

МК1 та МК2 складається з випадкових 7 описових питань теоретичного курсу. Максимальна кількість балів за кожне питання – 10 балів. Максимальна оцінка за модульний контроль – 100 балів. Якщо студент не був присутнім на модульному контролі, або бажає перездати - він має право його здати згідно розроблених процедур в Положенні про організацію освітнього процесу в ДВНЗ «Ужгородський національний університет».

### **Критерії оцінювання підсумкового семестрового контролю**

До складання екзамену допускаються здобувачі вищої освіти, які мають підсумковий доекзаменаційний рейтинговий бал не менше 35.

Здобувач вищої освіти, доекзаменаційний рейтинговий бал якого складає від 0 до 34 балів, зобов'язаний покращити його до початку екзамену під час чергування викладачів на кафедрі у строки, визначені викладачем дисципліни та погоджені деканатом факультету. В протилежному випадку, здобувач не допускається до екзамену, і у нього виникає академічна заборгованість.

Екзамен з навчальної дисципліни здобувач вищої освіти може не скласти, якщо він успішно пройшов усі модульні контролі та його влаштовує підсумкова доекзаменаційна рейтингова оцінка за навчальний рік. Здобувачі вищої освіти, рейтинговий бал яких становить від 35 до 59, екзамен складають обов'язково.

Здобувач освіти може підвищити на екзамені рейтинговий бал, при цьому, за результатами складання екзамену оцінка не може бути менша за доекзаменаційний рейтинговий бал.

Екзамен проводиться в усній формі. На екзамен вноситься навчальний матеріал семестру. Екзаменаційний білет складається з теоретичних питань та практичних завдань. Оцінювання результатів навчання на екзамені здійснюється за 100-бальною шкалою. Оцінка за екзамен вноситься у відомість обліку успішності.

**Переведення даних 100-бальної шкали оцінювання у оцінки за національною шкалою та шкалою ЄКТС**

Сума балів	Оцінка ЄКТС	оцінка за національною шкалою	
		екзамен, диф. залік	з а л і к
90 -100	A	відмінно	зараховано
82 - 89	B	добре	
74 - 81	C		
64 - 73	D	задовільно	
60 - 63	E		
35 - 59	FX	незадовільно	не зараховано
0 - 34	F		

**Оцінка відмінно (A)** виставляється, коли студент дає абсолютно правильні відповіді на теоретичні питання з викладенням оригінальних висновків, отриманих на основі програмного, додаткового матеріалу та нормативних документів. При виконанні практичного завдання студент застосовує системні знання навчального матеріалу, передбачені навчальною програмою.

**Оцінка добре (B)** виставляється студенту, який повністю розкрив теоретичні питання на основі програмного та додаткового матеріалу. При виконанні практичних завдань студент застосовує узагальнені знання навчального матеріалу, передбачені навчальною програмою.

**Оцінка добре (C)** виставляється студенту, який повністю розкрив теоретичні питання, а програмний матеріал викладено у відповідності до вимог. Практичні завдання виконані в цілому правильно, але мають місце окремі неточності.

**Оцінка задовільно (D)** виставляється, коли студент розкрив теоретичні питання, проте при викладенні програмного матеріалу допущені окремі помилки. При виконанні практичних завдань студент припускається помилок, за рахунок недостатнього розуміння програмного матеріалу.

**Оцінка задовільно (E)** виставляється, коли студент неповністю розкрив теоретичні питання, відповідь містить суттєві помилки. При виконанні практичних завдань студент припускається значних помилок, а виконання завдань викликає значні труднощі у студента.

**Оцінка незадовільно (FX)** виставляється студенту, який не розкрив теоретичні питання і не може виконати практичні завдання. Як правило такий студент виявляє здатність до викладення думки лише на елементарному рівні.

**Оцінка незадовільно (F)** виставляється студенту, який не виконав навчальну програму або якийсь серйозний елемент її складової, має фрагментарні знання, які не дозволяють розкрити теоретичні питання і виконати практичні

завдання. Такий студент не може викласти свою думку навіть на елементарному рівні. За результатами контролю знань студентів, дозволяється виставлення екзаменаційної оцінки (без підсумкового іспиту) – «відмінно», «добре», та «задовільно». Студент має право підвищити оцінку, складаючи іспит.

## **6. ПРОГРАМА НАВЧАЛЬНОЇ ДИСЦИПЛІНИ**

### **6.1. Зміст навчальної дисципліни**

#### **Модуль 1**

##### **Тема 1. Комп'ютерна лінгвістика: термінологічні засади, історія та напрями розвитку.**

Поняття «комп'ютерна лінгвістика», опрацювання текстів природною мовою (NLP). Напрями та принципи комп'ютерної лінгвістики. Етапи розвитку комп'ютерної лінгвістики. Структура та рівні опрацювання мовної інформації: морфологічний, синтаксичний, семантичний.

##### **Тема 2. Лінгвістична модель та моделі подання знань**

Характеристика моделей подання знань. Логічні моделі. Продукційні моделі. Фреймові моделі. Семантичні мережі. Онтологічні системи.

##### **Тема 3. Лінгвістичний алгоритм та лінгвістичний процесор**

Поняття лінгвістичного алгоритму. Види лінгвістичних алгоритмів. Лінгвістичні процесори в комп'ютерній лінгвістиці. Реалізація лінгвістичних алгоритмів та процесорів у сучасних інформаційних системах.

##### **Тема 4. Технології та інструменти NLP**

Використання Python-бібліотек для NLP (NLTK, SpaCy, Gensim). Інструменти для аналізу текстів. Введення до регулярних виразів та їх застосування в опрацюванні текстової інформації.

#### **Модуль 2**

##### **Тема 5. Автоматизований морфологічний аналіз тексту (АМА)**

Чинники результативності аналізу. Принципів «машинної» граматики. Експериментальні та промислові системи АМА. Використанням словника основ.

Графемний аналіз словоформ. Контекстний аналіз. Інструменти автоматичного морфологічний аналіз тексту.

### **Тема 6. Автоматизований синтаксичний аналіз тексту (АСА)**

Знання про синтаксичну структуру речення. Алгоритми аналізу речення. підходи щодо дослідження цього процесу. Аналіз за безпосередніми складникам. Аналіз за допомогою дерев залежностей (деревами синтаксичного підпорядкування). Інструменти автоматичного синтаксичний аналіз тексту. Парсинг дерев та побудова граматичних конструкцій. Алгоритми парсингу: CFG, PCFG, Dependency Parsing.

### **Тема 7. Автоматизований логіко-семантичний аналіз тексту**

Мова як система (мовних) знаків, що виконує функції означення та позначення. Дескриптивні (описові) й логічні категорії. Інструменти автоматичного логіко-семантичний аналіз тексту. Обробка значень та семантичних ролей у текстах. Виведення семантики через моделі знань: WordNet, FrameNet. Використання методів Word Embeddings: Word2Vec, GloVe, FastText.

### **Тема 8. Автоматизований сентиментальний аналіз тексту**

Поняття сентиментальний аналіз тексту. Чинники, що впливають на семантичний аналіз тексту. Способи покращення семантичного аналізу тексту. Інструменти автоматичного аналізу тональності текстів.

## **Модуль 3**

### **Тема 9. Електронні корпуси текстів розмітка текстів та формування електронних корпусів текстів. Бази даних, бази знань, сховища даних як інформаційні моделі.**

Поняття корпусної розмітки. Типи корпусної розмітки. Етапи створення корпусів текстів. Інструменти створення корпусів та роботи з ними. Корпусні менеджери (конкондансери). Використання корпусів текстів у лінгвістичних дослідженнях. Етапи проектування баз даних та знань. Способи організації лінгвістичних баз даних. Комп'ютерна копія та комп'ютерна версія документа. Правило modus ponens. Лінгвістичний аспект побудови інтерфейсів користувача.

### **Тема 10. Лінгвістичні проблеми створення баз знань**

Концепт «знання». Декларативні і процедурні знання. Представлення знань у системах штучного інтелекту. Організація баз знань та побудова експертного оцінювання в рекомендаційних системах.

### **Тема 11. Моделювання мовленнєвої діяльності в комп'ютерних діалогових системах**

Інтерфейси, їх різновиди: штучно- і природномовні інтерфейси. Діалогові системи. Форми діалогу з комп'ютером. Графічні та звукові інтерфейси.

### **Тема 12. Автоматизоване реферування та анотування текстів**

Методи автоматизованого реферування. Характеристика методів автоматизованого реферування. Методика формалізованого реферування. Система автоматичного реферування. Система квазіреферування. Автоматизоване анотування текстів. Постредагування автоматично згенерованих текстів. Алгоритми та методи покращення текстів. Використання AI-технологій для постредагування та корекції мовних помилок у автоматично згенерованих текстах.

## **Модуль 4**

### **Тема 13. Середовища розробки програмних продуктів для NLP**

Інтегровані середовища розробки (IDE) (PyCharm, Jupyter Notebook, VSCode). Використання бібліотек для побудови веб-додатків та API для NLP (Flask, FastAPI). Автоматизація процесів опрацювання. Платформи для тестування та налаштування NLP-моделей (Hugging Face, AllenNLP). Використання інструментів CI/CD та хмарних сервісів для автоматизації процесів у NLP. Комп'ютерно- лінгвістичні системи.

### **Тема 14. Великі дані в лінгвістиці**

Принципи опрацювання текстових корпусів великого розміру. Використання інструментів для роботи з Великими даними: Hadoop, Apache Spark. Програмні засоби та інструменти для збору, очищення та аналізу лінгвістичних даних.

## **Тема 15. Застосування систем ШІ та Великих даних у лінгвістиці**

Побудова мовних моделей на базі великих корпусів текстів. Вивчення частотних словників та автоматичне видобування знань з текстових даних. Методи аналізу значень та семантики за допомогою статистичних та глибинних підходів.

## **Тема 16. Методи машинного навчання для NLP**

Підходи до опрацювання мовних даних: класифікація, кластеризація, розпізнавання образів. Використання алгоритмів машинного навчання для аналізу текстів: логістична регресія, дерево рішень, SVM. Використання бібліотек машинного навчання: Scikit-learn, TensorFlow, Keras. Основи нейронних мереж: архітектури MLP, CNN, RNN. Моделі глибинного навчання: BERT, GPT, Transformers. Використання попередньо навчених моделей для різних NLP-завдань (класифікація тексту, генерація тексту, переклад). Класифікація методів машинного навчання для опрацювання україномовних текстів. Схема процесу лінгвістичного аналізу тексту природньою мовою засобами комп'ютерних лінгвістичних систем. Методи визначення стилю автора тематичного україномовного текстового контенту.

## 6.2. Структура навчальної дисципліни

Назви змістових модулів і тем	Кількість годин											
	Форма навчання: денна					Форма навчання: заочна						
	Усього	у тому числі				Усього	у тому числі					
		лекції	практичні	лабораторні	індивідуальна робота		самостійна робота	лекції	практичні	лабораторні	індивідуальна робота	самостійна робота
<b>Модуль 1</b>												
Тема 1 Комп'ютерна лінгвістика: термінологічні засади, історія та напрями розвитку		2	1			8		1	1			8
Тема 2. Лінгвістична модель та моделі подання знань		2	1			8		1	1			10
Тема 3. Лінгвістичний алгоритм та лінгвістичний процесор		2	1			8						10
Тема 4. Технології та інструменти NLP		2	2			8						10
Модульна контрольна робота		2										2
Разом за модуль	47	10	5			32		2	2			
<b>Модуль 2</b>												
Тема 5. Автоматизований морфологічний аналіз тексту (АМА)		2	2			8		2	2			10
Тема 6. Автоматизований синтаксичний аналіз тексту (АСА)		2	1			8		1	1			10
Тема 7. Автоматизований логіко-семантичний аналіз тексту		2	1			6						10
Тема 8. Автоматизований сентиментальний аналіз тексту		2	1			6						8
Модульна контрольна робота		2										2
Разом за модуль	53	20	5			28		3	3			80

<b>Модуль 3</b>											
Тема 9. Електронні корпуси текстів: розмітка текстів та формування електронних корпусів текстів. Бази даних, бази знань, сховища даних як інформаційні моделі		2	2			6		1	1		10
Тема 10. Лінгвістичні проблеми створення баз знань		2	2			6		1	1		10
Тема 11. Моделювання мовленнєвої діяльності в комп'ютерних діалогових системах		2	2			6					10
Тема 12. Автоматизоване реферування та анотування текстів		2	2			6					8
Модульна контрольна робота		2	2								2
Разом за модуль	40	10	6			24		2	2		40
<b>Модуль 4</b>											
Тема 13. Середовища розробки програмних продуктів для NLP		2	2			6		2	2		10
Тема 14. Великі дані в лінгвістиці		2	2			6		1	1		10
Тема 15. Застосування систем III та Великих даних у лінгвістиці		2	2			6					10
Тема 16. Методи машинного навчання для NLP		2	2			6					8
Модульна контрольна робота		2	2								2
Разом за модуль	40	10	6			24		3	3		40
<b>Разом за рік</b>	<b>180</b>	<b>20</b>	<b>32</b>			<b>108</b>		<b>10</b>	<b>10</b>		<b>160</b>

### 6.3. Теми практичних (семінарських) занять

№ з/п	Назва теми	Кількість годин	
		денна	заочна
1	Моделі подання знань для побудови систем опрацювання природної мови	2	2
2	Принципи функціонування лінгвістичного алгоритму та лінгвістичного процесора	4	2
3	Аналіз можливостей інформаційних технологій та інструментів для опрацювання природної мови	4	2
4	Використання спеціалізованих інструментів для створення та аналізу електронних корпусів текстів	4	2
5	Інструменти Великих даних у лінгвістичних дослідженнях	4	2
6	Використання методів штучного інтелекту у лінгвістичних дослідженнях	2	
7	Використання методів машинного навчання для NLP	2	
8	Використання бази даних для зберігання та опрацювання лінгвістичної інформації, моделювання знань.	2	
9	Інструменти автоматичного морфологічного аналізу тексту (АМА).	2	
10	Інструменти автоматичного синтаксичного аналізу тексту	2	
11	Інструменти автоматичного логіко-семантичний аналізу тексту	2	
16	Інструменти сентиментального аналізу	2	
	<b>Разом</b>	<b>32</b>	<b>10</b>

### 6.4. Самостійна робота

№ з/п	Назва теми	Кількість годин	
		денна	заочна
	Комп'ютерна лінгвістика: термінологічні засади, історія та напрями розвитку.	8	10
	Концепти інформація, знання	8	10
	Підходи до побудови інформаційних та концептуальних моделей лінгвістичних об'єктів	8	10
	Моделі подання знань для побудови систем опрацювання природної мови	8	10
	Бази даних для зберігання та опрацювання лінгвістичної інформації, моделювання знань.	8	10
	Принципи функціонування лінгвістичного алгоритму та лінгвістичного процесора	8	10
	Використання методів машинного навчання для NLP	8	10
	Технології розмітки текстів та формування електронних корпусів текстів	8	10

	Комп'ютерно-лінгвістична «екосистема» української мови	8	10
	Моделювання мовленнєвої діяльності в комп'ютерних діалогових системах	8	10
	Системи штучного інтелекту	8	10
	Лінгвістичні проблеми створення баз знань	8	10
	Автоматизований морфологічний аналіз тексту (АМА)	8	10
	Автоматизований синтаксичний аналіз тексту (АСА)	8	10
	Автоматизований логіко-семантичний аналіз тексту	10	10
	Автоматизований сентиментальний аналіз тексту	10	10
	<b>Разом</b>	<b>132</b>	<b>160</b>

## **7. ІНСТРУМЕНТИ, ОБЛАДНАННЯ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ, ВИКОРИСТАННЯ ЯКИХ ПЕРЕДБАЧАЄ НАВЧАЛЬНА ДИСЦИПЛІНА**

Технічні засоби: мультимедійний проектор.

Обладнання: персональні комп'ютери, ноутбуки.

Програмне забезпечення: Microsoft Office, сервіс Google Meet, комунікаційна платформа Moodle, Microsoft Excel, графічні редактори для специфікації та візуалізації онтологій.

## **8. РЕКОМЕНДОВАНІ ДЖЕРЕЛА ІНФОРМАЦІЇ**

### **ОСНОВНА ЛІТЕРАТУРА**

1. Austermeuhl, F. (2014). *Electronic Tools for Translators*. New York, NY: Routledge
2. *Corpora in cognitive linguistics : corpus-based approaches to syntax and lexis* / Ed. by S.Th. Gries, A. Stefanowitsch. – Berlin et al: Mouton de Gruyter, 2006.
3. *Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse* / Ed. by E. Fitzpartick. – Amsterdam, New York: Rodopi, 2007.
4. *Corpus Linguistics: An International Handbook* / Ed. by A. Lüdeling, M. Kytö. – Vols. 1-2. – Berlin et al: Walter de Gruyter, 2008.
5. *Dataset for NLP-Enhanced Image Classification* / Dmytro Dashenkov, Kirill Smelyakov, Nataliia Sharonova// CEUR Workshop Proceedings. – 2023. – Vol. 3396.-P. 88-101
6. *Developing linguistic corpora: a guide to good practice* / Ed. by M. Wynne. – Режим доступу: URL: <http://www.ahds.ac.uk/creating/guides/linguistic-corpora/index.htm>
7. *Enhanced LSA Method with Ukraine Language Support*/Nataliia Kunanets, Yurii Oliinyk, Dmytro Myhal, Khrystyna Shunevych, Antonii Rzheuskyi, Yuriy Shcherbyna// CEUR Workshop Proceedings. – 2021. – Vol. 2870.-P. 129-140

8. Haladzhun Z., Datsyshyn K., Kunanets N., Veretennikova N., Bidzilya Y. “Anti-vaccinationists&anti-vax”: linguistic means of actualizing assessment in the headlines and leads of Ukrainian text media // CEUR Workshop Proceedings. – 2023. – Vol. 3396 : Computational linguistics and intelligent systems 2023 : proceedings of the 7th International conference on computational linguistics and intelligent systems. Vol. II: Computational linguistics workshop. Kharkiv, Ukraine, April 20-21, 2023. – P. 118–129.
9. Haladzhun Z., Snitsarchuk L., Kunanets N., Veretennikova N. Legal aspects of journalist activity: successes and problems in project implementation // IEEE Access. – 2022. – Vol. 10. – P. 422–425.
10. Karp M., Burtnyk A., Bekhta I., Kunanets N., Melnychuk O., Shainer I. Study of linking words in political speeches of Bill Clinton using Python // IEEE 17th International conference on computer science and information technologies : proceedings, 10–12 November 2022, Lviv, Ukraine. – 2022. – C. 77–82.
11. Karp M., Kunanets N., Bekhta I. Hyperbole in To Kill a Mocking Bird by Harper Lee: reduction data vs representing data vs verification // CEUR Workshop Proceedings. – 2022. – Vol. 3171 : Computational Linguistics and Intelligent Systems 2022 : Proceedings of the 6th International conference on computational linguistics and intelligent systems (COLINS 2022). Vol. 1 : Main conference, Gliwice, Poland, May 12-13, 2022. – P. 501–513.
12. Leech G. Corpus Annotation Schemes // Literary and Linguistics Computing. 2003. Vol. 8. N. 4. P.275–281
13. Mas-Jones, E. (2016). Developments in Interpreting Technologies. MultiLingual, Issue January/February, 26-29.
14. Meiosis and litotes in The Catcher in the Rye by Jerome David Salinger: Text Mining /Marta Karp, Nataliia Kunane// CEUR Workshop Proceedings. – 2021. – Vol. 2870.-P. 166-178
15. Muegge, U. (2012). The silent revolution: Cloud-based translation management systems. TC World, 7, 17-21. Retrieved from: <http://www.tcworld.info/e-magazine/contentstrategies/article/the-silent-revolution-cloud-based-translation-memory-systems>.
16. Pym, A. (2013). Translation Skill-Sets in a Machine-Translation Age. Meta: Translator's Journal, 58(3), 487-503.
17. Qualitative and Quantitative Markers of Individual Authorial Conceptualization1-19 Olena Levchenko, Marianna Dilai Pyrogov// CEUR Workshop Proceedings. – 2023. – Vol. 3396.-P.

18. Reflex Systems of Natural Language Processing in Educational Management Information Systems /Mykola Kubiavka, Serhii Lienkov, Yurii Khlaponin, Liubov Kubiavka, Andrii Bernaz, Kyrylo Pyrogov// CEUR Workshop Proceedings. – 2023. – Vol. 3403.-P. 196-205
19. Routledge Handbook of Interpreting (2015). H. Mikkelsen, R. Jourdenais (Eds.), London, England & New York, NY: Routledge.
20. Usage of Sentiment Analysis to Tracking Public Opinion /Zoia Kochuieva, Natalia Borysova, Karina Melnyk, Dina Huliieva// CEUR Workshop Proceedings. – 2021. – Vol. 2870.-P. 272-285
21. Басюк Т.М., Литвин В.В., Захарія Л.М., Кунанець Н.Е. Машинне навчання: Навчальний посібник. Львів: Видавництво «Новий Світ - 2000», 2019. - 335 с.
22. Бідюк, Н. М. Основи прикладної лінгвістики: вступ до спеціальності : навч. посіб. / Н. М. Бідюк, О. О. Комочкова. – Хмельницький : ХНУ, 2017. – 191 с. <http://elar.khnu.km.ua/jspui/handle/123456789/6085>
23. Бобкова Т. В. До визначення корпусної лінгвістики в сучасному мовознавстві. // Наукові записки Національного університету "Острозька академія. Серія "Філологічна". Випуск 45. – Острог, 2014. – С. 3–6.
24. Денисюк В. В. Інформаційні технології в лінгвістиці. Умань : ФОП Жовтий О. О., 2015. 135 с.
25. Жуковська В.В. Вступ до корпусної лінгвістики. Навчальний посібник. — Житомир: Вид-во ЖДУ ім. І. Франка, 2013. — 142 с.
26. Карпіловська Є.А. Вступ до прикладної лінгвістики: комп'ютерна лінгвістика: Підручник.— Донецьк: ТОВ «Юго-Восток, Лтд», 2006.— 188 с
27. Климченко Н.Ф., Карпіловська Є.А., Комарова Л.І. та ін. Морфемно-словотвірний фонд української мови як дослідницька та інформаційно-довідкова система // Мовознавство. – 1990. - № 6. – С.41-50;
28. Методичні вказівки до виконання практичних робіт з курсу «Основи інформатики та прикладної лінгвістики зі змістовим модулем: копірайтинг» для студентів напряму підготовки 035 «Філологія» денної форми навчання / укладачі: Л. І. Гарцунова, А. В. Прокопенко. – Суми : Сумський державний університет, 2019. – 43 с. <http://lib.sumdu.edu.ua/library/docs/rio/2019/m4643.pdf>

## Інформаційні інтернет ресурси :

1. Британський Національний корпус: <http://www.natcorp.ox.ac.uk/>
2. Корпус сучасної американської англійської мови: <http://www.americancorpus.org/>
3. XML Corpus Encoding Standard Document <http://www.xces.org/>
4. Professor Tony McEnery (Lancaster University) delivers the annual John Sinclair Lecture at the University of Birmingham's English department on Thursday 9 May 2013. "Primed for Violence? A corpus analysis of jihadist discourse"  
<https://www.youtube.com/watch?v=dj5SXGgT11U>
5. <http://linguistlist.org/sp/get-software.cfm?submissionid=12195716> – PRESEMT програма (безкоштовна) для перекладачів.
6. <http://www.fon.hum.uva.nl/praat/>- програма-аналізатор фонетичної складової мовлення..
7. <http://www.speech.kth.se/snack/> - аналізатор звукових хвиль.
8. <http://wordnet.princeton.edu/> - лексична база знань англійської мови.
9. <http://www.lx.ugent.be/toolbox/> - зовнішній каталог програм для лінгвістичних досліджень.
10. [http://www.bmanuel.org/clar2\\_tt.html](http://www.bmanuel.org/clar2_tt.html) - каталог посилань з корпусної лінгвістики.
11. <http://www.langsoft.ch/> - програмне забезпечення для обробки й розбору природних мов, перевірка орфографії, машинний переклад, тезауруси, атрибуції текста для англійської, німецької, французької, італійської мов.
12. <http://similis.org/linguaetmachina.www/index.php>- similis повнофункціональний інструмент машинного перекладу для менеджерів проектів і перекладачів. Similis аналізує попередні переклади, генерує пам'ять перекладів(ТМ) і застосовує його до всіх нових проектів.
13. McEnery T., Wilson A. Corpus Linguistics. – Edinburgh University Press, 2001. – Режим доступу: URL: [http://books.google.com/books?id=nwmgdvN\\_akAC](http://books.google.com/books?id=nwmgdvN_akAC)
14. Bookmarks for corpus-based linguistics. [Электронный ресурс]. – Режим доступу: URL: <http://www.devoted.to/corpora>