



УТВЪРДИЛ: .....

Декан

Дата .....

## СОФИЙСКИ УНИВЕРСИТЕТ "СВ. КЛИМЕНТ ОХРИДСКИ"

Факултет: .....

Специалност: (код и наименование)

--	--	--	--	--	--	--	--	--	--

Магистърска програма: (код и наименование)

С	Л	Б	3	0					
---	---	---	---	---	--	--	--	--	--

Компютърна лингвистика. Интернет технологии в хуманитаристиката  
(редовна форма на обучение)

### УЧЕБНА ПРОГРАМА

Дисциплина: 

--	--	--	--	--

(код и наименование)

Увод в компютърната лингвистика

Преподавател: доц. д-р Иван Держански

Асистент:

Учебна заетост	Форма	Хорариум
Аудиторна заетост	Лекции	30
	Семинарни упражнения	
	Практически упражнения (хоспетиране)	
<b>Обща аудиторна заетост</b>		<b>30</b>
Извънаудиторна заетост	Реферат	
	Доклад/Презентация	
	Научно есе	
	Курсов учебен проект	
	Учебна екскурзия	
	Самостоятелна работа в библиотека или с ресурси	60
<b>Обща извънаудиторна заетост</b>		<b>60</b>
<b>ОБЩА ЗАЕТОСТ</b>		<b>90</b>
Кредити аудиторна заетост		1
Кредити извънаудиторна заетост		2

ОБЩО ЕКСТ	3
-----------	---

№	Формиране на оценката по дисциплината <sup>1</sup>	% от оценката
1.	Workshops (информационно търсене и колективно обсъждане на доклади и реферати)	
2.	Участие в тематични дискусии в часовете	25
3.	Демонстрационни занятия	
4.	Посещения на обекти	
5.	Портфолио	
6.	Тестова проверка	
7.	Решаване на казуси	
8.	Текуща самостоятелна работа /контролно	
9.		
10.		
11.		
12.	Изпит	75

**Анотация на учебната дисциплина:**

Курсът има за цел да запознае студентите-филолози с основите на компютърната лингвистика като теоретична и приложна област, да им даде представа за нейната история, основни задачи, методи и резултати и да ги подготви за по-задълбочено навлизане в нея самостоятелно или в други курсове.

**Предварителни изисквания:**

Познаване на основните понятия от езикознанието, елементарна представа от математика и търпимост към формалните методи. Знаене на английски език на начално равнище.

**Очаквани резултати:**

Студентите ще получат представа за мястото на компютърната лингвистика сред теоретичните и практическите дисциплини, занимаващи се с езика, ще знаят как е възникнала и как се е развивала досега, ще са запознати с нейните основни понятия, методи, задачи и резултати и ще имат желание и готовност да продължат запознанството си с нея.

<sup>1</sup> В зависимост от спецификата на учебната дисциплина и изискванията на преподавателя е възможно да се добавят необходимите форми, или да се премахнат ненужните.

### *Учебно съдържание*

<b>№</b>	<b>Тема:</b>	<b>Хорариум</b>
1	Същност и задачи на компютърната лингвистика	2
2	Езикова нееднозначност: видове и методи за разрешаване	2
3	Езикови корпуси: класификация и технологии	2
4.	Формални езици, граматика и автомати. Йерархия на Чомски	4
5.	Крайни автомати и регулярни граматика	2
6.	Приложения на крайните автомати. Лексикален разбор	4
7.	Тагиране по части на речта	2
8.	Компютърна морфология	2
9.	Безконтекстни граматика и магазинни автомати	2
10.	Синтактичен разбор	4
11.	Автоматично откриване на колокации	2
12.	Лексикално-семантични мрежи	2

### *Конспект за изпит*

<b>№</b>	<b>Въпрос</b>
1	Същност и задачи на компютърната лингвистика
2	Езикова нееднозначност: видове и методи за разрешаване
3	<b>Езикови корпуси: класификация</b>
4.	Езикови корпуси: подравняване и аотиране
5.	Формални езици, граматика и автомати
6.	Йерархия на Чомски
7.	Крайни автомати и регулярни граматика
8.	Приложения на крайните автомати
9.	Лексикален разбор
10.	Тагиране по части на речта
11.	Компютърна морфология
12.	Безконтекстни граматика и магазинни автомати
13.	Синтактичен разбор
14.	Автоматично откриване на колокации
15.	Лексикално-семантични мрежи

### *Библиография*

*Основна:*

Jurafsky and Martin, SPEECH and LANGUAGE PROCESSING: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Second Edition, McGraw Hill, 2008.

Светла Коева (съст.), Езикови ресурси и технологии за български език, Акад, изд. "Марин Дринов", 2014.

*Допълнителна:*

<https://people.cs.umass.edu/~>

<http://www.computing.dcu.ie/~>

**Дата: февруари, 2022 г.**

**Съставил:**

**доц. д-р Иван Держански**