



Утвърдил:

Декан

Дата

СОФИЙСКИ УНИВЕРСИТЕТ "СВ. КЛИМЕНТ ОХРИДСКИ"

Факултет: Славянски филологии

Специалност: (код и наименование)

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

Магистърска програма: (код и наименование)

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

УЧЕБНА ПРОГРАМА

Дисциплина:

--	--	--	--

(код и наименование) Приложение на големите езикови модели за лингвистично моделиране

Преподавател: гл. ас. д-р Валентина Стефанова

Асистент:

Учебна заетост	Форма	Хорариум
Аудиторна заетост	Лекции	30
	Семинарни упражнения	
	Практически упражнения (хоспетиране)	
Обща аудиторна заетост		30
Извънаудиторна заетост	Реферат	
	Доклад/Презентация	
	Научно есе	
	Курсов учебен проект	30
	Учебна екскурзия	
	Самостоятелна работа в библиотека или с ресурси	30
Обща извънаудиторна заетост		60
ОБЩА ЗАЕТОСТ		90
Кредити аудиторна заетост		1
Кредити извънаудиторна заетост		2
ОБЩО ЕКСТ		3

№	Формиране на оценката по дисциплината ¹	% от оценката
1.	Workshops {информационно търсене и колективно обсъждане на доклади и реферати)	
2.	Участие в тематични дискусии в часовете	
3.	Демонстрационни занятия	
4.	Посещения на обекти	
5.	Портфолио	
6.	Тестова проверка	
7.	Решаване на казуси	
8.	Курсов проект	50
9.		
10.		
11.		
12.	Изпит	50

Анотация на учебната дисциплина:

Курсът представя средствата за абстрактно семантично описание и описание на граматичните конструкции, които могат да се разглеждат като относително езиково независими. За илюстрация на формалното семантично и синтактично описание ще се използват български и английски език. Ще бъде показано как големите езикови модели и по-точно, чатботовете, които работят с тях, могат да бъдат използвани за генериране на метаезиково описание на семантично и синтактично ниво. Ще бъдат дискутирани начините за оценка на работата на големите езикови модели, както и рисковете и ограниченията при използването на големите езикови модели за лингвистични задачи.

Предварителни изисквания:

Възможност за четене на научна литература на английски език.

Очаквани резултати:

Студентите ще получат знания за нивата на абстрактно семантично описание и граматичните конструкции, които могат да се използват за произволен език, но ще бъдат илюстрирани с български и английски език. Ще получат практически умения за работа с големи езикови модели за извличане на езикова информация на семантично и синтактично ниво.

Учебно съдържание

№	Тема:	Хорариум
---	-------	----------

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

1	Абстрактно семантично описание (семантични фреймове)	4
2	Езикови (граматични) конструкции	4
3	Използване на големите езикови модели за изучаване на граматичните конструкции	4
4	Създаване на семантични фреймове с помощта на големи езикови модели: за и против	4
5	Оценка на описанието в семантични фреймове с помощта на големи езикови модели	4
6	Извличане на модели на граматични конструкции с помощта на големи езикови модели: за и против	4
7	Оценка на граматични модели с помощта на големи езикови модели	4
8	Рискове и ограничения при използването на големите езикови модели за лингвистични задачи	2

Конспект за изпит

№	Въпрос
1	Абстрактно семантично описание (семантични фреймове)
2	Езикови (граматични) конструкции
3	Използване на големите езикови модели за изучаване на граматичните конструкции (интерфейс на чатботове)
4	Създаване на семантични фреймове с помощта на големи езикови модели (интерфейс на чатботове)
5	Оценка на описанието в семантични фреймове с помощта на големи езикови модели
6	Извличане на модели на граматични конструкции с помощта на големи езикови модели: за и против
7	Оценка на граматични модели с помощта на големи езикови модели
8	Рискове и ограничения при използването на големите езикови модели за лингвистични задачи

Библиография

Основна:

- Bender, E. M. & Koller, A. 2020, July. Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5185–5198).
- Boas, H. C. 2005. Determining the productivity of resultative constructions: A reply to Goldberg and Jackendoff. *Language*, 81.2: 448–464.
- Boas, H. C. 2013. Cognitive construction grammar. In Hoffmann, T. & Trousdale, G. (eds.), *The Oxford Handbook of Construction Grammar* (pp. 233–252). Oxford: Oxford University Press.
- Boas, H. C. & Ziem, A. 2018. Constructing a constructicon for German. In Lyngfelt, B., Borin, L., Ohara, K., & Torrent, T. T. (eds.), *Constructicography: Constructicon Development across Languages* (pp. 83–228). Amsterdam: John Benjamins.
- Bond, F. & Foster, R. 2013. Linking and extending an open multilingual wordnet. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 1352–1362). Association for Computational Linguistics. Google Scholar
- Diessel, H. 2019. *The Grammar Network: How Linguistic Structure Is Shaped by Language Use*. Cambridge: Cambridge University Press.
- Fillmore, C. J. 1977. Scenes-and-frames semantics. In Zampolli, A. (ed.), *Fundamental Studies in Computer Science*, number 59 (pp. 55–88). Amsterdam: North Holland.
- Fillmore, C. J. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6.2: 222–254.
- Fillmore, C. J. 2013. Berkeley Construction Grammar. In Hoffmann, T. & Trousdale, G. (eds.), *The Oxford Handbook of Construction Grammar* (pp. 111–132). Oxford: Oxford University Press.
- Fillmore, C. J., Johnson, C. R., & Petruck, M. R. 2003. Background to FrameNet. *International Journal of Lexicography*, 16.3: 235–250.
- Goldberg, A. E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure Constructions*. Chicago: University of Chicago Press.
- Goldberg, A. E. 2006. *Constructions at Work*. Oxford: Oxford University Press.
- Goldberg, A. E. 2019. *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton: Princeton University Press.
- Hilpert, M. 2019. *Construction Grammar and Its Application to English*. 2nd ed. Edinburgh: Edinburgh University Press.
- Hoffmann, T. 2022. *Construction Grammar: The Structure of English* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Hoffmann, T. & Bergs, A. 2024. Constructions all the way! Text types as constructions. In Hennemann, A. & Tacke, F. (eds.), *Diskurstraditionen – Konstruktionen – Genres* (Sprache in kulturellen Kontexten/Language in Cultural Contexts). Bonn: Bonn University Press.
- Pavlick, E., Wolfe, T., Rastogi, P., Callison-Burch, C., Dredze, M., & Van Durme, B. 2015. FrameNet+: Fast paraphrastic tripling of FrameNet. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), Beijing, China (pp. 408–413). Association for Computational Linguistics.

- Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., Baker, C. F., & Scheffczyk, J. 2016. FrameNet II: Extended theory and practice. Berkeley, CA: ICSI.
- Torrent, T. T., Matos, E. E. D. S., Belcavello, F., et al. 2022. Representing context in FrameNet: A multidimensional, multimodal approach. *Frontiers in Psychology*, 13: 573.
- Torrent, T. T., Thomas Hoffmann, Arthur Lorenzi Almeida, Mark Turner, Copilots for Linguists AI, Constructions, and Frames, 2024.
- Touvron, H., Lavril, T., Izacard, G., et al. (2023). Llama: Open and efficient foundation language models. arXiv preprint. arXiv:2302.13971.
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. 2023. ChatGPT: Five priorities for research. Conversational AI is a game-changer for science. Here's how to respond. *Nature*, 614: 224–226.
- Yamada, K., Sasano, R., & Takeda, K. 2021. Semantic frame induction using masked word embeddings and two-step clustering. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 811–816). Association for Computational Linguistics.
- Yong, Z. X. & Torrent, T. T. 2020. Semi-supervised deep embedded clustering with anomaly detection for semantic frame induction. In *Proceedings of the 12th Language Resources and Evaluation Conference, Marseille, France* (pp. 3509–3519). European Language Resources Association (ELRA).

Допълнителна:

Дата: 15.04.2024

Съставил:

гл. ас. д-р Валентина Стефанова