| Module title | Abbreviation |
|---|---|
| **Multilingual NLP** | 10-I=MNLP-222-m01 |

| Module coordinator | Module offered by |
|---|---|
| holder of the Chair of Computer Science XII | Institute of Computer Science |

| ECTS | Method of grading | Only after succ. compl. of module(s) | |
|---|---|---|---|
| 5 | numerical grade | -- | |
| **Duration** | **Module level** | **Other prerequisites** | |
| 1 semester | graduate | -- | |

**Contents**

Languages of the world: language families, typology, etymology. Linguistic universals: words, morphology, parts-of-speech, syntax. Alphabets (scripts), encoding, and language identification. Multilingual word representation spaces (aka cross-lingual word embeddings). Transformer architecture and Pretrained (multilingual) Language Models. Machine translation. Multilingual resources: unlabeled corpora, lexico-semantic networks and word translations, parallel corpora. Cross-lingual transfer: from word alignment and label projection, over MT-based transfer to zero-shot and few-shot transfer with multilingual Transformer-based language models. Advanced topics: curse of multilinguality, modularization and language adaptation, multilingual sentence encoders, contextual parameter generation, multi-source transfer, gradient manipulations.

**Intended learning outcomes**

Students will acquire theoretical and practical knowledge on modern multilingual natural language processing and also get an insight into cutting edge research in (multilingual) NLP. They will learn how to represent texts from different languages in shared representation spaces that enable semantic comparison and cross-lingual transfer for various NLP tasks. Upon successful completion of the course, the students will be well-equipped to solve practical NLP problems regardless of the language of the text data, and to determine the optimal strategy to obtain best performance for any concrete target language.

**Courses** (type, number of weekly contact hours, language — if other than German)

V (2) + Ü (2)
Module taught in: German and/or English

**Method of assessment** (type, scope, language — if other than German, examination offered — if not every semester, information on whether module is creditable for bonus)

written examination (approx. 60 to 120 minutes)
If announced by the lecturer at the beginning of the course, the written examination may be replaced by an oral examination of one candidate each (approx. 20 minutes) or an oral examination in groups of 2 candidates (approx. 15 minutes per candidate).
Language of assessment: English
Creditable for bonus

**Allocation of places**

--

**Additional information**

--

**Workload**

150 h

**Teaching cycle**

--

**Referred to in LPO I** (examination regulations for teaching-degree programmes)

--

**Module appears in**

Master's degree (1 major) eXtended Artificial Intelligence (xtAI) (2020)

Master's degree (1 major) Artificial Intelligence & Extended Reality (2024)