# NLP and Text Mining

<table>
<thead>
<tr>
<th>Module title</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLP and Text Mining</td>
<td>10-I=STM-162-m01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Module coordinator</th>
<th>Module offered by</th>
</tr>
</thead>
<tbody>
<tr>
<td>holder of the Chair of Computer Science VI</td>
<td>Institute of Computer Science</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ECTS</th>
<th>Method of grading</th>
<th>Only after succ. compl. of module(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>numerical grade</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Duration</th>
<th>Module level</th>
<th>Other prerequisites</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 semester</td>
<td>graduate</td>
<td>--</td>
</tr>
</tbody>
</table>

## Contents
Foundations in the following areas: definition of NLP and text mining, properties of text, sentence boundary detection, tokenisation, collocation, N-gram models, morphology, hidden Markov models for tagging, probabilistic parsing, word sense disambiguation, term extraction methods, information extraction, sentiment analysis. The students possess theoretical and practical knowledge about typical methods and algorithms in the area of text mining and language processing mostly for English. They are able to solve problems through the methods taught. They have gained experience in the application of text mining algorithms.

## Intended learning outcomes
The students possess theoretical and practical knowledge about typical methods and algorithms in the area of text mining and language processing. They are able to solve practical problems with the methods acquired in class. They have gained experience in the application of text mining algorithms.

## Courses
V (2) + Ü (2)

## Method of assessment
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: German and/or English

## Allocation of places
--

## Additional information
Focuses available for students of the Master's programme Informatik (Computer Science, 120 ECTS credits): AT, IT, HCI.

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

## Module appears in
- Master's degree (1 major) Computer Science (2016)
- Master's degree (1 major) Computer Science (2017)
- Master's degree (1 major) Computer Science (2018)
- Master's degree (1 major) Computational Mathematics (2019)
- Master's degree (1 major) Mathematics (2019)
- Master's degree (1 major) Information Systems (2019)
- Master's teaching degree Gymnasium MINT Teacher Education PLUS, Elite Network Bavaria (ENB) (2020)
- Supplementary course MINT Teacher Education PLUS, Elite Network Bavaria (ENB) (2020)
- Master's degree (1 major) Computer Science (2021)
- Master's degree (1 major) Computational Mathematics (2022)
- Master's degree (1 major) Information Systems (2022)
- Master's degree (1 major) Mathematics (2022)