The course is focused on the study of the technologies for the development of advanced applications of spoken and written language. It focuses on new technologies based on deep learning and its application to automatic speech recognition, text to speech, and machine translation.

The final project gives students additional information about a particular topic, and also aims to help boost their own skills in the development of applications or in research.
### Study load

<table>
<thead>
<tr>
<th>Total learning time: 125h</th>
<th>Hours large group: 39h</th>
<th>31.20%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours medium group: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Hours small group: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Self study: 86h</td>
<td>68.80%</td>
</tr>
</tbody>
</table>
# Content

| 1. Introduction to speech processing | Learning time: 6h  
Theory classes: 6h |
|--------------------------------------|---------------------|
| **Description:**   
Introduction to speech technology and applications.  
Fundamentals of speech perception and speech production.  
Speech spectrum: STFT, Spectrogram.  
Features: Cepstrum, MFCC, Pitch.  
Techniques: Vector Quantizers, Gaussian Mixture Models. |

| 2. Deep Learning | Learning time: 9h  
Theory classes: 9h |
|-----------------|---------------------|
| **Description:**   
Introduction to Deep Learning techniques and their application to speech and language processing.  
**Specific objectives:**   
Introduction to deep learning, neural nets, learning algorithms, momentum and rmsprop, regularization, word embeddings, recursive neural networks. Applications to Natural Language Processing. |

| 3. Automatic speech recognition | Learning time: 10h  
Theory classes: 10h |
|---------------------------------|---------------------|
| **Description:**   
Pattern matching, Dynamic time warping  
Hidden Markov models. Isolated word recognition  
Large vocabulary continuous ASR: Acoustic modeling. Language modeling  
Deep Learning for language modelling and automatic speech recognition.  
Toolkits |

| 4. Speech synthesis | Learning time: 7h  
Theory classes: 7h |
|---------------------|---------------------|
| **Description:**   
Linguistic processing. Prosody modeling  
Waveform generation.  
Concatenation and statistical methods.  
Deep learning in speech synthesis. |
5. Machine Translation

Description:
Introduction to Machine Translation. Statistical Machine Translation
Neural Machine Translation
Speech translation.

Learning time: 7h
Theory classes: 7h

Qualification system
Midterm exam: 25%, assignments: 30%, research work: 45%

Bibliography

Basic:
Huang, X.; Acero, A.; Hon, H-W. Spoken language processing: a guide to theory, algorithm and system development. Upper

Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning [on line]. Cambridge, Massachusetts: MIT Press, 2016 [Consultation: