

---

## Master Thesis

---

Natural Language Generation  
and Dialogue Systems

Otto-Friedrich-University of Bamberg  
Gutenbergstraße 13  
96050 Bamberg

[www.uni-bamberg.de/ds/en](http://www.uni-bamberg.de/ds/en)

# Automatic Extraction of POI Categories and Attributes from Travel Guide Texts

### DESCRIPTION:

Within the travel guide domain, digital content is key to success. This content is often provided through a web application where places, so called point of interests, are assigned to (sub-)categories and tagged with semantic attributes, e.g., a restaurant with the label “environmentally sustainable”. This information, for example, may be used to better adjust trips to the preferences of the users.

However, assigning these categories and labels is usually a highly manual process executed by a set of expert labellers. Hence, it is desirable to automate this process.

The goal of this thesis is to build a system that can extract this information from the full text as provided by a travel guide book. This work will be executed together with the Michael Müller Verlag, a travel guide publisher from Erlangen, Germany.

The thesis comprises the following items:

1. Task analysis: what type of information is to be extracted.
2. Literature review of state-of-the-art techniques to extract respective information from text.
3. Selection of a set of techniques and application of this to data provided by the Michael Müller Verlag
4. Evaluation of the proposed techniques

### PREREQUISITES:

- Solid programming skills
- Profound interest in machine learning and text processing

**CONTACT:** Prof. Dr. Stefan Ultes (  
Natural Language Generation and Dialogue Systems  
room: GU13/02.27  
email: [stefan.ultes@uni-bamberg.de](mailto:stefan.ultes@uni-bamberg.de)