



# Bachelor Thesis Project

## » Automatic Generation of Regular Expressions «

**Roman Kern**

*Knowledge Discovery  
Knowledge Technologies Institute  
rkern@tugraz.at*

2015 / 16

## Overview

The task is to automatically generate regular expressions from lists of different types of names, e.g. product names. For example for names of countries such a regular expression could be found:

```
\w+land
```

## Target Group

Students interested in text analysis and software engineering.

- ▶ **Implementation:** Build a system to automatically infer regular expressions that capture certain groups of names with high quality.
- ▶ **Evaluation:** Measure the effectiveness of the regular expressions on a test data set.
- ▶ **Written:** Overview of regular expressions, the implementation and the evaluation.

... more details upon request!

## Contact



Roman Kern

rkern@tugraz.at

Knowledge Technologies Institute, TU Graz & Know-Center GmbH