

## Problem description

**Group:** Thomas Skardal, Thomas Jakobsen.

**Project nr. 1:** Readability index

### [Problems to solve]

In this project we are going to develop a readability index plugin for the Natural Language ToolKit (NLTK). A readability index tool, is a tool to determine how readable or understandable a text is. Some factors in this calculation can be: Syllables per word, characters per word, number of sentences and so on. This depends on what formula is used.

Our first step will be to implement a plugin that supports English texts only. There are already several methods for determining readability of English texts, so hopefully we will accomplish this pretty quick. Then we can start implementing support for Norwegian language as well.

### [Research]

We have to explore the NLTK, and get familiar with what services this already provide that can help us developing the plugin. We also have to study each of the methods for determining readability, and choose which of them that are suitable for use with other languages than English.

### [Motivation]

What we find very interesting in this project is the opportunity to contribute to an existing project. To develop something that may be used by others is a huge motivation-factor for us. Another motivation-factor is getting more into Python development.

### [High-level requirements]

There are few requirements. The program should be able to determine if the text is written in a supported language, and then determine the readability index using one or more known algorithms. While coding we're aiming at following the coding standards given in the NLTK developer guide, so that if this project has a successful result, it will be no problem to contribute with our code.

### [Solution strategy]

We aim at using an agile development strategy. We're planning to set up many smaller tasks that we try to solve one by one. After we've done some research about the topics, and getting to know NLTK, we'll start implementing the features into our plugin.