

# Semantic Web Technology – Assignment 1

Leon F.A. Wetzel  
s3284174  
`l.f.a.wetzel@student.rug.nl`

September 9, 2020

## 1 Introduction

In this report, we take a closer look at two systems designed for the task of named entity linking. We use different English corpora to evaluate the performance of both systems and to determine if they perform well or not. Although there are a number of systems capable of named entity linking, we only take a look at **TextRazor** and **OpenTapioca** as this fits in the scope of the actual assignment.

This report covers several aspects of these systems. We take a look at the data which we will use for the evaluation, we discover more about the metrics used for the evaluation and we take an in-depth look at the actual annotations that were generated by these systems.

## 2 Data

Our test data for evaluating the entity linking systems is fairly simple. We use two English news articles which will be annotated by both systems independently. The first news article is a sample from McGee (2020), which covers the situation in Britain regarding new, controversial legislation for Brexit. The second news article is a sample from McKeever (2020) and explains more about upcoming COVID-19 vaccines. Both articles have been picked based on presence of named entities and differing topic.

## References

Luke McGee. Boris johnson’s government is threatening to breach international law. it could backfire spectacularly, Sep 2020. URL <https://edition.cnn.com/2020/09/09/uk/boris-johnson-rule-of-law-brexit-intl-gbr/>.

Amy McKeever. Dozens of covid-19 vaccines are in development. here are the ones to follow., Sep 2020. URL <https://www.nationalgeographic.com/science/health-and-human-body/human-diseases/coronavirus-vaccine-tracker-how-they-work-latest-developments-cvd/>.