Results

After running the two slightly different variations of the program (named system 1 and system 2), some results were gathered. These results can be found in table INSERT NUMBER HERE.

*Table INSERT NUMBER HERE: System results comparison*

|  | System 1 | System 2 |
| --- | --- | --- |
| Amount of triples | 320 | 241 |
| Accuracy | 92,81% | 94,19% |
| Runtime (in minutes) | 25 | 0,10 |

Aside from the difference in amount of triples found by both variations, the biggest difference between the variations is the runtime. While the first system took 25 minutes after several improvements were already made, the second system only took 10 seconds. The difference between these runtimes can be explained by the methods being used. Both systems use dictionaries as datatypes for storing the information, but the first system has to make several more iterations for every triple in the file, which takes a long time. By altering system 1 into system 2 making use of optimization with several iterations,the runtime was drastically decreased.

However, by using this new method the amount of triples also decreased. The reason for this is that while reviewing the results from system 1, it became clear that there were many triples that only occurred once in the dataset. These triples were therefore excluded from the results.

One thing worth mentioning is that the results above are achieved using only a small portion of the total data dump provided by DBpedia - only the date properties were used. Seeing as the first system used 25 minutes for only those properties, it is unlikely that such a system is suited for mapping all the Dutch properties. The second system however, with higher accuracy and a lower runtime shows promise for mapping all the Dutch properties to their English equivalents.

Seeing as both systems do not achieve a 100% accuracy score, a closer look at the errors is required. While doing that, two different categories of errors have been discovered. The first category contains errors of triples in which the two properties have nothing to do with each other. Some examples of this kind of error are:

**SoccerPlayer <http://nl.dbpedia.org/property/clubupdate> dbpedia-owl:sameAs <http://dbpedia.org/ontology/deathDate>**

**Aircraft <http://nl.dbpedia.org/property/eerstevlucht> dbpedia-owl:sameAs <http://dbpedia.org/ontology/retired>**

The second category contains errors for which the property is simply to vaguely translated. These errors occur more often than errors from the first category. Examples from this category are:

**Settlement <http://nl.dbpedia.org/property/datum> dbpedia-owl:sameAs <http://dbpedia.org/ontology/populationAsOf>**

**SubMunicipality <http://nl.dbpedia.org/property/datum> dbpedia-owl:sameAs <http://dbpedia.org/ontology/populationAsOf>**

Conclusion

To summarize, we have tried to map Dutch DBpedia properties to their English equivalents using a sameAs property to link them. This has been done with two slightly different approaches to see which approach worked better. Because it became clear that the first method had a relatively high runtime, we chose to focus on a particular subset of all the properties in the Dutch DBpedia. Using this approach we achieved a runtime of 25 minutes using one approach, while the other could do it in merely 10 seconds. Both systems achieved a relatively high accuracy, but the system does not work perfectly. The main reason for not scoring 100% accuracy is the vaguely defined Dutch properties, which are not clear enough to map them by simply using a sameAs property.

Using this approach, different language properties can be mapped to their English equivalent. Our approach has been tested with the Dutch DBpedia, but because our approach is language independent it can be extended to other languages. The main advantage of doing this is making it possible to use queries in native languages instead of having to use English properties for the queries.