

---

## Information Retrieval and Text Mining: Assignment 3

### Problem 1.

In this exercise we will get acquainted with Galago Query Language and construct new queries using the query language

1. Think of 5 examples of queries where `#od:1` operator will come in handy. Can you think of 3 examples where `#od:1` is not operated on proper/common nouns?
2. You saw how `#syn` is used in the presentation, come up with 3 more examples of how you can do more efficient retrieval with this. e.g I want search for documents which have either "east berlin" or "west berlin" my query would be:

```
#combine( #syn( #od:1(east berlin) #od:1(west berlin)))
```

3. How will you search for the query which always has "black" in the title and "high heart" in the document. Note the first document of the result.
4. Think of 5 interesting information needs that you can execute with Galago but are more difficult than doing them with keyword search in search engines like Google. Write the Galago query for the same.

Try to use as many of the operators of `#od`, `#combine`, `#weight`, `#syn` in interesting ways. Try to construct complex queries with nested operators.

e.g. Information Need: I want to search for Abraham Lincoln's speeches. But Abraham Lincoln is often addressed as a President Lincoln in documents. So I query Galago:

```
#combine( #od:1(abraham lincoln) #od:1(president lincoln) speech)
```

### Problem 2.

Find explanations and examples for the following statements

1. In a tfidf-ranked system, adding a term that is already part of a query will change the ranking
2. In a tfidf-ranked system, the length of a document is important for the ranking

**Exercise date: Friday, November 25, 2011, 14:00**