

# Foundations of Information Retrieval

## Exercise 6

Exercise session: 10.12.2014  
Questions to: Markus Rokicki, rokicki@L3S.de

### 1. Relevance feedback & query expansion

- a. Jon looks for information about populations of banana slugs and issues a query „banana slug“ to the search engine *RFSearch*, which supports relevance feedback. For this query, *RFSearch* delivers the following top-3 results:

$D_1$ ="banana slug Ariolimax columbianus"  
 $D_2$ ="Santa Cruz mountains banana slug"  
 $D_3$ ="Santa Cruz Campus Mascot"

Jon marks the documents  $D_1$ ,  $D_2$  as relevant and  $D_3$  as not relevant for his information need. *RFSearch* uses term frequency and the Rocchio RF algorithm with  $\alpha = \beta = \gamma = 1$ . Which search query will be executed next based on the Jon's feedback? Specify the weights for each query term. Sort the terms in alphabetical order.

|  |  |  |  |  |  |
|--|--|--|--|--|--|
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

- b. Specify  $\alpha$ ,  $\beta$ , and  $\gamma$  values that facilitate a query retrieving pages similar to a relevant page within the search result using the Rocchio RF algorithm.