**Проекти за курс „Извличане на Информация” 2012**

1. Да се направи търсене на книги, като освен TF/IDF(Lucene) се използва и профил на потребителя(ключовите думи).
2. Да се направи търсене на книги, като освен TF/IDF(Lucene) се използва и онтология която ни помага за семантично търсене.
3. Professional search engines.
4. Open Source SE:
   1. <http://terrier.org/>
   2. <http://lucene.apache.org/solr/>
   3. Lemur Toolkit <http://www.lemurproject.org/>
   4. IndexTank <http://indextank.com/>
   5. Elastic search
   6. Apache Nutch
5. Правене на разюме на книга (summarization)
6. Правене на разюме на книга (въз основа на ключови думи и семантика) – да се ползват topic maps.
7. Да се модифицира горния алгоритъм за да може да работи с променящи се данни и да се използва за препоръки на нови продукти на потребителя.
8. Summarization - MEAD toolkit <http://kavita-ganesan.com/mead>
9. Правене на разюме на книга въз основа на заявка за търсене от потребителя (query based sumarization)
10. Машинен превод <http://www.statmt.org/moses/> или <http://www.euromatrixplus.net/> - IR approach.
11. Natural Language processing - OnenNLP, UIMA