

МОСКОВСКИЙ АВИАЦИОННЫЙ ИНСТИТУТ  
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)  
Факультет информационных технологий и прикладной математики  
Кафедра вычислительной математики и программирования

**Отчет по лабораторной работе №2**

**«Оценка качества поиска»**

**по курсу**

**«Информационный поиск»**

Группа: 80-106М

Выполнил: Забелин М.К.

Преподаватель: Калинин А.Л.

Москва, 2019

## ЛР2: Оценка качества поиска

Необходимо оценить качество своего поиска и сравнить их с двумя альтернативами (для Википедии можно собственный поиск по Википедии, поиск Google или Яндекс с ограничением по сайту Википедии). Как минимум, нужно измерить P, DCG, NDCG и ERR уровней @1, @3 и @5, приветствуется использование дополнительных метрик качества. Для оценки качества необходимо придумать 30 запросов, отражающих интересы пользователей или, если есть доступ к настоящим запросам пользователей, то выбрать репрезентативную подборку.

### Примеры запросов по тематикам «Техника» и «Информация»:

история AMD  
USB-накопитель  
агенство Эврика  
алгоритм Витерби  
преимущества метаанализа  
стандарт Stereo  
Домашний каталог  
функция MD5  
барабан и жесткий диск  
электронное удостоверение личности  
Технологическая школа IBM в России  
Когда появились карманные часы  
булева алгебра  
измерительные механизмы  
Калькулятор МК 56  
катоды косвенного накала  
смартфон Palm Treo  
формат wmv  
инструмент цитра  
типографское клише  
чем наполнен Шарльер  
вертолет Як  
влияет ли ЭМИ на людей  
чем является процесс с независимыми приращениями  
примеры профессиональных стандартов  
процесс сквозного шифрования  
список маломассивных звезд  
что такое тезурус  
что символизируют три обезьяны  
примеры трилогии

#### С точки зрения сложности интересными являются запросы:

- Домашний каталог (т. к. это редкое словосочетание, а по отдельности слова встречаются часто)
- влияет ли ЭМИ на людей (добавление частицы «ли»)
- Когда появились карманные часы (в документе есть утверждение об этом, но слова там в другом порядке)
- чем является процесс с независимыми приращениями (в оригинале обратный порядок слов без местоимения «что»)

Оценивание производилось следующим образом. Максимальную оценку «5» я ставил, если запрашиваемое словосочетание было найдено в нужном контексте (например, слово «барабан» в сочетании со словами «жесткий диск» было употреблено в смысле устройства компьютерной памяти, а не музыкального инструмента). «4» ставилась, если запрос содержал уточнение о запрашиваемом понятии (например, «чем заправляются автобусы?»), но в документе есть просто описано, что из себя это понятие представляет без ответа на уточнение. «3» ставилась, если запрашиваемое слово было в неправильном контексте или содержал не всю часть словосочетания, однако поисковик понял, какая часть словосочетания содержит основную смысловую нагрузку. «2» ставилась, если документ просто содержит какое-то слово из запроса. «1» - если документ вообще не подходит.

Запросы, написанные выше, производились поисковиками Google и Яндекс с ограничением по сайту Википедии.

Оценки, на основании которых считались метрики:

[https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/yandex\\_queries\\_scores.csv](https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/yandex_queries_scores.csv)

[https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/google\\_queries\\_scores.csv](https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/google_queries_scores.csv)

При поиске в Google проблемы были со следующими запросами:

1. Калькулятор МК-56. В результаты попали и другие калькуляторы серии «МК», но была статья про калькуляторы вообще
2. Что символизируют три обезьяны. В выдачу попали просто статьи об обезьянах.
3. стандарт Stereo. В выдачу хоть и попали статьи, касающиеся звука, но так как Stereo было латиницей написано, то ожидалась статья про Stereo 8.
4. вертолет Як — в выдачу попала статья про вертолеты в общем.

Однако стоит заметить, что Google очень хорошо привлекает именно объект поиска от уточняющих запросов слов. Например, на запрос «Чем наполнен Шарльер» все статьи содержали ответ на этот вопрос.

Что касается поиска Яндекса, то мне он понравился меньше. Очень часто документы не содержали всех слов словосочетания, что привело к более низким оценкам. Например, по запросу «История AMD» большая часть документов была об AMD, но не содержала историю компании. Либо была не об AMD, а о дочерней компании.

Для оценки качества поиска были выбраны следующие оценки P, Err, Dcg, Ndcg. Их определение можно найти в файле <https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/metrics.py>:

Результаты метрик P, DCG, NDCG и ERR уровней @1, @3 и @5 можно для поисковиков Яндекс и Google можно увидеть в файлах:

[https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/out\\_yandex\\_queries\\_scores.csv](https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/out_yandex_queries_scores.csv)

[https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/out\\_google\\_queries\\_scores.csv](https://github.com/Mikhail-Z/MAI/blob/master/sem10/Informational%20Search/is2/out_google_queries_scores.csv)

Средние показатели поисковиков по 30 запросам:

Яндекс:

P@1: 4.9

P@3: 4.444466666666667

P@5: 4.226666666666667

DCG@1: 29.4

DCG@3: 53.675133333333335

DCG@5: 68.22213333333333

NDCG@1: 0.9828

NDCG@3: 0.9457666666666666

NDCG@5: 0.9803

ERR@1: 0.9659

ERR@3: 0.9712999999999999

ERR@5: 0.9721333333333333

Google:

P@1: 4.966666666666667

P@3: 4.633366666666666

P@5: 4.453333333333333

DCG@1: 30.466666666666665

DCG@3: 57.60193333333333

DCG@5: 75.00173333333333

NDCG@1: 0.9828

NDCG@3: 0.9466666666666667

NDCG@5: 0.984933333333333

ERR@1: 0.9679666666666666

ERR@3: 0.9743666666666666

ERR@5: 0.9746666666666667

По всем метрикам Google чуть лучше Яндекса при поиске по Википедии.

## **Вывод**

Познакомился с различными метриками для оценки релевантности результатов поиска.

## **Ссылка на Git репозиторий**

<https://github.com/Mikhail-Z/MAI/tree/master/sem10/Informational%20Search/is2>