

# ***1<sup>η</sup> Προγραμματιστική Εργασία***



## Σχολιασμός

Η εμπειρία μου με το Lucene ήταν πολύ θετική. Μπορώ με σιγουριά να αναφέρω ότι αποτελεί ένα πολύ σημαντικό εργαλείο για την ευρετηριοποίηση και αναζήτηση σε σύνολα μεγάλων δεδομένων. Αξιοσημείωτο είναι και το γεγονός πως το Lucene περιλαμβάνει ένα εξαιρετικά καλό documentation, το οποίο περιέχει όλες τις πληροφορίες που θα χρειαστεί κάποιος προγραμματιστής για την εγκατάσταση και χρήση αυτού.

Αρχικά, ακολουθώντας τις οδηγίες της εκφώνησης χρησιμοποίησα και τον EnglishAnalyzer ο οποίος κάνει tokenizing, stemming και καλύτερο «φιλτράτισμα» τα κείμενα στην αγγλική γλώσσα βελτιώνοντας έτσι τη διαδικασία της ευρετηριοποίησης και αναζήτησης. Η χρήση αυτού με τις οδηγίες του documentation αποδείχθηκε αρκετά εύκολη. Έπειτα χρησιμοποιήθηκε και ο SimpleQueryParser ο οποίος επιτρέπει στους χρήστες να δημιουργούν σύνθετα ερωτήματα χρησιμοποιώντας μια απλή σύνταξη, διευκολύνοντας την κατασκευή ακριβών κριτηρίων αναζήτησης χωρίς να απαιτείται βαθιά κατανόηση της γλώσσας ερωτημάτων του Lucene.

Έτρεξα τα παρακάτω queries (ακολουθώντας το [https://lucene.apache.org/core/5\\_0\\_0/queryparser/index.html](https://lucene.apache.org/core/5_0_0/queryparser/index.html))

- Query: rat – (ίδιο αποτέλεσμα και με “rat”)

```
104879 documents indexed.
Query: rat
Title: Got 1794+ hits!
Abstract: Got 2567+ hits!
78322 2.4471707
AGGREGATION AND TRANSFORMATION OF RAT LYMPHOCYTES ON RAT EMBRYO MONOLAYERS
9248 2.3751268
Folate catabolism in tumour-bearing rats and rats treated with methotrexate.
22589 2.3751268
Establishment of Rat Embryonic Stem Cells and Making of Chimera Rats
94054 2.3157487
Determination of rat vertebral bone compressive fatigue properties in untreated intact rats and zoledronic-acid-treated, ovariectomized rats
9308 2.3072033
Comparison of cell-surface glycoproteins of rat hepatomas and embryonic rat liver.
15297 2.3072033
THE TRANSMISSION OF THE VIRUS OF MEXICAN TYPHUS FROM RAT TO RAT BY POLYPLAX SPINULOSUS
39803 2.3072033
THE INFLUENCE OF MOUSE-RAT PARABIOSIS ON THE GROWTH IN RATS OF A TRANSPLANTABLE MOUSE SARCOMA
78876 2.3072033
Mobility of surface proteins on normal rat macrophages and on a "macrophagelike" rat tumor
946 2.281942
Mesovarian leiomyomas in the rat.
12155 2.281942
ANTIGENICITY OF RAT COLLAGEN
Query:
```

- Query: rat liver (όταν δίνουμε δύο όρους θεωρείται αυτόματα ότι ενδιάμεσα υπάρχει το OR)

```
Query: rat liver
Title: Got 1021+ hits!
Abstract: Got 1016+ hits!
51107 4.953596
STUDY OF MITOCHONDRIA IN RAT LIVER
79359 4.953596
PLASMA MEMBRANES OF THE RAT LIVER
69559 4.701856
The Cellular Transport of Magnesium in Rat Liver
70390 4.701856
EFFECTS OF CORTISONE ON RAT LIVER MITOCHONDRIA
72595 4.701856
Effect of Thioacetamide on Rat Liver Regeneration
77057 4.701856
INORGANIC PYROPHOSPHATASE OF RAT LIVER MITOCHONDRIA
77069 4.701856
THE SYNTHESIS AND TURNOVER OF RAT LIVER PEROXISOMES
77427 4.701856
THE NUCLEOIDS OF RAT LIVER CELL MICROBODIES
77602 4.701856
QUANTITATIVE STUDY OF MITOCHONDRIA IN RAT LIVER
77711 4.701856
ISOLATION OF RAT LIVER PLASMA MEMBRANES
Query:
```

- Query: rat+liver (το + είναι το AND operator)

```

Query: rat+liver
Title: Got 325 hits!
Abstract: Got 1048+ hits!
51107 4.953596
STUDY OF MITOCHONDRIA IN RAT LIVER
79359 4.953596
PLASMA MEMBRANES OF THE RAT LIVER
69559 4.701856
The Cellular Transport of Magnesium in Rat Liver
70390 4.701856
EFFECTS OF CORTISONE ON RAT LIVER MITOCHONDRIA
72595 4.701856
Effect of Thioacetamide on Rat Liver Regeneration
77057 4.701856
INORGANIC PYROPHOSPHATASE OF RAT LIVER MITOCHONDRIA
77069 4.701856
THE SYNTHESIS AND TURNOVER OF RAT LIVER PEROXISOMES
77427 4.701856
THE NUCLEOIDS OF RAT LIVER CELL MICROBODIES
77602 4.701856
QUANTITATIVE STUDY OF MITOCHONDRIA IN RAT LIVER
77711 4.701856
ISOLATION OF RAT LIVER PLASMA MEMBRANES

```

- Query: -rat-liver (το – υποδηλώνει το NOT, Παρατηρούμε επίσης ότι το scoring όλων των επιστρεφόμενων αποτελεσμάτων είναι 1,0)

```

Query: -rat-liver
Title: Got 1001+ hits!
Abstract: Got 1001+ hits!
0 1.0
Health inequality in adolescence. Does stratification occur by familial social background, family affluence, or personal social position?
1 1.0
Subaortic and mid-ventricular obstructive hypertrophic cardiomyopathy with an apical Aneurysm: a case report
2 1.0
Commentary on effects of anthropogenic and natural organic chemicals on development, swimming behavior, and reproduction of Daphnia, a key member of aquatic ecosystems.
3 1.0
Association between air pollution and low birth weight: a community-based study.
4 1.0
A test of the model that RNA polymerase III transcription is regulated by selective induction of the 110 kDa subunit of TFIIC
5 1.0
The DNA polymerase λ is required for the repair of non-compatible DNA double strand breaks by NHEJ in mammalian cells
6 1.0
Review of new evidence regarding the relationship of gasoline exposure to kidney cancer and leukemia.
7 1.0
Defining the role of pollutants in the disruption of reproduction in wildlife.
8 1.0
Carbon dioxide kinetics and capnography during critical care
9 1.0
Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data

```

- Query: "rat liver" (με τα "" δημιουργούμε φράσεις με όρους)

```

Query: "rat liver"
Title: Got 257 hits!
Abstract: Got 506 hits!
51107 4.9535966
STUDY OF MITOCHONDRIA IN RAT LIVER
79359 4.9535966
PLASMA MEMBRANES OF THE RAT LIVER
69559 4.7018566
The Cellular Transport of Magnesium in Rat Liver
70390 4.7018566
EFFECTS OF CORTISONE ON RAT LIVER MITOCHONDRIA
72595 4.7018566
Effect of Thioacetamide on Rat Liver Regeneration
77057 4.7018566
INORGANIC PYROPHOSPHATASE OF RAT LIVER MITOCHONDRIA
77069 4.7018566
THE SYNTHESIS AND TURNOVER OF RAT LIVER PEROXISOMES
77427 4.7018566
THE NUCLEOIDS OF RAT LIVER CELL MICROBODIES
77602 4.7018566
QUANTITATIVE STUDY OF MITOCHONDRIA IN RAT LIVER
77711 4.7018566
ISOLATION OF RAT LIVER PLASMA MEMBRANES
Query:

```

- Query: micro\* (με το \* δηλώνουμε ένα prefix. Ουσιαστικά μας επιστρέφει αποτελέσματα που ξεκινάνε από αυτόν τον όρο π.χ. microarray, micronutrient κλπ.)

```

Query: micro*
Title: Got 1001+ hits!
Abstract: Got 1001+ hits!
9 1.0
Identification of frequent cytogenetic aberrations in hepatocellular carcinoma using gene-expression microarray data
136 1.0
Inhibition of highly productive HIV-1 infection in T cells, primary human macrophages, microglia, and astrocytes by Sargassum fusiforme
164 1.0
Challenging the concept of microinvasive carcinoma of the vulva: report of a case with regional lymph node recurrence and review of the literature
167 1.0
A randomized cross over trial of tolerability and compliance of a micronutrient supplement with low iron separated from calcium vs high iron combined with calcium in pregnant women [ISRCTN56071145]
194 1.0
Inference of Disease-Related Molecular Logic from Systems-Based Microarray Analysis
196 1.0
Signaling mechanisms in tumor necrosis factor alpha-induced death of microvascular endothelial cells of the corpus luteum
216 1.0
Transport of particles of colloidal gold within and from rat lung after local deposition by alveolar microinjection.
241 1.0
Micronucleus assay in lymphocytes as a tool to biomonitor human exposure to aneuploidogens and clastogens.
260 1.0
Microbial identification by mass cataloging
271 1.0
Epidemiologic evidence relevant to radar (microwave) effects.
Query:

```

- Query: "rat liver"~5 (το ~ προσδιόριζει κοντινά queries)

```
Query: "rat liver"~5
Title: Got 294 hits!
Abstract: Got 646 hits!
51107 4.9535966
STUDY OF MITOCHONDRIA IN RAT LIVER
79359 4.9535966
PLASMA MEMBRANES OF THE RAT LIVER
69559 4.7018566
The Cellular Transport of Magnesium in Rat Liver
70390 4.7018566
EFFECTS OF CORTISONE ON RAT LIVER MITOCHONDRIA
72595 4.7018566
Effect of Thioacetamide on Rat Liver Regeneration
77057 4.7018566
INORGANIC PYROPHOSPHATASE OF RAT LIVER MITOCHONDRIA
77069 4.7018566
THE SYNTHESIS AND TURNOVER OF RAT LIVER PEROXISOMES
77427 4.7018566
THE NUCLEOIDS OF RAT LIVER CELL MICROBODIES
77602 4.7018566
QUANTITATIVE STUDY OF MITOCHONDRIA IN RAT LIVER
77711 4.7018566
ISOLATION OF RAT LIVER PLASMA MEMBRANES
```