



APACHE LUCENE

Μεντόνκα Ιζαμπέλ – 3100109
Πίτσιος Σταμάτης - 3100153

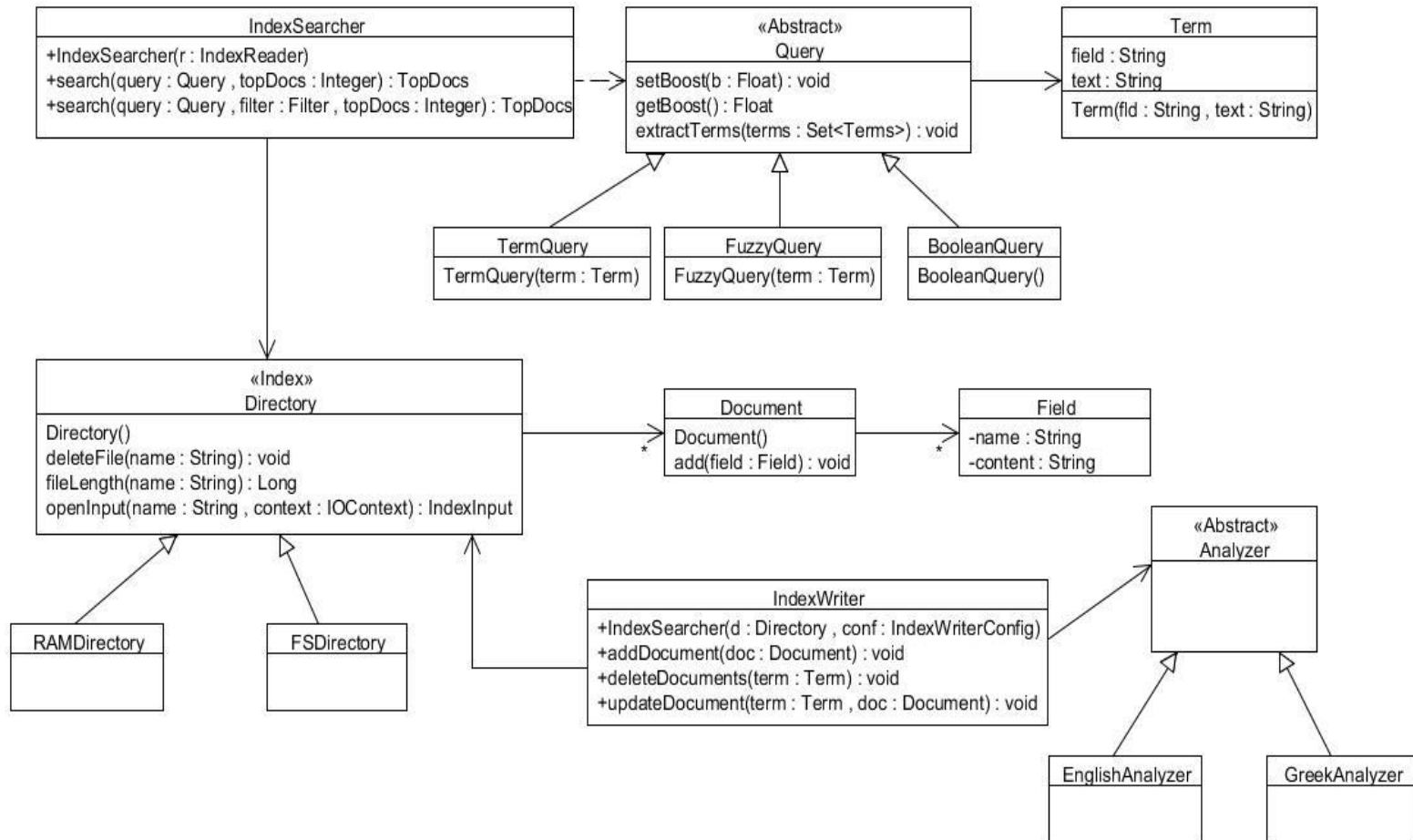
Λίγα λόγια για τη LUCENE (1)

- Η LUCENE είναι μια υψηλής απόδοσης μηχανή αναζήτησης **κειμένου** .
- Γλώσσα υλοποίησης : JAVA
- Διατίθεται ως ΕΛ/ΛΑΚ
- Άδεια : Apache License v.2

Λίγα λόγια για τη LUCENE (2)

- Δεν είναι μια πλήρης εφαρμογή αλλά χρησιμοποιείται ως **βιβλιοθήκη** ή **API** .
- Χρησιμοποιείται κυρίως σε εφαρμογές ανάκτησης πληροφοριών
- Αρχιτεκτονική : Component- Based

Βασική Αρχιτεκτονική



Σημαντικά έργα που χρησιμοποιούν τη LUCENE

□ Apache Solr

□ Elasticsearch

□ Apache Nutch

□ Compass

□ DocFetcher



LUCENE στο Github

- Στο Github, μπορεί κανείς να βρει τη LUCENE μαζί με το έργο SOLR
- Μέγεθος κοινότητας : 16 μέλη
- Μέγεθος έργου (πλήθος γραμμών κώδικα)
 - ❖ LUCENE → 666.142
 - ❖ SOLR → 335.390
- 10.000+ Commits
- 25 Releases

Scenes from our next episode...

Σκοπός μας είναι να συνεισφέρουμε σε τουλάχιστον ένα από τους παρακάτω τομείς :

- Προσθήκη νέας συνάρτησης ομοιότητας
- Προσθήκη /Τροποποίηση για επεξεργασία ελληνικών κειμένων
- Έλεγχος αν γίνεται χρήση της ίδιας συνάρτησης ομοιότητας κατά την ευρετηρίαση και την αναζήτηση
- Ανάλυση και αναζήτηση κειμένου αξιοποιώντας N-γράμματα

Πηγές

- Apache Lucene
<http://lucene.apache.org/core/>
- Wikipedia
<http://en.wikipedia.org/wiki/Lucene>
- Github
<https://github.com/apache/lucene-solr.git>
- SWEBOK v.3

Thank you!!

ENESOLRNETWORKSEARCHSYST
NNESSIAMPSONAVIVAYEATSBECH
KADO **REV** SHARPLUCENESC
EYCA TRINITYGUINNES
AFTONHES **S** RUGBYMOLL
RCHSYST **OLR** LEINSTEROIF
IESONAVIVA **L** UCENE YE
GBYMOLLYT **UTION** DC
TEMLEINSTE OIRE
YAYEATSBECH SHAV
ODDERFLEADHSHARPLUCENESOLR
ISTEROIREACHTASLIFFEYCASTLES