

# Retrospective Document

Sprint 6  
08.05.2016

## Work & Test Progress

---

Selectivity Algorithm: Selection mechanism for determining whether a multigram term is useful or useless - 100%

Optimization of the System: Optimizing each component in the system by increasing their performance and decreasing its memory and possible CPU usage. For test purposes, index gets created at each compilation. Index's being created twice is a Lucene implementation manner, that's why not dug very deep. 80%

Integration of components: Remaining percentage is completed; bugs are issued and cleaned as much as the tests showed.

Test milestones done but not listed in openproject;  
Working in c++ to compile and run Lucene which is written in Java

---

## Team Progress

Fatih Burak Belce - 25%  
Mustafa Güven - 25%  
Oğuzhan Demir - 25%  
Özgür Baskın - 25%

## Left-overs (Backlog)

---

Optimization of the System: For better or for worse, we managed to decrease the time for constructing final filtered index. Furthermore, as we applied the correct filters and selectivity on the index, final index structure used less memory. Ensuing tests also showed that search time also decreases due to shrinking document list on which the regular expression is made. We have not implemented prefix-freeness and suffix-freeness of the index yet. Though, as the physical access plan tree is being constructed, prefix-freeness was getting implemented by itself. But then, after the meeting we had with our instructor, we decided to generate our grams starting with smaller pieces. This led terms to have prefixes in the index if they also meet the requirements of usefulness. At the same meeting, we decided not to implement suffix-freeness since it has a questionable performance increase. This part of the milestones is on hold. Thus, Leftover percentage: 20%

## Next Sprint

---

### Simple UI for Searching

Finding a reasonable index selectivity constant. That would require a big data gathering and tests that could last a day long index creation process.

Comprehensive testing by time memory and against other contemporary products

---

## Team Comments

In this sprint, we accomplished a remarkable work and, had an effective communication and exchange of ideas both within ourselves and with the assistant and supervisor of the project.

We are almost done with the implementation of the total project. After this point there will be tests, more and more optimizations and UI for more understandable search results.

## Assistant Evaluation

---

*Assistant's (Team Leader's) comments regarding to this completed sprint.*

Write down your comments if the team has not made satisfactory progress in this sprint.

Also indicate whether the team's report on their progress (i.e. the reported completion percentages in the retrospective document) are accurate.

*This section should be filled in by Assistant (Team Leader)*

## Supervisor Evaluation

---

*Supervisor's (Team Leader's) comments regarding to this completed sprint.*

*This section should be filled in by Supervisor*