

PROJECT PHASE-II

SMART SEARCH ENGINE

TEAM MEMBERS:

YELISETTI KRISHNA TEJA (NET-ID : KXY200016)
SATYA SOMEPALLI (NET-ID : SXS190436)
SANJANA PENMETSA (NET-ID : SXP190149)
PREETHAM RAO GOTTUMUKULA (NET-ID : PXG210001)

HARSHITH RAVIPROLU (NET-ID : HXR180005)
SAKETH DASAVATHINI (NET-ID : SXD190016)
B MOUNIKA (NET-ID : MXB210007)
THOTA JAYASHREE SANTHOSHI (NET-ID : JXT210011)

SAI PRANAV REDDY DONTIDI (NET-ID : SXD200125)
APUROOP PARAVADA (NET-ID : AXP210033)
YOGESH BALA (NET-ID : YXB200007)
SRINATH REDDY MAVILLAPALLY (NET-ID : SXM210047)
KIRAN DEVARAJ RAJ (NET-ID : KXR190038)

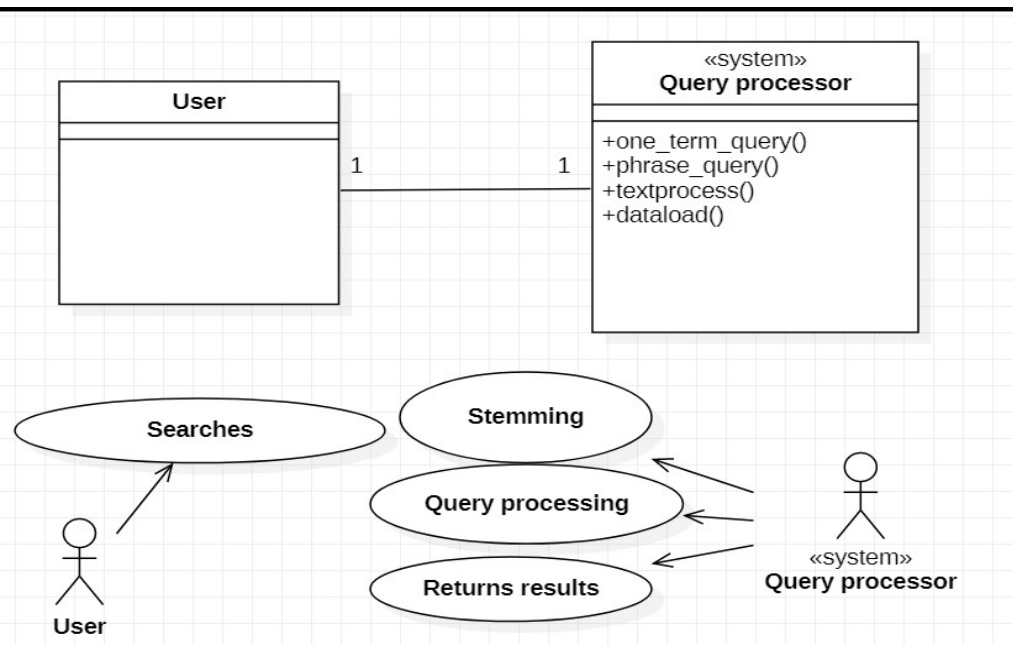
MAJOR RESPONSIBILITIES:

- Expanded the database
- PPT preparation
- Updated Sequence Diagrams
- State Transition-Diagrams
- User-Interface implementation
- Python module changes

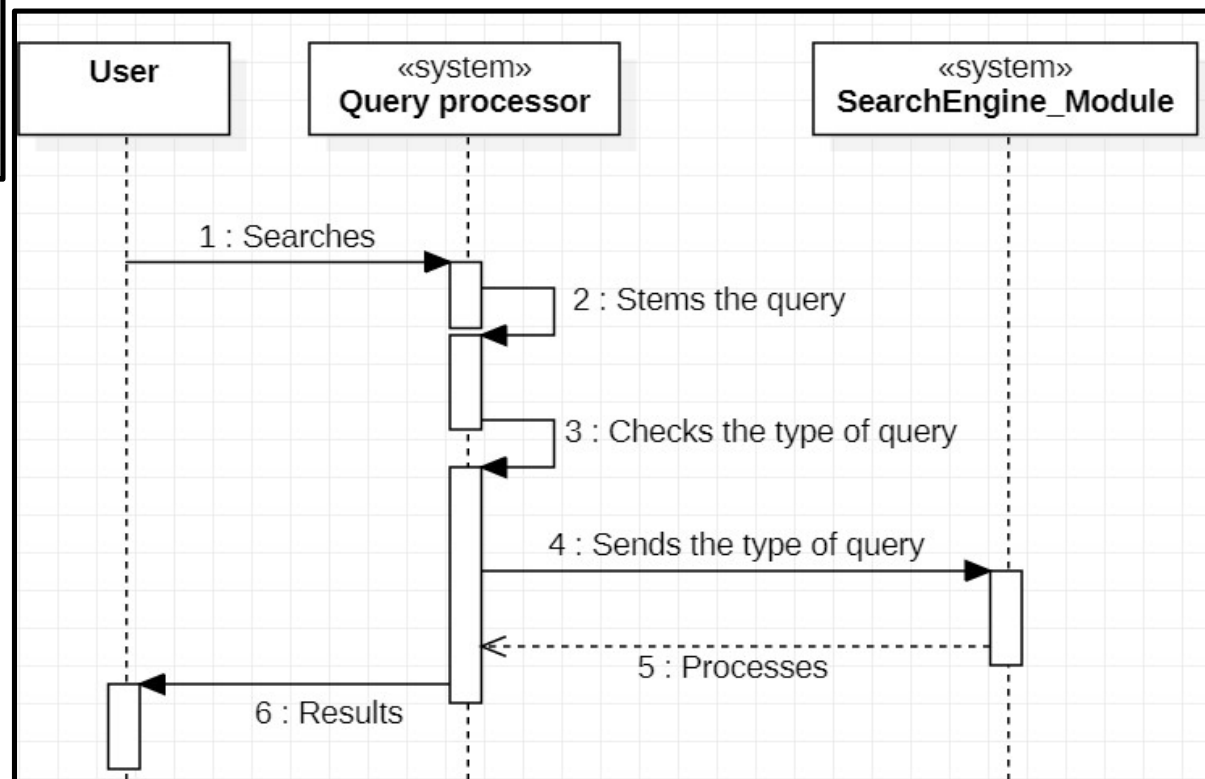
- Updated class-diagram
- Updated sequence diagrams
- Python module changes
- Expanded the database
- Activity Diagrams
- State-Transition diagrams

- Updated usecase diagrams
- Updated class-diagram
- Pre-processor python module changes
- PPT preparation
- User-Interface implementation

WHERE WERE WE
IN PROJECT PHASE-I ?



Just a simple search engine with few basic use cases in Phase-I



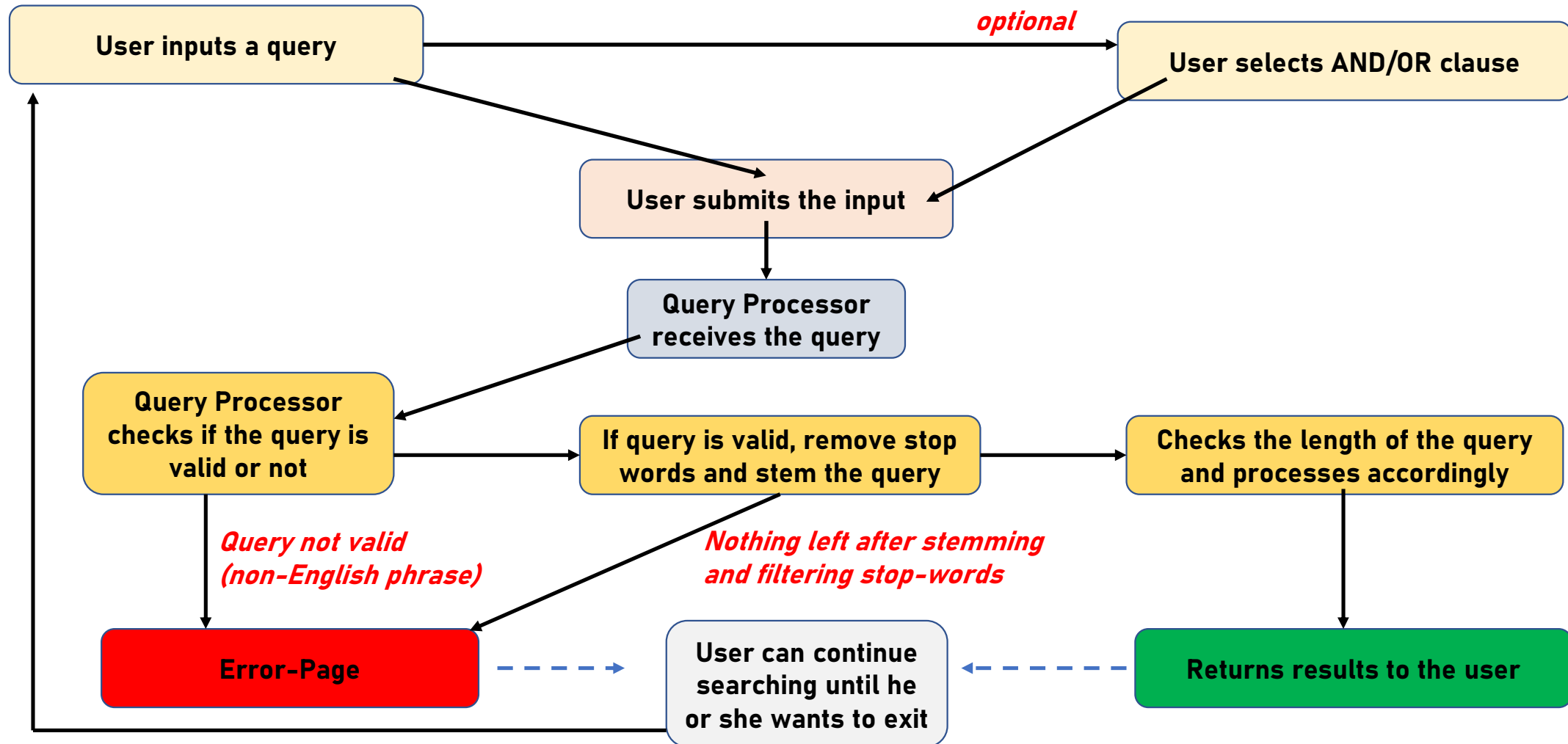
Main sequence diagram in Phase-I



SUMMARY OF SMART-SEARCH ENGINE IN PHASE-II

- **Smart Search Engine is a movie-based search engine where a user can search for movies using genre(s) or movie name or actor name(s) or any word related to movies.**
- **The most relevant results (in this case: most relevant movie related Wikipedia-URLs based on the word(s) searched) will be displayed at the top.**
- **This search engine also supports multi-phrase queries/words (example: multiple-genres based search).**
- **User can also click on a specific link displayed in the results and navigate to that link for more details on that result.**
- **User will be displayed with a two-line description below every Wikipedia-URL in the results.**
- **User will have an option of navigating through multiple results using the pagination functionality implemented in the phase-two.**

BASIC FLOW OF PROJECT IN PHASE-II



WHAT HAVE WE ADDED IN
PHASE-II OF PROJECT

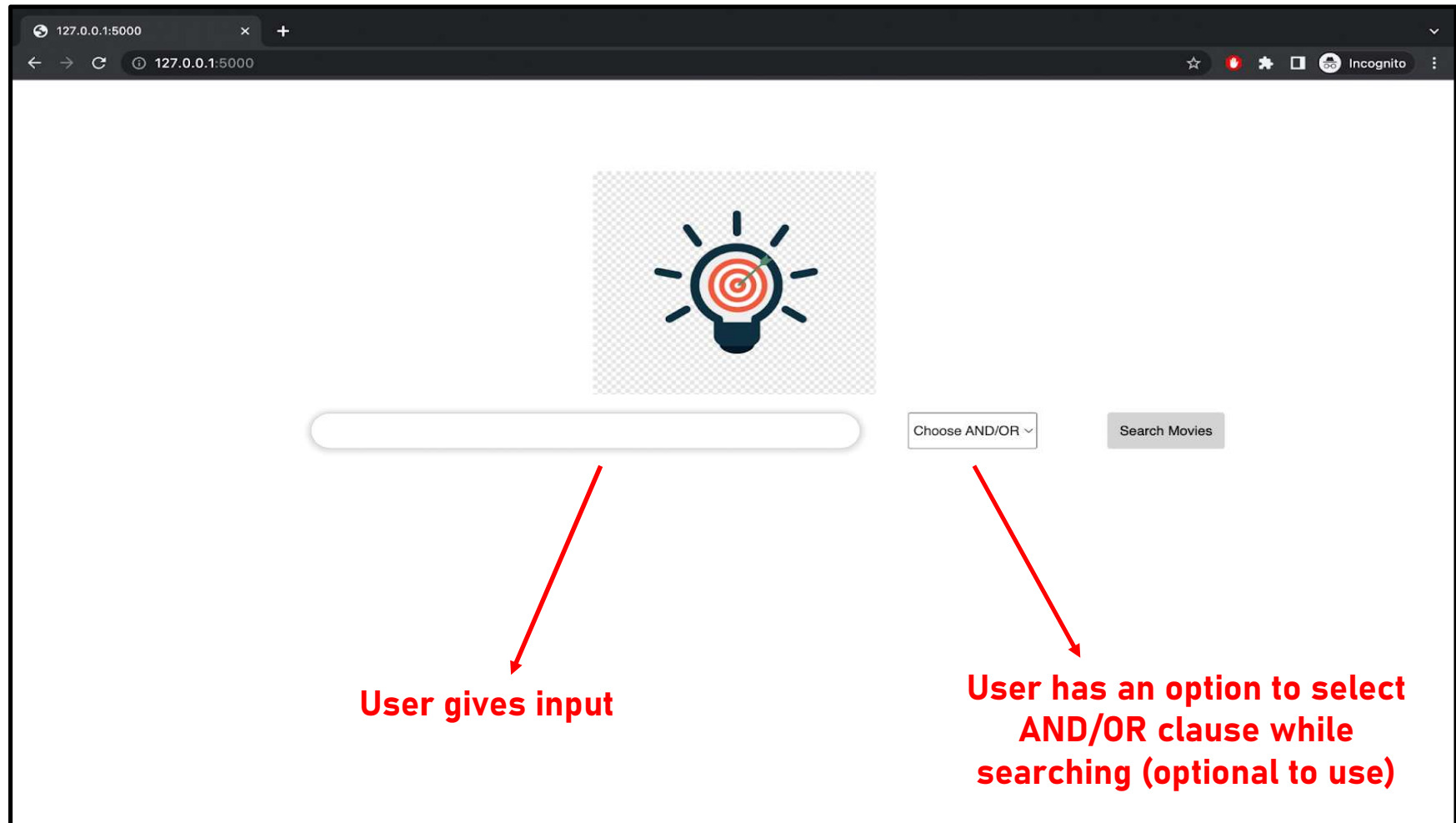
PHASE-2 OF THE PROJECT

- **Advancements in User Interface**
- **Auto Suggestion of input query**
- **AND/OR Clause implementation (*Multi Phrase Query*)**
- **Hyperlink enforcement and Description addition below URLs**
- **Updated Page-Ranking algorithm (*Relevancy of Results*)**
- **Low Latency & High Throughput**
- **Pagination for better navigation**
- **Error 404 Page Implementation (*Handling non-English input and unmatched query*)**
- **Addition of URLs (*Expanded the database*)**
- **Updated previously drawn UML diagrams**
- **State transition & Activity Diagrams**

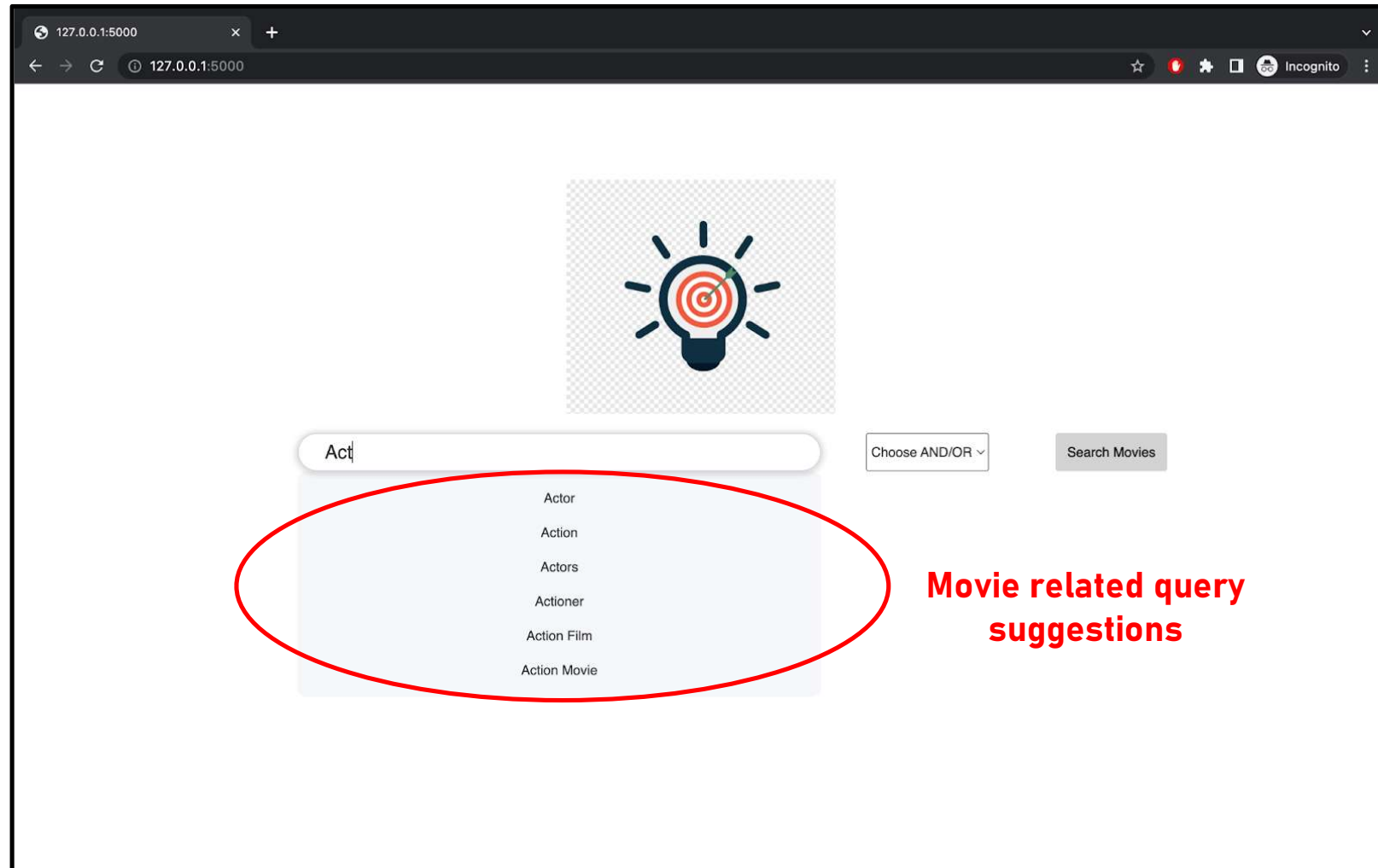
**ADDED
(OR)
UPDATED**

**ADDITIONAL OR UPDATED
FUNCTIONALITIES IN PHASE-II**

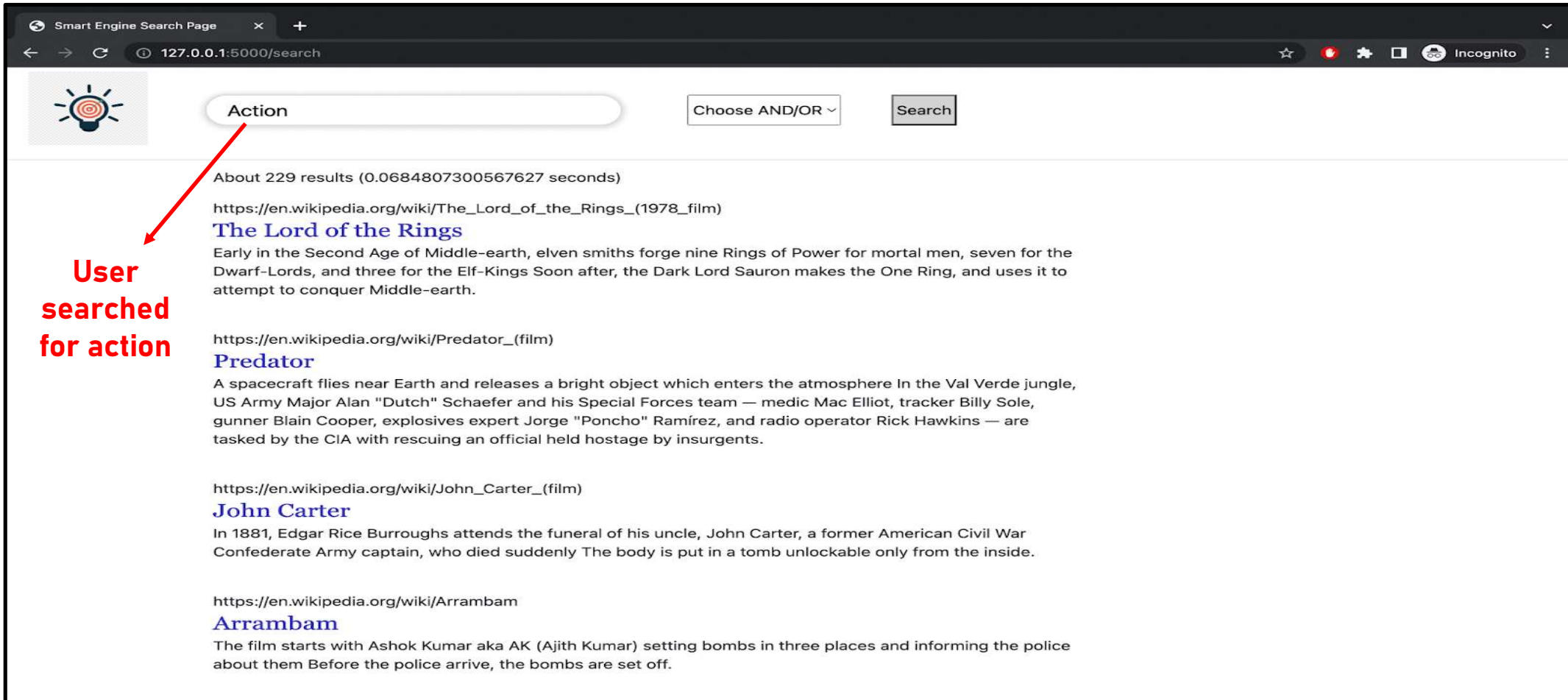
INDEX-PAGE OF SMART-SEARCH ENGINE IN PHASE-II



AUTOSUGGESTION IMPLEMENTATION IN PHASE-II



RESULT-PAGE OF SMART-SEARCH ENGINE IN PHASE-II




The screenshot displays a web browser window with the address bar showing "127.0.0.1:5000/search". The page features a search bar with the text "Action" and a "Search" button. Below the search bar, the results are listed, each with a URL, a title, and a brief description. A red arrow points from the text "User searched for action" to the search bar.

Smart Engine Search Page x +

127.0.0.1:5000/search

Incognito

 Action Choose AND/OR Search

About 229 results (0.0684807300567627 seconds)

[https://en.wikipedia.org/wiki/The_Lord_of_the_Rings_\(1978_film\)](https://en.wikipedia.org/wiki/The_Lord_of_the_Rings_(1978_film))
The Lord of the Rings
Early in the Second Age of Middle-earth, elven smiths forge nine Rings of Power for mortal men, seven for the Dwarf-Lords, and three for the Elf-Kings. Soon after, the Dark Lord Sauron makes the One Ring, and uses it to attempt to conquer Middle-earth.

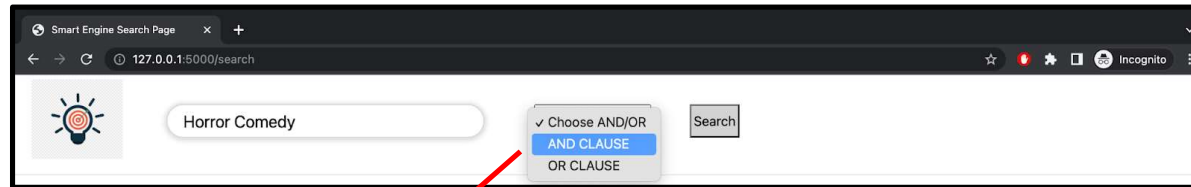
[https://en.wikipedia.org/wiki/Predator_\(film\)](https://en.wikipedia.org/wiki/Predator_(film))
Predator
A spacecraft flies near Earth and releases a bright object which enters the atmosphere. In the Val Verde jungle, US Army Major Alan "Dutch" Schaefer and his Special Forces team — medic Mac Elliot, tracker Billy Sole, gunner Blain Cooper, explosives expert Jorge "Poncho" Ramírez, and radio operator Rick Hawkins — are tasked by the CIA with rescuing an official held hostage by insurgents.

[https://en.wikipedia.org/wiki/John_Carter_\(film\)](https://en.wikipedia.org/wiki/John_Carter_(film))
John Carter
In 1881, Edgar Rice Burroughs attends the funeral of his uncle, John Carter, a former American Civil War Confederate Army captain, who died suddenly. The body is put in a tomb unlockable only from the inside.

<https://en.wikipedia.org/wiki/Arrambam>
Arrambam
The film starts with Ashok Kumar aka AK (Ajith Kumar) setting bombs in three places and informing the police about them. Before the police arrive, the bombs are set off.

User searched for action

AND-CLAUSE IMPLEMENTATION IN PHASE-II

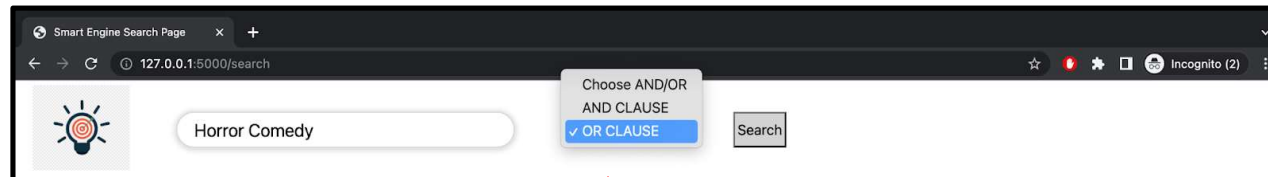


The main search results page shows the search bar with "Horror Comedy" and the dropdown menu set to "AND CLAUSE". Below the search bar, it displays "About 14 results (0.009011983871459961 seconds)". The results are listed as follows:

- https://en.wikipedia.org/wiki/Killer%27s_Moon
Killer's Moon
A coach full of schoolgirls breaks down in the Lake District, forcing the girls to take shelter for the night in a remote hotel. Meanwhile, strange and macabre things are happening to the locals (and their pets) and it is revealed that four escaped mental patients- Mr.
- https://en.wikipedia.org/wiki/Piranha_II:_The_Spawning
Piranha II: The Spawning
Off the coast of a Caribbean island, a young couple flee a hotel to have sex in the sea. But they swim into a sunken wreck which is also a piranha lair and they are both killed and eaten by the unseen piranha.
- https://en.wikipedia.org/wiki/The_Texas_Chainsaw_Massacre_2
The Texas Chainsaw Massacre 2
In 1986, thirteen years after the events of the first film, two high school seniors, Buzz (Barry Kinyon) and Rick (Chris Douridas) , race along a desolate stretch of Texas highway, en route to the Texas-OU football game at the Dallas Cotton Bowl. Heavily intoxicated, they use their car phone to call and harass on-air radio DJ Vanita "Stretch" Brock (Caroline Williams).
- [https://en.wikipedia.org/wiki/Burke_%26_Hare_\(1971_film\)](https://en.wikipedia.org/wiki/Burke_%26_Hare_(1971_film))
Burke & Hare
In Edinburgh in the 1820s, surgeon Dr Knox (Harry Andrews) employs graverobbers Burke and Hare (Derren Nesbitt and Glynn Edwards) to supply fresh corpses for his anatomical lectures at the medical academy. When graveyard supplies run low, the industrious pair turn to murder to keep the business going.

A red bracket on the right side of the results points to the text: "Movies that come under both Horror and comedy genres".

OR-CLAUSE IMPLEMENTATION IN PHASE-II



This screenshot shows the full search results page for the query 'Horror Comedy'. The results are displayed as a list of Wikipedia entries, each with a title, a URL, and a brief description. A red arrow points from the 'OR CLAUSE' dropdown in the top screenshot to the search input field in this screenshot. A large red bracket on the right side of the results list points to the text 'Movies that come under either of Horror or comedy genres'.

About 382 results (0.11728787422180176 seconds)

[https://en.wikipedia.org/wiki/Cat_People_\(1942_film\)](https://en.wikipedia.org/wiki/Cat_People_(1942_film))
Cat People
At the Central Park Zoo in New York City, Serbian-born fashion designer Irena Dubrovna makes sketches of a black panther. She catches the attention of marine engineer Oliver Reed, who strikes up a conversation.

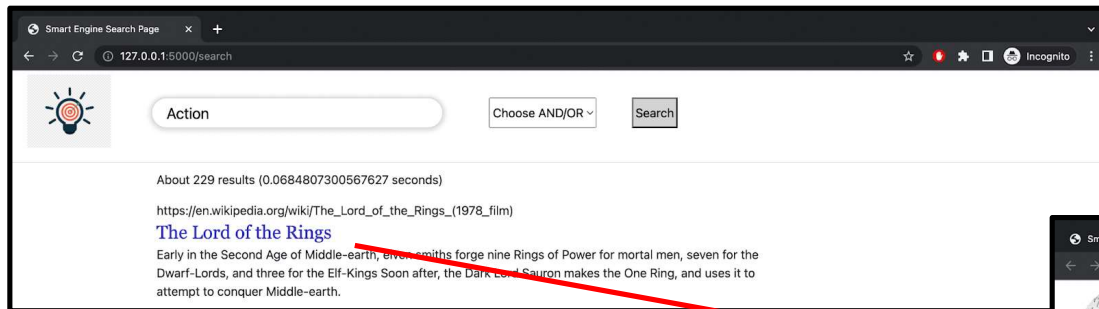
[https://en.wikipedia.org/wiki/Halloween_\(1978_film\)](https://en.wikipedia.org/wiki/Halloween_(1978_film))
Halloween
In 1963, on Halloween night in Haddonfield, Illinois, Michael Myers, dressed in a clown costume and mask, inexplicably stabs his older sister Judith to death with a kitchen knife in their home. He is subsequently hospitalized at Warren County's Smith's Grove Sanitarium.

https://en.wikipedia.org/wiki/Killer%27s_Moon
Killer's Moon
A coach full of schoolgirls breaks down in the Lake District, forcing the girls to take shelter for the night in a remote hotel. Meanwhile, strange and macabre things are happening to the locals (and their pets) and it is revealed that four escaped mental patients- Mr.

https://en.wikipedia.org/wiki/House_of_Horrors
House of Horrors
Struggling sculptor, Marcel De Lange (Martin Kosleck) is depressed about the events going on in his life, and decides to commit suicide. But just as he's about to kill himself, he spots and saves a madman, named "The Creeper" (Rondo Hatton) from drowning.

Movies that come under either of Horror or comedy genres

HYPERLINK & DESCRIPTION OF URL IMPLEMETATION IN PHASE-II

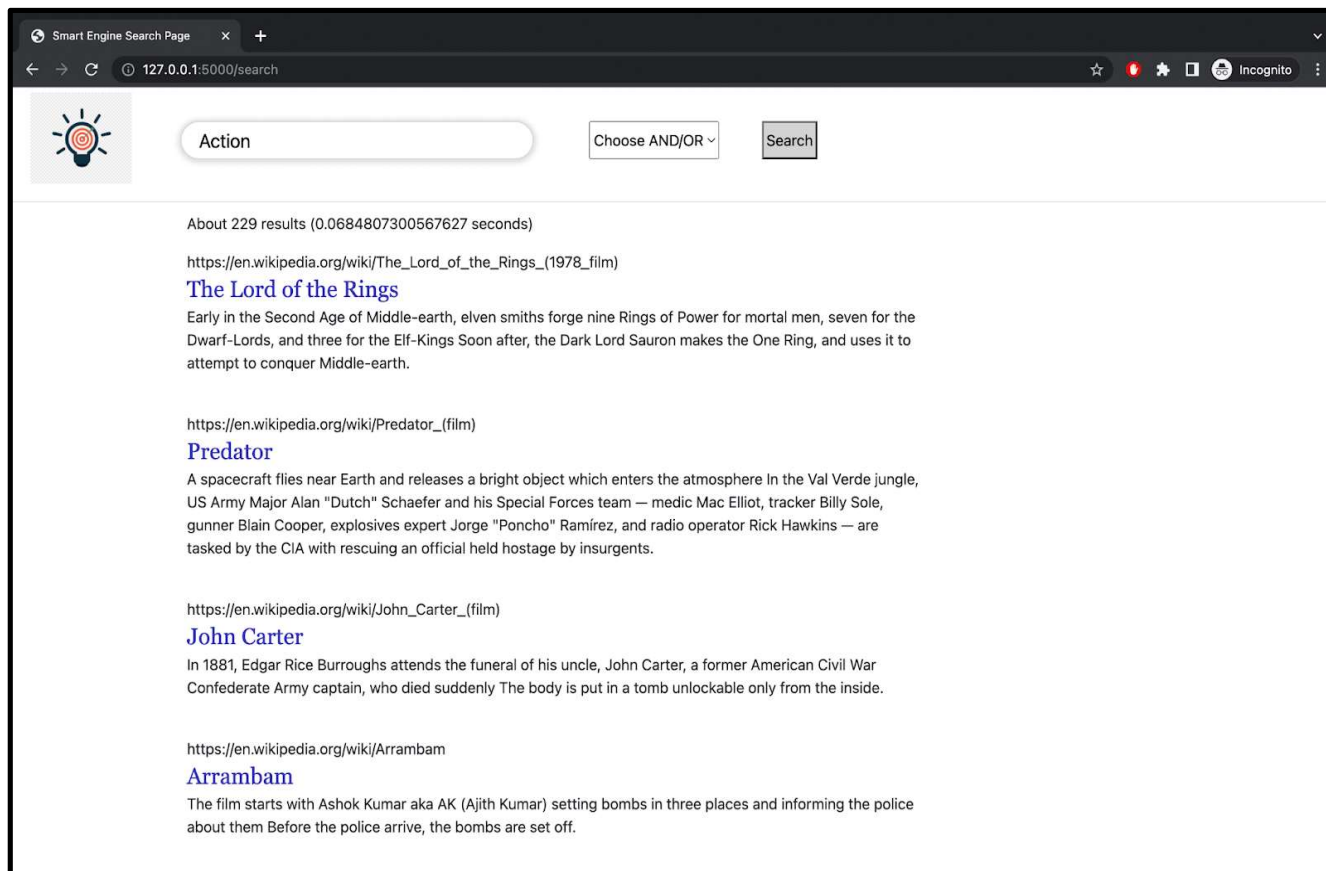


User has an option to navigate
to the URL by clicking on it



Two line
description

PAGE-RANKING ALGORITHM IMPLEMENTATION IN PHASE-II



**How do we conclude
that Lord of Rings is
most relevant result
for the query
Action?????**

PAGE-RANKING ALGORITHM IMPLEMENTATION IN PHASE-II

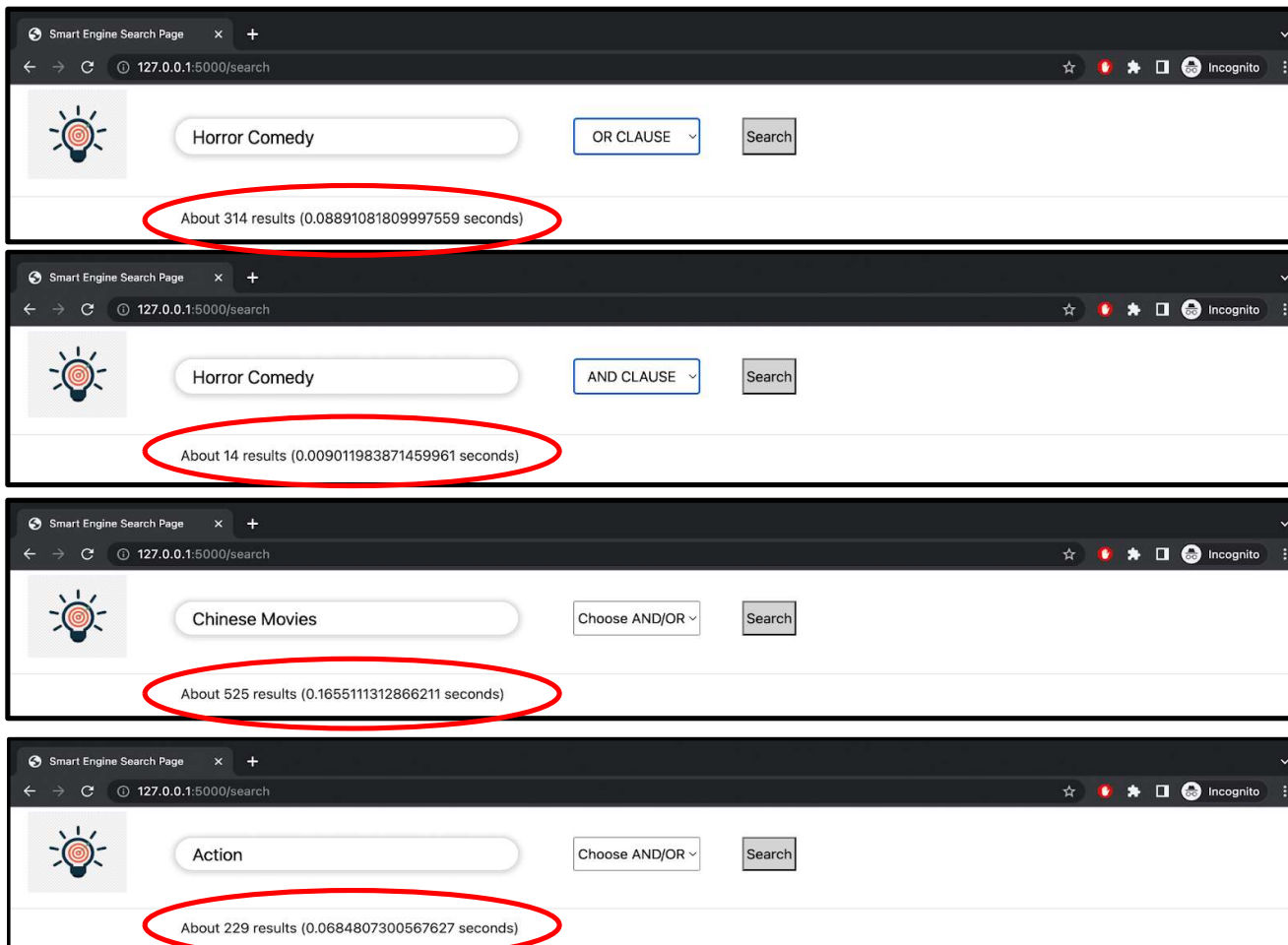
Single Phrase Query

```
Action
[('https://en.wikipedia.org/wiki/The_Lord_of_the_Rings_(1978_film)', [2.7733236106612713, 'The Lord of the Rings', 'Early in the
Second Age of Middle-earth, elven smiths forge nine Rings of Power for mortal men, seven for the Dwarf-Lords, and three for the E
lf-Kings Soon after, the Dark Lord Sauron makes the One Ring, and uses it to attempt to conquer Middle-earth']), ('https://en.wik
ipedia.org/wiki/Predator_(film)', [2.6189402188811672, 'Predator', 'A spacecraft flies near Earth and releases a bright object wh
ich enters the atmosphere In the Val Verde jungle, US Army Major Alan "Dutch" Schaefer and his Special Forces team – medic Mac El
liot, tracker Billy Sole, gunner Blain Cooper, explosives expert Jorge "Roncho" Ramirez, and radio operator Rick Hawkins – are ta
sked by the CIA with rescuing an official held hostage by insurgents']], ('https://en.wikipedia.org/wiki/John_Carter_(film)', [2.
230913584552052, 'John Carter', 'In 1881, Edgar Rice Burroughs attends the funeral of his uncle, John Carter, a former American C
ivil War Confederate Army captain, who died suddenly The body is put in a tomb unlockable only from the inside']], ('https://en.w
ikipedia.org/wiki/Arrambam', [2.230913584552052, 'Arrambam', 'The film starts with Ashok Kumar aka AK (Ajith Kumar) setting bombs
in three places and informing the police about them Before the police arrive, the bombs are set off']], ('https://en.wikipedia.o
rg/wiki/Aparichit_(2006_film)', [2.230913584552052, 'Aparichit', 'Ramanujam Iyengar alias Ambi is an orthodox Brahmin and a strai
ghtforward consumer protection advocate living in Triplicane, Chennai He expects everyone to follow the law and prosecutes those
who violate it']], ('https://en.wikipedia.org/wiki/The_Killer_(1989_film)', [2.1141977930880094, 'The Killer', 'Hitman Ah Jong is
on his last job for the Hong Kong Triad, but accidentally damages the eyes of a young nightclub singer named Jennie with a muzzl
e flash in a shootout After the attack, Ah Jong begins to watch Jennie perform at the nightclub and escorts her home when she is
attacked by thugs']], ('https://en.wikipedia.org/wiki/Rowdy_(2014_film)', [2.1141977930880094, 'Rowdy', 'The film opens in a stre
et in Hyderabad in a night with Krishna (Vishnu Manchu) saving Sirisha (Shanvi Srivastava) from some eve teasers and accompanying
her to the hostel she stays In the journey, they both come to know that they hail from Rayalaseema region of Andhra Pradesh']],
('https://en.wikipedia.org/wiki/Cardfight!!_Vanguard', [1.9713492091919835, 'Cardfight!! Vanguard The Movie', 'Season 1\r\nAichi
Sendou is a timid young boy in his third year of junior high school The one thing that keeps him going is his trading card Blaste
r Blade from Cardfight!! Vanguard, a trading card game that takes place on a different planet called "Cray" and is popular throug
hout the world']], ('https://en.wikipedia.org/wiki/Jin_Roh', [1.8858671496400232, 'Jin Roh', 'The story is set in a parallel 1950
s Japan, in which Germany has conquered Japan It focuses on Kazuki Fuse, a member of the elite Kerberos Panzer Cops, a metropolit
an antiterror unit']], ('https://en.wikipedia.org/wiki/Hana_Yori_Dango_Final', [1.8858671496400232, 'Hana yori Dango Final', 'Tsu
kasa Domyoji (Jun Matsumoto) and Tsukushi Makino (Mao Inoue) are driving through Nevada Weeks prior, Tsukasa announced to the wor
ld that he was engaged to Tsukushi']])
127.0.0.1 -- [24/Apr/2022 19:23:42] "POST /search HTTP/1.1" 200 -
127.0.0.1 -- [24/Apr/2022 19:23:42] "GET /static/css/search.css HTTP/1.1" 200 -
```

We can see that Lord of Rings has the highest page-rank and we can also see that the results are sorted in decreasing order of page-rank, i.e., results are displayed in decreasing order of relevancy (page-rank in our case).

Reference - Screenshot of the terminal

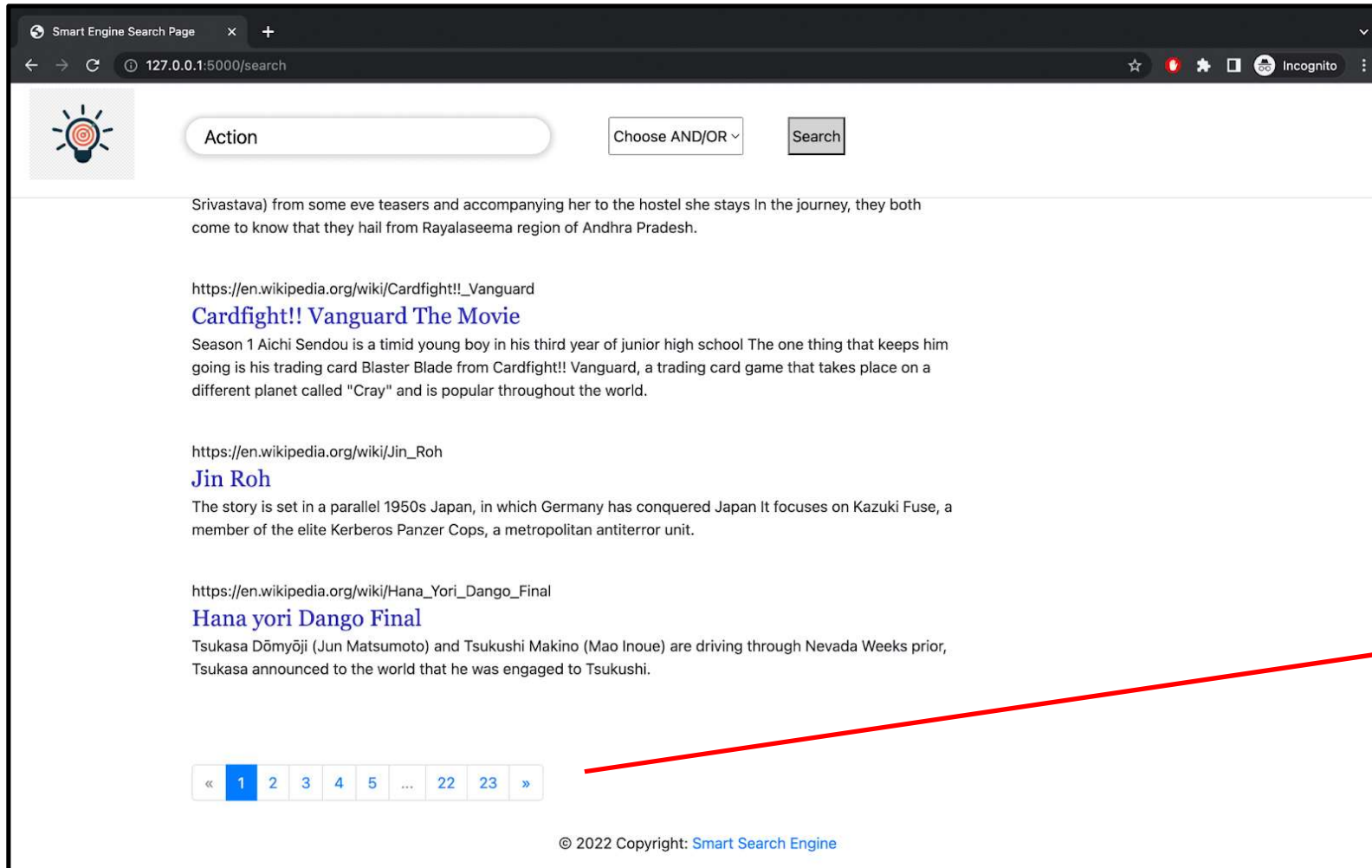
LOW LATENCY & HIGH THROUGHPUT IMPLEMENTATION IN PHASE-II



An additional functionality showing number of results as well as time taken to fetch those results is being displayed in Phase-II of the project.

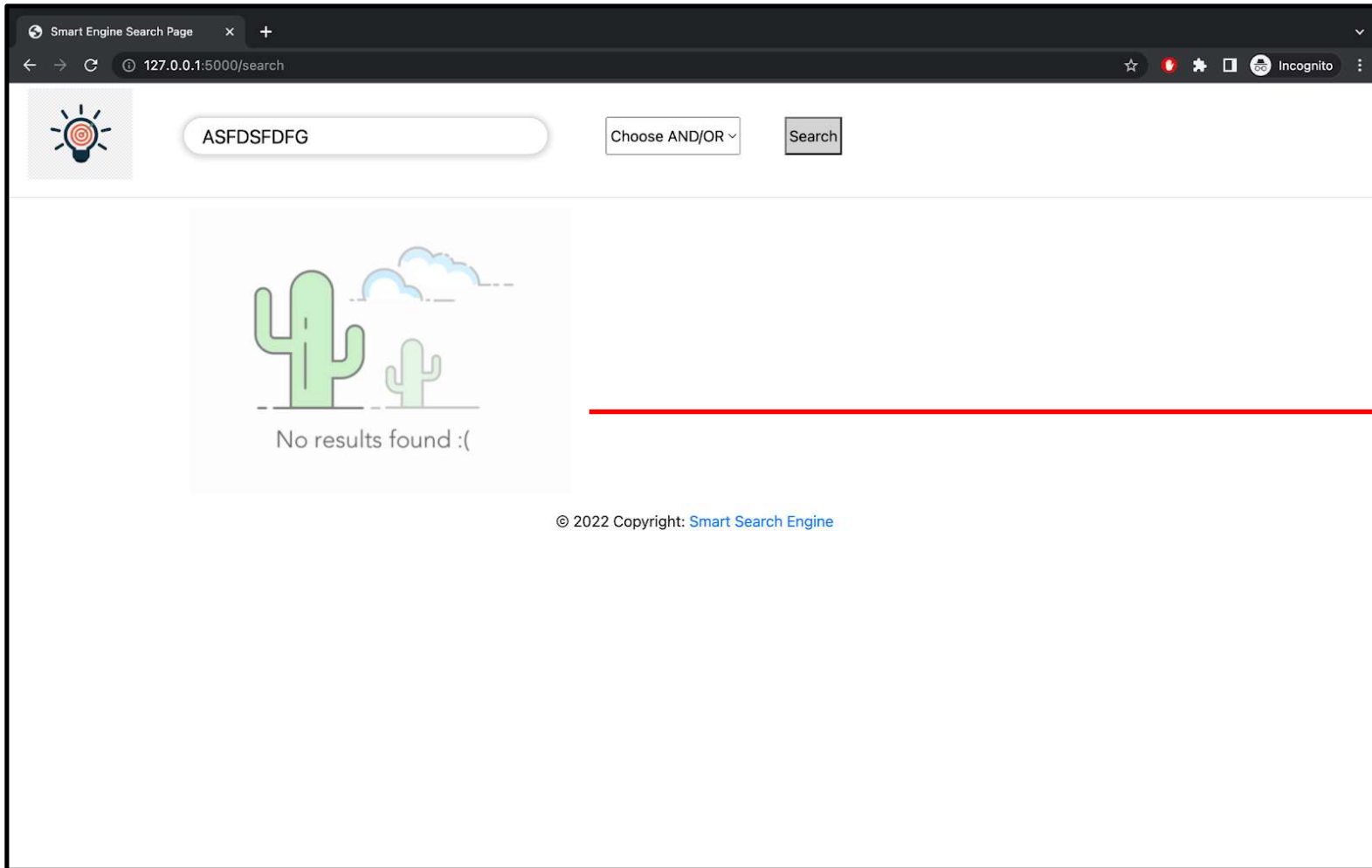
Note: We can observe the low latency(time taken) and the number of results being fetched after trying to find a match from 1000 URLs(high throughput)

PAGINATION IMPLEMENTATION IN PHASE-II



Pagination for better navigation of results (Maximum 10 results per page only)

ERROR-404 PAGE IMPLEMENTATION IN PHASE-II



An error 404 page has been implemented in Phase-II of the project which is displayed when user submits an invalid query or searches for something unrelated to the Movies.

USECASE TEMPLATE

Following are few of the most important usecases in the Smart Search Engine:

- **User** inputs the query
- **User** performs search
- **Query processor** receives the query

Note: Here the highlighted part in yellow are 'usecases' and highlighted part in blue are respective 'actors' to which those usecases are associated with.

Let's have a look at use-case templates for the above important usecases.

USECASE TEMPLATE-1

Usecase	Inputs the query
Description	User will give input in search box
Preconditions	Smart Search Engine Application is running
Postconditions	User will be displayed with autosuggestions
Actors Involved	User
Assumptions	Application has been locally hosted
Steps	1.User will give a query input 2.Auto Suggestion feature suggests the input query 3.User can either select a suggestion or manually type the whole query
Issues	User can search only for English words preferably related to Movies. (Since our search engine's specific domain is Movie URLs)

USECASE TEMPLATE-2

Usecase	Performs search
Description	User clicks enter or clicks on the submit button to perform search.
Preconditions	1.User has preferably given an input(English words) 2.User has selected AND/OR clause(optional)
Postconditions	User will be displayed with results or error-404 page based on the query
Actors Involved	User, Query Processor
Assumptions	1.User has given a set of English word(s) related to our movie database. 2.Locally hosted.
Steps	1.User will click on submit button or press enter after query search 2.Query Processor will process the query. 3.Query Processor will return the results. 4.User can view the results through the User Interface, i.e., browser.
Variations	User can specifically select AND/OR clause before searching multi-phrase query
Issues	User can search only for English words preferably related to Movies. (Since our search engine's specific domain is Movie URLs)

USECASE TEMPLATE-3

Usecase	Receive the query
Description	Query Processor processes the input given by user and returns search results taking into consideration AND/OR clause.
Actors Involved	User, Query Processor, FileSystem.
Preconditions	1.User has preferably given an input(English words) 2.User has selected AND/OR clause(optional) 3.Preprocessor has been run at least once and the database is being maintained in FileSystem.
Postconditions	1. Query processor renders either error-404 page or results page based on query 2.User will be displayed with results or error-404 page based on the query
Assumptions	1.Preprocessor has already processed the documents and created a database for query processor. 2.User has given english words. 3.Project has been hosted locally.
Steps	1.Receives the query from user. 2.After removal of stop words and stemming, it decides the type of query. 3.Depending on type of query, it performs either one_word_query operation or multi_phrase_query operation. 4.Returns results after successful processing. 5.User views the results through browser.
Variations	If user would have given AND/OR clause then query processor will fetch results based on that specifically in multi phrase query operation only.
Issues	1.User can search only for English words preferably related to Movies. (Since our search engine's specific domain is Movie URLs) 2.Some movie URL's (wiki links) might have expired.
Notes	To maintain this database, admin will/can preprocess regularly

UML DIAGRAMS

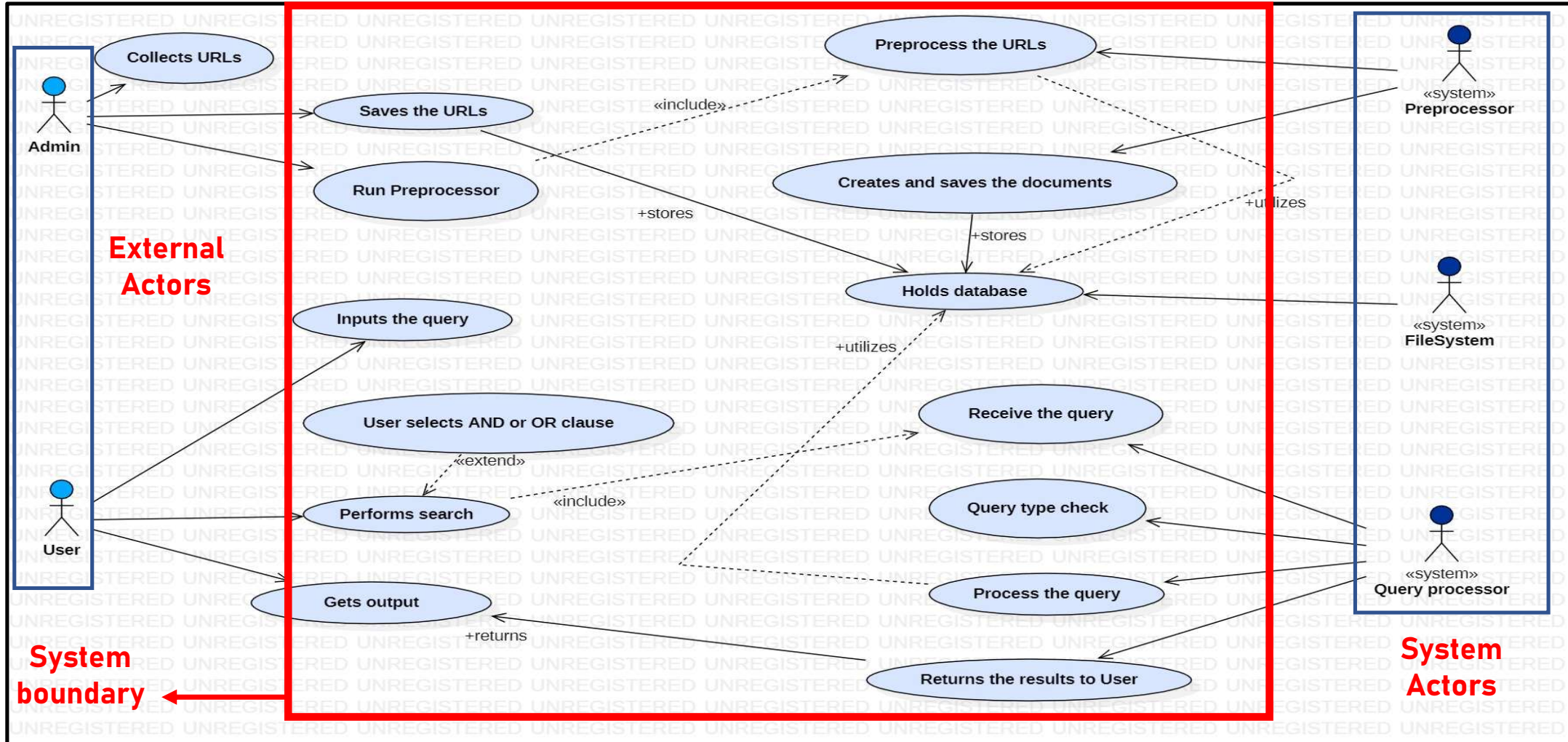
Following diagrams have been updated from Phase-I of the project:

- Usecase Diagram
- Class Diagram
- Sequence Diagram(s) (*Multiple scenarios shown*)

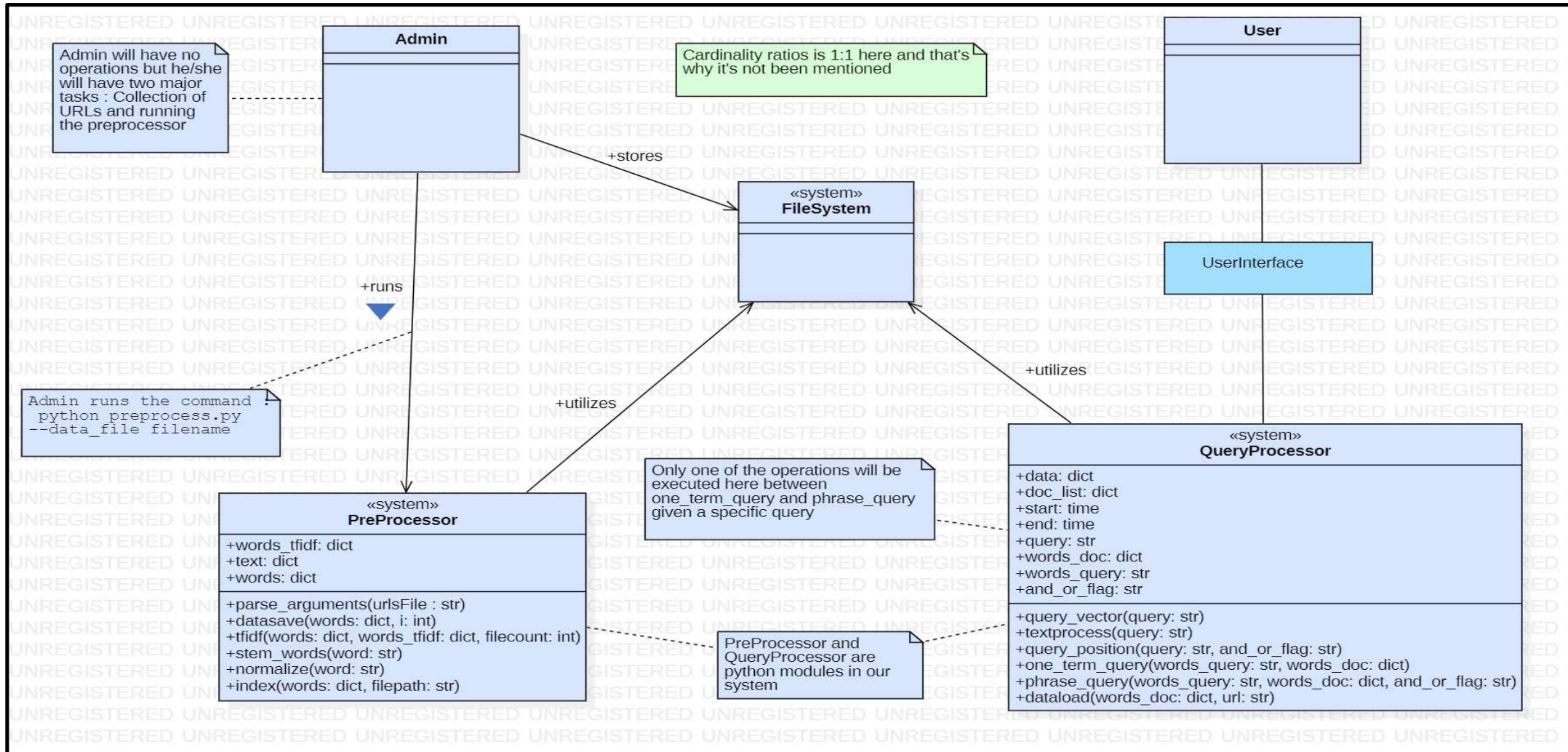
Following diagrams have been drawn in Phase-II of the project:

- State Transition Diagram(s)
- Activity Diagram(s)

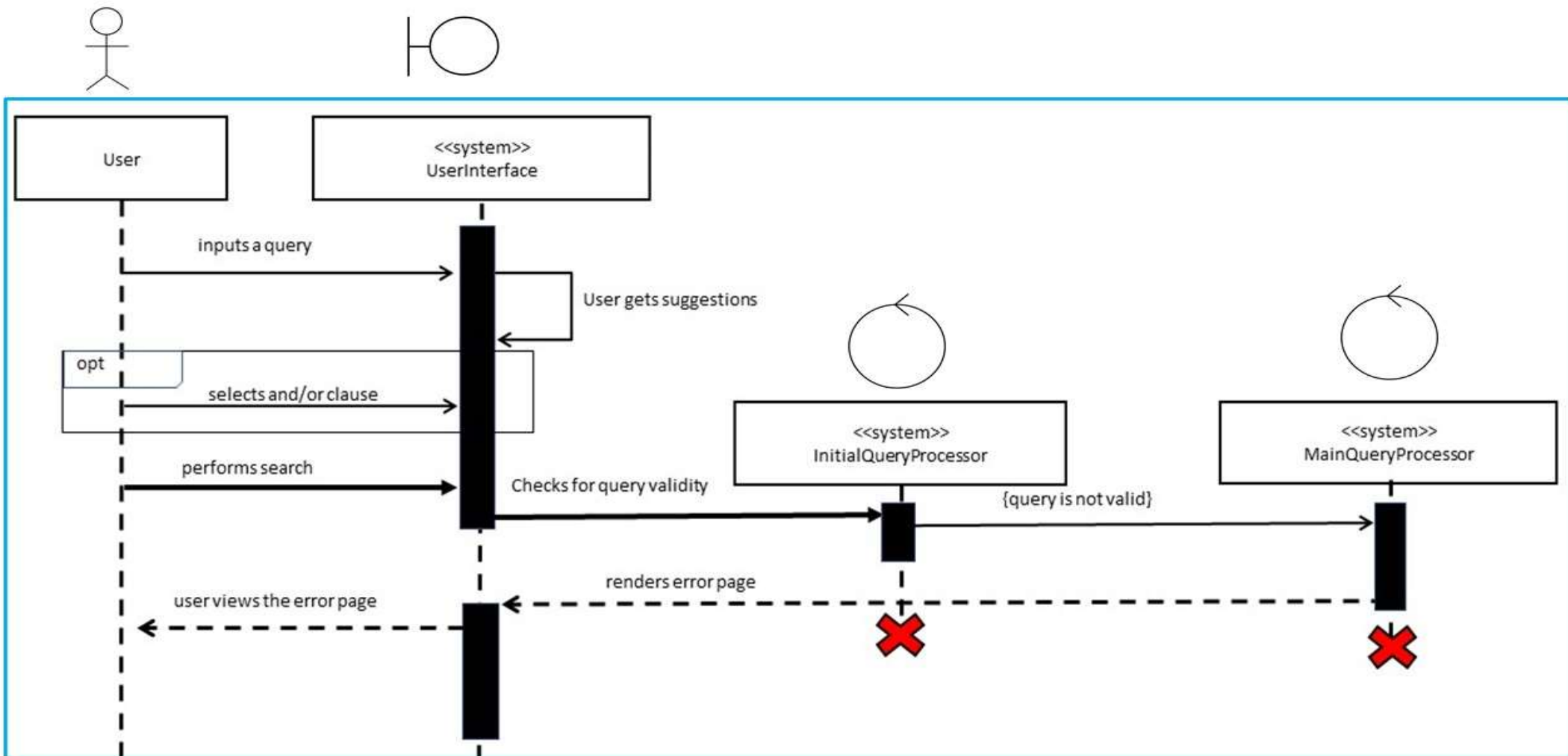
USECASE DIAGRAM



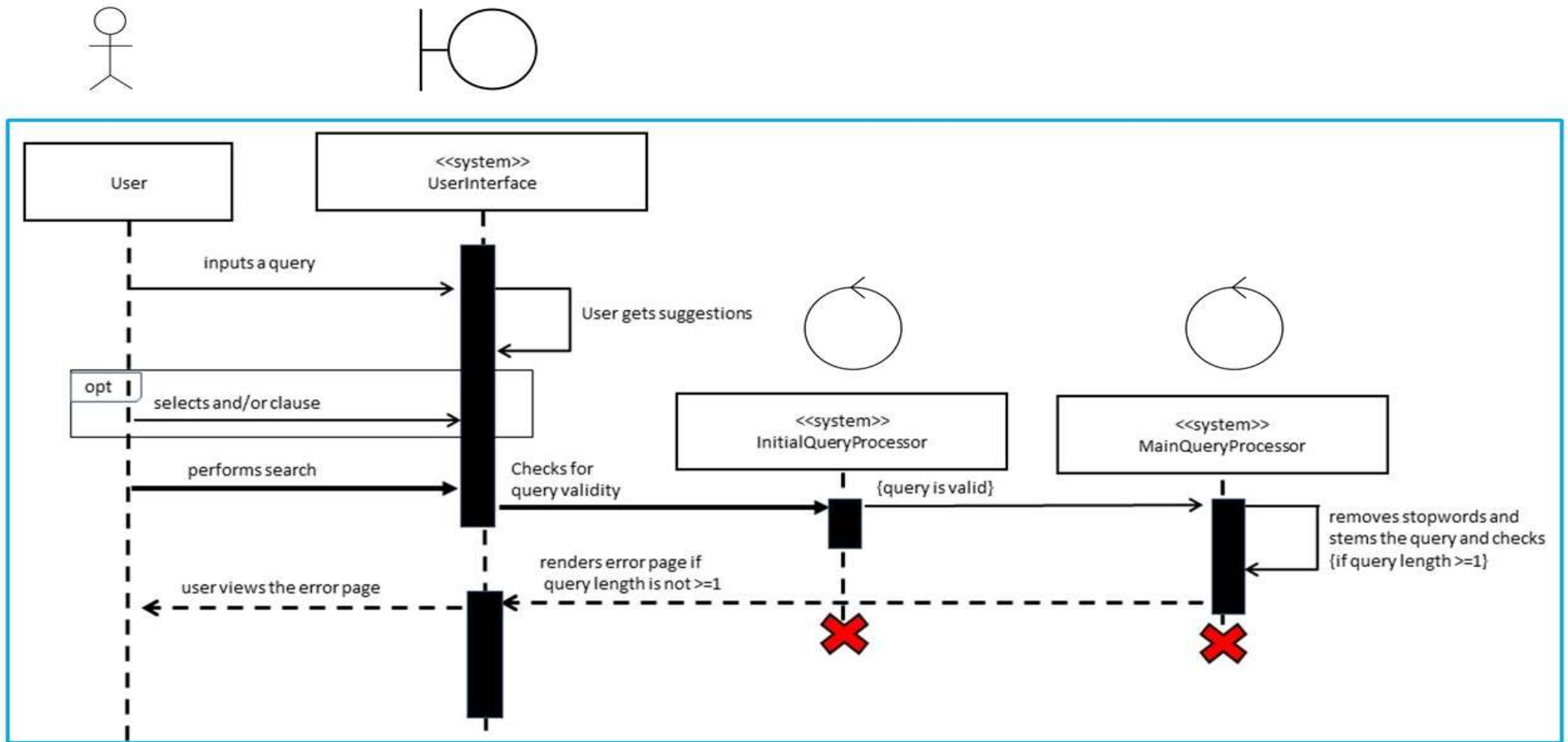
CLASS DIAGRAM



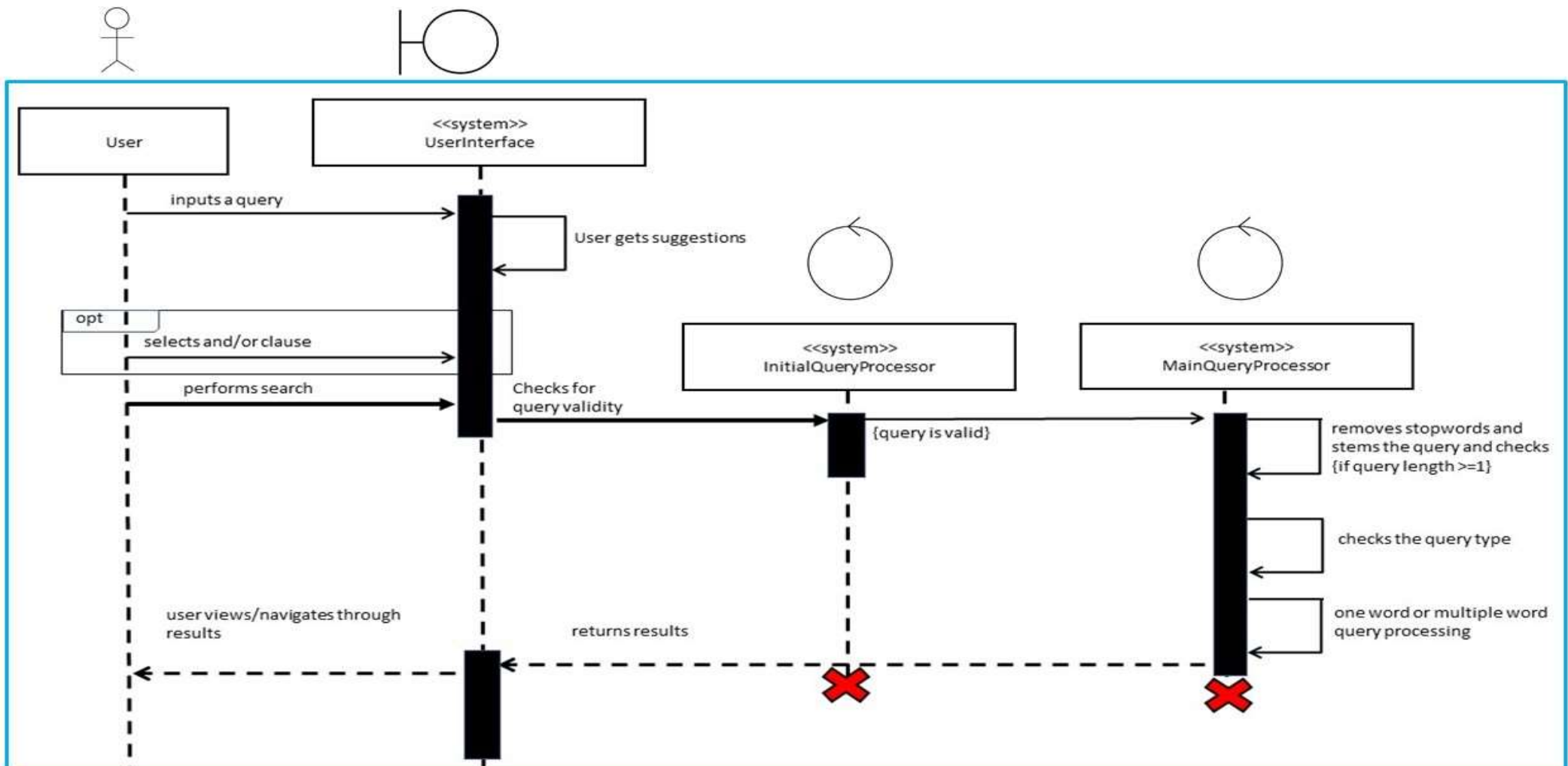
SEQUENCE DIAGRAM-1(INVALID INPUT)



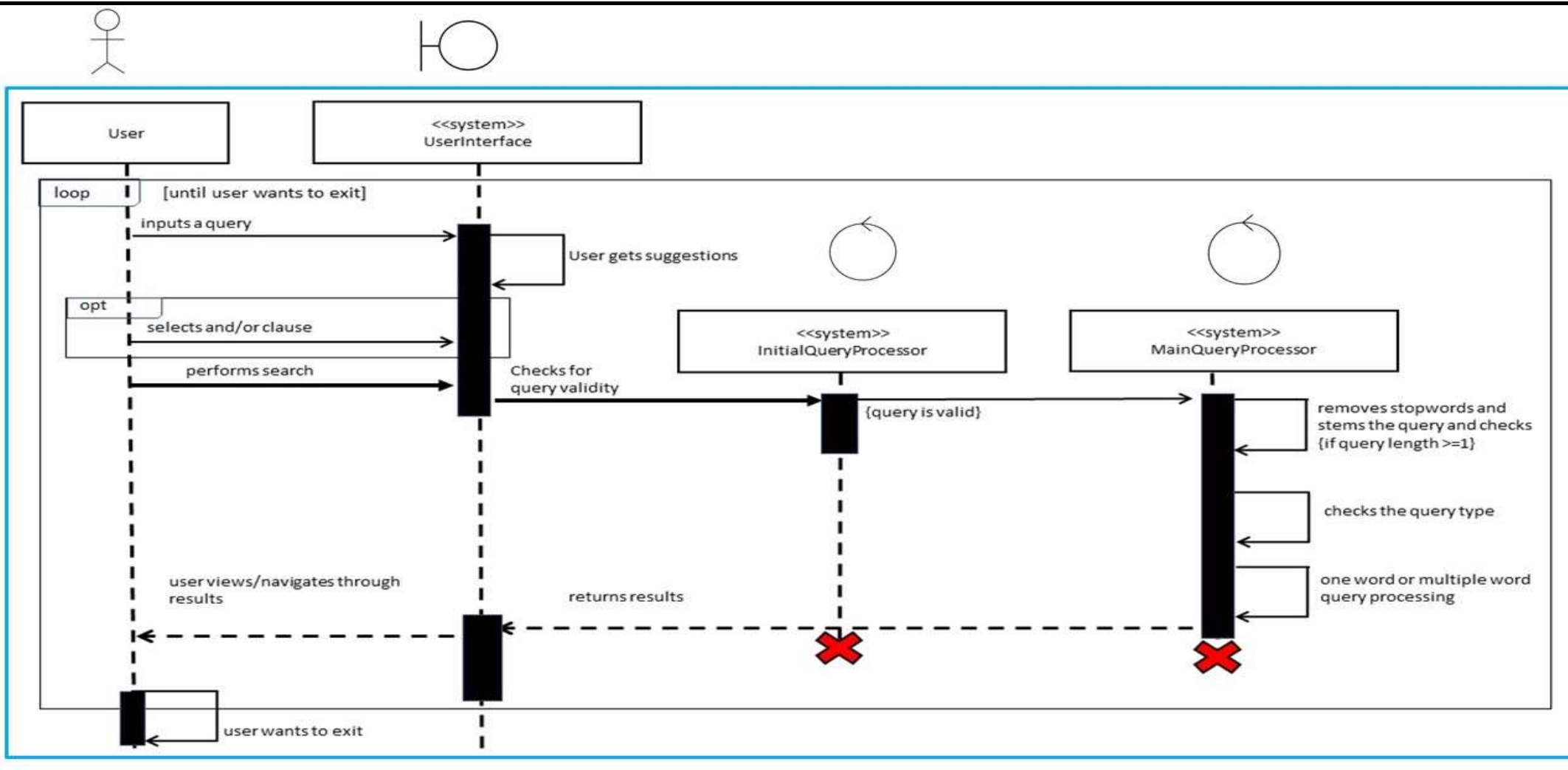
SEQUENCE DIAGRAM-2(IRRELEVANT INPUT)



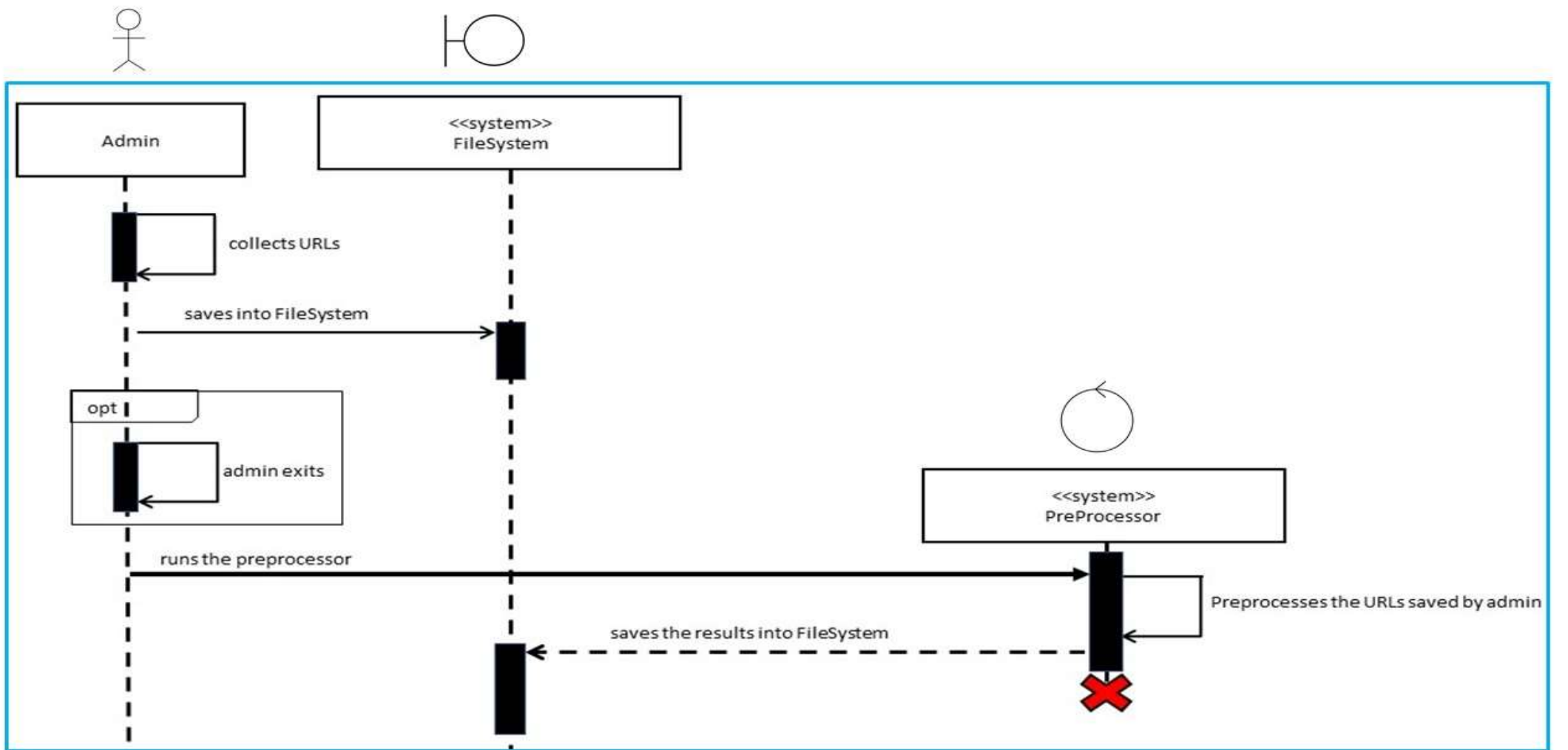
SEQUENCE DIAGRAM-3(SUCCESSFUL SEARCH)



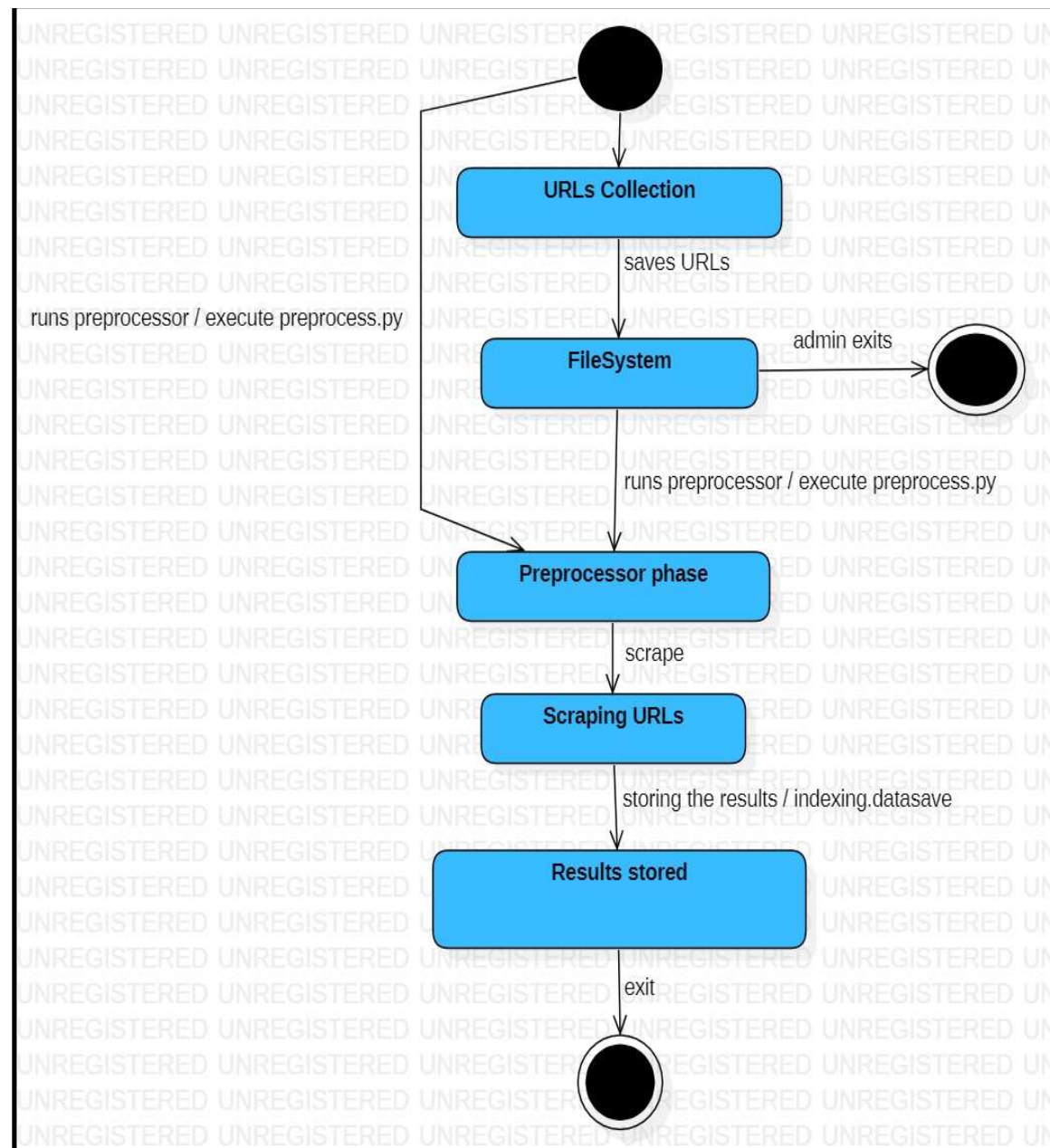
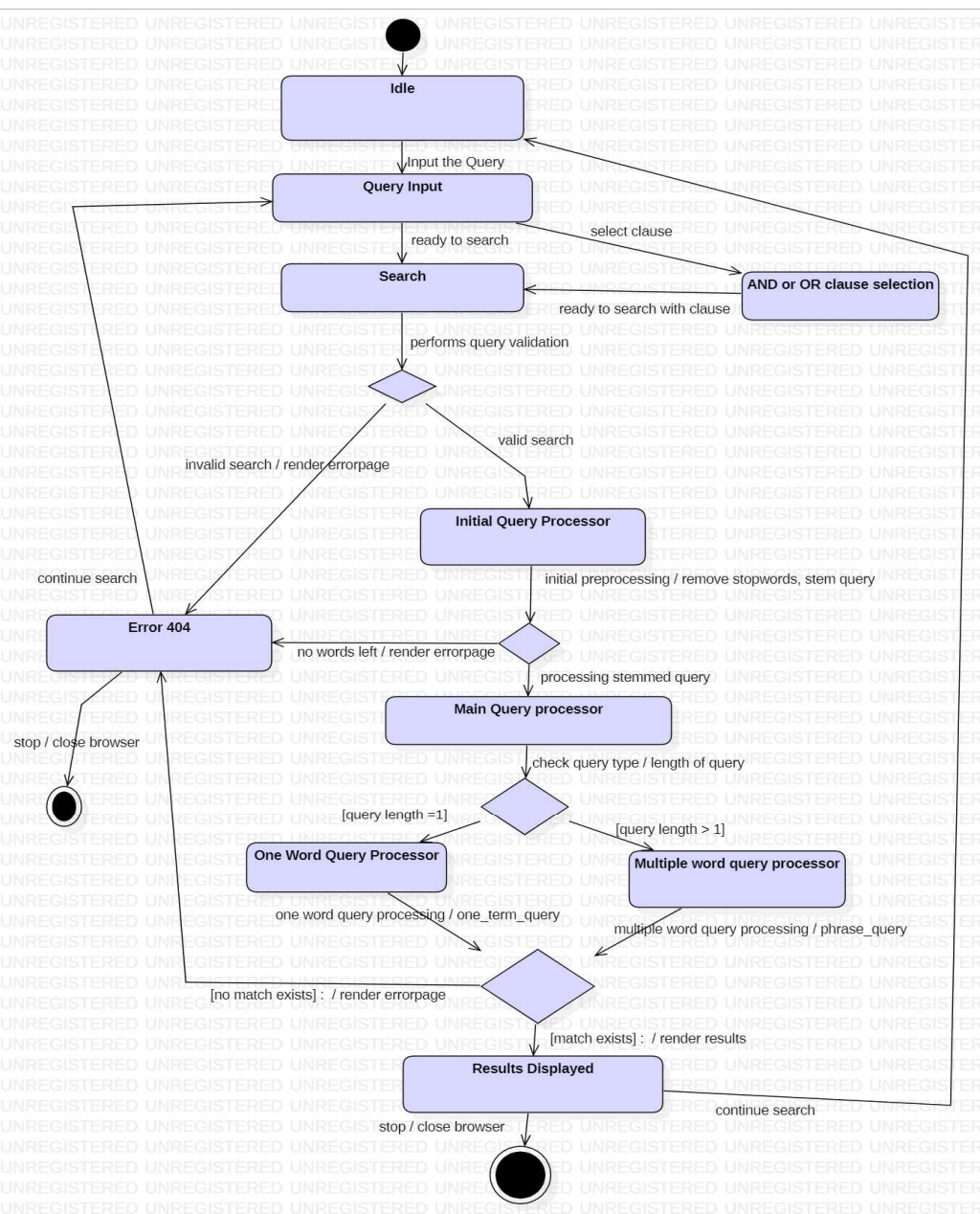
SEQUENCE DIAGRAM-4(LOOP FRAME ADDITION)



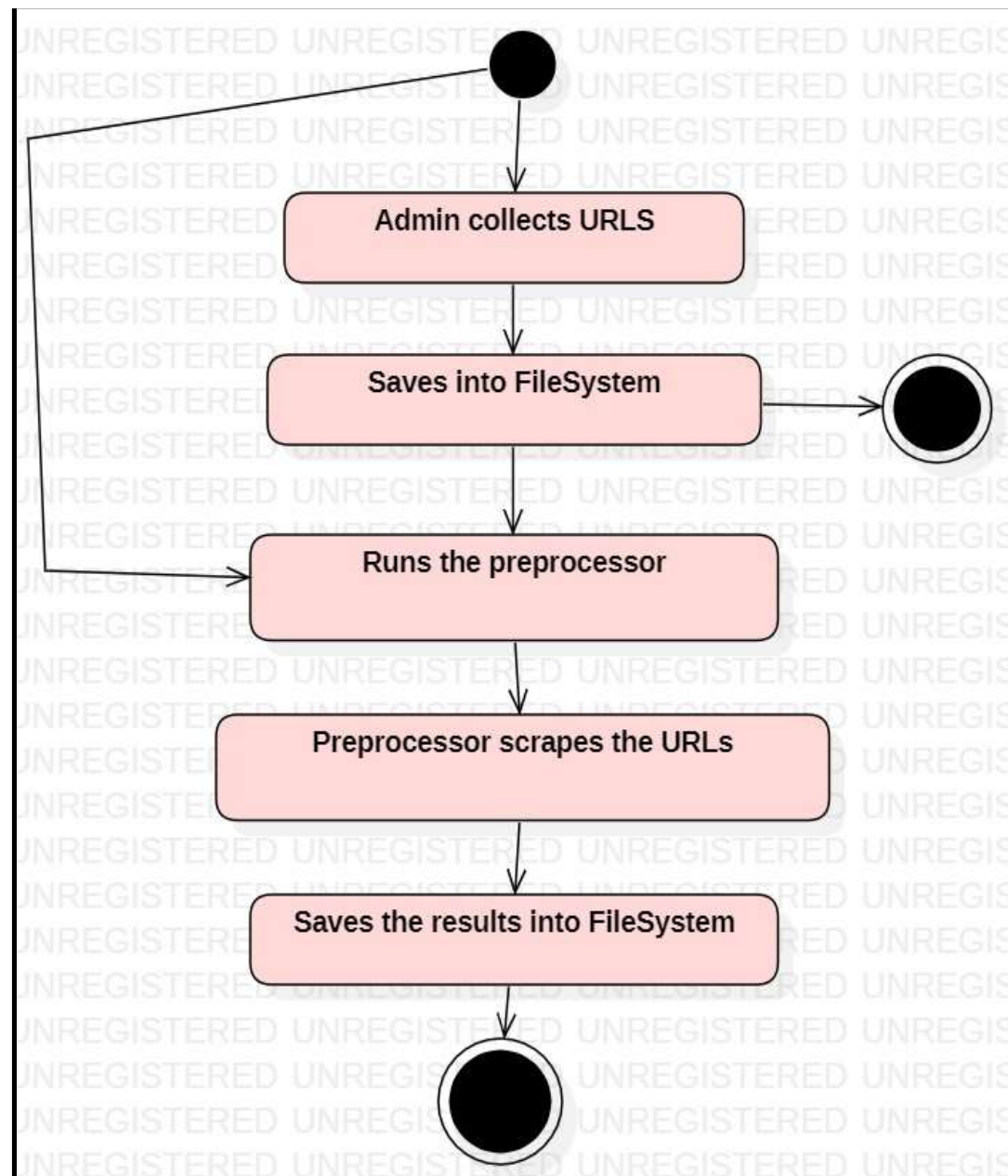
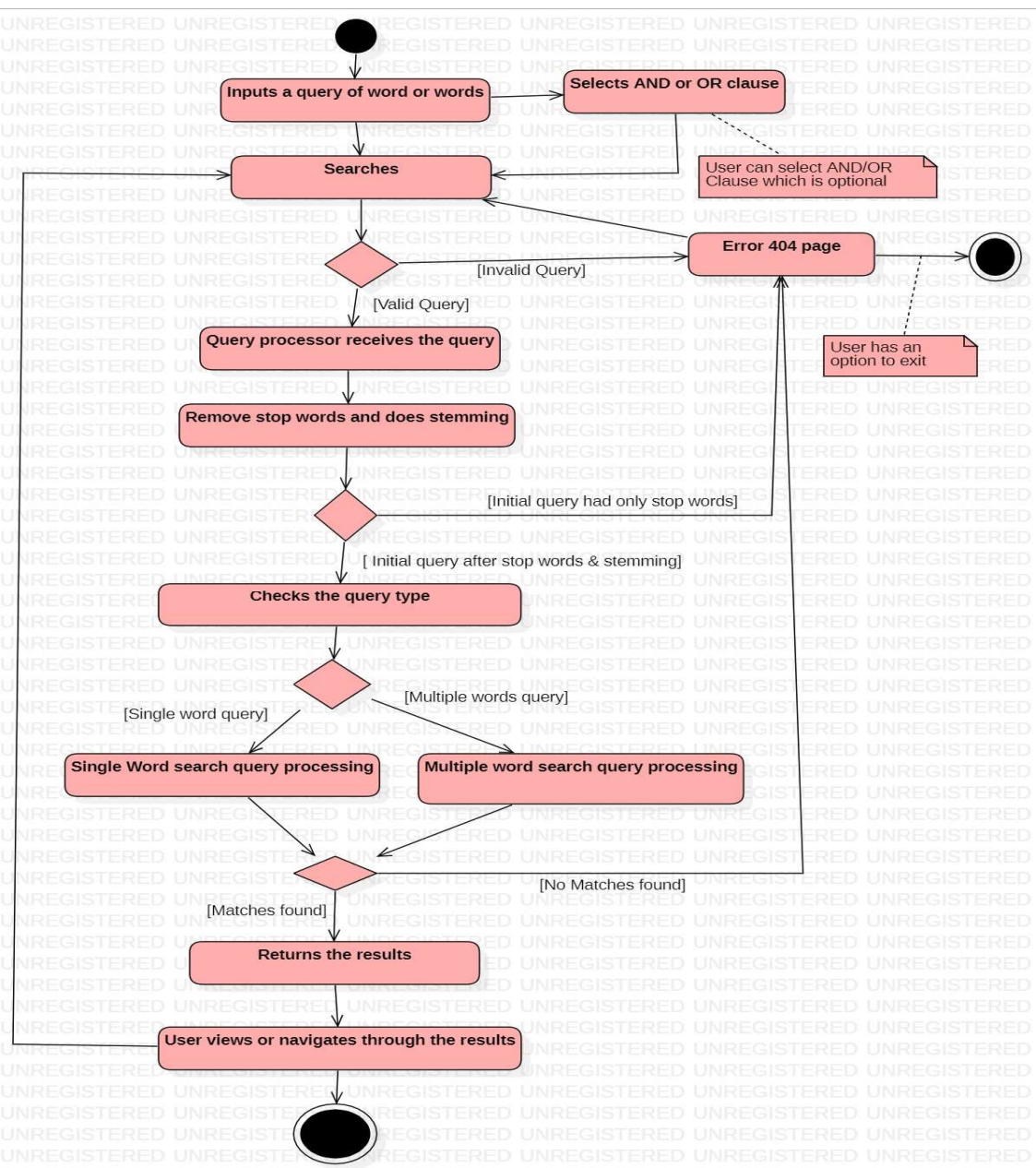
SEQUENCE DIAGRAM-5(ADMIN & PREPROCESSOR)



STATE TRANSITION DIAGRAM



ACTIVITY DIAGRAM



QUESTIONS ?

THANK YOU!