

Group members

Natnael Abebaw	UGR/7747/13
Nathnael Yimer	UGR/6855/13
Meron Abebe	UGR/9559/13
Hanna Legesse	UGR/4973/13
Robel Yemane	UGR/0067/13

News scraper

Objective:

Building a news search engine that searches news from different news feeds such as news streamer broadcasts (CNN, BBC others) uses concurrency and synchronisation techniques that different users can
Do search the news with better utilisation of the resources.

Procedures:

- Reading the news feeds from json data in the data folder. A data feed is an object with a broadcast url link, type, site name which uses the url link to make a network request concurrently.
- There is the server running on main.go which accepts the request from the server.
- On the server when a search request is made search.run() from the search module is called.

- This function taking the search query and make a search engine functionality by reading the news feeds from the Json data and make and apply different matchers for different types of feeders for example a CNN news with rss type will implement there rss matcher from the match module
- Each matcher has their own search implementation for rss matcher there is xml representation of the data that comes from the news feeder. Then we parse the xml and Json data from each implementation of the rss server.
- So whenever a search request is made the our app takes the search query and makes a synchronised concurrent search on each Network feeder url and the synchronisation makes it wait all the feeds send the search results to as and display to the screen.
- Our app has a caching capability that first checks the search query in the cache and then if it is not in the cache, searches the query in the news feeds. But if it in the cache takes the data from the cache without making a request in the network feeds this makes our app faster in some occasions.