horizontal line

**FrontPage Custom News**

II A3

Back-end

Plan de lucru

# DESCRIERE

Scopul echipei de back-end este de a **asigura colectarea eficienta a datelor** de pe anumite website-uri de interes (sportive, mondene, stiintifice, cu continut care starneste rasul - bancuri, informationale - stiri, vreme), date ce ulterior se vor concretiza in **articole**. **Colectarea datelor** va fi **automatizata** pe partea de server astfel: la un anumit interval de timp, se va face o cautare de articole noi cu ajutorul **parserelor** specifice fiecarui website.

Articolele vor fi stocate intr-o baza de date de tip **MongoDB**. Vom folosi acest tip de baza de date deoarece este **potrivita pentru un volum mare de date**, se poate **organiza** si **structura** intr-un mod facil (orientata pe documente **JSON**),este **rapida** si **usor utilizabila**.

**In final, ne asteptam ca serverul sa raspunda cat mai rapid la cererile clientului(browser-ului), astfel incat experienta utilizatorului(UX) sa fie cat mai placuta.**

**ECHIPA**

Echipa care se va ocupa de partea de back-end a proiectului este formata din:

* Andrei Iulian
* Bodnar Ana Maria
* Carp Paula
* Corneanu Cosmina
* Gherman Dan Gabriel
* Huzum George
* Ionesei Gabriel
* Maftei Ervin
* Popa Anca Teodora
* Popa Andrei Petrut
* Unghianu Anda

# OBIECTIVE SI ACTIVITATI

1. Impartirea sarcinilor
   1. Stabilirea pasilor necesari
   2. Repartizarea sarcinilor in functie de dorinta si experienta fiecarui membru
2. Gasirea site-urilor de unde vor fi preluate datele
   1. Stabilirea continutului ziarului (a categoriilor de articole posibile)
   2. Hotararea asupra unui site sursa pentru fiecare categorie in parte
3. Analizarea si crearea parserelor pentru articole in functie de site-ul de provenienta
   1. Plecand de la sursa HTML a unei pagini web, se preiau articolele necesare
   2. Articolele alese sunt inserate in baza de date in colectia corespunzatoare lor
4. Crearea serverului
   1. Stabilirea conexiunii intre aplicatia scrisa si baza de date
   2. Scrierea codului in care se primesc si se rezolva cererile de la client
5. Comunicarea in mod constant cu celelalte module
   1. Asigurarea legaturii intre front-end si back-end
   2. Discutarea impreuna cu cei de la front-end despre request-urile necesare clientului pentru a afisa datele in mod corespunzator

**TEHNOLOGIA UTILIZATA**

Pentru realizarea aplicatiei web, vom folosi framework-ul Spring din Java.

Pe partea de server, vom avea mai multe pachete pentru a imparti codul cat mai eficient. Unul dintre pachete va fi cel ce contine cate un parser pentru fiecare site de unde luam articolele din ziar. Actualizarea datelor se va face constant, la un interval de cateva minute, folosind java,util.Timer, care va crea la inceput cate un thread in background pentru fiecare parser.

Pe langa acest pachet, vom avea si pachetele controllers, models, repositories, services. In controllere va fi implementat REST API. Cererile HTTP de la client vor fi preluate, si dupa ce se fac cautarile necesare in baza de date si apoi prelucrate informatiile, se va trimite inapoi un raspuns.

# SPECIFICATII

## Baza de date

Baza de date utilizata va fi de tip MongoDB. Vom avea mai multe colectii cu urmatoarele campuri:

* Administrator: ’\_id’, 'username' <string>, ‘password’<string>, 'email' <string> (colectia ‘admins’)
* Utilizator: ‘\_id’, ‘username’ <string>, ‘password’ <string>, ‘email’ <string>, ‘preferred\_tags’ <list(string)>, ‘preferred\_sources’ <list(string)> (colectia ‘users’)
* Articol: ‘\_id’ <string> - titlul, ‘content’ <string>, ‘imageUrl’ <string>, 'tags' ["string"..], 'source' { ‘site’ <string>, ‘author’ <string>, ‘date’ <Date>}, ‘videoUrl’ <string> - optional (colectia ‘users’)

Administratorul are drept de editare,adaugare,modificare a articolelor din baza de date prin intermediul interfatei, pentru a evita lucrul la linia de comanda.

Articolele vor fi niste obiecte ce urmeaza structura de mai sus, continand o lista de taguri prin care front-end-ul va putea afisa utilizatorului articole in functie de preferintele personale.

In momentul conectarii la site se vor alege prin intermediul unor checkbox-uri, de exemplu, niste categorii de stiri, iar aceste categorii vor duce la alegerea articolelor ce au in lista de tag-uri categoriile respective.

Se vor stoca si informatiile din articolele ce contin imagini sau audio(stiri de exemplu), pentru a putea fi redate si prin intermediul site-ului nostru. De asemenea se vor pastra si informatii despre sursa articolului parsat in obiectul sursa ce contine site-ul sursa, autorul si data la care a fost postat articolul.