

Technical University of Cluj-Napoca Faculty of Automation and Computer Science Computer Science Department

Financial Efficiency Boosting System For Consumer-Grade Products

Student **Péter-Tibor ZAVACZKI**

Advisor

Assoc. Prof. Dr. Eng. Delia Alexandrina MITREA

A The problem

- People focus on efficiency, especially financially
- The popularization of the internet has brought along ecommerce
- Being able to make money online presented an opportunity to companies
- The ecommerce market is flooded with web shops

The solution

An application needs to be created that centralizes the options across the ecommerce market

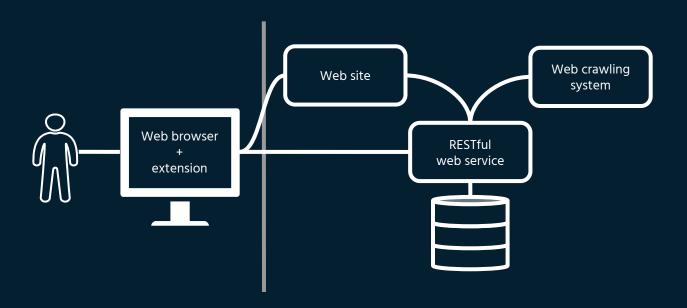
The application needs to have the qualities:

- Very accessible
- Easy to use
- Can track prices of multiple web sites
- Easily extensible

"The more we can organize, find and manage information, the more effectively we can function in our modern world." Vint Cerf - co-inventor of TCP/IP

The plan

Create a system with a modular architecture

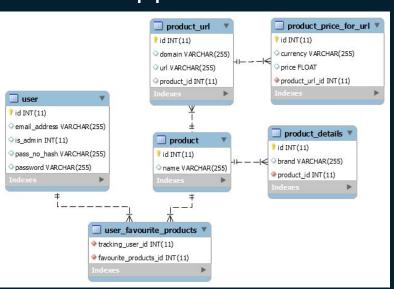


∃ The database

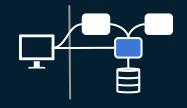


MySQL was chosen as the database management system for this application

due to the fact that I need to store the product data in a structured manner



* The RESTful web service



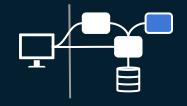
The RESTful web service is created using the Spring framework for Java 8

It acts as a uniform interface connecting the modules of the application

It is split into a DAL and a BLL



≭ The crawling system



The crawling system was build using Python3 and Scrapy

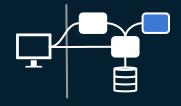
It is the key element of this application

The development of this module consisted of:

- Creating a script that starts the crawling process based on a simple command
- Creating the necessary spiders
- Creating an Item pipeline section to persist the extracted data

The crawling system

* The crawler caller



The crawler caller is the script I've created for starting the process

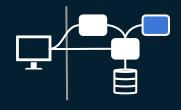
It works as follows:

- > It requests the URLs of products belonging to the given domain
- > It sets the received URLs as starting URLs for the spider
- It starts a crawling process with the previously mentioned URLs and the spider that has the name that has been given to it

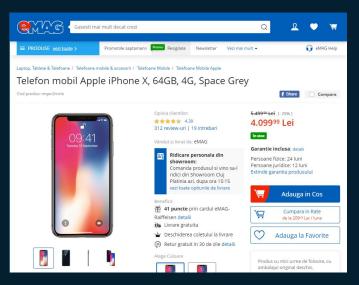


The crawling system

* The spiders



Spiders are scripts for Scrapy, which extract data from a fetched web page





{'brand': 'Apple', 'crawled_name': 'Telefon mobil Apple iPhone X, 64GB, 4G, Silver', 'currency': 'RON', 'domain': 'emag', 'price': 4099.99, 'productId': 1} The crawling system

* The Item pipeline



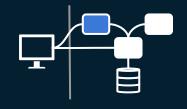
The Item pipeline is the post-processing element for extracted data

It is composed of multiple small-scale scripts, which perform some action with the data they receive

Its components are activated and sorted in the project's settings



The web site



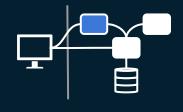
The web site of the application was built with the AngularJS framework

It represents the largest part of the application's front-end

It can act as a kind of centralized web shop interface

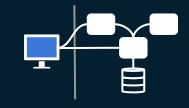
The web site

W User permissions



Operation	Guest	User	Admin
Login			
Register			
Search for a product			
See the details of a product			
Logout			
Add a product to favourites			
Remove a product from favs			
View their favourites list			
View the full list of products			
Add a new product			
Update a product			
View the list of a product's URLs			
Add a new URL			
Update a URL			

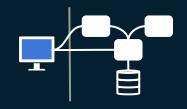




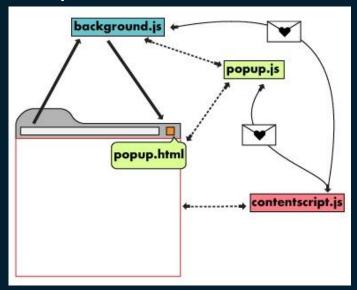
The Google Chrome browser extension is the most important element of the application's front-end

It is the 'one-click-away' accessible solution for the presented problem

9 Under the hood

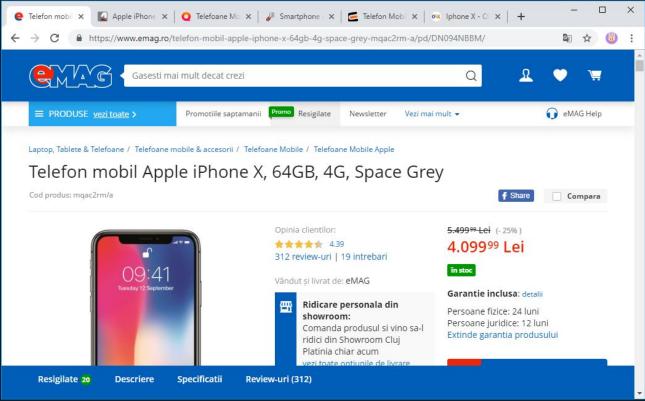


A Chrome extension has a simplistic, but powerful composition

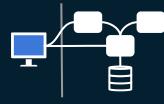


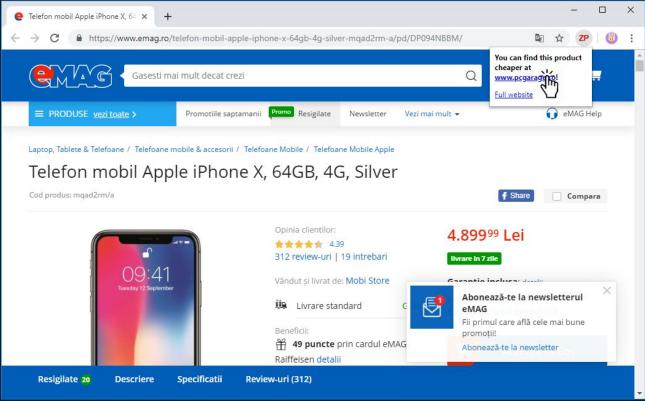




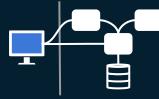


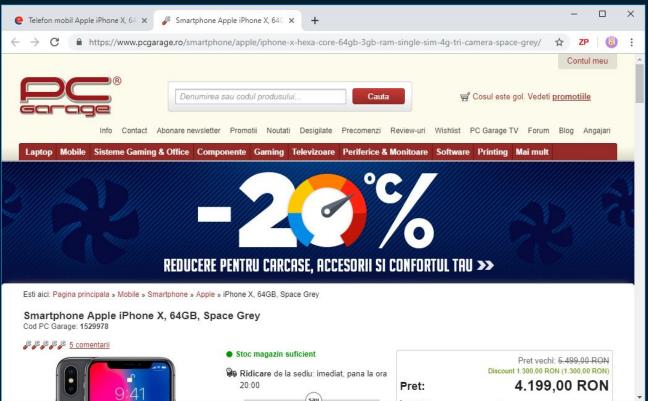














- Modular, extensible application
- Up to date database of products
- Easy to use interface
- One-click-away tool for saving money

≺ Future development

- ## Support more domains & products
- C Automate the crawling process
- Port the extension to other browsers
- Implement stronger security measures



THANK YOU FOR YOUR ATTENTION!