



**PAISII HILENDARSKI UNIVERSITY  
OF PLOVDIV**



---

FACULTY OF MATHEMATICS AND INFORMATICS  
DEPARTMENT OF COMPUTER TECHNOLOGY

**STANISLAV MINCHEV DAKOV**

**TOOLS FOR IMPROVED USER INTERACTION IN  
ELECTRONIC COMMERCE**

**ABSTRACT**

of a dissertation  
for obtaining the educational and scientific degree "PhD"

Area of higher education: 4. Natural sciences, mathematics and informatics; professional direction: 4.6. Informatics and computer science; doctoral program: Informatics

**Scientific Supervisor:** Associate Professor Veselin Kyurkchiev, PhD

**Plovdiv, 2026**

The dissertation work has been discussed and directed for defense to the departmental council of 12.05.2026 from the Department of Computer Technologies at the Faculty of Mathematics and Informatics at Paisii Hilendarski University of Plovdiv.

The dissertation has a total volume of 179 pages. Contains 62 figures and 4 tables. It consists of an introduction, three chapters, a conclusion, a list of the author's publications on the dissertation work, noted citations, approbation, and references. Each of the chapters ends with a conclusion summarizing the findings. The literature used includes 204 sources in English and 2 in Bulgarian.

The list of the author's publications consists of 6 titles

The defense of the dissertation will take place on 30.06.2026 at 14:00 p.m. in the Meeting Hall of Paisii Hilendarski University of Plovdiv - New building of Paisii Hilendarski University of Plovdiv.

The materials for the defense are available to those interested in the dean's office of FMI, the New Building of PU, room 330, every working day from 8:30 AM to 5:00 PM

**Author: Stanislav Minchev Dakov**

**Title: Tools for Improved User Interaction in E-Commerce**

# **I. GENERAL CHARACTERISTICS OF THE DISSERTATION**

## **1. Relevance of the dissertation**

Nowadays, the Internet is a fast and convenient way of distributing a variety of up-to-date information, allowing to overcome time, space, and language barriers. E-commerce is implemented on its basis. Its essence is the purchase and sale of goods and services. Users can compare prices and make informed product choices. Different and convenient payment methods are available. All this saves both customers and businesses time and effort. Therefore, it is necessary for e-commerce sites to be attractive, offer favorable conditions for purchase and guarantee security.

According to research [Dakov, 2024b] in Bulgaria, about 50% of people living in large settlements prefer to shop online. About 67% buy clothes, shoes and accessories, followed by (13%) electronics, souvenirs and food. The presented results confirm data from the National Statistical Institute (NSI).

The desire to establish and develop any business in the modern highly competitive digital environment of e-commerce is closely related to effective and sustainable interaction with customers, based on security and mutual trust. Convenient and attractive user interfaces, personalized product recommendations, convenient and responsive customer support, as well as secure payment methods have a significant impact on user activity and loyalty. All this creates the need for continuous development and implementation of tools that enable more pleasant and effective user interaction [Varbanov, 2008], [Shishmanov, 2020].

That is why the creation of attractive and convenient e-commerce applications, through which interaction with customers is carried out, is particularly relevant. Technologies and relevant tools are being developed through which products and payment methods are offered, and the wishes and opinions of users are monitored.

## **2. Purpose and objectives of the dissertation**

*The main objective* of this dissertation is: Researching possibilities and developing prototypes of software tools for improving user interaction in e-commerce.

To achieve this objective, the following tasks must be solved:

1. To analyze existing tools and the main factors influencing the interaction between users and businesses in e-commerce.
2. To develop a model of innovative opportunities for improving user interaction in e-commerce.
3. To develop a prototype of an online platform with a set of tools satisfying the improvement of user interaction in e-commerce.

In the context of this dissertation, the term "e-commerce" refers to websites and online platforms for selling standard consumer goods and advertisements for such goods (e.g. clothing, electronics, cosmetics, household goods), without covering specialized platforms for trading financial instruments, precious metals or cryptocurrencies.

## **II. STRUCTURE OF THE DISSERTATION THESIS**

This dissertation is developed in a volume of 179 pages. It contains 62 figures and 4 tables. It consists of an introduction, three chapters, a conclusion, a list of author's publications on the dissertation work, noted citations, approbation and used literature. Each of the chapters ends with a conclusion, which summarizes the conclusions.

The Introduction justifies the relevance of the topic under consideration, defines the goal and main tasks of the dissertation work.

The first chapter "Main tools and technologies for interaction between businesses and consumers in e-commerce" examines the main factors influencing the interaction between consumers and businesses in the field of e-commerce, as well as some existing theories and research in this area. An overview is made of the most widely used tools and platforms that contribute to improving the user experience in e-commerce. Some main problems related to the security and protection of user data are indicated. Conclusions are drawn regarding the effectiveness of applying various tools to improve user interaction.

In the second chapter "Innovative approaches to data mining in e-commerce: Hybrid model" we present a model for data mining from the Internet, improving user interaction in e-commerce. Based on the integration of three fundamentally different approaches to data processing (deterministic symbolic approach (rule-based), sub-symbolic deep learning (deep learning), linguistic-semantic interpretation (large language models)), a formal theoretical framework and conceptual architecture of the model have been created. In the practical implementation of the hybrid model, technical details regarding the implementation of each component have been specified. It is built on the basis of its three main components - a Java-based system, a YOLOv8 model and integration with ChatGPT.

The design of the proposed model allows them to work both autonomously and synchronously. The advantages of each of the components are used, with the disadvantages of a given tool being compensated by another, thus achieving an effective balance between accuracy and speed of data mining.

In the third chapter, "Online platform with tools for improved user interaction", a system developed by us is presented, in which the hybrid model is integrated. Its main components included in the architecture, the technologies used and the challenges overcome for its development, as well as the opportunities for future development, are described. Additional opportunities that the proposed model provides to users are presented. The online system developed by us provides a set of tools through which it is possible to automatically extract publicly available information about products on the Internet, analyze and summarize this information and send the relevant notifications to users.

In this chapter, two additional modules to the online platform are presented and developed by us, which support its work. The proposed Telegram bot provides user-friendly management of part of the platform's functionalities. A Chrome plugin can search and check for changes in data for a specific product in any content. Their architecture, technologies used and their functionality are presented, the challenges overcome for their development and the opportunities for their future development are described.

The Conclusion presents an analysis of the implementation of the tasks set and the achievement of the main goal of this dissertation. Relevant conclusions are drawn from the work on the topic, achieved results, contributions of the doctoral student and guidelines for future development are outlined.

A list of the doctoral student's publications, on the basis of which the dissertation was developed, is attached, and the noted citations are indicated. The Literature lists the literary sources used, which include 206 titles, of which 2 are in Cyrillic and 204 in English. The main content of the dissertation is presented in three chapters.

### **III. SHORT SUMMARY OF THE DISSERTATION**

#### **CHAPTER 1. MAIN TOOLS AND TECHNOLOGIES FOR INTERACTION BETWEEN BUSINESS AND CONSUMERS IN ELECTRONIC COMMERCE**

The first chapter of this dissertation is of an overview nature. An overview of the most widely used tools and platforms in e-commerce is provided. For this purpose, the main factors influencing the interaction between consumers and businesses and the existing theories in this area were initially examined. According to [Laudon, 2023], these are: personalizing customer interaction, building trust and reliability, building a messaging system and mechanisms for receiving feedback, adapting to changing trends.

Some of the existing theories and models related to the interaction between businesses and consumers are briefly described, for example, the Technology Acceptance Model (TAM); Social Commerce and Social Integration theory; Personalization and Recommender Systems theory; Gamification and Engagement theory, etc. There are also theories that consider cross-cultural interaction, i.e. how cultural norms, values and communication styles influence customer expectations and preferences. Research is also being conducted in the field of Human-Computer Interaction. Usability, accessibility, and user interface are studied to optimize the way customers interact with e-commerce platforms. This study does not include a detailed analysis of all existing theoretical models, as the emphasis is on applied aspects and technological solutions.

This chapter presents some of the main tools and technologies for interaction between businesses and consumers in e-commerce that improve

the user experience. They are software applications that facilitate and satisfy the needs and expectations of users. Special attention is paid to good user interface design; live chat tools, chatbots, and virtual assistants; augmented reality (AR) and virtual reality (VR); forums and platforms for feedback and surveys; customer data platforms, tools for tracking and comparing prices, personalization, and user analysis; payment and transaction portals, etc.

The issue of data protection is also considered. We will note that the greater the functionality of a site, the more opportunities for attacks against its security exist. That is why e-commerce sites must follow certain rules to protect themselves and their customers from attacks. They must apply all the imposed standards for the security of user data. Significant attention is paid to the encryption of information. Finally, examples from world practice for the effective use of e-commerce tools and platforms are presented.

### **Chapter 1 Summary and Conclusions:**

1. Appropriate tools collect information about customer preferences and behavior, analyze it, and help personalize and engage users.

2. The user interface should function across devices and offer clear navigation, product descriptions, customer reviews, and a convenient and secure payment process.

3. The integration of AI, chatbots, and personal assistants into e-commerce platforms can lead to more effective real-time customer interactions and personalized product recommendations.

4. Augmented and virtual reality (AR and VR) technologies create immersive interactive experiences, but are also used to fully present products.

5. Particular attention is paid to protecting user data and payment security.

6. Language processing algorithms are effectively used to overcome language barriers.

The analysis of existing approaches for automated data extraction in e-commerce shows that most of them have limited applicability in real conditions. The main reasons for this are the high degree of heterogeneity of web content, the possibilities for dynamic generation of interfaces, frequent changes in the DOM structure, as well as the lack of unified standards for presenting information about the products offered. Therefore, we believe that there is a need for a new type of theoretical model that combines the advantages of existing methods for processing information while minimizing their individual limitations.

## CHAPTER 2. INNOVATIVE APPROACHES TO DATA MINING IN ELECTRONIC COMMERCE: HYBRID MODEL

In the second chapter of the dissertation, a hybrid model for data extraction from the Internet with an integrated set of tools developed by us is presented. The corresponding theoretical model and its practical implementation in the real conditions of e-commerce are described. In the framework of this study, a “hybrid model” is understood as a multi-component architecture built on the basis of the integration of three conceptually different, but at the same time mutually complementary approaches: deterministic (rule-based), visual (deep learning) and linguistic-semantic (large language models). Each of them has specific advantages, which in an appropriate combination contribute to the formation of a convenient and adaptive system, working in a coordinated manner for the purpose of automated extraction and structuring of data from Internet environments. The results of this research of ours are published in the article [Dakov, 2025].

### 2.1. Theoretical framework of the proposed hybrid model

Theoretically, the hybrid data mining model, denoted by  $H$ , can be defined as a formal triplet of modules:

$$H = \{ R, V, L \},$$

where:

- $R$  is a deterministic data extraction module based on predefined rules and structural analysis of HTML/DOM;
- $V$  is a visual module for detecting information objects using neural networks;
- $L$  is a linguistic-semantic module for interpreting and structuring information through linguistic analysis.

The functionality of the proposed hybrid model is formally described by functions defined on appropriate sets and taking values in the set formed by the extracted data. For this purpose, it is convenient to introduce the following sets:

- Initial data set - the set  $X$  with elements  $x$ , consisting of all possible input web pages, can be defined as follows:

$$X = \{ x \mid x = (D, I, M) \},$$

where:

- $D$  represents the DOM structure of the page (a tree-like hierarchical organization of HTML elements);
- $I$  is the visual image of the page (a screenshot with a certain resolution);
- $M$  is additional information (URL address, screen resolution, language, viewport parameters, timestamps).

Typically, input web pages represent a multidimensional combination of structural, visual and textual information. Different components of the model work with different projections of the input space:

- Module  $R$  works mainly with components  $D$  and  $M$ ;
- Module  $V$  works mainly with component  $I$ ;
- Module  $L$  works with component  $I$ .

- Set of retrieved objects - we define a set  $O$  of target information objects as follows:

$$O = \{ o_1, o_2, o_3 \},$$

where:

- $o_1 \in \text{String}$  - product title. This is a variable-length text;
- $o_2 \in \mathbb{R}^+$  - product price. This is a positive real number;
- $o_3 \in \text{Image}$  - basic product image. It is represented by a two-dimensional matrix of pixels.

These objects represent the minimum semantic set necessary for automated processing and comparison of products in the field of e-commerce.

- Set of extracted data results - we define the set  $Y$  of all possible results of the extraction process:

$$Y = \{ y \mid y = (v, c, s, b) \},$$

where:

- $v$  is the extracted value (text, number, or image);
- $c \in \{o_1, o_2, o_3\}$  is the object class;
- $s \in [0,1]$  is the confidence score of the result;
- $b = (x, y, w, h)$  is the bounding box (coordinates and dimensions of the localization in the image).

Introducing a parameter  $s$ , characterizing the reliability, allows the results to be evaluated and sorted at subsequent processing stages. The threshold value for accepting a result is usually determined empirically. In this work, a threshold  $s \geq 0.5$  was used.

The functioning of the individual components of the proposed hybrid model can be represented as formal functions in the following way:

- The action of the deterministic module R is defined as a *Rule-based* function:

$$F_R: X \rightarrow P(Y),$$

where  $P(Y)$  is the power set of  $Y$  (the set of all subsets of  $Y$ ). The  $F_R$  function offers a set of candidates, extracted through structural rules and CSS selectors.

- The functionality of the visual module V is represented by a Visual function:

$$F_V: I \rightarrow P(Y).$$

It is a probability function that allows each result to be associated with a degree of confidence.

This module compensates for the weaknesses of R in visually dominant interfaces and dynamically generated pages. The module V is activated when R fails and analyzes the interaction between elements.

- The activity of the linguistic-semantic module L is defined by the Semantic function:

$$F_L: I \rightarrow P(Y).$$

The  $F_L$  function is defined for each input, as it always generates a structured output, even in the absence of clear structural or visual patterns.

The hybrid model can be viewed as a composite function that integrates the work of the three components in an optimal way:

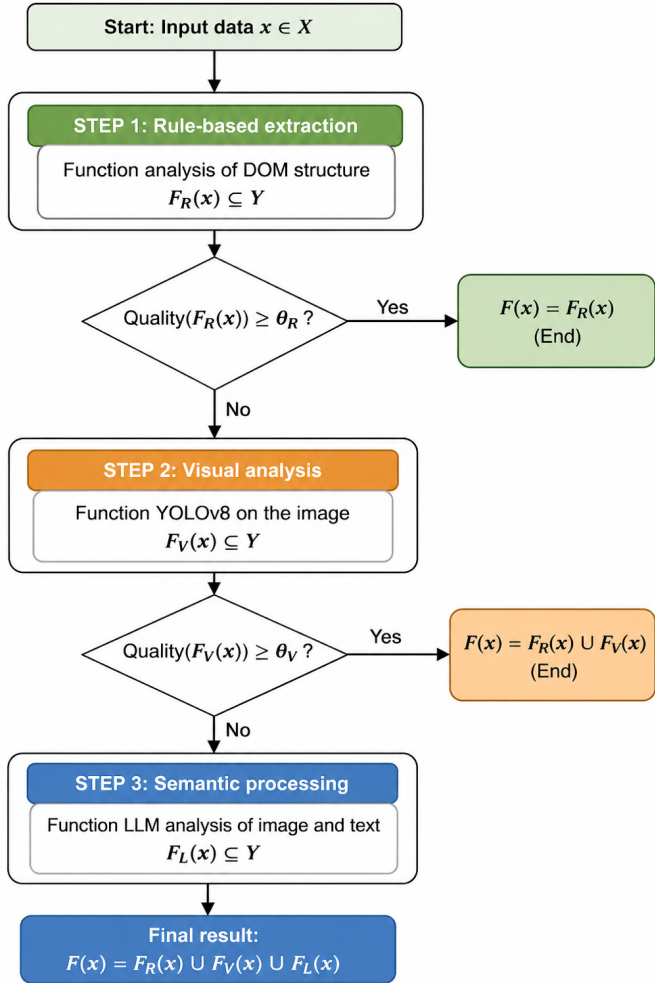
$$F: X \rightarrow P(Y).$$

This formalization allows the hybrid model to be considered as an adaptive system in which each component performs a clearly defined role depending on the characteristics of the input data.

The algorithm of the cascade call strategy is presented in Fig. 1. The following notations are used in it:

- $\text{Quality}(\cdot)$  - a function for assessing the quality of the results (based on confidence scores, completeness of the data),
- $\theta_r, \theta_v$  - threshold values for the reliability of the results from the activities of modules R and V, respectively,
- $\cup$  - an operation of unifying results from different modules (removal of duplicates, prioritization by confidence),
- $\emptyset$  - an empty set (when the module has not shown results).

Fig. 2 presents the conceptual architecture of the hybrid model and the data flow between the different components. The architecture demonstrates the possibility of processing the input data sequentially or in parallel by the three main modules, the goal being to generate the final result.



*Fig. 1. Algorithm of the composite function  $F(x)$*

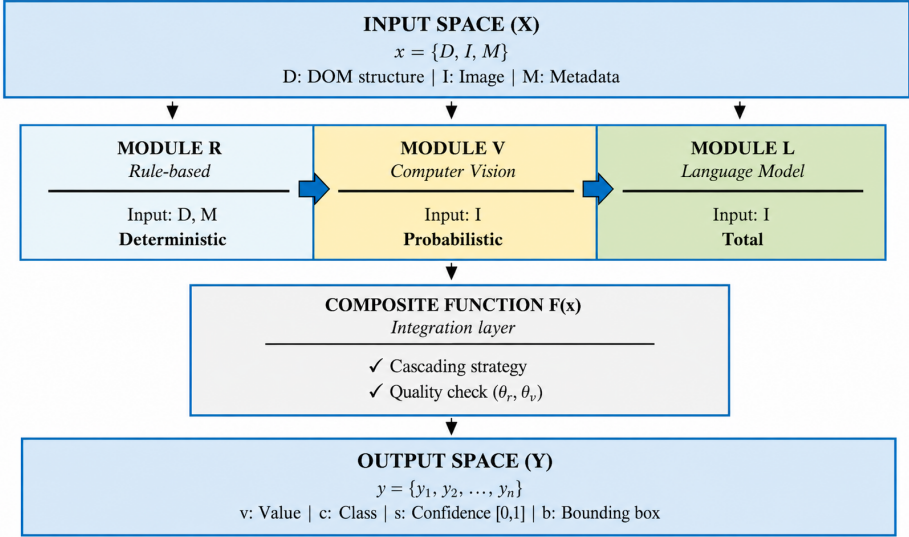


Fig. 2. Conceptual architecture of the hybrid model

**Legend:**

Blue	Blue Deterministic module - fastest, but limited
Orange	Orange Visual module - robust, high accuracy for images
Green	Green Semantic module - total function, maximum adaptability
Purple	Purple Composite function - integration layer with cascading logic

The information flow in the system follows a cascade logic with the possibility of parallel processing. The main stages are:

1. Initial processing - R module. Performs structural analysis and generates a screenshot. If the results are valid, they are final. If R fails, a second stage follows.

2. Visual processing - V module. Analyzes the screenshots. If the results are insufficient, stage 3 follows.

3. Semantic interpretation - L module, performs semantic analysis of the screenshots.

4. Combining and validating the result from the result of the activated modules. A final output Y is generated.

The proposed theoretical hybrid model is designed so that each component compensates for the limitations of the others. The proposed cascade strategy provides the following advantages:

- Resource optimization - the fastest module (R) is used first;
- Minimization of computational complexity - more complex modules are activated only when needed;
- Accuracy maximization - combining results increases the overall reliability;
- Guaranteed completeness - the total FL function always provides a result.

The presented theoretical hybrid model does not claim to be a universal solution to all problems in the field of data mining. Rather, it provides a structured framework for systematically addressing the specific challenges in the field of e-commerce. The model is designed to be extensible - additional components can be integrated without breaking the basic architecture.

## **2.2. Practical implementation of the hybrid model**

The developed hybrid system consists of three main components, each of which corresponds to a module of the theoretical model and performs a specific function within the data extraction process:

- Module R is based on Java, built on rules (Java tool). It has the following features:

- Uses a headless browser, controlled by CSS paths.
- Human intervention is required to determine the path to the required element.
- Extracts key elements from websites under predefined rules.
- Generates a dataset intended for training the YOLOv8 model.
- Provides a REST API, which allows for the integration of functionalities from external applications.

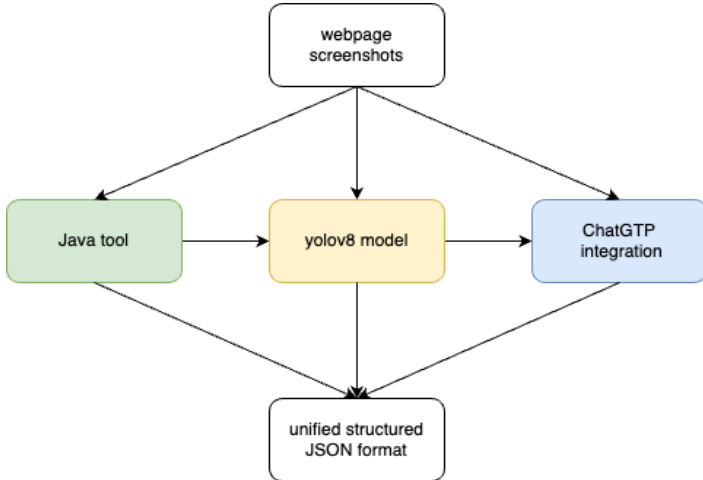
- Module V is implemented using YOLOv8, which is additionally trained. Its main focus is on identifying product titles, prices and images through the analysis of screenshots. Although a newer version of the model is available - YOLOv11 [YOLOv11, 2025], we preferred to work with YOLOv8, due to its relatively shorter training time. Its characteristics are:

- Relatively easy training
- Convenient integration
- High efficiency
- Uses few resources
- Despite the quality of training, the model can make mistakes

- Module L is implemented by a ChatGPT-based extraction tool. In situations where YOLOv8 fails to correctly identify the necessary information, ChatGPT takes over the task of extracting the data. GPT-4o-mini was chosen for the integration, due to its relatively high speed and adequate accuracy for solving the task. It is characterized by high efficiency. However, it has the following limitations:

- Paid solution
- Dependence on an external system

Figure 3 illustrates the interaction between the three main components of the hybrid system

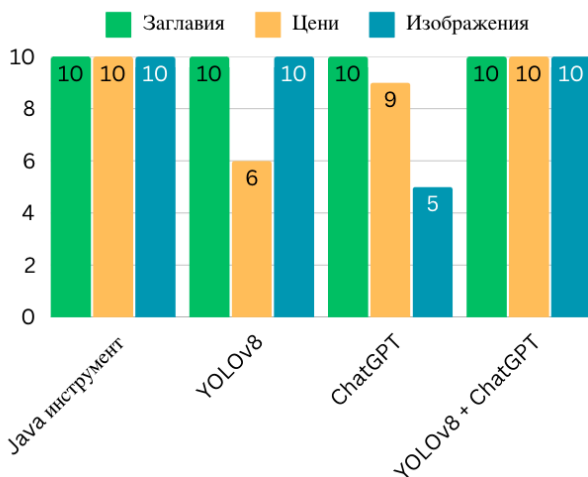


*Fig. 3. Main components of the hybrid system*

### 2.3. Results of testing the three methods

Tests of the three methods (Java tool, YOLOv8, and ChatGPT) were conducted on 10 websites (not participating in YOLOv8 training) to evaluate their efficiency in extracting images, titles, and prices. In addition, a test of their combined use (YOLOv8 + ChatGPT) was conducted. The test results are summarized in the table presented in Fig. 4.

The results show that YOLOv8 + ChatGPT is the most reliable approach for data extraction, demonstrating a 100% success rate. This proves the effectiveness of the hybrid approach, in which the shortcomings of one method are compensated by the other. The Java tool also demonstrates high accuracy, but the dependence on manually specified paths makes it less flexible for new websites. ChatGPT shows significant flexibility, but is not completely accurate in extracting images. YOLOv8 is highly efficient in extracting images and titles, but has limited accuracy in extracting price data.



Фиг. 4. Резултати от тестване на трите метода

## Summary and conclusions from Chapter 2:

### Main theoretical contributions of the model:

1. A formal mathematical definition of the input and output space is presented, allowing for precise analysis and reproducibility of the results.
2. A justification is given for formulated statements about completeness, robustness, and optimality, which guarantee the applicability and effectiveness of the model.
3. A cascade strategy is proposed, achieving satisfactory results.

### Main practical contributions of the model:

1. The advantages and disadvantages of each element are considered and options for compensating for them are proposed.
2. The presented results of the training of the YOLOv8 model show that the model gradually becomes more precise in object recognition.
3. An experiment on the operation of the YOLOv8 model is carried out - for a set of 10 websites that are not included in the initial training data set. Theoretical estimates show a total success rate of 86.7%. The presented results demonstrate a high success rate in image and title extraction. The relatively lower success rate in price extraction suggests an area for improvement.
4. Combining the three tools in a hybrid system leads to a high success rate in solving data recognition and extraction tasks. In this approach, the shortcomings of one module are compensated by the other.

## CHAPTER 3. ONLINE PLATFORM WITH TOOLS FOR IMPROVED USER INTERACTION

In order to improve the user interaction in e-commerce, we implemented the theoretical model presented in Chapter 2 within a platform with e-commerce tools. It enables automatic retrieval of publicly available information about products on the Internet, analysis and sending relevant notifications to users. It represents a set of different components that allow:

- Monitoring data on the Internet: prices and list of objects;
- Saving images of receipts and warranty cards;
- Tracking expenses;
- Tracking shipments by Bulgarian Post.

Some of the results we obtained are published in the articles [Dakov, 2021b], [Dakov, 2024a], [Dakov, 2023], [Dakov, 2025b]. All security standards discussed in [Dakov, 2021a] are met.

### 3.1. Model of an online platform for extracting data from the Internet

Fig.5 presents the general scheme of the work of all components of the platform we offer.

The process begins with the placement of tasks for execution. To place a data extraction task, we created a backend application that can be easily used and integrated, including in third-party applications. After the user sends a task for execution, it is registered in the task processing module, saved in the database and sent to the task buffer.

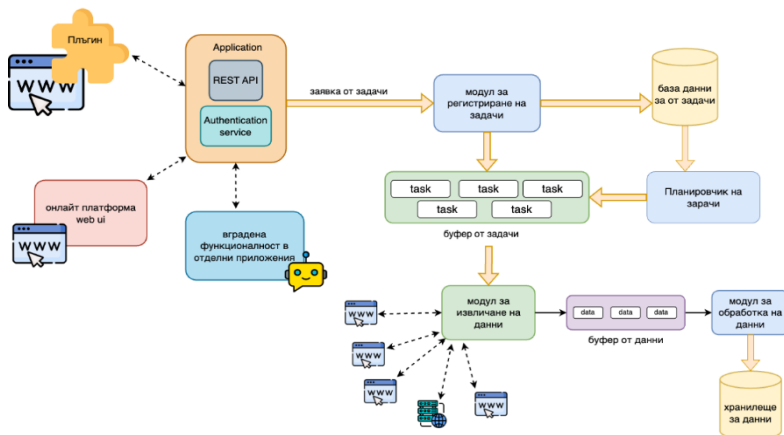


Fig. 5. General diagram of the workflow between the platform components

A task scheduling module takes care of the execution of each task. It periodically checks which task should be started when, retrieves it from the database and adds it to the task buffer. Once a task is in the buffer, it reaches the task execution module from there. This module plays the most important role, namely, it extracts data from the Internet.

In the process of developing the basic model, we went through different options. The first option of this model used an adapter to scan each different site. In case we wanted to observe data in a new e-store, new adapters had to be constantly updated and added. To solve this problem, we made a single module working with three components that can work with all sites, without the need for additional human intervention.

After the data extraction module receives a given task for processing, it generates data that is sent to the data processing module. The data is stored in a buffer so that the processing module does not become overloaded and there is no possibility of data being missed.

After the data is processed, it is saved in the data warehouse. For this purpose, we have chosen a relational database, since the data we are saving is of a simple type. Otherwise, we could also use non-relational databases [Khan, 2023], which can easily save unstructured data. In addition to saving data, this module also plays an important role in reporting data changes. When it detects a change, the module sends a request to the module for sending notifications.

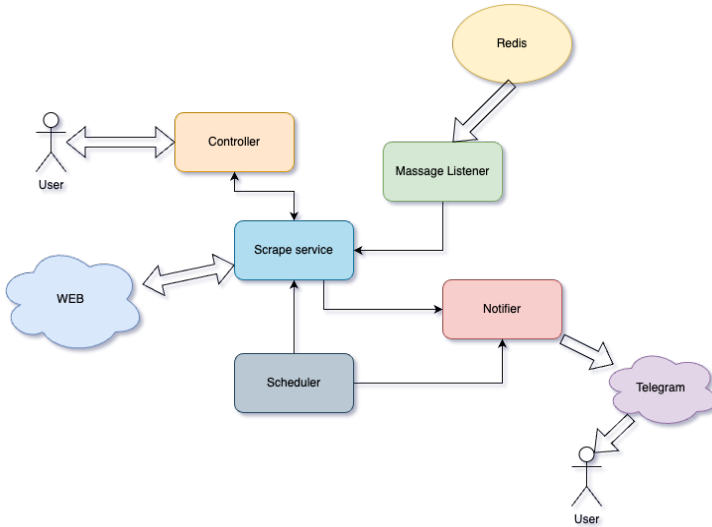
These notifications can reach the user through different flows, such as receiving an email about the change, a notification in the browser plugin itself or in third-party applications. The system has implemented a set of tools that we believe improve user interaction in e-commerce - product tracking, product list tracking, saving receipts, registering and tracking expenses, tracking products in the Bulgarian Post Office.

### **3.2. Internet data extraction module**

The module is developed based on the hybrid model presented in the second chapter of the dissertation. It is a backend service designed to scan web pages and extract data from them. It is built using a microservices-based architecture [Hasselbring, 2017] and has several modules, including Scraper Service, Controller, Queue Listener, Notifier and Scheduler Service. The

application is written in Java using the Spring Boot framework, which provides a wide range of features for building web applications, including automatic configuration, dependency injection and database connectivity.

The application is used to monitor websites for new products or pricing changes and notify the user when these changes occur. In addition, it can be used to aggregate and analyze data from multiple sources to extract valuable business information. Fig.6 presents a block diagram of the proposed application.



*Fig. 6. Block diagram of the proposed application*

The main elements of its architecture are:

- Data Extraction Service - This module is responsible for extracting data from web pages based on the provided URL and CSS path. The module can be used to scrape dynamic web pages that require user interaction, such as filling out forms and clicking buttons. The extracted data is then processed and can be stored in a database or passed to other modules for further processing. The Scraper service is designed to handle errors gracefully and retry data retrieval in case of failure. In case of failure, the module makes a record of the undiscovered data and information about the specific error so that the problem with the given task can be solved later.

- Controller - The module provides an API interface to trigger the web page scraping operation. The API endpoint accepts the URL and CSS path as parameters and passes them to the Scraper service to extract the data. After extracting the required information, it returns the data back to the user.

- Listener - This module monitors a Redis queue for new tasks to be processed. When a new task is added to the queue, the module passes it to the Scraper service for data exhaustion.

- Notifier - The module sends notifications to the user via a Telegram bot when changes are detected in web pages.

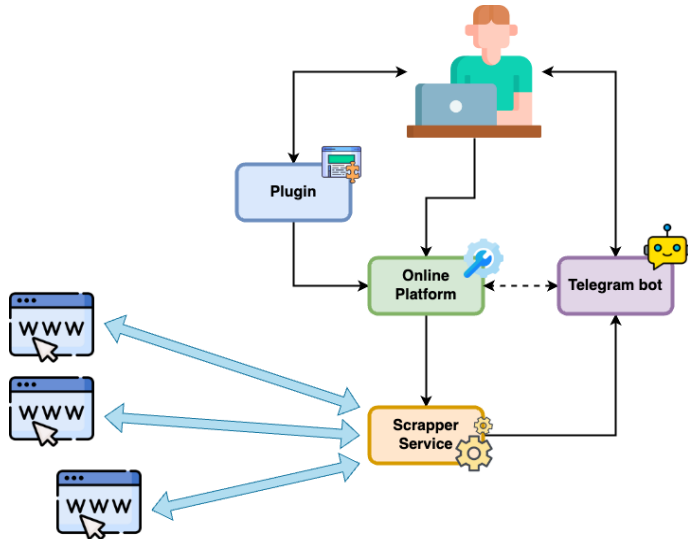
- Task Scheduler Module - This module regularly launches tasks to scan the network for changes based on a predefined schedule. It sends tasks to the Scraper module for execution, then, if it detects changes, they are written to the database and a message is sent to the end user.

### **3.3. Additional modules to the platform**

We have developed additional modules to the proposed online platform that improve user interaction in e-commerce. These are a Telegram bot for web scraping and a Chrome plugin. The functional relationship between the platform and the Telegram bot and the plugin, respectively, is presented in Fig. 7.

### **3.5. Telegram bot for web scraping**

The application we developed is a bot for web scraping that uses Telegram messages. The entire process is controlled by the user through commands. Its task is to provide the user with convenient and easy management of the online platform in order to improve the user experience.



*Fig. 7. Functional dependency between the online platform, data extraction module, Telegram bot and plugin*

The bot we offer is based on a “command listener”, which is responsible for managing the Telegram bot. It periodically checks and processes commands sent by the user. Based on what command is executed, the command listener gives an appropriate response to the user.

The architecture of the Telegram bot is built on the following main blocks:

- Telegram Bot API - used to interact with the Telegram bot. The bot can receive messages from users, send messages to users and receive updates from Telegram servers;
- User interface - users interact with the bot using the messaging interface provided by the Telegram application. The bot interprets the command entered by the user and responds appropriately based on the functionality it supports;
- Command handler - responsible for processing incoming messages from users and executing the corresponding commands. It analyzes incoming messages, checks whether they match certain commands and executes the command if it matches;

- Backend services - provide the necessary functionality to support the bot's functions. This includes image storage services, command processing, database management, and other third-party services.

The bot's functionality includes:

- Command embedding - Based on key commands, the Telegram bot can provide a powerful tool for tracking prices and notifying users. Using the Telegram API and a programming language (such as Java), functionality is provided to users in a convenient and easy way to be informed about the desired products and stores;

- Embedded buttons - Adding a button is done by sending additional options to the command. These buttons improve the interaction with the user.

### 1.3.1. Chrome Plugin

In order to track the price and other characteristics of a given product in e-commerce, we developed a Chrome plugin. It has the ability to search and check for changes in any content, such as description, title, product value or quantity of the product.

The plugin we offer provides seamless user authentication via REST API, securely stores the authentication token in the browser and retrieves data from the API to display it in its pop-up interface. It is necessary to ensure seamless communication with the API, well-protected token storage and efficient data retrieval. Fig. 8 schematically presents the communication flow.

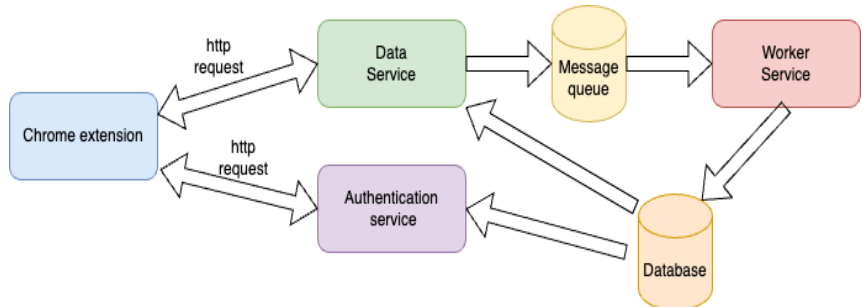


Fig. 8. Communication flow

The plugin functionality includes the following important operations:

- Authentication - When a user clicks on the Chrome extension icon, it triggers an event that signals the start of the authentication process. This event is typically captured by a background script in the plugin that makes an HTTP request to the authentication service endpoint;

- Product Price Selection - A button allows users to interactively highlight the desired element on the screen. The process of highlighting the element with a visible frame when hovering the mouse is initiated. This button serves as a trigger to start the element selection process. As the user moves the mouse cursor over the web page, the plugin dynamically highlights elements that could potentially represent the price of the product.

When the cursor is hovered over an element, the plugin visually marks it with a frame or overlay, making it easily distinguishable from other elements on the page;

- Screenshot - The user interface provides visual feedback to indicate that the feature is enabled and offers a checked checkbox;

- Displaying product data - When the user goes to a specific product page and clicks on the plugin, it receives the URL of the currently active browser tab and requests the REST API to retrieve already stored information about this product.

### **1.3. Practical application of the developed tools**

Some of the individual modules and tools developed within the framework of the dissertation have been successfully implemented in real operation and continue to function as stand-alone solutions serving an active user audience:

- Bulgarian Post parcel tracking tool - has been operating as a stand-alone web application since 2021 at <https://bulgarskiposhti.com>. The system generates stable traffic from 2,000 to 4,000 visits per month, which confirms its practical value and the sustainable user need for a similar type of service.

- Personal Expense Tracking Tool - has been in real operation since 2025 at <https://lolydash.com>, currently serving between 100 and 200 visits per month. The application demonstrates the applicability of the developed technologies in neighboring domains outside of e-commerce.

- Telegram web scraping bot - was successfully implemented in a real user scenario for tracking real estate listings (apartments and plots). Through automated monitoring of listings and timely notifications of new offers, the bot assisted real users in choosing and purchasing a property, which is a direct indication of its practical usefulness outside of the e-retail sphere.

These real-world results complement the controlled tests and confirm that the developed tools are not only theoretically sound, but also viable in a production environment with real users and real tasks.

### **Summary and conclusions from Chapter 3:**

Based on the developed tools and achieved results, we can draw the following conclusions:

1. The developed hybrid model has been successfully integrated into a product tracking tool in the platform.

2. The user interface allows users to add, view, update and delete information.

3. The proposed platform allows the client to track the delivery of packages through Bulgarian Post.

4. The proposed platform also includes a module for tracking costs.

5. The developed Telegram bot successfully manages a set of operations in the online platform.

6. The developed Chrome plugin facilitates the addition of a product for tracking.

7. The practical applicability of the developed technologies has been confirmed by real implementation of individual modules in a real environment.

When developing the platform and the module for extracting data from the Internet, a combination of technologies was used and a number of factors such as scalability, database management security, network stability were taken into account. The proposed system can process large volumes of product data and provide up-to-date and secure information to users. The proposed Telegram bot and plugin facilitate the operation of the already described online platform, offer browsing advantages and help improve user interaction in e-commerce.

## **CONCLUSION**

The presented results correspond to and satisfy the set goals and objectives. In the final part of the dissertation work, the results of the conducted research are summarized and its contributions are presented. More and more customers prefer Internet shopping to save time and money. There is a growing need for effective and long-term cooperation and mutual assistance between consumers and businesses in the field of e-commerce, based on mutual interest, trust and secure protection of customer data and transactions. The issue of retaining consumer interest is particularly relevant, including improving the user experience when browsing on different platforms, providing sufficient information about products, as well as offering accessible ways of communication, assistance and additional services.

## **CONTRIBUTIONS OF THE DISSERTATION**

In accordance with the goals and objectives of the dissertation, research has been conducted and new systems and modules have been developed to improve user interaction in e-commerce. The contributions of the achieved results can be defined as:

1. A formal conceptual-methodological framework of a hybrid model for automated data extraction and structuring from Internet environments has been proposed. The model is built by integrating three conceptually different, but at the same time complementary approaches - deterministic (rule-based), visual (deep learning) and linguistic-semantic (large language models). Based on a cascade strategy, the specific advantages of each of them contribute to the

formation of an adaptive system that achieves a balance between speed and accuracy of extraction, as well as precise analysis and reproducibility of the results. The model is designed so that, if necessary, additional components can be integrated into it without violating its basic architecture.

2. Based on the theoretical hybrid model, a real model for data extraction from the Internet has been developed. It has three main components - a Java-based system, a YOLOv8 model and integration with ChatGPT, which can work autonomously and synchronously. Its design allows the advantages of individual components to be used and combined, and the disadvantages of one tool are compensated by the other. This achieves an effective balance and high success rate in solving data recognition and extraction tasks.
3. The proposed hybrid method has been successfully integrated into an online platform containing a set of tools that contribute to improving user interaction. Through it, users can easily and conveniently monitor changes in product characteristics, change product lists, including tracking shipments by Bulgarian Post, manage personal expenses and save receipts.
4. A special module for extracting information from the Internet has been developed, which is connected to the proposed platform and significantly supports its work. Its functionality includes processing tasks sent from the online platform and processing commands sent by users via the Telegram application. The module provides effective web scraping and data extraction for new products, current price changes and user notification.
5. A Chrome plugin has been developed, through which users can easily and conveniently add products to monitor certain characteristics in the online system. The functionality of the plugin includes the ability for users to take a screenshot of a web page at a given time. In this way, the client can save the current state of product data.

## DEVELOPMENT PROSPECTS

- Regarding the theoretical framework of the proposed hybrid model
  - expanding the set of extracted objects and including additional products; integrating additional modules for specific data types; optimizing the composite function F through machine learning of the component selection strategy itself;
- Regarding the future development of the proposed online platform - expanding the data sources for products integrated into the system, such as APIs or databases; applying machine learning algorithms; developing a mobile application; improving the user interface, including functions such as data visualizations; adding a function for scanning and analyzing images of receipts.
- Regarding the special module for web scanning and data extraction
  - overcoming limitations in data extraction through techniques such as IP address rotation; adding and testing other models; integrating models to overcome CAPCHA; advanced monitoring and sending real-time notifications for data changes.
- Regarding the developed Telegram bot - adding photos of receipts and their storage in the online platform; adding commands for managing a list of products in the online platform; adding commands for tracking shipments by Bulgarian Post.
- Regarding the proposed plugin - adding new functionality such as entering a list of products in the system; adding all CRUD operations for a product and a list of products; displaying all monitored products and product lists; displaying a notification when the price of a product changes or a change in a list of products.

## LITERATURE

1. **[Dakov, 2024b]** Dakov S., Dakova., M., Studying the need for new tools in e-commerce development, (2024). Scientific works of the Union of Scientists in Bulgaria-Plovdiv, series C. Technics and Technologies., Vol. XXI, ISSN 1311-9419 (Print), ISSN 2534-9384 (On-line).
2. **[Varbanov, 2008]** Varbanov R., „Elektronnata targovia-klyuchov faktor za konkurentosposobnost na malkite i sredni predpriyatia“, Stopanska akademia “D. A. Tsenov, 2008.
3. **[Shishmanov, 2020]** Shishmanov K., Tashkova M., Markova, M., Savremenni tendentsii v sazhdavaneto na prilozhenia za elektronna targovia, Almanah „Nauchni izsledvania“, Tom 28, 2020, 247, <http://hdl.handle.net/10610/4300>. [Laudon, 2023] Laudon, K., Traver, C., E-commerce 2023: Business, Technology, and Society, 17th edition, 2023, Pearson.
4. **[YOLOv11, 2025]** Ultralytics YOLOv11, 2025, from <https://docs.ultralytics.com/models/yolov11/>.
5. **[Dakov, 2021a]** Dakov, S., Malinova, A., A survey of E-commerce security threats and solutions, CBU International Conference Innovation in Science and Education, Prague, Czech Republic, Vol. 2 (2021): pp. 1-9. ISSN: 2695-0758, Proceedings of CBU in Natural Sciences and ICT, DOI: <https://doi.org/10.12955/pns.v2.135>.
6. **[Khan, 2023]** W. Khan, T. Kumar, C. Zhang, K. Raj, A. Roy, B. Luo, SQL and NoSQL Database Software Architecture Performance Analysis and Assessments—A Systematic Literature Review, Big Data and Cognitive Computing, 2023, May 12 7(2):97, <https://doi.org/10.3390/bdcc7020097>.
7. **[Hasselbring, 2017]** Hasselbring, W., Steinacker, G., „Microservice Architectures for Scalability, Agility and Reliability in E-Commerce“, IEEE International Conference on Software Architecture Workshops (ICSAW), Gothenburg, Sweden, 2017, pp. 243-246, <https://doi.org/10.1109/ICSAW.2017.11>.

## DECLARATION OF ORIGINALITY

**I declare that this dissertation is entirely my own work and that no third-party publications have been used unlawfully or in violation of copyright in its development.**

## ACKNOWLEDGEMENTS

I express my sincere gratitude to my scientific supervisor Assoc. Prof. Dr. Vesselin Kyurkchiev for the trust, opportunities, inspiring example and constant support. I express special gratitude for the professional support and valuable consultations of Prof. Dr. Nikolay Pavlov, as well as Prof. Dr. Assen Rahnev for his continuous support and assistance from my student years to the present.

I express my gratitude to Prof. Dr. Asya Stoyanova-Doycheva and Assoc. Prof. Dr. Angel Golev for their meaningful recommendations and advice. I sincerely thank Prof. Dr. Anna Malinova for her overall assistance. I thank all colleagues from the Department of Computer Technologies at the

Faculty of Mathematics and Physics at the Paisiy Hilendarski University, who contributed to the realization of this work. I thank the scientific jury for the time spent reviewing and evaluating this dissertation.

Last but not least, I express my deep gratitude to my family for their unwavering support in all my endeavors.

## **PARTICIPATION IN INTERNATIONAL AND NATIONAL CONFERENCES**

1. International Conference “Innovation in Science and Education”, Czech University of Life Sciences Prague, Prague, Czech Republic, 17 March 2021.
2. International Journal “Differential Equations and Applications” (REMIA 2021), Plovdiv, Bulgaria, 22 - 24 October 2021.
3. Scientific Session “Science Days 2023”, Union of Scientists in Bulgaria - Plovdiv, Bulgaria, 23 - 24 November 2023.
4. International Scientific Conference IMEA23 - Informatics, Mathematics, Education and Their Applications, Pamporovo, Bulgaria, 29 November - 1 December 2023.
5. International Scientific Conference TECHSYS 2025 - Engineering, Technologies and Systems, Technical University of Sofia - Plovdiv Branch, Plovdiv, Bulgaria, 15 - 17 May 2025.
6. 11th International Scientific-Business Conference LIMEN 2025: Leadership, Innovation, Management, and Economics: Integrated Politics of Research, 10 - 12 December, 2025.

## SCIENTIFIC PUBLICATIONS ON THE TOPIC OF THE DISSERTATION

1. **Dakov, S.,** Malinova, A., A survey of E-commerce security threats and solutions, CBU International Conference Innovation in Science and Education, March 17, 2021, Prague, Czech Republic, (2021): Vol. 2, Proceedings of CBU in Natural Sciences and ICT, pp. 1-9. ISSN: 2695-0758, <https://ojs.cbuic.cz/index.php/pns/article/view/135/314>, **(WoS)**
2. **Dakov, S.,** Malinova, A., Automated product information retrieval in E-commerce, REMIA 2021, International Journal of Differential Equations and Applications, (2021): Vol. 20, No.2, pp.157-168, ISSN(Online):1314-6084, <https://www.scopus.com/record/display.uri?eid=2-s2.0-85119899434&origin=resultslist>, **(Scopus, SJR 2021=0.214, Q4)**
2. **Dakov S.,** Communication between micro web services and web browser extension, Scientific works of the Union of Scientists in Bulgaria-Plovdiv, series C. Technics and Technologies., (2024): Vol. XXI, pp. 18-22, ISSN 1311-9419 (Print), ISSN 2534-9384 (On-line),[https://usb-plovdiv.org/wp-content/uploads/2024/05/2024\\_tehnika\\_i\\_tehnologii\\_tom\\_XXI.pdf](https://usb-plovdiv.org/wp-content/uploads/2024/05/2024_tehnika_i_tehnologii_tom_XXI.pdf)
3. **Dakov, S.,** Malinova, A., Enhancing the E-commerce experience: A wishlist browser extension with price-detection, IMEA23, Proceedings of the International Scientific Conference „Informatics, Mathematics, Education and Their Applications“, 29 Nov - 01 Dec 2023, Pamporovo, Bulgaria, (2023): ISBN:978-619-7663-79-2, <https://www.fmi-plovdiv.org/GetResource?id=4652>
4. **Dakov, S.,** Dakova, M., Training a model for automated information retrieval from the internet, TEM Journal, (2025): Vol. 14, No 3, ISSN: 2217-8309 (Print)., ISSN: 2217-8333 (Online), [https://www.temjournal.com/content/143/TEMJournalAugust2025\\_258\\_8\\_2598.html](https://www.temjournal.com/content/143/TEMJournalAugust2025_258_8_2598.html), **(Scopus, SJR 2025=0.258, Q3)**
5. **Dakov, S.,** Dakova, M. (2025). Event-Driven Data Orchestration: A Modular Approach for High-Volume Real-Time Processing. Engineering Proceedings, 100(1), 48. <https://doi.org/10.3390/engproc2025100048>, **(Scopus, SJR 2025=0.254, Q3)**