Comparing job descriptions offered by the Romanian labor market with the occupational classification ESCO. IT support solution

Moldovan Andreea-Diana-Ariana[1]
[1]*Faculty of Economics and Business Administration, West University of Timișoara, Romania*

Abstract: In this paper, we aim to develop a framework for scanning job offers displayed on recruitment websites and analyzing their descriptions by comparing them to the ESCO classification. Specifically, we are focusing on two occupational profiles: software developers and web developers, whose appearances we have monitored on the ejobs.ro website. Our goal is to identify to what extent employers use ESCO formulations in describing the required skills and competencies for these occupations. The framework is flexible and allows for the inclusion of other occupations, as well as the scanning of other websites containing job offers.

Keywords: ESCO, web scraping, NoSQL – MongoDB, Python, MS Power BI

1. Introduction

ESCO (European Skills, Competences, Qualifications, and Occupations) represents the European classification of skills, competences, qualifications, and occupations and is a component of the Europe 2020 strategy. The ESCO classification establishes a bridge between the labor market and the education and vocational training sector by providing standard terminology (Figure 1).

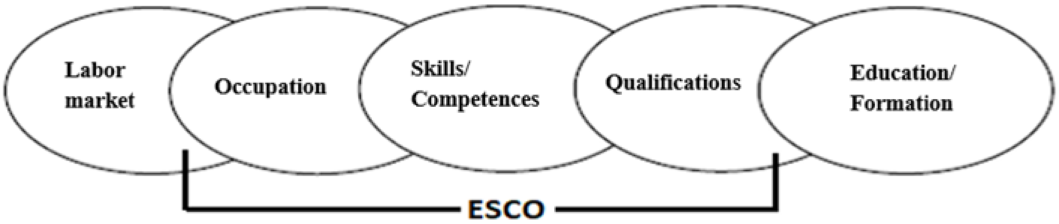


Figure 1: ESCO between the labor market and educational services

The most relevant abilities, competencies, and qualifications required for each occupation are defined in ESCO. In the development of ESCO, a number of international classifications were taken into account, namely: ISCO-08 (International Standard Classification of Occupations), ISCED (International Standard Classification of Education). Both employers and job seekers can benefit from ESCO. Employers can use ESCO to define the necessary abilities and competencies when formulating job descriptions, while job seekers can use ESCO elements regarding abilities and competencies when creating their CV and job application form. At the same time, educational service providers can refer to ESCO when setting learning outcomes for different study programs. Learning outcomes, expressed in terms of knowledge, skills, and responsibilities, represent the supporting elements of qualifications. The ESCO classification ensures the integration of national occupational classifications into a unified European framework. ESCO is continuously updated to reflect changes in the European labor market, as well as those in education and training. These modifications are reflected in different ESCO versions.

In Romania, the classification of occupations is established through COR (Classification of Occupations in Romania). It is updated quarterly or whenever necessary, for example, when the occupation is included in ESCO. Various job offers are posted on specialized websites such as ejobs.ro, bestjobs.ro, or hipo.ro. They consist of the company description, job description, and requirements imposed on applicants. Job descriptions differ from site to site, and even within the same site, we have various formulations. Both educational offers and corresponding descriptions from the occupational taxonomy ESCO can be extracted through web scraping processes. The paper proposes a web scraping solution implemented in Python, with the extracted and transposed data in JSON (JavaScript Object Notation) format stored in a MongoDB database. Based on the collections of documents stored in MongoDB, the intended analyses are carried out regarding the way job offer formulations are reported in the ESCO occupational taxonomy.

2. Literature Review

The purpose of this case study is to compare the job descriptions offered by IT companies to ESCO formulations when posting a job on a recruitment portal. The dataset is composed of data and information found on the analyzed website, ejobs.ro, and extracted through a web scraping process, then brought for analysis in a Business Analytics tool. Additionally, the aim is to compare and detect similarities between ESCO standards and the curriculum of the Economic Informatics specializationfrom the Faculty of Economics and Business Administration (West University of Timișoara, Romania), these formulations being a bridge between the labor market and education. The purpose of higher education is to prepare future specialists, and for this, the curricula and course outlines must be continuously adapted. The labor market and the requirements of a dynamic environment demand continuous modification to prepare students for the future job market.

2.1 NoSQL databases. Mongo DB

NoSQL databases, unlike traditional databases, use different approaches and techniques for storing data, such as column group-based storage, document-based storage, graph-based data storage, etc.In MongoDB, collections of documents correspond to tables in the relational model, and documents correspond to records. In relational databases, records have a fixed structure, whereas in MongoDB, documents in a collection can have different fields.In the absence of relationships in traditional databases, data (documents) modeling can be done in two ways: using references or embedded data. Embedded data represents a denormalized form of data and is preferred when modeling relationships of the inclusion or one-to-many type.

MongoDB databases offer the following features (Hows, 2013): document-oriented storage; complex indexing capabilities; flexible queries of documents; replication and availability; partitioning; spatial data management; advanced backup and restore options.

2.2 Web scraping

Web scraping represents the process of extracting data from web sources. Nowadays, an increasing number of companies are using web scraping to collect a large volume of data in the shortest amount of time possible, to serve the decision-making process. This process is composed of two stages: extracting data in an unstructured format and transforming it into a structured, tangible format that is understandable by humans and useful for the company's decision-making analyses. There are two options for creating a web scraper: using a programming language or dedicated tools. The most commonly used programming languages for web scraping are Python, Ruby, Node.js, C++, PHP, and Java (Gheorghe, Mihai &Dârdală, 2018). The tools available on the market for web scraping are generally paid, based on a monthly subscription. These solutions make it easy to develop a web scraper, even for non-programmers, by providing automation platforms and no-code scraping.

Among the software for automating business processes that offer the possibility of developing a web scraper, we mention two products, namely UiPath and Power Automate. UiPath is a company founded in Romania that develops a platform for process automation. It has an intuitive interface that is easy to work with even for non-programmers, with drag-and-drop elements. Creating and debugging robots is easy to do, and the platform even offers the possibility of inserting Python scripts. Additionally, UiPath offers advanced data scraping options. Microsoft Power Automate is a task automation solution offered by Microsoft. It has both a web and desktop version. Power Automate offers the possibility of web scraping without writing code, which can be a great advantage for certain users, but only in the desktop version.The two solutions presented above are possible choices for implementing a web scraping process, but each has its strengths and weaknesses. UiPath offers more functionalities than Power Automate and contains more advanced options for extracting and manipulating data.

Most web scrapers are created in Python, which is the most widely used programming language for web scraping, having dedicated libraries for this process, such as BeautifulSoup and Requests (Mitchel, 2018). This object-oriented language is powerful and a highly efficient web scraper can be developed in a short time. Unlike UiPath, developing a web scraper in Python takes less time, runs much faster, and the obtained data has higher accuracy. Using this language and its associated libraries leads to an increase in data extraction speed, obtaining a large volume of data in a relatively short time.

The complex structure of the data and the need to save it in certain nested data structures have led to choosing to write code for the data extraction process from the web. Furthermore, data processing has been performed on the collected data, extracting keywords for the described technologies from job postings. The complexity of the processing and the method of data storage led to the choice of writing a web scraper in Python instead of using UiPath.

Python is an ideal language for web data extraction because it can be automated and combines the two stages of web scraping (Patel, 2020). Automation comes from the possibility of using certain libraries that allow the web scraper to run daily, which increases the speed of data extraction. Python combines the two stages of a web scraping process: scraping in an unstructured format and importing it into a structured one (Norcross, Schuessler, & Jones, 2020). There are also dedicated tools that only perform the first stage, while others make the second stage a laborious process. A web scraper built in Python has several functionalities, such as: performing scraping, adding data, analyzing, importing, and saving it, but also the possibility of visualizing the data with Matplotlib.

2.3 Business Analytics tools. MS Power BI

Currently, disruptive technologies generate a large volume of data, and Business Analytics tools come to the aid of companies by facilitating the efficient collection, analysis, and presentation of data in real-time. This type of tool contains a variety of software applications that can be used to analyze and visualize data. MS Power BI is one of the most commonly used Business Analytics tools, allowing for the creation of dashboards and dynamic data visualization (Wade, 2022). MS Power BI enables easy connection and retrieval of data sources, visualization of data, and creation of reports and charts to aid in decision-making processes, as well as their distribution within the company.

3. Methodology

The proposed framework in this paper involves designing and implementing an information repository that will contain job descriptions taken from online job postings, as well as descriptions of similar occupations from ESCO. In addition to being able to store a large volume of Big Data, NoSQL databases allow for rapid, real-time access to this data. The information repository, in the form of a NoSQL database, is populated with data resulting from the execution of two web scraping processes, namely a process of retrieving job descriptions from the job portal, and another process of extracting descriptions of similar occupations from ESCO. To complete the analysis, a process of investigation using a Business Analytics tool will be conducted based on the information in the repository (Figure 2).

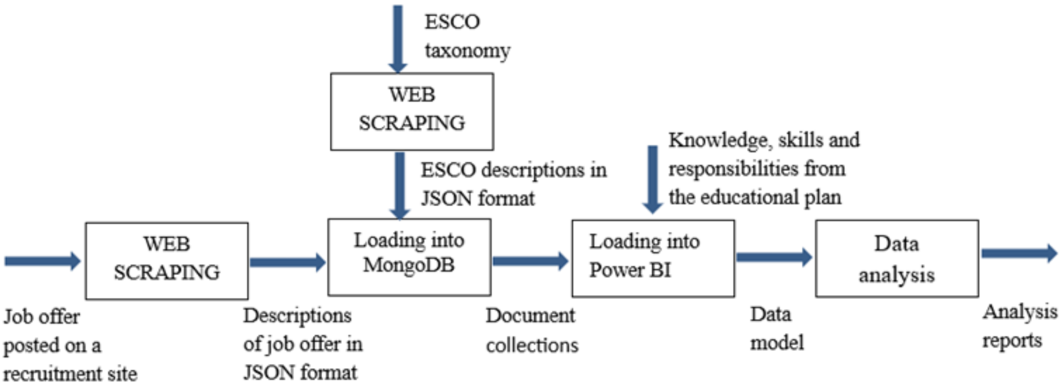
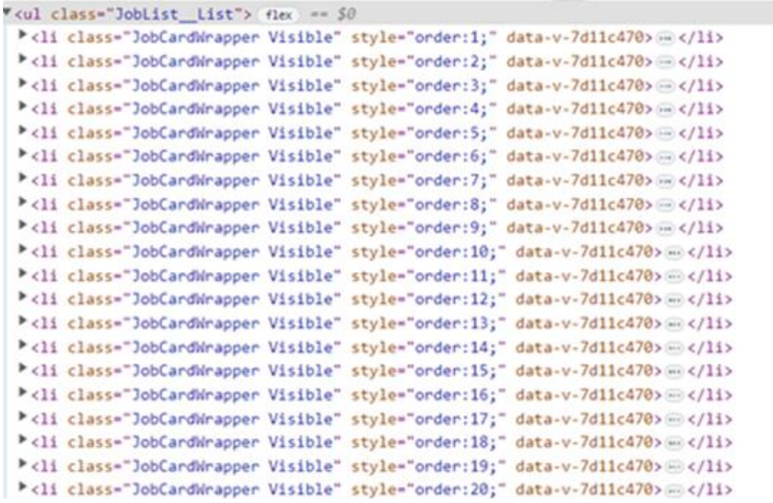

ESCO taxonomy
WEB SCRAPING
ESCO descriptions in JSON format
Knowledge, skills and responsibilities from the educational plan
WEB SCRAPING
Loading into MongoDB
Loading into Power BI
Data analysis
Job offer posted on a recruitment site
Descriptions of job offer in JSON format
Document collections
Data model
Analysis reports

Figure 2: Analysis framework. Processes.

3.1 Web scraping. The takeover of the offered jobs

The practical part of this work consists of developing a web scraping algorithm for extracting job-related data from a specialized site. The data of interest to be extracted are those related to the job title, posting date, company offering the job, salary if mentioned, and job requirements. Two categories of IT jobs were chosen, namely web development and software development. In studies or analyses conducted by various specialized sites, such as LinkedIn and Indeed, the two are among the top IT jobs, with high demand from employers. According to an analysis by a recruiting site, Best Jobs, in the first half of 2022, one of the most sought-after types of IT jobs was that of software developer. Additionally, IT jobs were ranked among the top domains with the most applications.

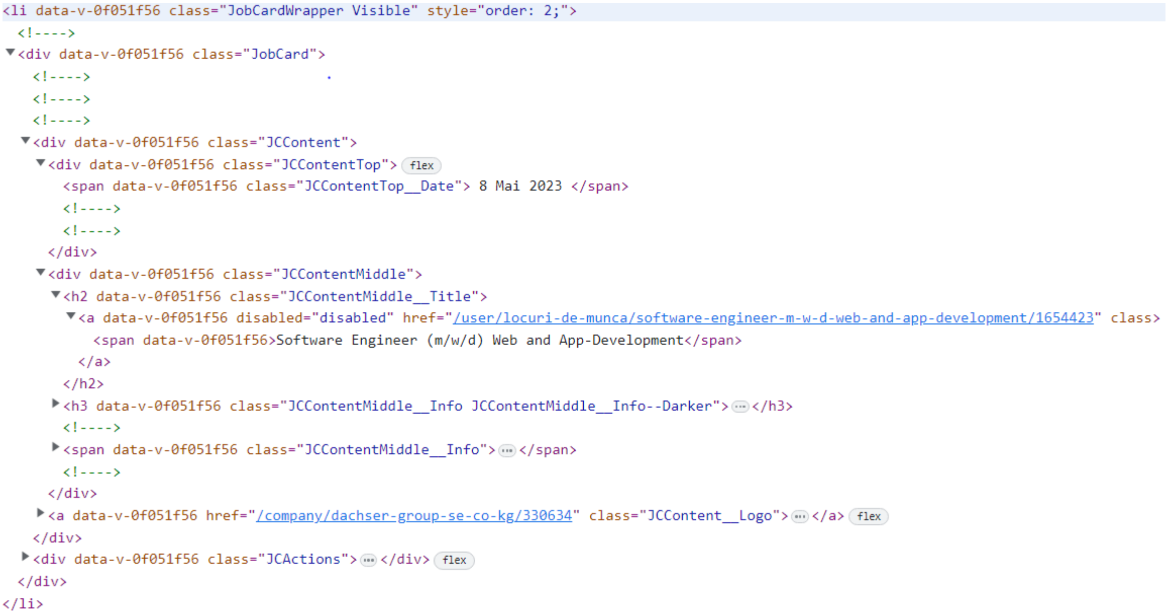
The first step in creating a web data extraction algorithm is analyzing the structure of the respective page. This is possible by inspecting the HTML structure of each element separately (Figure 3).



```
▼<ul class="JobList__List"> flex == $0
  ▶<li class="JobCardWrapper Visible" style="order:1;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:2;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:3;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:4;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:5;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:6;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:7;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:8;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:9;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:10;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:11;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:12;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:13;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:14;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:15;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:16;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:17;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:18;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:19;" data-v-7d11c470>⋯</li>
  ▶<li class="JobCardWrapper Visible" style="order:20;" data-v-7d11c470>⋯</li>
```

Figure 3: The structure of an ejobs page

In Figure 3, it can be observed that the ads on the analyzed site are represented as <li> elements, stylized using cards, within an unordered list (<ul>). In each list item, job details are found (Figure 4). These are: posting date, job title, company, city, and salary if applicable. To see additional details, such as the requirements for each offered position, each ad must be accessed separately.



```
<li data-v-0f051f56 class="JobCardWrapper Visible" style="order: 2;">
  <!---->
 ▼<div data-v-0f051f56 class="JobCard">
    <!---->
    <!---->
    <!---->
   ▼<div data-v-0f051f56 class="JCContent">
     ▼<div data-v-0f051f56 class="JCContentTop"> flex
        <span data-v-0f051f56 class="JCContentTop__Date"> 8 Mai 2023 </span>
        <!---->
        <!---->
      </div>
     ▼<div data-v-0f051f56 class="JCContentMiddle">
       ▼<h2 data-v-0f051f56 class="JCContentMiddle__Title">
         ▼<a data-v-0f051f56 disabled="disabled" href="/user/locuri-de-munca/software-engineer-m-w-d-web-and-app-development/1654423" class>
            <span data-v-0f051f56>Software Engineer (m/w/d) Web and App-Development</span>
          </a>
        </h2>
       ▶<h3 data-v-0f051f56 class="JCContentMiddle__Info JCContentMiddle__Info--Darker">⋯</h3>
        <!---->
       ▶<span data-v-0f051f56 class="JCContentMiddle__Info">⋯</span>
        <!---->
      </div>
     ▶<a data-v-0f051f56 href="/company/dachser-group-se-co-kg/330634" class="JCContent__Logo">⋯</a> flex
    </div>
   ▶<div data-v-0f051f56 class="JCActions">⋯</div> flex
  </div>
</li>
```

Figure 4: Structure of a job offer

To create a web scraper that extracts the aforementioned data, the following libraries are used (Oprea & Bâra, 2022):

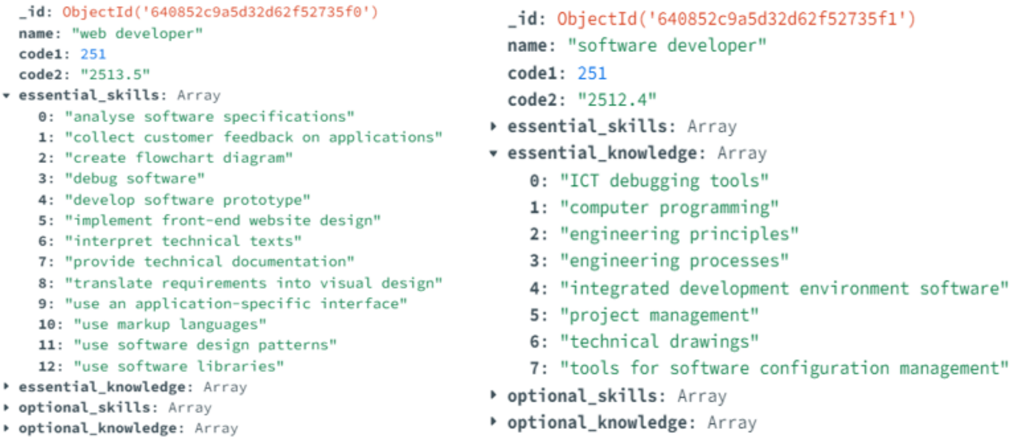
- Beautiful Soup - used to retrieve HTML code for data extraction;
- DateTime - used to extract the date when data was extracted;
- Json - used for saving data in JSON format;
- Selenium.

To write a web scraper, the page link of a Web Driver will be specified and a variable soup will be created, which will transform the page's content into a Beautiful Soup, which can be used to navigate HTML and search for the desired data (Oprea & Bâra, 2022). Within the transformed text, the <ul> tag will be searched for, and for greater accuracy, the class (class_='JobList__List') will also be mentioned. Within the unordered list, the following elements are to be parsed and extracted from each <li>: job name, posting date, company, city, and salary (if mentioned). To extract job descriptions, it is necessary to access each ad's links through the Web Driver, and this is possible by concatenating the site's root (https://www.ejobs.ro/) with the link associated with the ad's title. Next, the page content will be transformed into a Beautiful Soup, and the job description and requirements found on the page will be extracted. This data is saved as a dictionary consisting of a key-value pair, and then added to an empty list using the append() method. The last step is to save the list in JSON format so that it can be brought into a database. Another way to save data can be to use the pandas library, which can be used to save data in a csv file.

From the web page of each ad, the requirements necessary for the job offered by the companies have been extracted. The requirements section differs from one employer to another, and its structure is not exact, which is why it was desired to write an algorithm in Python to detect the main technologies and knowledge desired for each job. The first step was to create a list of all the knowledge and technologies that could be required for an IT job.

3.2 Web scraping. Retrieval of ESCO classifications

Two job categories were analyzed within ESCO: web developer and software developer. The acquisition of specific competencies and skills for each occupation was achieved by creating a web scraper, similar to the one used for job extraction, using Python and libraries dedicated to this process. Each occupation within ESCO has sections for knowledge, skills, programming languages, and technologies required for such an occupation. Comparing the ESCO requirements to those of Romanian employers, it can be observed that the latter place emphasis on optional knowledge related to programming languages and technologies. From each occupation, its name, category code, occupation code, as well as the aforementioned knowledge and skills were extracted in the form of arrays. Figure 5 illustrates the skills required by ESCO for a web developer and software developer.



```
_id: ObjectId('640852c9a5d32d62f52735f0')
name: "web developer"
code1: 251
code2: "2513.5"
▾ essential_skills: Array
    0: "analyse software specifications"
    1: "collect customer feedback on applications"
    2: "create flowchart diagram"
    3: "debug software"
    4: "develop software prototype"
    5: "implement front-end website design"
    6: "interpret technical texts"
    7: "provide technical documentation"
    8: "translate requirements into visual design"
    9: "use an application-specific interface"
    10: "use markup languages"
    11: "use software design patterns"
    12: "use software libraries"
▸ essential_knowledge: Array
▸ optional_skills: Array
▸ optional_knowledge: Array
```

```
_id: ObjectId('640852c9a5d32d62f52735f1')
name: "software developer"
code1: 251
code2: "2512.4"
▸ essential_skills: Array
▾ essential_knowledge: Array
    0: "ICT debugging tools"
    1: "computer programming"
    2: "engineering principles"
    3: "engineering processes"
    4: "integrated development environment software"
    5: "project management"
    6: "technical drawings"
    7: "tools for software configuration management"
▸ optional_skills: Array
▸ optional_knowledge: Array
```

Figure 5: ESCO classifications for web developer and software developer brought to MongoDB

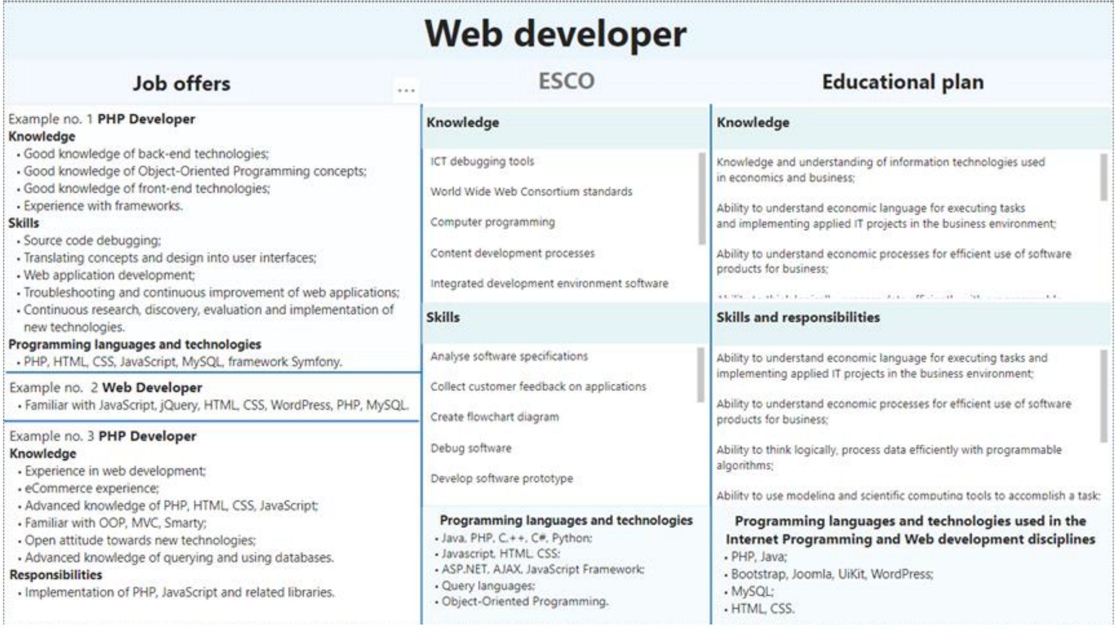
4. Results and Discussions

The dataset obtained using a web scraping process was then brought into MongoDB, then taken into Microsoft Power Bi in order to analyze ESCO requirements and their compliance by the employers. Following the web scraping process, jobs from a certain period have been drawn, namely between January 9 and February 7. 250 jobs have been analyzed in total, 130 ads have been selected for the web developer position and 120 ads for the software developer position. This dataset may be developed in the future by conducting weekly draws of job offers and tracking their evolution over time.

The two entities retrieved from the database are the jobs collection, respectively ESCO collection. In order to make the connection between the two and the educational plan, the creation of a static entity in MS Power Bi was necessary, which incorporates the knowledge, the skills and the responsibilities from the FEAA curriculum for the Economic Informatics specialization.

Within the educational plan, four disciplines, that correspond the two analyzed job profiles, have been evaluated. Thus, the disciplines of Internet Programming and Web Development combine the knowledge and skills necessary for a web developer, but also the technologies and programming languages essential to such a position within an organization. Moreover, the disciplines of Visual Programming Environments and Object Oriented Programming prepare the future graduate for a software developer position.

Therefore, by using the Business Analytics tool, we wanted to make two comparisons between the employers' requirements, the ESCO classifications and the curriculum for the two types of analyzed jobs. The analysis structure for the two professions is similar, centered around the ESCO skills and abilities, which are in fact, as previously mentioned, the bridge between the labor market and education. The first analysis that was carried out is for the web developer profession and its structure can be seen in Figure 6.



Web developer

Job offers

Example no. 1 **PHP Developer**
Knowledge
- Good knowledge of back-end technologies;
- Good knowledge of Object-Oriented Programming concepts;
- Good knowledge of front-end technologies;
- Experience with frameworks.

Skills
- Source code debugging;
- Translating concepts and design into user interfaces;
- Web application development;
- Troubleshooting and continuous improvement of web applications;
- Continuous research, discovery, evaluation and implementation of new technologies.

Programming languages and technologies
- PHP, HTML, CSS, JavaScript, MySQL, framework Symfony.

Example no. 2 **Web Developer**
- Familiar with JavaScript, jQuery, HTML, CSS, WordPress, PHP, MySQL.

Example no. 3 **PHP Developer**
Knowledge
- Experience in web development;
- eCommerce experience;
- Advanced knowledge of PHP, HTML, CSS, JavaScript;
- Familiar with OOP, MVC, Smarty;
- Open attitude towards new technologies;
- Advanced knowledge of querying and using databases.

Responsibilities
- Implementation of PHP, JavaScript and related libraries.

ESCO

Knowledge

ICT debugging tools
World Wide Web Consortium standards
Computer programming
Content development processes
Integrated development environment software

Skills

Analyse software specifications
Collect customer feedback on applications
Create flowchart diagram
Debug software
Develop software prototype

Programming languages and technologies
- Java, PHP, C++, C#, Python;
- Javascript, HTML, CSS;
- ASP.NET, AJAX, JavaScript Framework;
- Query languages;
- Object-Oriented Programming.

Educational plan

Knowledge

Knowledge and understanding of information technologies used in economics and business;

Ability to understand economic language for executing tasks and implementing applied IT projects in the business environment;

Ability to understand economic processes for efficient use of software products for business;

Skills and responsibilities

Ability to understand economic language for executing tasks and implementing applied IT projects in the business environment;

Ability to understand economic processes for efficient use of software products for business;

Ability to think logically, process data efficiently with programmable algorithms;

Ability to use modeling and scientific computing tools to accomplish a task;

Programming languages and technologies used in the Internet Programming and Web development disciplines
- PHP, Java;
- Bootstrap, Joomla, UiKit, WordPress;
- MySQL;
- HTML, CSS.

Figure 6: Power BI analysis for the web developer professions

After analyzing the ads and applying some filters, we wanted to select three examples of job offers for each of the two occupational profiles analyzed. The first drawn example respects the ESCO formulations, focusing on the knowledge and skills desired by the future employee. A skill specified by the employer, conform to the standard formulations, is related to the developer's interaction with the application's user or client, requiring him the ability to transpose the user's vision and requests into the application's design.

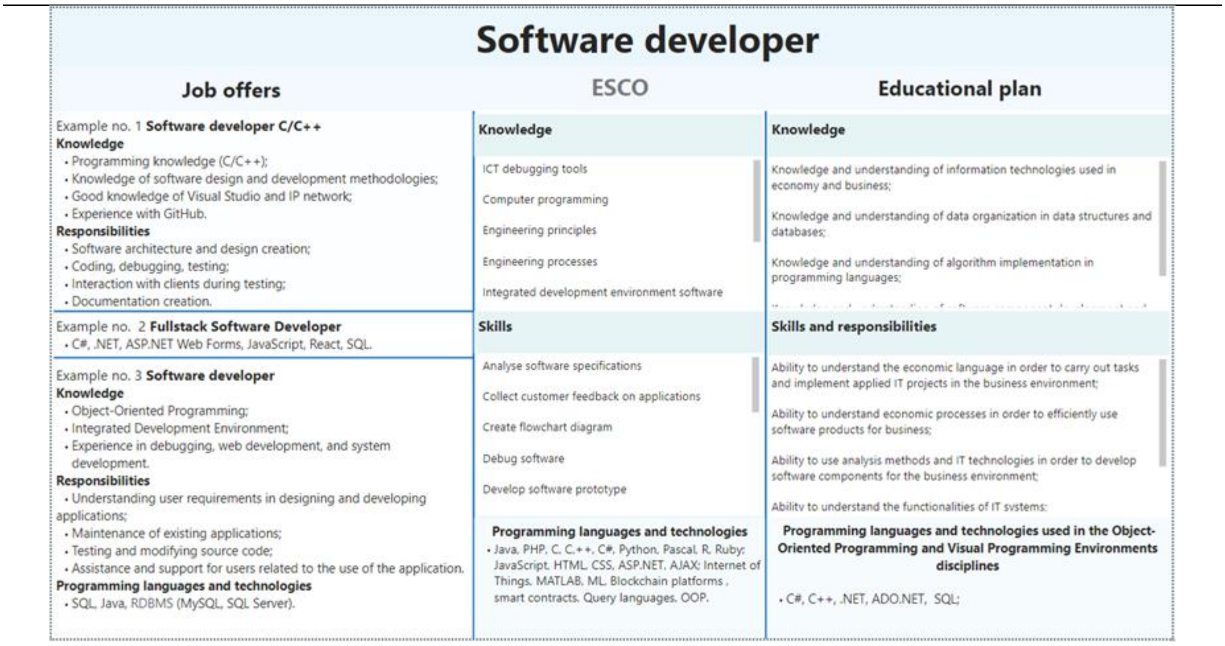
The second example is an ad that doesn't comply with ESCO's standard formulations, by listing only the technologies necessary to apply to the job, followed by a brief description of the company. The last ad shown combines the structure of the two examples presented previously, with particular emphasis on technologies and programming languages required within the knowledge category.

The ads' formulations varies and most employers do not follow the ESCO ones. But looking at the dataset in Power BI, we can see a common element between the ESCO standard and the Romanian employers' formulations, namely the emphasis placed by the latter for programming languages and technologies required for the offered position, these being found in ESCO within the last section. Also, because the emphasis of the employers falls on technologies and languages, an analysis of them was carried out in Power BI, namely about the ratio filled by each programming language of the total job offers.

From graphical views realized, we could see that the largest share is owned by the Java language, being the most required programming language for web development jobs, then followed by JavaScript, HTML, CSS and PHP.

The main languages desired for a specialist in the web development field is also covered in the undergraduate program, exception JavaScript, which is in the top of the most known and used languages for the front-end development of web applications. In the offer made by the Economic Informatics program, within the discipline sheets for the two analyzed subjects, we can see their emphasis on CMSs (website content management systems), but also on other frameworks, that fill a very small ratio in the analysis.

The second analysis carried out is for the software developer profession. In a similar way, we have drawn three ads with different structures for the job offers part, the ESCO's requirements and skills have been adapted according to this type of job, while in the education plan we included the knowledge, skills and responsibilities for the two disciplines that prepare a graduate to occupy a software developer position (Figure 7).



Software developer

Job offers

Example no. 1 **Software developer C/C++**
Knowledge
- Programming knowledge (C/C++);
- Knowledge of software design and development methodologies;
- Good knowledge of Visual Studio and IP network;
- Experience with GitHub.

Responsibilities
- Software architecture and design creation;
- Coding, debugging, testing;
- Interaction with clients during testing;
- Documentation creation.

Example no. 2 **Fullstack Software Developer**
- C#, .NET, ASP.NET Web Forms, JavaScript, React, SQL.

Example no. 3 **Software developer**
Knowledge
- Object-Oriented Programming;
- Integrated Development Environment;
- Experience in debugging, web development, and system development.

Responsibilities
- Understanding user requirements in designing and developing applications;
- Maintenance of existing applications;
- Testing and modifying source code;
- Assistance and support for users related to the use of the application.

Programming languages and technologies
- SQL, Java, RDBMS (MySQL, SQL Server).

ESCO

Knowledge

ICT debugging tools
Computer programming
Engineering principles
Engineering processes
Integrated development environment software

Skills

Analyse software specifications
Collect customer feedback on applications
Create flowchart diagram
Debug software
Develop software prototype

Programming languages and technologies
- Java, PHP, C, C++, C#, Python, Pascal, R, Ruby; JavaScript, HTML, CSS, ASP.NET, AJAX; Internet of Things, MATLAB, ML, Blockchain platforms , smart contracts, Query languages, OOP.

Educational plan

Knowledge

Knowledge and understanding of information technologies used in economy and business;
Knowledge and understanding of data organization in data structures and databases;
Knowledge and understanding of algorithm implementation in programming languages;

Skills and responsibilities

Ability to understand the economic language in order to carry out tasks and implement applied IT projects in the business environment;
Ability to understand economic processes in order to efficiently use software products for business;
Ability to use analysis methods and IT technologies in order to develop software components for the business environment;
Ability to understand the functionalities of IT systems;

Programming languages and technologies used in the Object-Oriented Programming and Visual Programming Environments disciplines
- C#, C++, .NET, ADO.NET, SQL;

Figure 7: Power BI analysis of the software developer profession

The first example of ad includes knowledge that can also be found in the ESCO formulations, such as use of debugging tools and integrated development environments, programming knowledge programming, but also certain skills, such as: identifying user requirements and carrying out technical documentation. The second announcement is a modest one, listing only the languages and the technologies desired by the employer. The third example includes ESCO formulations, and the part of abilities is called responsibilities, which actually correspond to the standard abilities. While analyzing this category of jobs as well, we can see that there is no homogeneous structure, each ad complying more or less with the standards, but, similarly to the ads for web development, the common element is the indication of the principal technologies and programming languages required.

Most of the ads feature a string of such technologies and languages, followed by the presentation of the facilities and benefits offered by the company. In the ads, the emphasis falls on the practical side of these professions, which is why several graphs representative of these have been made in Power Bi, providing information about their ratio within the ads. The main programming language required by companies for such positions is C, then followed by Java and, with a smaller ratio, C# and C++. There is also an emphasis on databases knowledge and their query language. Within those two FEAA disciplines that prepare a graduate for such a job, the studied languages are C++, C# and the SQL query standard.

5. Conclusion

The job offer analysis highlighted the fact that, within the same job portal, offers do not have a consistent structure for their descriptions. Emphasis is placed on programming languages and technologies, with this category of information being included in all job descriptions. Referring to ESCO, these types of requirements are found among the items in the Languages and Technologies category. Even though the languages and technologies covered are not specified in the formulations for competencies, skills, and responsibilities of the curriculum, they are explicitly mentioned in the discipline sheets. The ESCO occupational taxonomy is the universal reference for formulating descriptions of different occupations. Although we do not encounter the same syntax in the job offer formulations, semantically, they correspond to the ESCO formulations.

The implementation was based on the Python, MongoDB, and MS Power BI triad, which made it possible to carry out the process of retrieving data from the targeted websites, storing it in an information repository that constituted the data source for MS Power BI visual analyses. Python is recognized for its increased capacity for processing web data, MongoDB allows for storing large volumes of Big Data, and allows for rapid real-time access to this data, and MS Power BI is a leader in the family of support tools for Business Analytics.

The analysis carried out presents limitations, in that only two occupational profiles were considered, and the job postings were analyzed on a single profile site. In the future, we intend to approach multiple occupational profiles in the IT field and apply scraping algorithms to other job offer portals. Furthermore, there is a need to perform a semantic analysis of jobs extracted through web scraping processes and classify them into ESCO classifications using artificial intelligence. This will be achieved by applying Natural Language Processing algorithms, which will analyze and categorize the skills and competencies required for a job according to ESCO standards.

References

[1]. Gheorghe, M., Mihai, F.C., & Dârdală, M. (2018). Modern techniques of web scraping for data scientists, Revista Română de Interacțiune Om-Calculator, 11 (1), 63-75

[2]. Guilmette, A. (2020). Workflow Automation with Microsoft Power Automate: Achieve digital transformation through business automation with minimal coding, Packt

[3]. Hows, D., Plugge, E., Membrey, P., & Hakins, T. (2013). The Definitive Guide to MongoDB, Apress

[4]. Mitchel, R. (2018). Web Scraping with Python: Collecting More Data from the Modern Web, O'Reilly Media Inc.

[5]. Mullakara, N., & Asokan, A.K. (2020). Robotic Process Automation Projects: Build real-world RPA solutions using UiPath and Automation Anywhere. Packt

[6]. Norcross, J.S., Schuessler, J. H., & Jones, D. G. (2020). Web Scraping to Inform Curriculum Decisions, Journal of Higher Education Theory and Practice, 20(10), 100-112

[7]. Oprea, S., & Bâra, A. (2022). Why Is More Efficient to Combine BeautifulSoup and Selenium in Scraping For Data Under Energy Crisis. "Ovidius" University Annals, Economic Sciences Series, 146-152. Accesibil la: https://stec.univ-ovidius.ro/html/anale/RO/2022-issue2/Section%201%20and%202/19.pdf

[8]. Patel, J. M. (2020). Getting Structured Data from the Internet: Running Web Crawlers/Scrapers on a Big Data Production Scale, Apress

[9]. Wade, R. (2020). Advanced Analytics in Power BI with R and Python: Ingesting, Transforming, Visualizing, Apress