

Ministry of Education and Science of Ukraine  
Black Sea Universities Network

# ODESA NATIONAL UNIVERSITY OF TECHNOLOGY

International Competition of  
Student Scientific Works

# BLACK SEA SCIENCE 2022 PROCEEDINGS



ODESA, ONUT 2022

Ministry of Education and Science of Ukraine

Black Sea Universities Network

Odesa National University of Technology

International Competition of Student Scientific Works

# **BLACK SEA SCIENCE 2022**

**Proceedings**

Odesa, ONUT 2022

**Editorial board:**

**Prof. B. Iegorov**, D.Sc., Professor, Rector of the Odesa National University of Technology, Editor-in-chief

**Prof. M. Mardar**, D.Sc., Professor, Vice-Rector for Scientific and Pedagogical Work and International Relations, Editor-in-chief

**Dr. I. Solonytska**, Ph.D., Associate Professor, Director of the M.V. Lomonosov Technological Institute of Food Industry, Head of the jury of «Food Science and Technologies»

**Dr. Yu. Melnyk**, D.Sc., Associate Professor, Director of the G.E. Weinstein Institute of Applied Economics and Management, Head of the jury of «Economics and Administration»

**Dr. S. Kotlyk**, Ph.D., Associate Professor, Director of the P.M. Platonov Educational-Scientific Institute of Computer Systems and Technologies “Industry 4.0”, Head of the jury of «Information Technologies, Automation and Robotics»

**Prof. O. Titlov**, D.Sc., Professor, Head of the Department of Oil and Gas Technologies, Engineering and Heat Power Engineering, Head of the jury of «Power Engineering and Energy Efficiency»

**Prof. G. Krusir**, D.Sc., Professor, Head of the Department of Ecology and Environmental Protection Technologies, Head of the jury of «Ecology and Environmental Protection»

**Dr. V. Kozhevnikova**, Ph.D., Associate Professor, of the Department of Hotel and Catering Business, Technical Editor

**Black Sea Science 2022:** Proceedings of the International Competition of Student Scientific Works / Odesa National University of Technology; B. Iegorov, M. Mardar (editors-in-chief) [*et al.*]. – Odesa: ONUT, 2022. – 749 p.

Proceedings of International Competition of Student Scientific Works «Black Sea Science 2022» contain the works of winners of the competition.

The author of the work is responsible for the accuracy of the information.

### **Organizing committee:**

**Prof. Bogdan Iegorov**, D.Sc., Rector of Odesa National University of Technology, Head of the Committee

**Prof. Maryna Mardar**, D.Sc., Vice-Rector for Scientific and Pedagogical Work and International Relations of Odesa National University of Technology, Deputy Head of the Committee

**Prof. Baurzhan Nurakhmetov**, D.Sc., First Vice-Rector of Almaty Technological University (Kazakhstan)

**Prof. Michael Zinigrad**, D.Sc., Rector of Ariel University (Israel)

**Prof. Plamen Kangalov**, Ph.D., Vice-Rector for Academic Affairs of “Angel Kanchev” University of Ruse (Bulgaria)

**Prof. Heinz Leuenberger**, Ph.D., Professor of the Institute of Ecopreneurship of University of Applied Sciences and Arts (Switzerland)

**Prof. Edward Pospiech**, Dr. habil., Professor of the Institute of Meat Technology of Poznan University of Life Sciences (Poland)

**Prof. Lali Elanidze**, Ph.D., Professor of the Faculty of Agrarian Sciences of Iakob Gogebashvili Telavi State University (Georgia)

**Dr. Dan-Marius Voicilas**, Ph.D., Associate Professor of the Institute of Agrarian Economics of Romanian Academy (Romania)

**Prof. Stefan Dragoev**, D.Sc., Vice-Rector for Scientific Work and Business Partnerships of University of Food Technologies (Bulgaria)

**Prof. Jacek Wrobel**, Dr. habil., Rector of West Pomeranian University of Technology (Poland)

**Dr. Mei Lehe**, Ph.D., Vice-President of Ningbo Institute of Technology, Zhejiang University (China)

**Dr. V. Kozhevnikova**, Ph.D., Associate Professor of the Department of Hotel and Catering Business of Odesa National University of Technology, Secretary of the Committee

## INTRODUCTION

International Competition of Student Scientific Works “Black Sea Science” has been held annually since 2018 at the initiative of Odesa National University of Technology (formerly Odesa National Academy of Food Technologies) with the support of the Ministry of Education and Science of Ukraine. It has been supported by Black Sea Universities Network (the Association of 110 higher education institutions from 12 countries of the Black Sea Region) since 2019, and by Iseki-FOOD Association (European Integrating Food Science and Engineering Knowledge into the Food Chain Association) since 2020.

The goal of the competition is to expand international relations and attract students to research activities. It is held in the following fields:

- Food science and technologies
- Economics and administration
- Information technologies, automation and robotics
- Power engineering and energy efficiency
- Ecology and environmental protection

The jury includes both Ukrainian and foreign scientists. In the 4 years that the competition has been held, the jury included scientists from universities of 24 countries: Angola, Azerbaijan, Benin, Bulgaria, China, Czech Republic, France, Georgia, Germany, Greece, Israel, Italy, Kazakhstan, Latvia, Lithuania, Moldova, Pakistan, Poland, Romania, Serbia, Slovakia, Switzerland, Turkey, USA.

At the same time, every year the geography has expanded and the number of foreign jury members has increased: from 46 jury members representing 25 universities from 12 countries in 2018, to 73 jury members of the 46 universities from 19 countries in 2022.

More than a thousand student research papers have been submitted to the competition from both Ukrainian and foreign institutions from 25 countries: China, Poland, Mexico, USA, France, Greece, Germany, Canada, Costa Rica, Brazil, India, Pakistan, Israel, Macedonia, Lithuania, Latvia, Slovakia, Romania, Kyrgyzstan, Kazakhstan, Bulgaria, Moldova, Georgia, Turkey, Serbia.

The interest of foreign students in the competition grew every year. In 2018, the students representing 15 institutions from 7 countries have submitted 33 works. In 2021 the number of submitted works increased to 73, authored by the students of 40 institutions from 18 countries.

The competition is held in two stages. In the first stage, student research papers are reviewed by members of the jury who are experts in the relevant fields. In the second stage of the competition, the winners of the first stage have the opportunity to present their work to a wide audience in person or online.

All participants of the competition and their scientific supervisors are awarded appropriate certificates, and the scientific works of the winners are included in the electronic proceedings of the competition. Every year the competition receives a large number of positive responses from Ukrainian and foreign colleagues with the desire to participate in the coming years.

# **3. INFORMATION TECHNOLOGIES, AUTOMATION AND ROBOTICS**

## DEVELOPMENT OF SOFTWARE MODULE FOR ANALYSIS OF IT SPECIALISTS' LABOR MARKET

**Author:** Anhelina Dub

**Advisor:** Anna Zhurba

Ukrainian State University of Science and Technologies (Ukraine)

***Abstract.** The paper describes how software module was developed to analyze the labor market of IT professionals using the Python programming language in the integrated programming environment. The software module screens, crawls, parses and exports data from specialised sites. The software module allows to find, structure and export data to CSV and TXT files. The sample of key parameters was investigated by means of business analytics.*

*The developed program can be expanded with additional functions, supplemented by a graphical interface, uploaded to the web hosting.*

*The software module provides processing of a large array of unstructured information from vacancy announcement sites, reduces the amount of routine manual operations and provides an opportunity to focus stakeholders' attention on key priorities.*

*The aim of the work is to develop a parsing software module for automated collection and analysis of open position data to determine which knowledge and skills are most important for employers in the IT industry.*

***Keywords:** web-scraping, data mining, business intelligence, programm module, Python programming language.*

### I. INTRODUCTION

Collection and analysis of information about the labor market is necessary constantly, and especially relevant in modern conditions. The economic crisis as a result of the spread of COVID-19 has affected entire industries, the number of employees has decreased in almost all countries. As of the beginning of 2022, the unemployment rate in France is 7.5%, Sweden – 8%, Ukraine – 10.4%, Turkey – 11.5%, Spain – 14.1%, Georgia – 18.5%, South Africa – 34.9% [1-3].

The information technology sector is able to bring the world economy out of the crisis and ensure social development. The situation in the IT sector is characterized by a parallel staff shortage (lack of highly qualified specialists) and the influx, and consequently, the difficulty of employment for junior beginners. Also, today the amount of information in general, and about the labor market in particular, causes both congestion in the number of sources, and the omission of important elements for individuals and organizations.

The aim of the work is to develop a parsing software module for automated collection and analysis of open position data to determine which knowledge and skills are most important for employers in the IT industry.

Stakeholders of such a project are:

1) active job seekers – for a more efficient and rational search (review of more vacancies); for faster selection of positions that correspond to the available

qualifications; to reduce mental stress and resource-intensive search operations, to improve the psychological state;

2) employers – to determine competitive salary, understanding of popular technologies;

3) institutions of higher education, centers of advanced training, trainers – for the formation and adjustment of training programs in accordance with market requirements;

4) students and graduates – to get acquainted with the existing requirements, choose the right disciplines, prepare relevant research topics; to study what is needed and relevant in the labor market;

5) IT specialists – to determine career opportunities (it is usually considered that every two years you should learn a new programming language to maintain competitiveness and improve professional level);

6) professionals who plan to change the field of activity – to understand the demands of the industry, focus retraining on the necessary skills;

7) government – increasing the number of employees reduces social tensions (the population becomes more self-sufficient, does not rely on social benefits), provides economic growth (potentially more employees provide more payments to the budget and GDP growth);

8) investors – as in a promising direction in specialized search portals, data aggregators, startups of analytical services, etc. invest in venture capital organizations, media holdings, recruitment firms, etc.

These parties are potential users of the software module.

The results of the work – a systematic list of vacancies and summary in tabular and graphical form of employers' requirements for a position (knowledge and skills of specialists, programming languages, years of experience, education, etc.) – can be the basis for creating a resume tracker, cover letters and interviews; formation of a (self) training plan, which will organize an effective job search and acquaint a wide range of people with popular requirements, increase the level of education, provide an understanding of trends in the IT industry.

## II. LITERATURE ANALYSIS

### 2.1. Basic concepts of web scraping

Web scraping technology is best suited to fulfill the task set in the paper – the development of a software module for the analysis of the IT professionals' labor market, from the initial collection of information to its processing and analysis.

Web scraping (also scraping, web collection, data retrieval from the web) is:

1) automated collection of information from the World Wide Web;

2) technology, when a computer program extracts data;

3) the practice of data collection in any way other than the program that interacts with the API. This is most often achieved by writing an automated program that sends a request to a web server, reads the data (usually in the form of HTML and other files that make up web pages), and then analyzes it to extract the necessary information.

Web scraping can be a stand-alone tool for targeted information retrieval, as well as a component of web development used for web indexing, web mining and data mining, online monitoring of price changes and their comparison, to monitor competition, collection of other data. It is also used for web automation.

Automated data gathering from the Internet exists almost as much as the Internet itself. Although web scraping is not a new term, in the past this practice was better known as screen scraping, data mining, web miner, web harvesting or similar variations. In general, the term web scraping is now more popular, although programs that go through multiple pages are usually called web crawlers, spiders, spiderbots, and scraping programs themselves are bots [4, c. ix preface].

The topic of scraping is complex, as it is not possible to be limited to one language or technology, you need knowledge of databases, web servers, HTTP, HTML, Internet security, image processing, data science and other tools. Data management is especially relevant in the context of big data, when only a small part of them is structured.

In practice, web scraping covers a wide range of software techniques and technologies, such as data analysis, natural language parsing and information security [4, p. x preface].

Scraping usually involves the use of two tools: a crawler, which navigates the website to extract (extraction, fetching, collecting, capturing) HTML from the page; and a parser that finds an HTML tag in pages to extract relevant information.

Parsing (from the Latin pars – part of speech) – the study of a predetermined sequence of characters (from natural or computer language or data structure), for further grammatical analysis of the structure.

Parser is a full-fledged program or part of it designed for parsing.

In the process of parsing text components are brought into a coherent form, design the data structure, most often – in the tree of the syntactic structure of a given sequence, which can be further processed. As a rule, the work of parsers goes through two stages: the identification of meaningful tokens (lexical analysis), and then a parsing tree is formed.

A token is an object formed from a lexeme (a word as an independent unit of meaning) in the process of lexical analysis. In applied programming, the concept of token and its lexeme may not differ [5].

The parser is an extension or part of the scraper.

Websites consist of 3 main components: HTML, which contains the content of the website; CSS, which contains its style, and Javascript, which contains elements that update themselves without refreshing the page, or that you can interact with (such as a drop-down menu).

For scraping each individual site requires its own code [6]. The general process of web scraping is shown in Fig.1. For each new resource, you need to go through the whole process from start to finish to get specific data.

After collecting data in the set, you can search, change the presentation format, copy to a table or database (DB).



Fig. 1. Process of data-mining

Source: author’s design

## 2.2. Usage

Web-scraping (web / data / text content mining, text analytics) is needed for research, analysis; formation of personal collections, personal or organizational knowledge bases; information retrieval, knowledge discovery, etc. [7, p. 33].

Common applications are collecting data from public users of social networks as potential customers, comparing prices, aggregating news, collecting ads for rent or sale of housing, weather, outlets, tracking competitors or reviews of their own brand, systematization of internal documentation (email, PDF files etc.); analysis of consumer sentiment, machine learning, search for cheap hotels / tickets, creation of additional or main content (blog posts, creation of own aggregator). In the table 1 we give specific examples.

Table 1. Examples of using web scraping

Company / product	Description
Wandex (World Wide Web Wanderer)	the first web robot; for measuring the size of the WWW (established in 1989), worked since June 1993, in 2011 resumed work
JumpStation	the first search engine based on crawler, not on the work of administrators (December 1993)
WebCrawler	the first full-text search engine (1994)
Salesforce, eBay	first Web API and API Crawler (2000)
Google Search	indexing the Internet with crawlers and later AI
Google Knowledge Panel	pulling up query data from Wikipedia, the official site, and other sources
Facebook Telegram	when you insert a URL into a message or post, the title and thumbnail of the linked page is pulled up
Google Alerts Visualping Wachete Service	monitoring changes on sites, new articles on request
Beautiful Soup	library for scraping in Python (2004), is considered the most complex and advanced

Source: author’s summary

## 2.3. Alternatives

Sources of structured data are:

1) official APIs – limited, expensive. Amazon AWS and Google (API Discovery service) provide end users with free tools, services, and publicly available parsing data.

As for the sites-virtual bulletin boards (Indeed, Work.ua, Rabota.ua), the API is available only to employers (i.e., for registered organizations, for a fee and have restrictions). Part of the functionality is not available for ua.indeed.com, and dou.ua does not have an API. In addition, APIs can be closed at any time;

2) ready-made databases – such kits are available in various formats (SQL, Excel, JSON online forms, log files of automated systems, etc.). But data is rapidly becoming obsolete and needs constant updating. In the case of vacancy announcements, a number of sites do not even publish the archive, removing already inactive publications from public access;

3) RSS or Atom feeds – subscriptions to site updates using XML formats, consists of full text or part and metadata, is read by special programs-readers / aggregators (such as Feedly, Feedbro, Inoreader, TinyTinyRSS, closed in 2012 Google Reader). Atom or RSS feeds save a lot of time tracking updates, and readers allow you to view content changes centrally without going directly to the sites themselves. The most common uses are frequently updated sites, blogs, news sites, web magazines, podcasts, etc. RSS is actually an XML document and belongs to the sources of structured data.

One reader allows subscriptions to many sites or sections (however, services often set a free usage limit above which a fee is charged). Also, depending on the settings, RSS usually displays only the last 20 articles, depending on the date of subscription to the feed. StackOverflow Jobs states that RSS is their API [8], but the page displays only the last 1001 posts, while jobs.dou.ua displays the latest 51 posts in RSS.

The format helps to disseminate information without e-mails or personal messages through bots, etc., and also helps to distribute publications on a “post once-syndicate everywhere” basis, when publications appear on a single official resource individuals or organizations and then distributed to all platforms used.

As this reduces the site traffic in some way (users do not come many times to check for updates), and the number of views / conversions decreases, the format is not very popular, many users do not know about it at all, and for notifications about news sites they subscribe to their pages at social networks. In addition, parsing RSS is often not appropriate, as it usually does not display the news archive, but only the last 20 posts.

In general, the advantages of using web scraping compared to these methods are:

- data structured according to queries;
- the owners of the sites pay considerable attention to data management and maintenance, and the work is done more carefully than with the API documentation;
- there is practically no limit of use, the main thing – not to overload servers;
- anonymous access;
- data is visible immediately – without separate requests to the organization or waiting to open access to the API.

We analysed existing tools (30 the most used and the like) and defined that they have a number of limitations: they are often proprietary, paid (or freemium), the settings are limited by the functionality of services or applications. As often in such cases, the best solution in modern conditions is to develop your own software product of a certain specification.

### III. OBJECT, SUBJECT, AND METHODS OF RESEARCH

The object of research is the process of collecting, processing and analyzing data on the labor market of IT professionals.

Subject is software tools for labor market analysis of IT professionals.

Research methods used in the paper are the following: comparative analysis, web scraping, structuring and organization of data, parsing of texts, statistical methods of analysis of poorly structured data of large volume, tabular and graphical.

To solve the problem, we choose the Python language with the appropriate libraries for scraping, parsing and visualization.

In 2021, 38% of the developers who identified Python as one of their three main programming languages used it to develop web parsers / scrapers / crawlers (survey [9]). In general, being the most popular language, Python is also the most popular for web scraping, because it:

- is well designed, considered easy to learn;
- is well documented on official resources;
- has good support, including due to popularity, there are a large number of publications, forums, tutorials;
- is suitable for rapid development;
- is free;
- has an active community;
- contains a large number of tools in the standard package and additional libraries;
- creates code that is readable, easy to further support, can be reused [10, p.13-17].

Python is a powerful automation tool, contains a rich collection of libraries, has an intuitive syntax. The 3 best known approaches to processing web pages in Python are regular expressions, BeautifulSoup and lxml modules [11, p.39-45]. All of them were used in this research to gain the most complete data.

Python is usually able to avoid various scraping restrictions by using multiple proxies, retrying downloads, and configuring a user agent. It is able to read data from web pages with dynamic content, which is downloaded not from the internal database of the server, but through Javascript. Uploading / exporting data can be done in a database or table (SQL, CSV, XLS, JSON, other formats).

### IV. RESULTS

#### 4.1. Project structure

For the project it is decided to choose 3 sites, the total number of ads – 500-1000, geography – the world, time period – April 19 – May 20, 2021.

In the process of preparation, an alternative approach was considered – scraping of those already employed with LinkedIn, but the profiles may not be complete, not updated, freelance or the wrong sector is indicated. Vacancy announcements are more reliable.

The project methodology is similar to the general scheme in Fig.1, but supplemented in accordance with the tasks (Fig.2).

The process of developing a software module corresponds to such steps:

- open the file, on the site with a web scraper collect links from the list from one page to individual pages of announcements, and using pagination – from the rest;
- collect data on the pages of the ads themselves, reading HTML-tags;
- using loops “for”, go to full descriptions;
- extract data for each tag and save to the table;
- analyze key parameters;
- determine the frequency of use of key terms;
- create elements of information panel by visualization with business analytics tools.

Having decided on the sequence of work, we move on to the next stage.

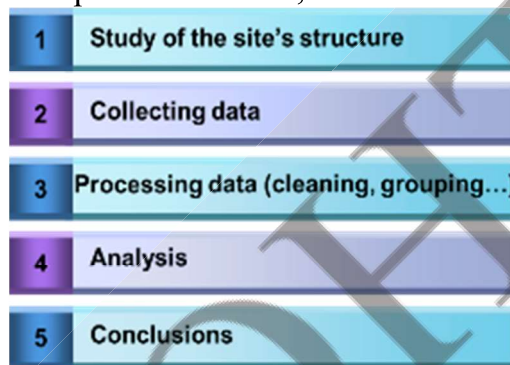


Fig. 2. Blocks of the project

#### 4.2. Selection of sources

At this stage, it is important to consider the internal documentation of the sites. Scraping is absolutely legal, if it does not contradict the rules of site owners [4, p. ix, 263-277]. In the lawsuits that actually took place (plaintiff LinkedIn), the decisions were in favor of the defendants: what is visible to users on the sites can be collected and automatically [12]. However, the number of LinkedIn ads in the IT sector as of May 4, 2021 in the United States alone for the last 24 hours is 5,792 positions, last week – 35,196, month – 115,436, any time – 159,340. The quality of such information is not very high – too many duplicate results, and information cannot be fully filtered at the collection stage on the site itself. Therefore, given these two factors, LinkedIn will not be included in the analyzed sites.

The terms of use of the largest aggregator of vacancies in the world – Indeed – explicitly prohibit the scraping of any information from the site with an indication of the possibility of appropriate lawsuits [13].

Considering other options, the choice is summarized in table 2.

Selection criteria are availability of data and a sufficient number of publications for analysis.

Table 2. Site selection and number of job ads

Site	Number
Rabota.ua	8 647
Rabota.ua – other countries	119
StackOverflow Jobs	1 337
Dou.ua	10 800
Усього	20 903

Source: author’s summary based on data from [14-16]

### 4.3. Project implementation

To obtain the most up-to-date set of vacancy announcements, we used scraping for all pages with categories “Information Technology” on selected sites: Rabota.ua, Dou.ua, StackOverflow Jobs by:

- code planning;
- inspect HTML;
- scrape – the first function – to collect links;
- pagination and
- export (data storage)

We saved the results (export data) in the CSV file, because it can be opened in both text and spreadsheet editors and has no line limit [17].

All pages with a list of links and relevant available information were successfully collected from three sites (Fig. 3).

title	company	salary	location	link
Tools Lead	kwalee		Royal Leai	<a href="https://stackoverflow.com/jobs/527954/">https://stackoverflow.com/jobs/527954/</a>
Senior Java Engineer - Remote	wallethub		No office	<a href="https://stackoverflow.com/jobs/527950/">https://stackoverflow.com/jobs/527950/</a>
(Senior) Full-Stack Web Developer	wunder mobility		Hamburg,	<a href="https://stackoverflow.com/jobs/527937/">https://stackoverflow.com/jobs/527937/</a>
(Senior) Software Engineer (Python)	delivery hero se		Berlin, Ge	<a href="https://stackoverflow.com/jobs/527921/">https://stackoverflow.com/jobs/527921/</a>
Blockchain Rust Developer	komodo platform		No office	<a href="https://stackoverflow.com/jobs/527914/">https://stackoverflow.com/jobs/527914/</a>
PHP Software Engineer	trivago nv		Düsseldorf	<a href="https://stackoverflow.com/jobs/527910/">https://stackoverflow.com/jobs/527910/</a>
Solution Engineer (m/f/d)	relayr		Berlin, Ge	<a href="https://stackoverflow.com/jobs/527907/">https://stackoverflow.com/jobs/527907/</a>
DevOps Engineer (AWS) Required	montash		No office	<a href="https://stackoverflow.com/jobs/527882/">https://stackoverflow.com/jobs/527882/</a>
DevOps Engineer (m/f/d) Cloud	covestro deutschland ag		Leverkuse	<a href="https://stackoverflow.com/jobs/527878/">https://stackoverflow.com/jobs/527878/</a>

Fig. 3. Fragment of the intermediate table

To collect a complete description of vacancies, we must go to each of the collected links and take data from the relevant fields.

Proceeding with project is to use:

- crawl – the second function – to collect complete descriptions of requirements
- cleaning data and merging descriptions (the next step is post-processing, data cleanup – removing duplicates, non-necessary spaces, posts that appear several times and so on);
- parse – the third function – the parsing of the text.

For convenience, we summarized all job descriptions in one text file. In the process of parsing job descriptions, we performed tokenization, deleted stop words (a, the, etc.), create n-grams (stable phrases) and identified the words that are used most often.





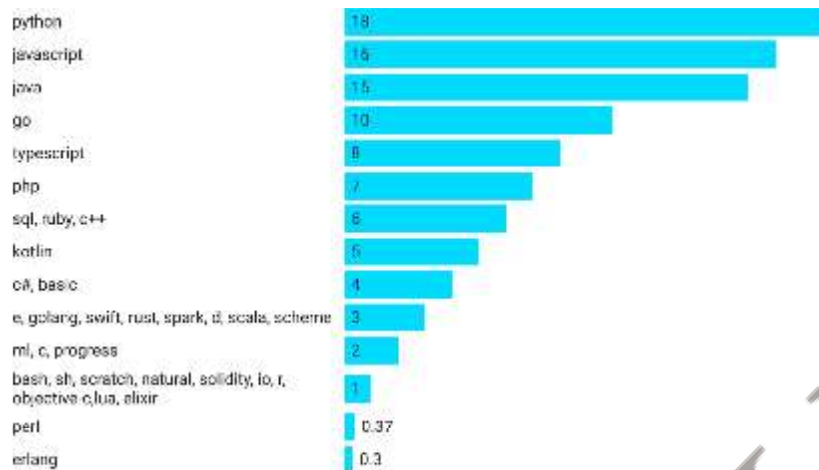


Fig. 7. Programming languages prevalence, %

Source: author’s design according to the sample [14-16]

The level of positions is shown in Fig. 8. In 44% of cases, Senior positions are open.



Fig. 8. Level of vacancies, %

Source: author’s design according to the sample [14-16]

*Location*

In Ukraine, most vacancy announcements in the IT sector were in Kyiv, Kharkiv, Lviv, Odesa and Dnipro (Fig.9).

In the distribution *by country*, the leaders in the number of open positions are Germany, Great Britain, the Netherlands, the United States, Austria and Japan (Fig. 10). It is worth noting that so far, most vacancies have been marked “remote” or “no office location”. This is due to both external (COVID-19) and internal factors (traditionally greater loyalty of employers in this sector to work remotely).



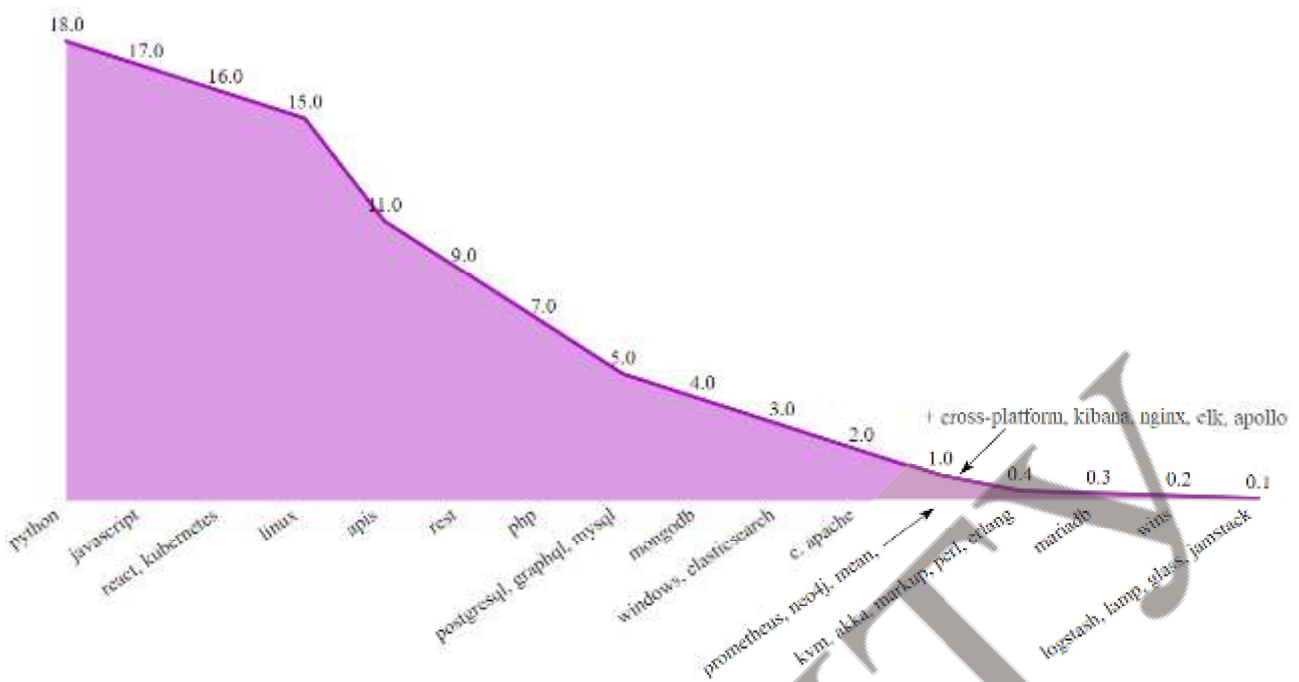


Fig. 11. 35 techstack tools, %  
 Source: formed by the author according to the sample

The key areas (specialization) in ads are machine learning, software as a service and artificial intelligence (Fig. 12).

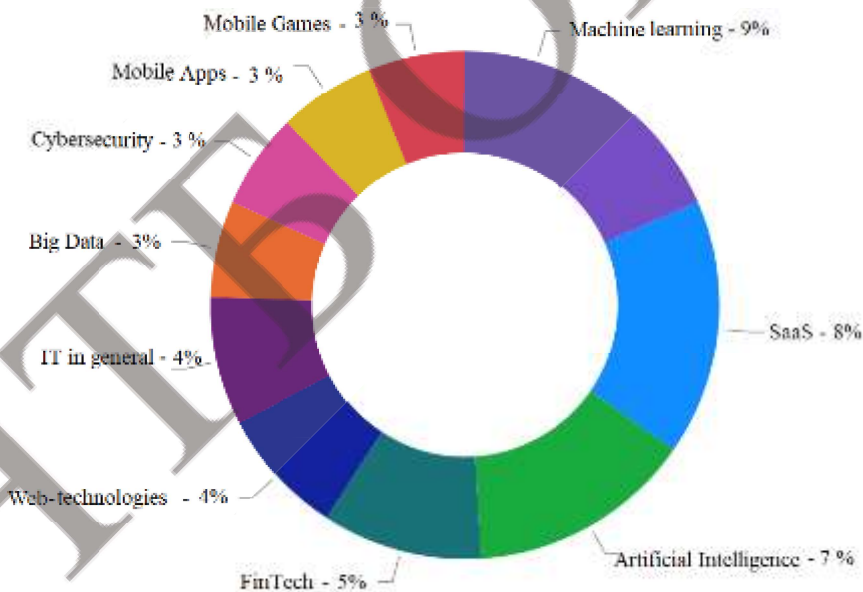


Fig. 12. Subsectors  
 Source: created by the author according to the sample

*Soft Skills*

Among the “flexible” skills in the ads there were not mentioned problem solving, public speaking and hard work (0 matches with the dictionary). Instead, sociability, leadership skills and creativity are leaders in demand (Fig. 13).

*Salary* for the unification of conclusions was analyzed in dollars. The exchange rates of the sampled currencies (Australian and Singapore dollars, Swiss francs, Danish

and Swedish krona, euros, British pounds, Japanese yen and Polish zlotys) to the dollar were taken from Google Finance [20].

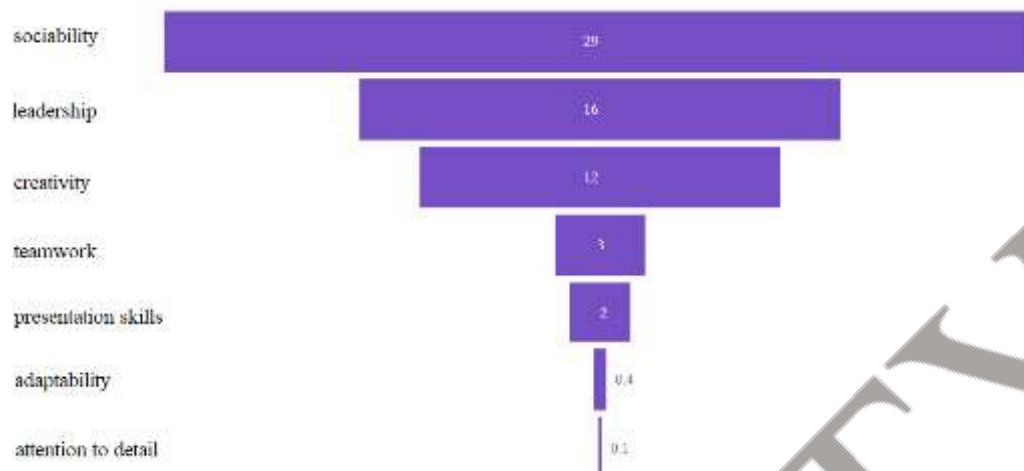


Fig. 13. Soft skills, %

Source: created by the author according to the sample data

After resolving recruiters mistakes, we saw that the minimum salary offered by employers to IT professionals on rabota.ua was \$90 (for the position of “Teacher for C++ courses”), a maximum – of \$48,600 (Golang Developer), the most common minimum salary was \$540 (10.5% of ads), the most common maximum – 720-1080 dollars (4.6%). At dou.ua, the minimum salary was \$300 (Senior PHP Engineer) and the maximum salary was \$70,000 (in the “Abroad” section for Senior Frontend Developer).

At StackOverflow Jobs, the minimum salary per month was \$4,500, the maximum was \$18,300, and the average was in the range of \$5,750-8300. It is noted that almost 11% of companies provide non-preferred shares.

Thus, after receiving, storing, organizing and analyzing data, we see that the most popular are the following:

- programming languages: Python, Javascript, Java;
- tools: React, Kubernetes, Linux;
- areas: machine learning, software as a service, artificial intelligence;
- soft skills: sociability, leadership and creativity.

## V. CONCLUSIONS

As a result of the research paper, the task of developing a software module for the analysis of IT specialists’ labor market was performed.

To analyze the labor market of IT specialists, it was decided to take the primary source – job vacancies from employers themselves, rather than information from various reports, statistics, surveys. The Python programming language was chosen for the software implementation. The final product is a module that performs scraping, crawling, parsing and exporting data from 3 sites.

The study examined *more than 20,000* vacancy announcements in the IT sector from Rabota.ua, Dou and StackOverflow Jobs in order to summarize the requirements of employers for knowledge and skills of specialists, programming languages, years of

experience, education, etc. Vacancy announcements were collected from the indicated sites, transferred to tabular form, and conclusions were analyzed and visualized. The software implementation facilitated the structuring of information and systematization of these requirements.

The most popular characteristics for positions are defined. The results are formalized with graphs, charts, summary tables, and other elements of visualization.

Software module allows to find, structure and export data to CSV and TXT files.

The final product meets all the specifications defined in the paper. The developed program can be extended with additional functions, supplemented by a graphical interface, uploaded to the web hosting.

The software module provides processing of a large array of unstructured information from vacancy announcement sites, reduces the amount of routine manual operations and provides an opportunity for stakeholders to focus on key areas.

## VI. REFERENCES

1. Portal MinfinMedia (2022). *Unemployment rate in Ukraine*  
<https://index.minfin.com.ua/ua/labour/unemploy/2021/>
2. Trading Economics (2022). *Unemployment Rate – Asia*.  
<https://tradingeconomics.com/country-list/unemployment-rate?continent=asia>
3. Countryeconomy.com (2022). *Unemployment Rate: 2022 Countries*.  
<https://countryeconomy.com/unemployment>
4. Mitchell, R. (2018). *Web scraping with Python. Collecting More Data from the Modern Web*. Sebastopol, CA: O'Reilly Media.
5. Wiki-TsDPU: Tokens (2022). <https://wiki.cuspu.edu.ua/index.php/Токени>
6. Breuss, M. (2021) *Beautiful Soup: Build a Web Scraper with Python*. Real Python, [realpython.com/beautiful-soup-web-scraper-python/](https://realpython.com/beautiful-soup-web-scraper-python/)
7. Chapagain, A. (2019). *Hands-on Web Scraping with Python: Perform Advanced Scraping Operations Using Various Python Libraries and Tools Such as Selenium, Regex, and Others*. Packt Publishing.
8. StackOverflow Careers API (2022). <https://meta.stackexchange.com/questions/158005/stack-overflow-careers-api>
9. JetBrains (2021). *Python Programming – The State of Developer Ecosystem in 2020 Infographic*. JetBrains, <https://www.jetbrains.com/lp/devecosystem-2020/python/>
10. Acodemy. (2015). *Learn web scraping with Python: The ultimate crash course to learning the basics of web scraping with Python in no time*. North Charleston, SC: Createspace Independent Publishing Platform.
11. Jarmul, K., Lawson, R. (2017). *Python Web Scraping: Fetching Data from the Web*. Packt Publishing.
12. hiQ Labs v. LinkedIn. <https://www.natlawreview.com/article/data-scraping-survives-least-now-key-takeaways-9th-circuit-ruling-hiq-vs-linkedin>
13. Indeed (2022). ToS. URL: <https://www.indeed.com/legal?hl=en&redirect=true>
14. Rabota.ua
15. Dou.ua
16. Stackoverflow.com
17. Python Documentation (2022). <https://docs.python.org/2/library/csv.html#csv.DictWriter>
18. Heydt, M. (2018). *Python Web Scraping Cookbook: Over 90 Proven Recipes to Get You Scraping with Python, Microservices, Docker, and AWS*. Birmingham, UK: Packt Publishing.
19. Index. TIOBE – The Software Quality Company (2021). <https://www.tiobe.com/tiobe-index/programming-languages-definition/>
20. Google Finance (2021). <https://www.google.com/finance/>

USE OF WEB-TECHNOLOGIES IN THE PROBLEM OF DIGITALIZATION OF THE DORMITORY Authors: Daria Liakhovska, Diana Kochuk Advisor: Tetyana Astistova Kiev National University of Technologies and Design (Ukraine).....	371
IMPROVING THE LEVEL OF DETAILING IN THE FORMATION OF REALISTIC THREE-DIMENSIONAL SCENES Author: Max Zakharchyk Advisor: Romanyuk Oksana Vinnytsia National Technical University (Ukraine).....	383
A REAL-WORLD CASE STUDY OF A VEHICLE ROUTING PROBLEM Authors: Arnas Matusevičius, Karolis Lašas Advisor: Tomas Krilavičius Vytautas Magnus University (Lithuania).....	399
DECISION SUPPORT SYSTEM FOR FORECASTING THE NUMBERS OF THE TROOP IN THE MIDDLE AGES Author: Andrei Kapeleshchuk Advisor: Oleksandr Melnykov Donbas State Engineering Academy (Ukraine).....	408
DEVELOPMENT OF SOFTWARE FOR AUTOMATION OF KNOWLEDGE TESTING Author: Maksym Kiyashko Advisor: Kateryna Kirei Petro Mohyla Black Sea National University (Ukraine).....	416
A COMPILER OF DOMAIN-SPECIFIC LANGUAGE FOR "SMART-HOME" APPLICATIONS: DESIGN PRINCIPLES AND IMPLEMENTATION ISSUES Author: Oleksandr Nelipa Advisor: Mykola Tkachuk V. N. Karazin Kharkiv National University (Ukraine).....	428
DEVELOPMENT OF SOFTWARE MODULE FOR ANALYSIS OF IT SPECIALISTS' LABOR MARKET Author: Anhelina Dub Advisor: Anna Zhurba Ukrainian State University of Science and Technologies (Ukraine).....	439
CONTROL SYSTEM OF CONDENSING DRYING PROCESS WITH ENERGY RECOVERY USING HEAT PUMP Author: Denis Chaplygin Advisor: Dmytro Kovalchuk Odessa National Academy of Food Technologies (Ukraine).....	454