

УДК 004.021

DOI <https://doi.org/10.32689/maup.it.2023.3.4>

**Роман ЗОЛОТУХА**

аспірант зі спеціальності 122 – Комп'ютерні науки, факультету інформаційних технологій, Національного Університету Біоресурсів та Природокористування України, вул. Героїв Оборони, 16А, Київ, Україна, індекс 03041 ([remko740@gmail.com](mailto:remko740@gmail.com))

ORCID: 0000-0003-3099-722X

**Олена ГЛАЗУНОВА**

доктор педагогічних наук, професор, декан факультету інформаційних технологій, Національного Університету Біоресурсів та Природокористування України, вул. Героїв Оборони, 16А, Київ, Україна, індекс 03041 ([o-glazunova@pubip.edu.ua](mailto:o-glazunova@pubip.edu.ua))

ORCID: 0000-0002-0136-4936

**Roman ZOLOTUKHA**

Postgraduate Student of 122 specialty – Computer sciences, Faculty of Information Technology, National University of Life and Environmental Sciences of Ukraine, 16A Heroiv Oborony str., Kyiv, Ukraine, postal code 03041 ([remko740@gmail.com](mailto:remko740@gmail.com))

**Olena GLAZUNOVA**

Doctor of Pedagogical Sciences, Professor, Dean of the Faculty of Information Technologies, National University of Life and Environmental Sciences of Ukraine, 16A Heroiv Oborony Str., Kyiv, Ukraine, postal code 03041 ([o-glazunova@nubip.edu.ua](mailto:o-glazunova@nubip.edu.ua))

**Бібліографічний опис статті:** Золотуха, Р., Глазунова, О. (2023). Алгоритм розпізнавання тексту з PDF-резюме для автоматизації підбору кандидатів в ІТ проекти. *Інформаційні технології та суспільство*, 3, 30–38. DOI: <https://doi.org/10.32689/maup.it.2023.3.4>

**Bibliographic description of the article:** Zolotukha, R., Glazunova O. (2023). Alhorytm rozpiznavannia tekstu z PDF-reziume dlia avtomatyzatsii pidboru kandydativ v IT proekty [Text recognition algorithm from PDF resumes to automate the selection of candidates for IT projects]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 30–38. DOI: <https://doi.org/10.32689/maup.it.2023.3.4>

**АЛГОРИТМ РОЗПІЗНАВАННЯ ТЕКСТУ З PDF-РЕЗЮМЕ  
ДЛЯ АВТОМАТИЗАЦІЇ ПІДБОРУ КАНДИДАТІВ В ІТ ПРОЕКТИ**

**Анотація.** У теперішній час, розвиток технологій та зростання індустрії інформаційних технологій супроводжується безпрецедентним попитом на висококваліфікованих інженерів та спеціалістів у галузі ІТ. У даній статті автори розглядають проблему автоматизації підбору кандидатів в ІТ проекти та пропонуємо алгоритм розпізнавання тексту з PDF-резюме, який, за допомогою мови Python, значно спрощує та прискорює процес відбору кандидатів. Алгоритм використовує сучасні інструменти обробки природної мови (Natural Language Processing, NLP) та бібліотеки для роботи з PDF-файлами з метою виділення ключової інформації з резюме кандидатів. Він розпізнає важливі дані, такі як освіта, навички, контактна інформація та інше, і структурує їх в легкозрозумілий формат. Результати нашого дослідження вказують на ефективність запропонованого алгоритму та його здатність до швидкого та точного аналізу великої кількості резюме. Це відкриває широкі можливості для впровадження автоматизованих систем підбору кандидатів в ІТ галузі, що збільшує продуктивність та сприяє ефективному використанню ресурсів у сфері HR. У статті ми також обговорюємо потенціал розвитку даного алгоритму, включаючи можливість розширення підтримуваних мов, вдосконалення процесу розпізнавання навичок та врахування специфіки окремих ІТ-галузей. Ми підкреслюємо важливість інтеграції машинного навчання для поліпшення точності та аналізу патернів у резюме кандидатів. Автори статті ставлять перед собою завдання вдосконалення та спрощення процесу підбору кандидатів в ІТ проекти, що допоможе підприємствам більш ефективно використовувати свій інтелектуальний потенціал та забезпечити наявність висококваліфікованих спеціалістів у сфері інформаційних технологій.

**Ключові слова:** розпізнавання тексту, PDF-резюме, автоматизація підбору кандидатів, ІТ, NLP, HR технології.

**TEXT RECOGNITION ALGORITHM FROM PDF RESUMES  
TO AUTOMATE THE SELECTION OF CANDIDATES FOR IT PROJECTS**

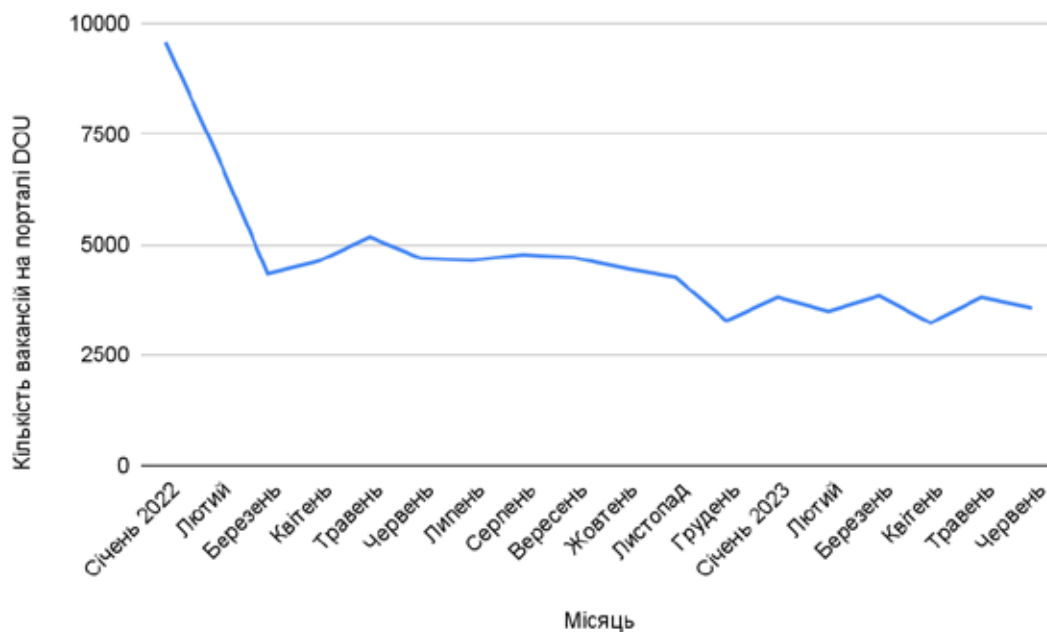
**Abstract.** Nowadays, the development of technology and the growth of the IT industry are accompanied by an unprecedented demand for highly qualified engineers and IT specialists. In this article, the authors consider the problem of automating the selection of candidates for IT projects and propose a text recognition algorithm from PDF resumes that, using the Python

language, greatly simplifies and speeds up the candidate selection process. The algorithm uses modern Natural Language Processing (NLP) tools and libraries to work with PDF files to extract key information from candidates' resumes. It recognizes important data, such as education, skills, contact information, etc., and structures it into an easily understandable format. The results of our study indicate the effectiveness of the proposed algorithm and its ability to quickly and accurately analyze a large number of resumes. This opens up wide opportunities for the implementation of automated candidate selection systems in the IT industry, which increases productivity and promotes the efficient use of HR resources. In the article, we also discuss the potential for the development of this algorithm, including the possibility of expanding the supported languages, improving the skill recognition process, and taking into account the specifics of individual IT industries. We emphasize the importance of integrating machine learning to improve the accuracy and analysis of patterns in candidates' resumes. The authors of the article aim to improve and simplify the process of recruiting candidates for IT projects, which will help enterprises to use their intellectual potential more efficiently and ensure the availability of highly qualified IT specialists.

**Key words:** text recognition, PDF resume, automation of candidate selection, IT, NLP, HR technologies.

**Актуальність проблеми.** У сучасному світі інформаційних технологій індустрія надзвичайно динамічна та швидкозмінна. З кожним роком з'являються нові технології, інструменти та професійні навички, які стають необхідними для успішної кар'єри в цьому секторі. Ця стійка динаміка вимагає від підприємств залучення висококваліфікованих ІТ-фахівців, що відповідають сучасним вимогам ринку.

В умовах повномасштабного вторгнення РФ український ІТ ринок показав неабияку стійкість, проте і він суттєво втратив, особливо в кількості активних вакансій. Ринок кандидата змінився на ринок мопаній, а з цією зміною і суттєво збільшилось навантаження на HR департаменти. Згідно статистики інтернет порталу DOU середня кількість вакансій в порівнянні з 2022 роком впала в більше ніж 2 рази, а кількість відгуків на вакансію зросла в 5 разів. [5]



**Рис. 1. Кількість ІТ вакансій на порталі DOU**

В умовах постійного зростання кількості заявок на робочі позиції в галузі ІТ, завдання відбору кращих кандидатів стає вкрай складним і вимагає нових підходів. Інтелектуальні системи, здатні аналізувати та оцінювати професійні навички та досвід кандидатів, стають ключовими інструментами для оптимізації цього процесу.

У даній науковій статті ми досліджуємо проблему автоматизації підбору кандидатів в ІТ проекти та пропонуємо алгоритм розпізнавання тексту з PDF-резюме. Цей алгоритм дозволяє ефективно виділяти ключову інформацію з резюме кандидатів, таку як освіта, навички та контактна інформація, та автоматично структурувати її для подальшого аналізу або зберігання в профільних HR базах.

Наша робота спрямована на вирішення важливої проблеми в сфері HR та рекрутингу в ІТ галузі. Ми розглянемо деталі алгоритму, його архітектуру, методи та результати використання. Крім того, ми проаналізуємо можливості подальшого розвитку та вдосконалення цього підходу для оптимізації відбору кандидатів у сучасному ІТ середовищі.

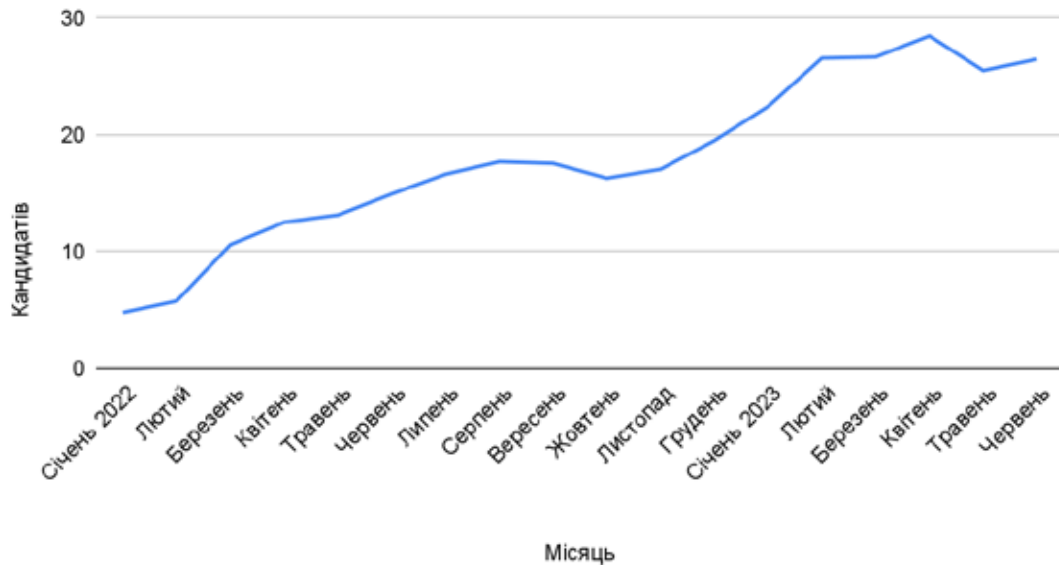


Рис. 2. Середня кількість надісланих резюме кандидатів

**Аналіз останніх досліджень та публікацій.** Проблемі формування команд для реалізації проєктів, а також управлінню персоналом присвячено багато вітчизняних та іноземних праць. Зокрема у роботі [1] розглянуто існуючі в науковій літературі методи розробки стратегій управління персоналом, а також запропоновано методичні положення розробки загальної стратегії управління персоналом. У статті [2] авторами досліджено та запропоновано прототип математичної моделі управління кадрами з урахуванням специфіки діяльності ІТ-компаній з цільовою функцією моделі спрямованою на оптимізацію часу, який витрачають HR-менеджери на роботу з підбору кадрів в ІТ-команди.

Дослідженню питанню автоматизації роботи з текстовими файлами присвячують свої роботи все більше науковців. У статті [3] визначає ефективність використання інструментів обробки природної мови та систем штучного інтелекту для аналізу тексту в інтернеті. Автор розглядає, які інструменти можуть бути використані для виявлення ключових слів у тексті оголошень, а також як ці інструменти можуть бути поєднані зі спеціалізованими моделями машинного навчання для виявлення шахрайської та зловмисної інформації. Стаття [4] присвячена визначенню області проведення досліджень, пов'язаних з підключенням зацікавлених осіб до процесу розробки програмного забезпечення через оцінювання сприйнятої якості сервіс-орієнтованих систем в контексті їхнього використання, коли початкова специфікація системи задана природною мовою. Пропонується використання технології аналізу природної мови для отримання інформації про область застосування з цієї специфікації у форматі спеціальних моделей перепроєктування, які є сумісними з основними модулями імітаційного рішення.

Також питанню автоматизації процесу підбору присвячена велика кількість наукових статей, зокрема у статті [6] розглянута можливість використання підходу з застосуванням скінченного автомата Мура для створення 3D резюме у процедурах підбору, сегментації та навчання кадрів. У даній роботі 3D резюме розглядається як документ, який представляє професійні досягнення особи у трьох вимірах. Використання автомата пропонується для автоматизації перебору питань та завдань інтерактивного резюме в залежності від попередніх відповідей респондента.

**Метою статті** є розробка та представлення алгоритму розпізнавання тексту з PDF-резюме для автоматизації процесу підбору кандидатів в ІТ проєкти.

#### **Виклад основного матеріалу.**

Для розробки алгоритму нами була розроблена модель потоків даних процесу подачі резюме кандидатом. Кандидат подає заявку на вакансію, прикріплюючи резюме у форматі PDF. Резюме обробляється системою і виділяє ключову інформацію про кандидата: контактну інформацію, місце навчання та навички. Ця інформація структурується і вноситься в базу даних HR, де HR-спеціаліст може бачити потрібних кандидатів, використовуючи задані фільтри. Варто зазначити, що запропонований процес руху даних від кандидата до HR значно оптимізує часові витрати порівняно зі стандартним підбором кандидатів. Він виключає механічне вивчення HR-фахівцем резюме всіх кандидатів, які подали заявки на вакансію, а також зменшує кількість контактів, які не відповідають вакансії. У цій статті ми зосередимося на розробці алгоритму, який зчитуватиме необхідну інформацію з PDF-резюме і конвертуватиме її в JSON-файли.

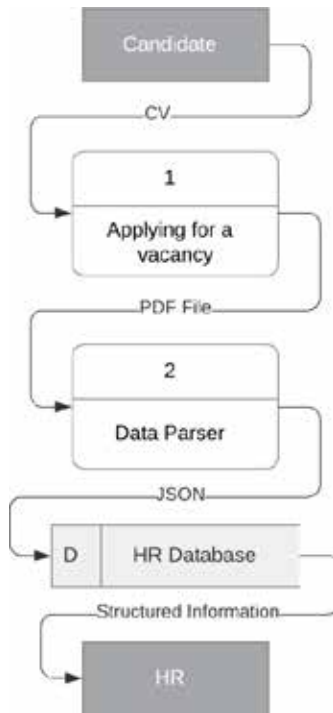


Рис. 3. Схема руху даних від кандидата до HR

Щоб виділити ключові розділи в резюме, ми згенерували 30 резюме, доступних на сайтах-агрегаторах резюме. Базове резюме містить контактну інформацію про кандидата, інформацію про його попередній досвід роботи, місце навчання, перелік технологій, якими володіє кандидат, та рівень володіння іноземними мовами. [7] Базовий вигляд резюме показано на рисунку 4.

## Roman Zolotukha

Product Analyst

### Profile

Product analyst with fintech, foodtech, saas experience.

### Employment History

**Product Analyst at Rocket, Kyiv**  
March 2021 — Present

- I joined the company as the first product analyst in the client area. At the moment I already have my sub-team in the structure of the whole analytical team.
- Complete building product analytics from scratch.
- Building non-standard funnel types (funnel to re-subscribe to the app, funnel with overflow traffic to retail, etc.)
- Launch and calculation of the economics of the subscription service model.
- Launch the web version of the service and its full analytical support.
- Automation of daily tasks, automation of communication, development and implementation of tools on health tracking analytics systems.
- Constant search for user's pains and problems in the app (restaurant output, user segmentation, path to purchase, navigation through the app, reasons for order cancellations, etc.) , P&L optimization ( redesign of the referral program, subscription model, order cancellation button

### Details

Kyiv, +380951333017  
remko740@gmail.com  
Date of birth  
07.02.1997

### Skills

- Data Studio, Tableau
- mySQL, Google BigQuery, Postgres
- Firebase, GA, GTM, Hotjar, DataDog
- Python
- Adjust, AppsFlyer
- MS Excel

### Languages

Ukrainian  
Russian  
English

Рис. 4. Базовий приклад вигляду резюме в експерименті

Для розробки такого алгоритму необхідне розуміння наборів сучасних ІТ навичок, які користуються високим попитом і на які часто звертають увагу особи, що шукають роботу в ІТ-індустрії [8; 9; 10].

Наша методологія передбачала ретельне вивчення цих наукових ресурсів, в ході якого ми визначили та каталогізували ключові слова та фрази, що вказують на технічні навички, мови програмування, інструменти, фреймворки та інші релевантні ІТ-навички. Такий системний підхід дозволив нам скласти вичерпний перелік потенційних навичок, що підлягають дослідженню. Проте варто зауважити, що список характеристик не є сталим явищем та змінюється в залежності від динаміки розвитку ІТ індустрії та появи там нових професій та сфер діяльності. На 2023 рік нами був підготовлений список характеристик навичок кандидата який включає в себе 38590 скілів, які може вказати кандидат у своєму CV.

Для досягнення мети дослідження і розробки алгоритму інструментом реалізації задуму була обрана мова програмування високорівнева інтерпретована мова програмування Python. Ця мова програмування використовується в різних галузях, включаючи веб-розробку, наукові дослідження, штучний інтелект, обробку природної мови, аналіз даних та багато інших. Вона є однією з найпопулярніших мов програмування у світі завдяки своїм перевагам та широкому спектру застосувань. Ми обрали її для розробки алгоритму на основі ряду ключових переваг, які роблять Python відмінним інструментом для вирішення задачі автоматизації підбору кандидатів в ІТ галузі:

- Популярність та спільнота розробників: Python є однією з найбільш популярних мов програмування у світі, що призводить до наявності великої та активної спільноти розробників. Це означає доступ до безлічі ресурсів, бібліотек та форумів для підтримки та вирішення проблем.
- Простота: Python відомий своєю простотою та читабельністю коду. Це дозволяє розробникам легко створювати та розуміти складні алгоритми, що є важливим для розробки алгоритму розпізнавання тексту.
- Багатофункціональність: Python має багату екосистему бібліотек та модулів для роботи з текстом, обробки PDF-файлів, обчислювального аналізу, та багато іншого. Ця багатофункціональність дозволяє ефективно реалізувати всі етапи алгоритму.
- Портативність: Python підтримується на багатьох платформах, що дозволяє легко переносити алгоритм на різні операційні системи та середовища.
- Велика кількість розширень: Python має багато сторонніх розширень та бібліотек, які спрощують роботу з текстом та PDF-файлами. Наприклад, бібліотеки `pdfplumber` та `srapy` дозволяють легко витягти текст та виконати обробку природної мови.
- Активний розвиток: Python постійно розвивається, і нові версії мови регулярно надають нові функції та покращення.

Завдяки цим перевагам мови програмування Python вона відмінно підходить для розробки алгоритму розпізнавання тексту з PDF-резюме для автоматизації підбору кандидатів в ІТ проекти.

Для реалізації даного алгоритму ми використали універсальні можливості мови програмування Python. Надійна екосистема бібліотек Python полегшила наші зусилля в розборі та вилученні релевантної інформації зі складної структури резюме кандидатів. Щоб розібратися в тонкощах PDF-документів, ми використовували бібліотеку `"pdfplumber"`, яка дозволила нам точно витягувати текстовий контент. Ця бібліотека надає можливість переглядати макет PDF-резюме та виокремлювати необхідні текстові сегменти для подальшого аналізу. Для обробки природної мови та розпізнавання текстових шаблонів ми використовували бібліотеку `"srapy"`. Цей потужний інструмент НЛП дозволив токенізувати, тегувати та аналізувати текст, що дало нам змогу виявити шаблони та сутності, важливі для виокремлення навичок та освіти. Модуль `"Matcher"` у складі `"srapy"` допоміг виявити конкретні лінгвістичні патерни, впорядкувавши наш процес визначення ключової інформації. Щоб полегшити організацію та зберігання наших результатів, ми використали модуль `"csv"` для створення та управління структурованими наборами даних. Бібліотека `"pandas"` запропонувала нам ефективний засіб для маніпулювання та аналізу цих наборів даних, що дозволило нам отримати уявлення та тенденції з отриманої інформації. Як невід'ємну частину нашої реалізації ми використали можливості вбудованої бібліотеки Python під назвою `"json"`. Ця бібліотека надала нам необхідні інструменти для легкого перетворення складних структур даних у формат JSON (JavaScript Object Notation), що дозволяє безперешкодно серіалізувати та зберігати дані. [11; 12; 13]

Запропонований нами процес перетворення PDF документу, яка містить інформацію про кандидата в JSON файл готовий до внесення в базу даних можна візуалізувати наступним чином:

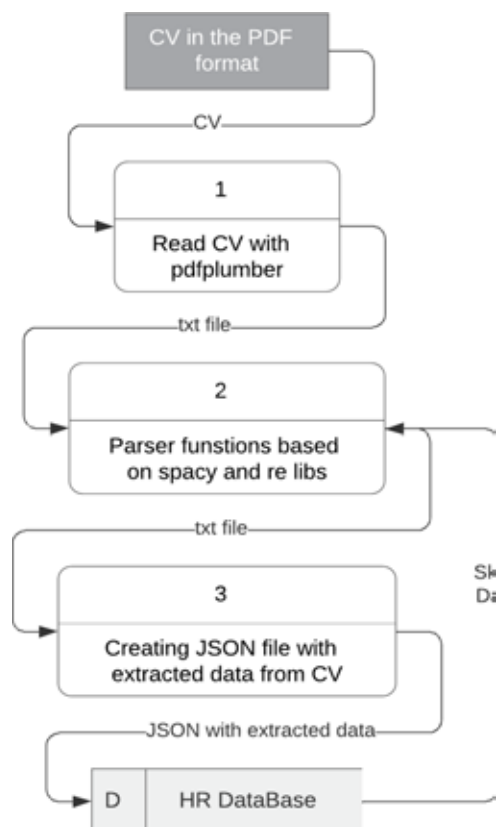


Рис. 5. Покрокова схема алгоритму

```

if __name__ == '__main__':
    resume_text = extracted_text

    name = extract_name(resume_text)
    if name:
        print("Name:", name)
    else:
        print("Name not found")

    contact_number = extract_contact_number_from_resume(resume_text)
    if contact_number:
        print("Contact Number:", contact_number)
    else:
        print("Contact Number not found")

    email = extract_email_from_resume(resume_text)
    if email:
        print("Email:", email)
    else:
        print("Email not found")

    skills_list = csv_data_list
    extracted_skills = extract_skills_from_resume(resume_text, skills_list)
    if extracted_skills:
        print("Skills:", extracted_skills)
    else:
        print("No skills found")

    extracted_education = extract_education_from_resume(resume_text)
    if extracted_education:
        print("Education:", extracted_education)
    else:
        print("No education information found")
  
```

```

Name: Roman Zolotukha
Contact Number: 380951333017
Email: remko740@gmail.com
Skills: ['computer science', 'mysql', 'health', 'p', 'api', 'mailchimp', 'economics', 'gmail', 'automation', 'python', 'subscribe', 'analytics', 'google analytics', 'firebase', '2014', 'communication', 'improvement', '.com', 'english', 'travel', 'british', 'data collection', 'web', 'digital', 'navioation', 'com', 'facebook', 'russian', 'analvt
  
```

Рис. 6. Частина коду реалізованого алгоритму в середовищі Jupyter Notebook

Для даного дослідження ми використали середовище розробки Jupyter Notebook. У нашому дослідженні ми використали середовище розробки Jupyter Notebook з кількох ключових причин:

- Інтерактивний аналіз даних: Jupyter Notebook дозволяє нам виконувати код блоками і бачити результати негайно після виконання. Це надає можливість проводити ітеративний аналіз даних та швидко реагувати на зміни та експерименти.
- Об'єднання коду та тексту: У Jupyter ми можемо поєднувати код з пояснювальним текстом, що робить наші дослідження більш зрозумілими та легкими для сприйняття.
- Візуалізація результатів: Ми можемо легко вставляти графіку та візуалізації результатів в наші документи, що робить наші дослідження більш інформативними.

Цей JSON-файл є наочним представленням результатів нашого дослідження і полегшує організоване зберігання та обмін відповідною інформацією. JSON-файл має кілька окремих полів, які відображають ключові аспекти профілю кандидата:

- Ім'я: Це поле містить повне ім'я кандидата, що дозволяє його чітко ідентифікувати та посилатися на нього.
- Контактний номер: Поле контактного номера містить наданий кандидатом номер телефону, що забезпечує безперебійний зв'язок та комунікацію.
- Електронна пошта: У полі електронної пошти вказується адреса електронної пошти кандидата, яка слугує основним засобом для листування.
- Навички: У цьому полі ретельно реєструється цілий ряд навичок. Ці навички охоплюють різноманітні технічні навички, мови програмування, інструменти, фреймворки та інші компетенції, пов'язані з ІТ, якими володіють кандидати.
- Освіта: Сфера освіти містить інформацію про академічні досягнення кандидата.

Результатом роботи нашого алгоритму є ретельно структурований JSON-файл, зображений на рисунку нижче:

```
{
  "Name": "Roman Zolotukha",
  "Contact Number": "380951333017",
  "Email": "remko74@gmail.com",
  "Skills": [
    "computer science",
    "mysql",
    "mailchimp",
    "economics",
    "gmail",
    "automation",
    "python",
    "analytics",
    "google analytics",
    "firebase",
    "data collection",
    "web",
    "digital",
    "facebook",
    "analytical",
    "excel",
    "data studio",
    "dashboards",
    "sql",
    "analyzing data",
    "tableau",
    "mysql",
    "search",
    "models",
    "data",
    "analysis",
    "computer science",
    "ms excel",
    "data collection",
    "output",
    "calculations",
    "hypothesis",
    "analytical support",
    "excel",
    "analytics",
    "api",
    "facebook api",
    "sql",
    "tableau",
    "python"
  ],
  "Education": [
    "PhD in Computer Science"
  ]
}
```

Рис. 7. Приклад JSON-файлу з інформацією з PDF CV

**Висновки.** Результатом нашого дослідження є розроблений алгоритм на мові програмування Python, який перетворює дані, витягнуті з PDF-резюме кандидатів, в агреговану, готову до використання інформацію. Цей трансформаційний процес усуває необхідність додаткових ручних зусиль, спрощуючи інтеграцію даних про кандидатів у бази даних HR. Використовуючи передові технології, такі як обробка природної мови та аналіз даних, наш алгоритм досягає автоматизації агрегації даних.

Цей алгоритм має величезний потенціал для HR-відділів, оскільки він не тільки прискорює процес рекрутингу, але й зменшує ручну працю, вишукування необхідної інформації серед великого тексту у резюме кандидата. Здатність алгоритму самостійно витягувати, класифікувати та структурувати інформацію про кандидатів дає змогу HR-фахівцям швидко оцінювати придатність кандидатів. Це означає оптимізацію використання ресурсів та більш ефективно прийняття рішень.

Подальшим розвитком дослідження цього алгоритму є розробка інтерфейсу для HR спеціалістів, а також створення та наповнення бази кандидатів та подальшої імплементації цього алгоритму в робочому середовищі та застосування на практиці.

#### Список використаних джерел:

1. Криворучко О. М., Водолазська Т. О. Розробка стратегій управління персоналом методом концептуального абстрагування. *Економіка транспортного комплексу*. 2016. № 28. С. 69–83.
2. Івченко І.Ю., Лінгур Л.М., Філатова Т.В. Моделювання управління кадрами на IT-ринку праці. *Вісник Харківського національного університету імені В. Н. Каразіна серія «Економічна»*. 2021. № 101. С. 101-112. URL: <https://doi.org/10.26565/2311-2379-2021-101-10>
3. Круціцький В. Я., Сугоняк І. І. Оцінка ефективності використання інструментів NLP та систем AI для аналізу рекламних оголошень у системах обміну інтернет-рекламою. *Технічна інженерія*. 2023. № 1(91). С. 161–165. URL: [https://doi.org/10.26642/ten-2023-1\(91\)-161-165](https://doi.org/10.26642/ten-2023-1(91)-161-165)
4. Шевцов В. А., Баженов Н. А. Використання НЛП для визначення обсягу оцінки стейкхолдерами імітованих якостей послуг. *Штучний інтелект*. 2010. № 3. С. 161–169.
5. Огляд IT-ринку праці, червень 2023. *DOU.ua*. URL: <https://dou.ua/lenta/articles/it-job-market-june-2023/>
6. Ющенко К.С. Підхід до автоматизації процесу підбору кадрів за допомогою 3D резюме. *Математичні машини і системи*. 2022. № 2. С. 29–39.
7. Eggert M. *Perfect CV (Perfect)*. Random House Books, 2007. 192 p.
8. Martin R. *Clean Code: A Handbook of Agile Software Craftsmanship*. – Pearson, 2008. 497 p.
9. Gamma E. *Design patterns: Elements of reusable object oriented software*. Reading, USA : Addison-Wesley, 1995. 395 p.
10. Peter A. *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press. 2012. 416 p.
11. McKinney W. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter*. O'Reilly Media, 2022. 550 p.
12. Muellera A *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, 2016.
13. Yu D., Deng L. *Automatic Speech Recognition: A Deep Learning Approach*. Springer-Verlag Longon, 2015.

#### References:

1. Krivoruchko, O. M., & Vodolazhska, T. O. (2016). Rozrobka stratehii upravlinnia personalom metodom kontseptualnoho abstrahuvannia [Development of personnel management strategies by the method of conceptual abstraction]. *Ekononika transportnogo kompleksa – Economy of the transport complex*, 28, 69–83. [in Ukrainian].
2. Ivchenko, I.Y., Lingur, L.M., Filatova T.V. (2021). Modeliuvannia upravlinnia kadramy na IT-rynku pratsi [Modeling of personnel management in the IT labor market]. *Visnyk Kharkivskoho natsionalnoho universytetu imeni V. N. Karazina seriia «Ekononichna» – Bulletin of V. N. Karazin Kharkiv National University, Economic Series*, 101, 101–112. Retrieved from <https://doi.org/10.26565/2311-2379-2021-101-10> [in Ukrainian].
3. Krutsitsky, V. Y., & Sugonyak, I. I. (2023). Otsinka efektyvnosti vykorystannia instrumentiv NLP ta system AI dlia analizu reklamnykh oholoshen u systemakh obminu internet-reklamoiu [Evaluation of the effectiveness of using NLP tools and AI systems for analyzing advertisements in online advertising exchange systems]. *Tekhnichna inzheneriia – Technical Engineering*, 1(91), 161-165. Retrieved from [https://doi.org/10.26642/ten-2023-1\(91\)-161-165](https://doi.org/10.26642/ten-2023-1(91)-161-165) [in Ukrainian].
4. Shevtsov, V. A., & Bazhenov, N. A. (2010). Vykorystannia NLP dlia vyznachennia obsiahu otsinky steikkholderamy imitovanykh yakosteï posluh [Using NLP to determine the extent of stakeholder evaluation of simulated service qualities]. *Shtuchnyi intelekt – Artificial intelligence*, 3, 161-169. [in Ukrainian].
5. Ohliad IT-rynku pratsi, cherven 2023 [IT labor market, June 2023]. Retrieved from <https://dou.ua/lenta/articles/it-job-market-june-2023/> [in Ukrainian].
6. Yushchenko, K. S. (2022). Pidkhd do avtomatyzatsii protsesu pidboru kadriv za dopomohoiu 3D reziume [An approach to automating the recruitment process using 3D resumes]. *Matematychni mashyny i systemy – Mathematical machines and systems*, 2, 29-39.
7. Eggert, M. (2007). *Perfect CV (Perfect)*. Random House Books. 192 p.
8. Martin, R. C. (2008). *Clean Code: A Handbook of Agile Software Craftsmanship*. Prentice Hall. 497 p.



9. Gamma, E. (1995). Design patterns: Elements of reusable object orientated software. Addison-Wesley. 395 p.
10. Peter, A. (2021). Machine Learning: The Art and Science of Algorithms that Make Sense of Data. 416 p.
11. McKinney, W. (2022). Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. 3rd edition. O'Reilly Media. 550 p.
12. Müller, A. C., Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. 1st edition. O'Reilly Media.
13. Yu, D., Deng, L. (2015). Automatic Speech Recognition: A Deep Learning Approach. Springer-Verlag London. DOI: 10.1007/978-1-4471-5779-3.