

Автоматизація оцінки навичок під час тоск-інтерв'ю в реальному часі

Науковий керівник: магістр комп'ютерних наук,

Старший викладач

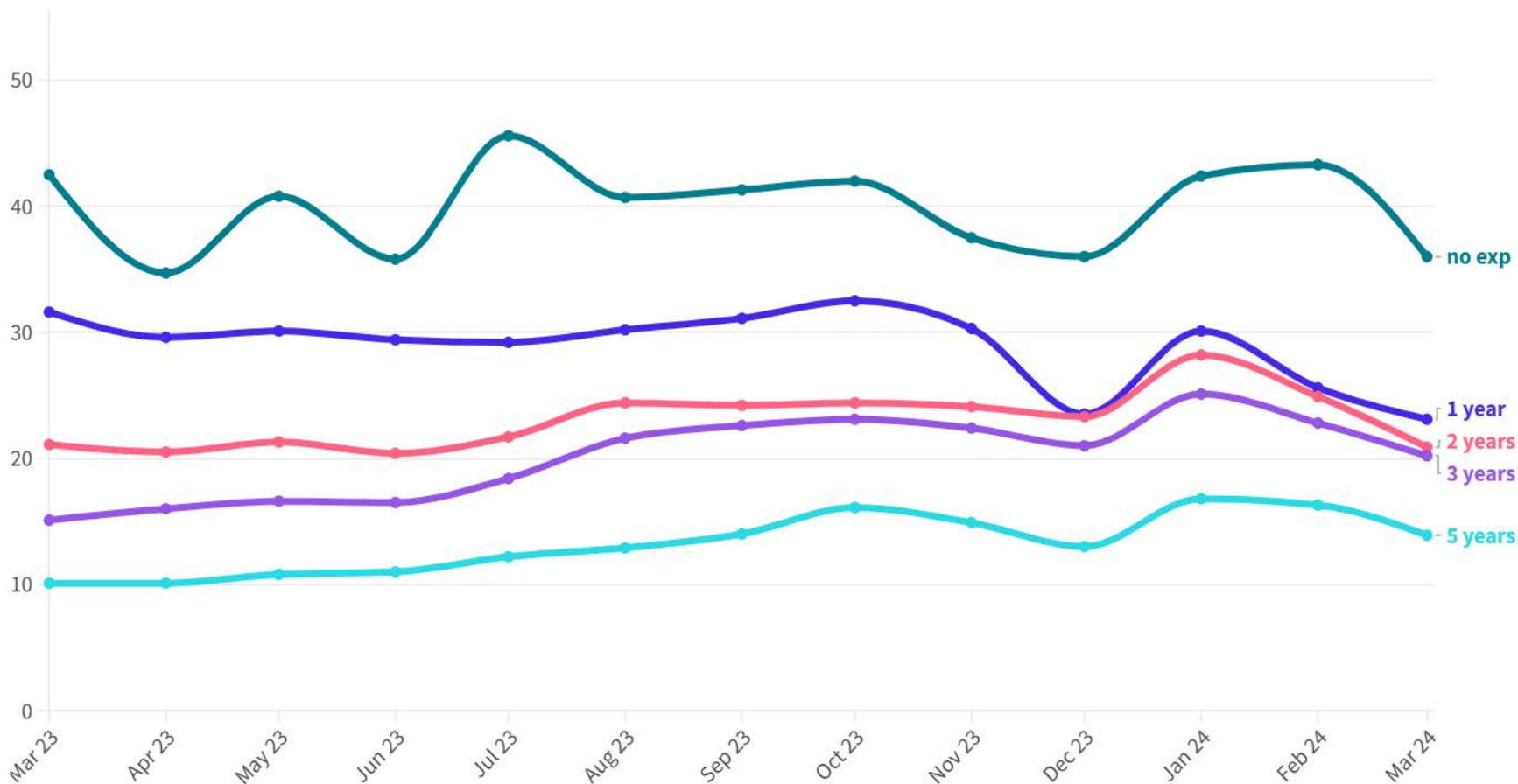
Медвідь С.О.

Виконав: студент 3-го року навчання,

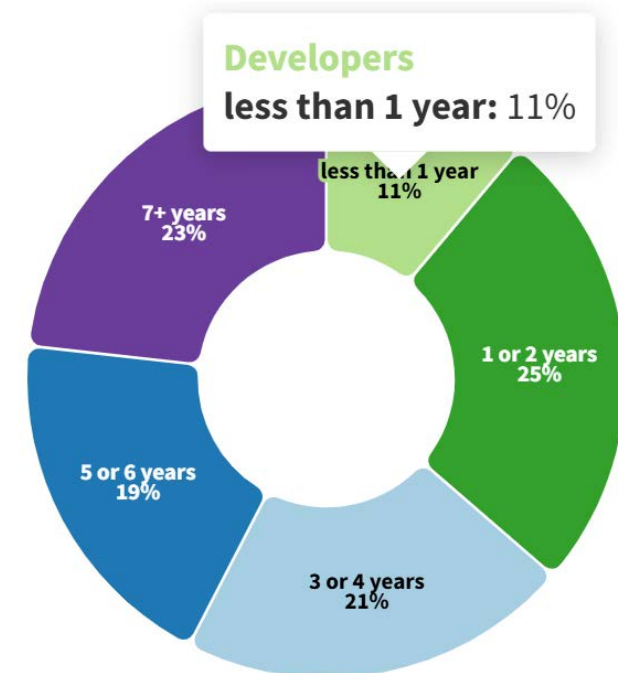
Спеціальності “Інженерія Програмного Забезпечення”

Шлапак Д.В.

Актуальність



Developers



Source: [Djini.co](https://djini.co)

Мета:

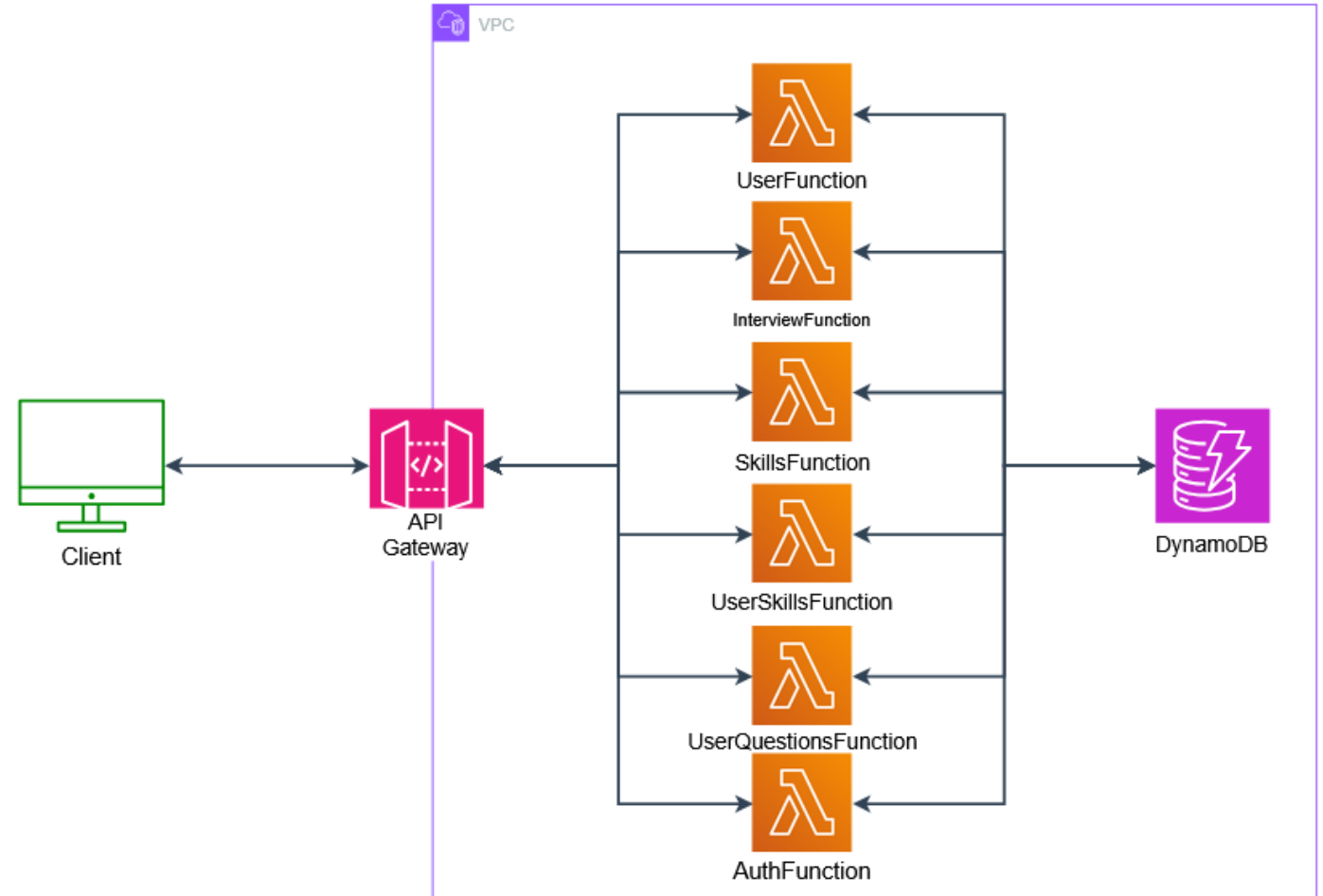
- Розробити систему для автоматизованої оцінки технічних навичок під час mock-інтерв'ю із застосуванням LLM для генерації запитань і аналізу відповідей.
- Забезпечити масштабованість і стабільність завдяки serverless-архітектурі.

Запропоноване рішення

- **Архітектура:** serverless (AWS Lambda, API Gateway, DynamoDB)
- **Інтеграція з OpenAI:** Whisper-1 (транскрипція), GPT-4o (генерація та аналіз)
- **Розгортання:** AWS SAM (Serverless Application Model)

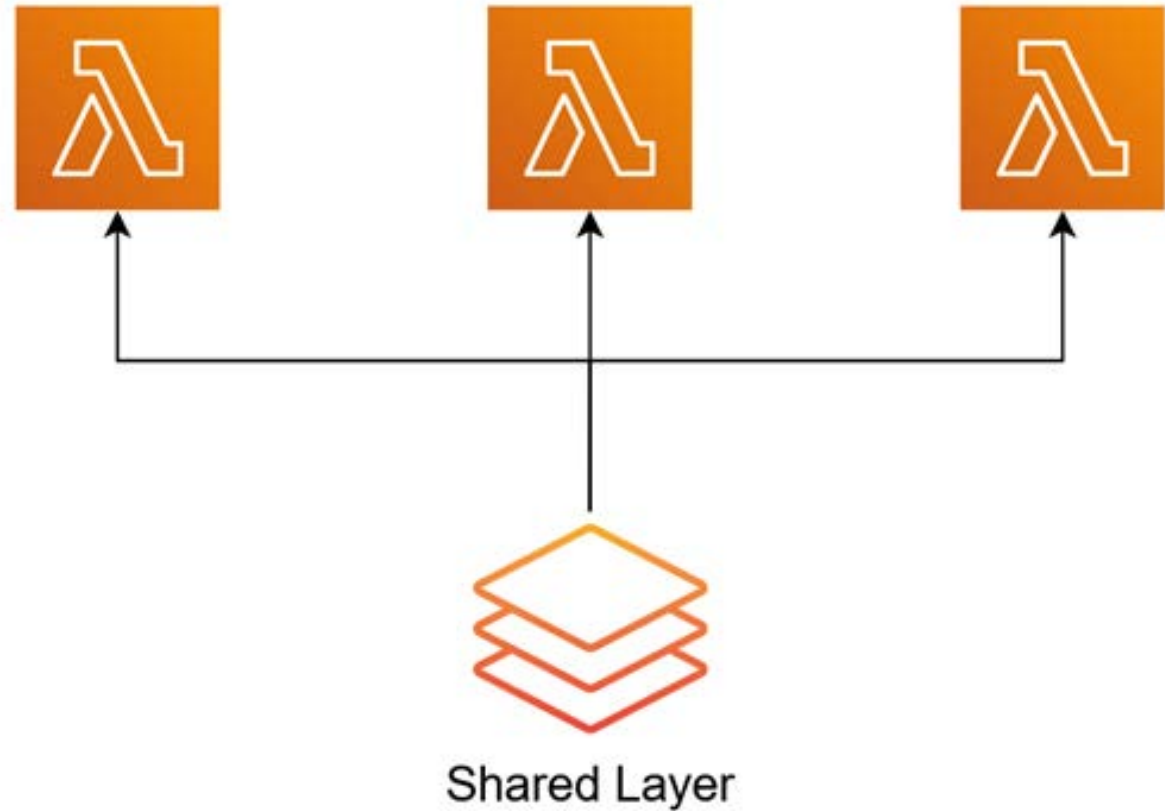
Архітектура системи

- Обробка логіки через AWS Lambda
- Обробка вхідних HTTP-запитів через API Gateway
- Збереження даних у NoSQL базі даних DynamoDB
- Уникнення дублювання коду завдяки AWS Lambda Layers



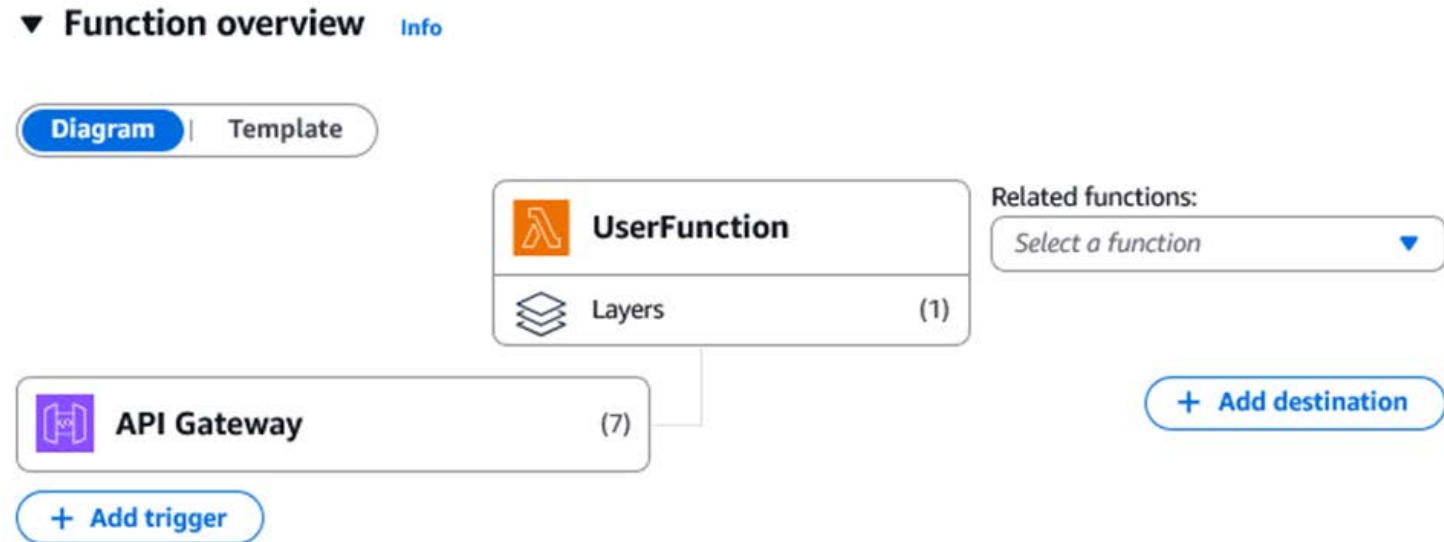
Архітектура системи

- AWS Lambda Layer



Особливості реалізації

- Mangum для інтеграції FastAPI з AWS Lambda
- Розв'язання проблеми Vendor Lock-In
- Застосування патернів: Dependency Injection, Facade
- Локальна розробка і тестування за допомогою контейнеризації



Розгортання і підтримка масштабованості

- AWS SAM: декларативне описання інфраструктури (config.yaml)
- Інтеграція основних сервісів AWS: Lambda, API Gateway, DynamoDB
- Автоматизоване розгортання

```
127 > AuthFunction:
128   Type: AWS::Serverless::Function
129   Properties:
130     FunctionName: AuthFunction
131     CodeUri: app/auth/
132     Handler: controller.auth_controller.lambda_handler
133     Runtime: python3.12
134     Architectures:
135       - x86_64
136     Layers:
137       - !Ref CommonLayer
138     Events:
139       GetAllSkillTrees:
140         Type: Api
141         Properties:
142           Path: /auth/register
143           Method: post
144       GetSkillTreeByRootId:
145         Type: Api
146         Properties:
147           Path: /auth/login
148           Method: post
149     Policies:
150       - AWSLambdaBasicExecutionRole
151       - AmazonDynamoDBFullAccess
```


LLM Integration

- Whisper-1: транскрипція голосових відповідей у текстовий формат
- GPT-4o: генерація адаптивних технічних питань і детальний аналіз відповідей



Реалізована система

Personal profile



Danylo Shlapak
CANDIDATE

Customize Profile

Skills

Full-stack Development

UI/UX Design

Interview Statistics

Total	Planned	Completed	Cancelled
1	1	0	0

Interview Journal

Schedule Interview

Date	Role	Title	Status
⋮ 23 May 2025, 17:00	Interviewer	Job Interview	PLANNED

Реалізована система

AI Assistant - Question Suggestions

Choose Skill:

Select a skill ▼

Ask AI

AI Assistant - Question Suggestions

Choose Skill:

Full-stack Development ▼

Suggested Questions:

Explain how you would optimize full-stack applications for performance. What tools or techniques would you use?

Describe the process of how a data request is

Реалізована система

Time: 10:00 • Paused Java Developer Interview Start Interview

Candidate: Danylo Shlapak

Full-stack Development UI/UX Design

+ Add question

Full-stack Development

Explain the concept of RESTful services and how you would implement them in a full-stack application.

Grade (1–10):

Start Recording

Наукова новизна

- Новизна роботи полягає в інтеграції великої мовної моделі з безсерверною архітектурою для створення гнучкої та масштабованої системи оцінки навичок, що поєднує адаптивне генерування запитань, транскрипцію відповідей і їх детальний аналіз у реальному часі.

Висновки

- Реалізовано систему автоматизованої оцінки технічних навичок із використанням LLM для генерації і аналізу відповідей.
- Використано `serverless`-архітектуру, досліджено переваги та недоліки під час роботи з нею.
- Виявлено обмеження: складність розробки `Lambda`-функцій та `cold-start`.

Перспективи подальших досліджень

- Впровадження Retrieval-Augmented Generation (RAG) для підвищення релевантності запитань.
- Розробка адаптивних алгоритмів підбору питань за результатами попередніх відповідей.
- Оптимізація cold start Lambda-функцій для зменшення затримок.

Дякую за увагу!