

Назва: Сервіс створення анімації на основі зображення та аудіо

Керівник (product owner):

доцент Турута Олексій Петрович

Проблема(и), яку вирішує програмний продукт:

1. Продукт вирішує задачу генеративного штучного інтелекту.
2. Сервіс створює анімоване зображення на основі статично (2D) зображення та аудіодоріжки.
3. Створене анімоване зображення виконує функцію розмовляючого аватара.

Анотація (мета, основні функції, складові):

Мета роботи створити сервіс, який дозволить замінити лектора (співрозмовника) шляхом анімації статичного зображення таким чином, щоб міміка обличчя відповідала аудіодоріжці.

Сервіс має виконати функцію завантаження статичного зображення, завантаження тексту, переворотення тексту на аудіо потік, анімацію статичного зображення з урахування аудіопотоку.

Сервіс планується реалізувати на основі клієн-серверної технології.

Початкові дані:

визначаються виконавцем із узгодженням із керівником відповідно до документа SRS

Джерела:

1. [Neural Natural Language Generation: A Survey on Multilinguality, Multimodality, Controllability and Learning](#)
2. [Extension Multi30K: Multimodal Dataset for Integrated Vision and Language Research in Ukrainian](#) (Saichyshyna et al., UNLP 2023)
3. [Audio-driven Talking Face Generation with Stabilized Synchronization Loss](#)
4. [JoyHallo: Digital human model for Mandarin](#)

Склад команди (ролі, завдання):

1 студент

Технології (методології) :

Технології GenAI, Natural Language Generation, Computer Vision, Text-to-Speech.

Критерії (метрики) успішності:

Виконання всіх вимог SRS. Контроль за виконанням завдань здійснюється на кожному етапі робіт, відповідно до затвердженого календарного плану.

Графік виконання:

Робота повинна виконуватися поетапно за узгодженим календарним планом відповідно до плану навчання у 2024-25 навчальному році.