

**РОЗРОБКА ІНТЕЛЕКТУАЛЬНОГО ВЕБ-ДОДАТКУ
МУЛЬТИМОДАЛЬНОЇ АДАПТАЦІЇ ТЕКСТОВОГО КОНТЕНТУ У
ВІДЕОФОРМАТ**

Шептуха І.В.

e-mail: illia.sheptukha@nure.ua

Науковий керівник – ст.викл. каф. ІСТ Ганшин Д.Г.

Харківський національний університет радіоелектроніки каф. ІСТ
м. Харків, Україна

This work presents ReadNet, an intelligent web application for multimodal adaptation of text content into video format. The system transforms books into immersive video experiences using AI-powered content reformatting via Large Language Models, text-to-speech synthesis with word-level timing synchronization, dynamic image display, and background audiovisual content. The platform is built with Next.js (React, TypeScript), PostgreSQL, and AWS S3, integrating LLM APIs through LiteLLM and Microsoft Edge TTS for narration. Personalization is achieved through customizable reading modes and user interest profiles that guide AI content transformation. The system supports PDF book upload, automatic text and image extraction, content caching, and seamless page-to-page transitions with preloading. Video mode synchronizes background video, TTS narration, word-by-word subtitles, and book illustrations into a single cinematic experience.

З кожним роком все більше людей надають перевагу відеоконтенту замість читання книг. Платформи коротких відео (TikTok, YouTube Shorts, Instagram Reels) формують звичку до швидкого споживання інформації, через що традиційне читання стає все менш привабливим. При цьому дослідження мультимедійного навчання [1] показують, що поєднання візуального та аудіального каналів сприйняття значно покращує засвоєння інформації порівняно з використанням лише одного каналу.

Після аналізу існуючих рішень для адаптації текстового контенту було виявлено їх суттєві обмеження: аудіокниги працюють лише з аудіальним каналом, сервіси синтезу мовлення (Speechify, NaturalReader) не переформатовують текст та не враховують інтереси користувача, а платформи генерації відео (Pictory, Lumen5) орієнтовані на маркетинг і не працюють з книгами. Для вирішення цих проблем було розроблено інтелектуальний веб-додаток ReadNet, який перетворює книги у відеоформат з аудіонарацією, послівними субтитрами та ілюстраціями, адаптованими під інтереси конкретного читача. На рисунку 1 наведено загальну архітектуру розробленої системи.

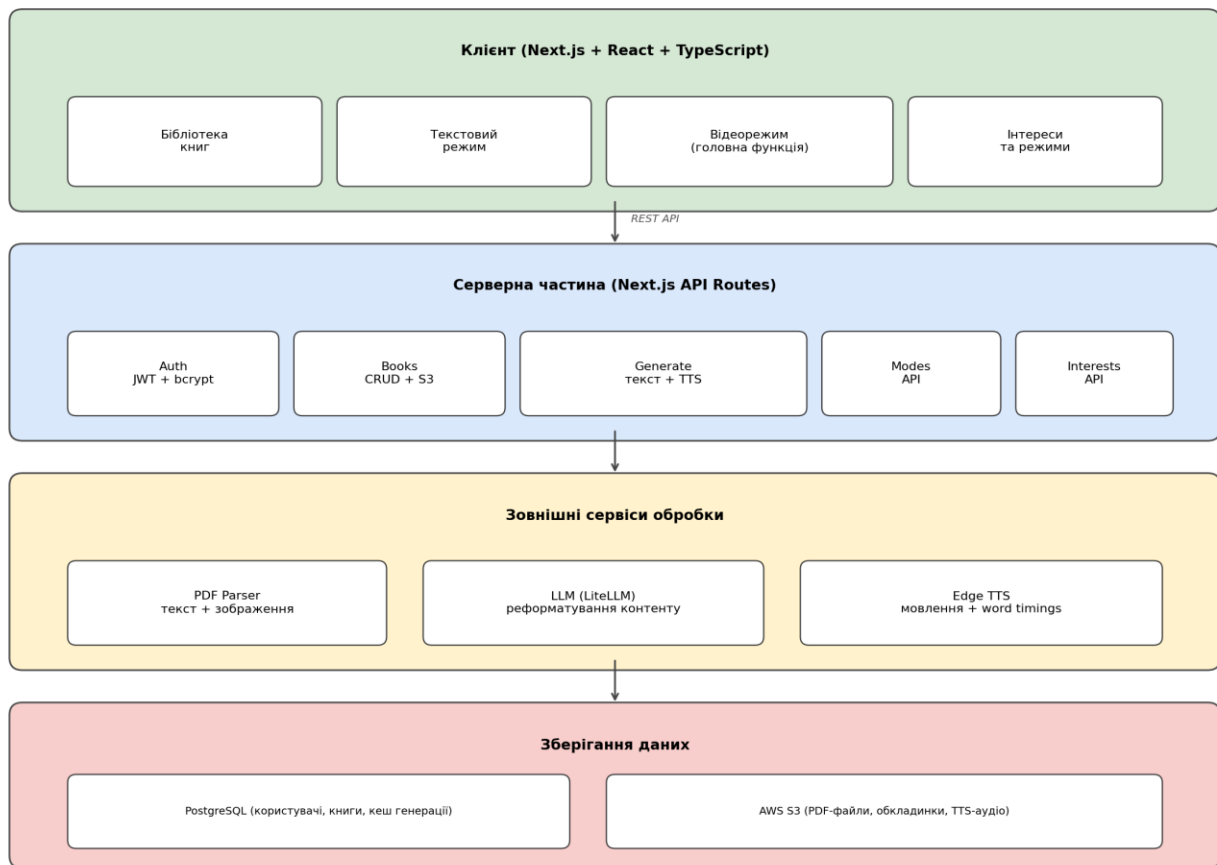


Рисунок 1 – Архітектура платформи ReadNet

Клієнтська та серверна частини побудовані на фреймворку Next.js [2] (React, TypeScript). Для зберігання даних використовується PostgreSQL (облікові записи, метадані книг, налаштування режимів, кеш згенерованого контенту) та AWS S3 (PDF-файли, обкладинки, аудіофайли). Авторизація працює через JWT-токени з механізмом оновлення сесій.

Головним елементом системи є генерація контенту. Коли користувач відкриває сторінку книги, система витягує з PDF текст та зображення. Далі цей контент відправляється до великої мовної моделі через API-шлюз LiteLLM [3], який підтримує різних провайдерів (Claude, GPT тощо). Модель переписує текст відповідно до обраного режиму читання та інтересів користувача. Після цього текст передається до Microsoft Edge TTS [4], який генерує аудіо з тайм-кодами для кожного слова. Всі результати кешуються, щоб не генерувати повторно.

У відеорежимі користувач обирає фонове відео, аудіо та режим читання. Під час відтворення на екрані одночасно показуються: фонове відео, субтитри, що підсвічуються послівно у такт з озвучкою, та ілюстрації з книги у відповідні моменти тексту. Наступна сторінка завантажується заздалегідь, тому перехід

між сторінками відбувається без затримок. Контент адаптується під кожного читача через систему режимів (шаблони промптів для мовної моделі) та профілі інтересів.

В результаті роботи створено веб-додаток, який за допомогою штучного інтелекту, синтезу мовлення та мультимедійної синхронізації перетворює книги у відеоформат. Dodatok адаптує контент під інтереси читача, що робить процес читання більш залучаючим порівняно зі звичайним переглядом тексту.

Список використаних джерел:

1. Mayer R. E. Multimedia Learning. 2nd ed. New York : Cambridge University Press, 2009. 304 p.
2. Next.js : The React Framework for the Web : вебсайт. URL: <https://nextjs.org/> (дата звернення: 14.03.2026).
3. LiteLLM : Call 100+ LLMs using the OpenAI Input/Output Format : вебсайт. URL: <https://docs.litellm.ai/> (дата звернення: 14.03.2026).
4. msedge-tts : Microsoft Edge text-to-speech library : вебсайт. URL: <https://www.npmjs.com/package/msedge-tts> (дата звернення: 14.03.2026).