

Makeathon 2025

UniAi

Challenge 2: ΕΥ - Δημιουργία Αι εργαλείου για αποδοτικό διάβασμα!

TEAM: LITTLEBITQUIRKY

TEAM MEMBERS: ΑΛΕΞΑΝΔΡΟΣ - ΕΜΜΑΝΟΥΗΛ ΛΟΥΔΑΡΟΣ, ANNA ΜΥΣΙΡΗ

UNI: NKUA

DEPARTMENT: Digital Industries

1. Data used and processing:

Είσοδος	Επεξεργασία	Τελικό κείμενο
Αρχεία από τον χρήστη σε μορφές: Εικόνες (png, jpg, jpeg) Ήχος (mp3) Παρουσιάσεις (pptx) PDF αρχεία Απλό κείμενο (txt)	1.OCR για εικόνες με pytesseract 2.Speech-to-text για mp3 με speech_recognition (Google API) 3.Ανάγνωση και εξαγωγή κειμένου από pptx (python-pptx) και pdf (PyPDF2) 4.Απλή ανάγνωση txt αρχείων	1.Καθαρίζεται (tokenization, αφαίρεση stopwords, filler words, stemming, lemmatization) 2.Περίληψη μέσω προεκπαιδευμένου NLP μοντέλου (transformers - BART) 3.Δημιουργούνται εκπαιδευτικές δραστηριότητες (q&a,quiz, flashcards)

ΣΗΜΕΙΩΣΗ:Χρησιμοποιήθηκαν οι βιβλιοθήκες από το MIT

2. Techniques and models applied:

Transcription	OCR	Parsing και text extraction	Summarization	NLP Text cleaning	Quiz generation	Flashcards	Q&A
Μετατροπή ήχου (mp3) σε κείμενο μέσω της βιβλιοθήκης speech_recognition (Google API).	Εξαγωγή κειμένου από εικόνες (png, jpg, jpeg) με pytesseract.	Εξαγωγή από αρχεία PPTX (με python-pptx), PDF (με PyPDF2), TXT (απλή ανάγνωση).	Χρήση προεκπαιδευμένου μοντέλου facebook/bart-large-cnn μέσω της βιβλιοθήκης transformers για περίληψη κειμένου.	Tokenization, stopwords removal, stemming (PorterStemmer) και lemmatization (WordNetLemmatizer) με τη βιβλιοθήκη nltk	Δημιουργία ερωτήσεων τύπου Ναι/Όχι βασισμένων σε ανάλυση προτάσεων.	Δημιουργία κάρτας με κενά (masked words) για επανάληψη.	Αναζήτηση προτάσεων σε περίληψη με βάση λέξεις-κλειδιά.

3.System architecture & workflow:

Data ingestion	Preprocessing:	NLP processing	UI update
Ανεβάζεις αρχεία (εικόνες, ήχο, pdf, pptx, txt).οντέλο μετασχηματιστή)	Εξαγωγή κειμένου ανάλογα με τύπο αρχείου (OCR, speech-to-text, parsing). Αποθήκευση κειμένου σε μεταβλητή και αρχείο Extracted.txt.	1.Καθαρισμός κειμένου (stopwords, filler words, stemming, lemmatization). 2.Δημιουργία περίληψης μέσω προεκπαιδευμένου μοντέλου. 3.Δημιουργία quiz και flashcards με βάση την περίληψη ή το καθαρισμένο κείμενο.	Προβολή κειμένου, περίληψης, flashcards, quiz. Επιλογές για λήψη κειμένου.

4.How the solution addresses the problem:

Problem:

Οι μαθητές συχνά έχουν δυσκολία:

- Να οργανώσουν πληροφορίες από διάφορες μορφές αρχείων
- Να κάνουν περίληψη και εντοπισμό βασικών σημείων
- Να επαναλάβουν αποτελεσματικά το υλικό

Solution:

- Αυτόματη εξαγωγή και επεξεργασία πληροφοριών από πολλαπλές μορφές αρχείων
- Δημιουργία flashcards & quiz για ενεργητική επανάληψη
- Περίληψη και εννοιολογική αποτύπωση (mind map) για κατανόηση
- Εργαλείο ερωτήσεων/σημειώσεων για εμπλοκή του χρήστη

5.The value it provides to students and education:

- **Βαθύτερη κατανόηση** μέσω concept maps και περιλήψεων
- **Αυτόνομη μάθηση & επανάληψη** με flashcards και quiz βασισμένα στο υλικό
- **Υποστήριξη διαφορετικών τύπων αρχείων** με δυνατότητα μελέτης από εικόνες, βίντεο, παρουσιάσεις κ.λπ
- **Διαδραστικότητα**, οι χρήστες συμμετέχουν γράφοντας ερωτήσεις και σημειώσεις
- **Εξοικονόμηση χρόνου** με αυτόματη δημιουργία περιλήψεων, βασικών εννοιών και ερωτήσεων

ΕΥΧΑΡΙΣΤΟΥΜΕ ΓΙΑ ΤΗ ΠΡΟΣΟΧΗ ΣΑΣ!

-LITTLEBITQUIRKY