



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ

Πτυχιακή Εργασία

Τίτλος Πτυχιακής Εργασίας	Εφαρμογή Android για Ιδιωτική Αλληλεπίδραση με LLMs σε Τοπικό Περιβάλλον Android Application for Private Interaction LLMs in a Local Environment
Όνοματεπώνυμο Φοιτητή	Θεόδωρος Κοζάνογλου
Πατρώνυμο	Νικόλαος
Αριθμός Μητρώου	Π20094
Επιβλέπων	Ευθύμιος Αλέπης, Καθηγητής

Ημερομηνία Παράδοσης

Σεπτέμβριος 2024

Copyright ©

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν αποκλειστικά τον συγγραφέα και δεν αντιπροσωπεύουν τις επίσημες θέσεις του Πανεπιστημίου Πειραιώς.

Ως συγγραφέας της παρούσας εργασίας δηλώνω πως η παρούσα εργασία δεν αποτελεί προϊόν λογοκλοπής και δεν περιέχει υλικό από μη αναφερόμενες πηγές

Ευχαριστίες

Με την ολοκλήρωση της πτυχιακής μου εργασίας, θα ήθελα να εκφράσω την ευγνωμοσύνη μου σε όσους με στήριξαν και με καθοδήγησαν όλα αυτά τα χρόνια μέχρι σήμερα.

Πρώτα απ' όλα, ευχαριστώ όλους τους καθηγητές του Τμήματος Πληροφορικής του Πανεπιστημίου Πειραιά για τις γνώσεις και τις πρακτικές εμπειρίες που μου πρόσφεραν στα μαθήματά τους. Η καθοδήγησή τους και η αφοσίωσή τους στη διδασκαλία διαμόρφωσαν σημαντικά τις γνώσεις και τις δεξιότητές μου.

Ιδιαίτερα, θέλω να ευχαριστήσω τον επιβλέποντα καθηγητή μου, κ. Ευθύμη Αλέπη, για την εμπιστοσύνη, τη υποστήριξη και τις πολύτιμες συμβουλές του κατά τη διάρκεια αυτής της πτυχιακής.

Επίσης, ευχαριστώ την οικογένειά μου που ήταν πάντα δίπλα μου να με υποστηρίξει. Η βοήθειά τους ήταν καθοριστική για να πετύχω τους στόχους μου.

Τέλος, ευχαριστώ τους φίλους και συμφοιτητές μου που με συνόδευσαν σε αυτό το ταξίδι. Η συνεργασία, οι συζητήσεις και η αμοιβαία υποστήριξη που μοιραστήκαμε έκαναν αυτή την εμπειρία ξεχωριστή.

Περίληψη

Η παρούσα πτυχιακή εργασία επικεντρώνεται στην ανάπτυξη μιας εφαρμογής που επιτρέπει στον χρήστη να αλληλεπιδρά με «Μεγάλα Γλωσσικά Μοντέλα» (LLMs) χωρίς τη χρήση υπηρεσιών νέφους, προσφέροντας μεγαλύτερη ασφάλεια και έλεγχο στα δεδομένα του. Η εφαρμογή «Home AI» αξιοποιεί πόρους από τον τοπικό υπολογιστή του χρήστη, προσφέροντας εξατομικευμένες απαντήσεις.

Οι κύριοι στόχοι της εφαρμογής περιλαμβάνουν την ιδιωτικότητα των δεδομένων μέσω τοπικής επεξεργασίας, την εξατομίκευση των απαντήσεων με βάση τα δεδομένα του χρήστη, τη δημιουργία φιλικού περιβάλλοντος και την αξιοποίηση τοπικών πόρων για την εξάλειψη της ανάγκης εξωτερικών υπηρεσιών.

Για την επίτευξη αυτών των στόχων, σχεδιάστηκε ένα σύστημα αρχιτεκτονικής client-server, προσφέροντας μια ασφαλή εμπειρία χρήστη. Ο server βασίζεται στον «Έλεγχος Πρόσβασης Βασισμένο σε Ρόλους» (RBAC) και υλοποιήθηκε με σύγχρονες τεχνολογίες, όπως FastAPI, ασφάλεια με JWT, Postgres, ChromaDB και Ollama για αλληλεπίδραση με LLMs.

Ο client είναι μια Android εφαρμογή που επιτρέπει την αλληλεπίδραση με τον server μέσω μιας φιλικής και λειτουργικής «Διεπαφής Χρήστη» (UI). Η ανάπτυξη βασίστηκε σε «Καθαρή Αρχιτεκτονική» (Clean Architecture), με τεχνολογίες όπως Retrofit, OkHttp, Jetpack Compose και Kotlin Coroutines, προσφέροντας ευέλικτη εμπειρία χρήσης.

Θεματική Περιοχή: Πληροφορική, Ιδιωτικότητα, Τεχνητή Νοημοσύνη

Λέξεις Κλειδιά: εφαρμογή android, προσωπικός βοηθός, ιδιωτικότητα, τοπική λύση, μεγάλα γλωσσικά μοντέλα

Abstract

The present thesis focuses on the development of an application that allows users to interact with Large Language Models (LLMs) without the use of cloud services, offering greater security and control over their data. The application, called "Home AI," utilizes resources from the user's local computer, providing personalized responses.

The main objectives of the application include data privacy through local processing, personalized responses based on the user's data, creating a user-friendly environment, and leveraging local resources to eliminate the need for external services.

To achieve these objectives, a client-server architecture system was designed, providing a secure user experience. The server is based on Role-Based Access Control (RBAC) and implemented using modern technologies such as FastAPI, security with JWT, Postgres, ChromaDB, and Ollama for interaction with LLMs.

The client is an Android application that allows interaction with the server through a friendly and functional User Interface (UI). The client development was based on Clean Architecture, utilizing technologies such as Retrofit, OkHttp, Jetpack Compose, and Kotlin Coroutines, ensuring a flexible user experience.

Thematic Area: Computer Science, Privacy, Artificial Intelligence

Keywords: android application, personal assistant, privacy, local solution, large language models

Πίνακας Περιεχομένων

Εισαγωγή.....	1
Περιγραφή Προβλήματος	1
Βασική Ερευνητική Στόχευση και μεθοδολογική προσέγγιση.....	1
Στόχοι εφαρμογής	2
Επιστημονική Συνεισφορά.....	2
Δομή της πτυχιακής εργασίας	2
Κεφάλαιο 1: Ανάλυση Απαιτήσεων	3
1.1 Ρόλοι Χρηστών.....	3
1.2 Λειτουργίες Χρηστών.....	3
1.3 Απαιτήσεις Εφαρμογής.....	3
1.4 Περιορισμοί Εφαρμογής	4
Κεφάλαιο 2: Λειτουργίες και Χαρακτηριστικά Εφαρμογής.....	5
2.1 Επισκόπηση Εφαρμογής.....	5
2.2 Λειτουργίες Server.....	5
2.3 Χαρακτηριστικά Server	6
2.4 Λειτουργίες Client	6
2.5 Χαρακτηριστικά Client	6
Κεφάλαιο 3: Αρχιτεκτονική Εφαρμογής	8
3.1 Λειτουργία Μοντέλου Client-Sever.....	8
3.2 Αρχιτεκτονική Server	8
3.2.1 Δομή Server.....	8
3.2.2 Βιβλιοθήκες.....	9
3.2.3 Τεχνολογίες.....	9
3.2.4 Διαχείριση και Βάση Δεδομένων	10
3.3 Αρχιτεκτονική Client	11
3.3.1 Δομή Client	11
3.3.2 Βιβλιοθήκες.....	13
3.3.3 Τεχνολογίες.....	13
3.3.4 Διαχείριση Δεδομένων	14
3.4 Επικοινωνία μέσω API	14
3.5 Αρχιτεκτονικής Συστήματος Home AI	14
Κεφάλαιο 4: Εγχειρίδιο Χρήσης.....	15
4.1 Εγκατάσταση της εφαρμογής Home AI.....	15
4.1.1 Εγκατάσταση Server.....	15
4.1.2 Εγκατάσταση Client	16
4.1.3 Σύνδεση της android εφαρμογής με τον server	17

4.2 Λογαριασμός Χρήστη	18
4.2.1 Είσοδος στην εφαρμογή.....	18
4.2.2 Δημιουργία Λογαριασμού.....	18
4.2.3 Επεξεργασία στοιχείων προφίλ.....	19
4.2.4 Αλλαγή κωδικού πρόσβασης	20
4.2.5 Έξοδος από την εφαρμογή	21
4.3 Συνομιλίες	22
4.3.1 Δημιουργία νέας συνομιλίας.....	22
4.3.2 Τρόποι εισαγωγής μηνυμάτων στην εφαρμογή	22
4.3.3 Συνέχιση προηγούμενης συνομιλίας	23
4.3.4 Διαγραφή παλαιότερων συνομιλιών	24
4.4 Έγγραφα	26
4.4.1 Ανέβασμα ενός έγγραφου	26
4.4.2 Επιλογή εγγράφων σε συνομιλία.....	27
4.4.3 Διαγραφή παλαιότερων εγγράφων.....	28
Κεφάλαιο 5: Αποτελέσματα, Οφέλη και Μελλοντικές Επεκτάσεις	30
5.1 Αποτελέσματα	30
5.2 Οφέλη.....	30
5.3 Μελλοντικές Επεκτάσεις	30
Κεφάλαιο 6: Συμπεράσματα.....	32
Κατάλογος Εικόνων	33
Κατάλογος Διαγραμμάτων	34
Πίνακας Ορολογίας.....	35
Πίνακας Συντμήσεων-Αρκτικόλεξων.....	36
Βιβλιογραφία.....	37

Εισαγωγή

Ζούμε σε μια εποχή όπου η πρόσβαση στην πληροφορία είναι πιο εύκολη και άμεση από ποτέ. Ακόμα και άτομα χωρίς εξειδικευμένες γνώσεις πληροφορικής μπορούν να διαχειρίζονται τις περισσότερες καθημερινές τους υποχρεώσεις, χρησιμοποιώντας αποκλειστικά το κινητό τους τηλέφωνο. Αυτή η αυξημένη προσβασιμότητα ενισχύεται από τη ραγδαία πρόοδο της τεχνολογίας, με ιδιαίτερη έμφαση στις καινοτομίες στον τομέα της Τεχνητής Νοημοσύνης (AI), οι οποίες συνεχώς βελτιώνουν τη λειτουργικότητα και τη χρηστικότητα των εφαρμογών και συστημάτων. Αυτό έχει γίνει ιδιαίτερα έντονο τα τελευταία χρόνια, καθώς η Τεχνητή Νοημοσύνη έχει επιφέρει σημαντικές αλλαγές στον τρόπο με τον οποίο οι άνθρωποι αλληλεπιδρούν με την τεχνολογία, καθιστώντας την ακόμη πιο προσιτή και αποτελεσματική (AT&T, 2024).

Τα τελευταία χρόνια, όροι όπως «Τεχνητή Νοημοσύνη», «Νευρωνικά Δίκτυα» και «Μεγάλα Γλωσσικά Μοντέλα» (LLMs) κυριαρχούν στα θέματα των ΜΜΕ, παρουσιάζοντας τα ως τους κύριους πυλώνες της μελλοντικής τεχνολογικής εξέλιξης. Αυτές οι τεχνολογίες, που κάποτε ανήκαν αποκλειστικά στα ερευνητικά εργαστήρια, είναι πλέον διαθέσιμες σε καθημερινούς χρήστες χάρη στην ανάπτυξη και τη διάθεση εργαλείων που χρησιμοποιούν μοντέλα Τεχνητής Νοημοσύνης (Michael Haenlein, 2019). Μεγάλες πολυεθνικές εταιρείες έχουν προσφέρει δωρεάν πρόσβαση σε εργαλεία τα οποία μπορούν σε αμελητέο χρόνο να δημιουργήσουν κείμενο (ChatGPT), εικόνες (DALL-E), ήχο (ElevenLabs) ακόμα και πολυμεσικά αρχεία, όπως βίντεο (Sora), με απλές εντολές. Αυτά τα εργαλεία αυξάνουν την παραγωγικότητα, αυτοματοποιώντας επαναλαμβανόμενες εργασίες και δίνοντας τη δυνατότητα στους ανθρώπους να επικεντρωθούν σε πιο δημιουργικές και στρατηγικές δραστηριότητες (Dell'Acqua, et al., 2023).

Περιγραφή Προβλήματος

Παρά τα σημαντικά πλεονεκτήματα που προσφέρουν αυτά τα νέα εργαλεία, η εξάρτησή τους από υπηρεσίες νέφους (cloud) και η ανάγκη απομακρυσμένης συλλογής, επεξεργασίας και αποθήκευσης των δεδομένων των χρηστών, για την βελτίωση των εργαλείων αυτών, δημιουργούν ανησυχίες σχετικά την διασφάλιση της ιδιωτικότητας και της προστασίας των χρηστών (Council, 2024). Σε αυτό το πλαίσιο, η ανάπτυξη λύσεων που επιτρέπουν την τοπική επεξεργασία δεδομένων με τη χρήση τεχνολογιών Τεχνητής Νοημοσύνης γίνεται όλο και πιο επιτακτική (Fei Liu, 2024) (Gruber, 2024).

Η παρούσα πτυχιακή εργασία εστιάζει στην ανάπτυξη μίας εφαρμογής για κινητά τηλέφωνα Android, ονόματι **Home AI**. Ο χρήστης μέσω της εφαρμογής θα μπορεί να αλληλεπιδρά με Μεγάλα Γλωσσικά Μοντέλα μέσω μιας τοπικής υποδομής. Η εφαρμογή θα είναι ικανή να προσφέρει εξατομικευμένες απαντήσεις χωρίς να θυσιάζεται παράλληλα η ιδιωτικότητα και η ασφάλεια των δεδομένων του χρήστη.

Βασική Ερευνητική Στόχευση και μεθοδολογική προσέγγιση

Η βασική ερευνητική στόχευση της παρούσας εργασίας είναι η ανάπτυξη μίας εφαρμογής που επιτρέπει τον χρήστη να εκτελεί τοπικά LLMs ώστε να λαμβάνει εξατομικευμένα αποτελέσματα με χρήση των προσωπικών του δεδομένων. Με αυτή την υλοποίηση, ο χρήστης θα διασφαλίζει την ιδιωτικότητά του χωρίς να αποκλείεται η πρόσβασή του στην πληροφορία.

Για να πραγματοποιηθεί η στόχευση αυτή, δημιουργήθηκε μια αρχιτεκτονική client-server. Ο server της εφαρμογής εκτελείται σε τοπικό υπολογιστή του ίδιου δικτύου που θα είναι παράλληλα και οι συσκευές των client της εφαρμογής. Η επεξεργασία και η αποθήκευση των δεδομένων του χρήστη παραμένουν τοπικά και δεν αποστέλλονται σε απομακρυσμένους διακομιστές. Ο client της εφαρμογής είναι μια εφαρμογή με φιλικό περιβάλλον χρήστη (UI), επιτρέποντας στους χρήστες να αξιοποιούν τις λειτουργίες του server με ευκολία, χωρίς να

απαιτούνται εξειδικευμένες γνώσεις για τον τρόπο λειτουργίας του συστήματος ή την τεχνολογική του υποδομή.

Στόχοι εφαρμογής

Η ανάπτυξη της εφαρμογής Home AI στοχεύει στην εκπλήρωση των παρακάτω στόχων:

1. **Αξιοποίηση Τοπικών Πόρων:** Για την ομαλή και αυτόνομη λειτουργία της εφαρμογής να χρησιμοποιηθούν υπολογιστικοί πόροι ενός τοπικού υπολογιστή στο δίκτυο του χρήστη, εξαλείφοντας την ανάγκη για χρήση υπολογιστικών υπηρεσιών νέφους (cloud services).
2. **Ιδιωτικότητα Δεδομένων:** Η συλλογή, επεξεργασία και η αποθήκευση των δεδομένων του χρήστη στην εφαρμογή θα γίνεται τοπικά στο κινητό και στον προσωπικό υπολογιστή του χρήστη.
3. **Διαχείριση Προσωπικών Δεδομένων:** Η εφαρμογή να δίνει την δυνατότητα στον χρήστη να επεξεργαστεί τα προσωπικά στοιχεία του λογαριασμού του.
4. **Εξατομίκευση Απαντήσεων:** Η εφαρμογή να παρέχει την δυνατότητα στον χρήστη να ανεβάσει τα δικά του προσωπικά αρχεία στην εφαρμογή, ώστε να λαμβάνει εξατομικευμένες απαντήσεις βάση των δικών του δεδομένων.
5. **Φιλικό Περιβάλλον για τον Χρήστη:** Το περιβάλλον της εφαρμογής να είναι εύχρηστο και απλό, ώστε ο χρήστης που ανοίγει για πρώτη φορά την εφαρμογή να ξέρει τις λειτουργίες της εφαρμογής και να περιηγηθεί εύκολα σε αυτές.

Επιστημονική Συνεισφορά

Η αναμενόμενη επιστημονική συνεισφορά της παρούσας πτυχιακής εργασίας είναι η δημιουργία μιας πλατφόρμας που αποδεικνύει ότι η ασφαλής και εξατομικευμένη επεξεργασία δεδομένων με χρήση LLMs μπορεί να γίνει τοπικά χωρίς τη χρήση υπηρεσιών νέφους. Η εργασία συνεισφέρει στην έρευνα γύρω από την προστασία της ιδιωτικότητας στα συστήματα Τεχνητής Νοημοσύνης (AI).

Δομή της πτυχιακής εργασίας

Η δομή των κεφαλαίων της πτυχιακής εργασίας είναι:

- **Κεφάλαιο 1: Ανάλυση Απαιτήσεων**
Αναλύονται οι τύποι χρηστών της εφαρμογής, οι διαθέσιμες λειτουργίες, καθώς και οι απαιτήσεις και οι περιορισμοί που θα προκύψουν από την υλοποίηση της εφαρμογής.
- **Κεφάλαιο 2: Λειτουργίες και Χαρακτηριστικά Εφαρμογής**
Αναλύονται οι βασικές λειτουργίες που θα έχει η εφαρμογή Home AI, εστιάζοντας στα δύο κύρια μέρη της: client και sever.
- **Κεφάλαιο 3: Αρχιτεκτονική Εφαρμογής**
Αναφέρονται οι βιβλιοθήκες και οι τεχνολογίες που χρησιμοποιήθηκαν, η διαχείριση των δεδομένων της εφαρμογής και τα δομικά μέρη του client και του server.
- **Κεφάλαιο 4: Εγχειρίδιο Χρήσης**
Αναλυτικό εγχειρίδιο χρήσης της εφαρμογής που απευθύνεται σε απλούς χρήστες.
- **Κεφάλαιο 5: Αποτελέσματα, Οφέλη και Μελλοντικές Επεκτάσεις**
Αναφέρονται τα αποτελέσματα, τα οφέλη και οι πιθανές μελλοντικές επεκτάσεις της πτυχιακής εργασίας.
- **Κεφάλαιο 6: Συμπεράσματα**
Αναγράφονται τα τελικά συμπεράσματα.

Κεφάλαιο 1: Ανάλυση Απαιτήσεων

Σε αυτό το κεφάλαιο αναλύονται οι τύποι χρηστών της εφαρμογής, οι λειτουργίες που θα διαθέτουν, καθώς και οι απαιτήσεις και οι περιορισμοί που θα παρουσιάζει η εφαρμογή.

1.1 Ρόλοι Χρηστών

Η εφαρμογή θα εξυπηρετεί δύο τύπους χρηστών: τον διαχειριστή (admin) και το μέλος σπιτιού (house member). Τα δικαιώματα του διαχειριστή θα είναι να βλέπει όλους τους εγγεγραμμένους χρήστες της εφαρμογής και να τροποποιεί τα στοιχεία και την κατάσταση των λογαριασμών τους. Το μέλος σπιτιού θα είναι ο βασικός χρήστης της εφαρμογής, με δυνατότητα να λαμβάνει εξατομικευμένες απαντήσεις με βάση τα δεδομένα του.

1.2 Λειτουργίες Χρηστών

Οι λειτουργίες των χρηστών θα είναι οι ακόλουθες:

Διαχειριστής (admin):

- Πρόσβαση και προβολή όλων των εγγεγραμμένων χρηστών της εφαρμογής.
- Επεξεργασία των στοιχείων και της κατάστασης των λογαριασμών των χρηστών.

Μέλος σπιτιού (house member):

- Εγγραφή στην εφαρμογή.
- Διαχείριση του προσωπικού λογαριασμού.
- Διαχείριση των συνομιλιών του.
- Διαχείριση των αρχείων του.

1.3 Απαιτήσεις Εφαρμογής

Για την απρόσκοπτη λειτουργία της εφαρμογής και την ευρεία χρήση της από διαφορετικούς χρήστες, η ανάπτυξη της θα βασιστεί σε μία αρχιτεκτονική διαχωρισμού μεταξύ του client και του server. Η επιλογή αυτή επιτρέπει στον client να αναλαμβάνει την παρουσίαση των δεδομένων που λαμβάνει από τον server και στον server την κατανομή των απαιτητικών υπολογιστικών λειτουργιών της εφαρμογής.

Ο client θα είναι λιγότερο απαιτητικός από πλευράς υπολογιστικών πόρων, καθώς η επεξεργασία των δεδομένων θα γίνεται κυρίως από τον server. Η συσκευή του client πρέπει να διαθέτει επαρκή μνήμη RAM και να τρέχει ένα σύγχρονο λειτουργικό σύστημα. Η επιλογή για την ανάπτυξη του client βασίζεται στην εμπειρία μου πάνω στη γλώσσα προγραμματισμού Kotlin. Ως εκ τούτου, η εφαρμογή client θα στοχεύει συσκευές Android και θα αναπτυχθεί με το Android SDK 33+ (Android 13+).

Η προτεινόμενη συσκευή client θα διαθέτει 8 GB μνήμη RAM και τουλάχιστον 300MB διαθέσιμο χώρο αποθήκευσης. Αυτές οι προδιαγραφές εξασφαλίζουν την ομαλή λειτουργία της εφαρμογής στην πλευρά του χρήστη.

Ο server θα αναλάβει την εκτέλεση των πιο απαιτητικών αλγοριθμικών λειτουργιών, όπως την επεξεργασία δεδομένων και τη χρήση μικρών LLMs. Η επιλογή μικρών μοντέλων LLMs εξασφαλίζει χαμηλότερες απαιτήσεις σε υπολογιστικούς πόρους, επιτρέποντας στον server να λειτουργεί αποδοτικά χωρίς την ανάγκη κάρτας γραφικών υψηλών επιδόσεων ή μεγάλης χωρητικότητας μνήμης RAM. Αυτή η προσέγγιση μειώνει σημαντικά το κόστος υλοποίησης και συντήρησης, καθώς οι χρήστες δεν χρειάζεται να επενδύσουν σε ακριβό υλικό για να εξασφαλίσουν την ομαλή λειτουργία της εφαρμογής.

Η βέλτιστη διαμόρφωση για τον server περιλαμβάνει υπολογιστή με σύγχρονο τετραπύρνο επεξεργαστή, 16 GB μνήμη RAM, 10 GB αποθηκευτικού χώρου και σύγχρονο

λειτουργικό σύστημα (Windows, Mac OS ή Linux). Αυτή η υποδομή θα διασφαλίζει την ικανοποιητική απόδοση του server σε ρεαλιστικούς χρόνους απόκρισης.

1.4 Περιορισμοί Εφαρμογής

Η συγκεκριμένη υλοποίηση της εφαρμογής εμφανίζει περιορισμούς λόγο της υλοποίησής της σε τοπικό περιβάλλον. Για παράδειγμα, ένας από τους πιο βασικούς περιορισμούς της αφορά την υπολογιστική ισχύ του server. Επειδή η εφαρμογή λειτουργεί τοπικά, βασίζεται αποκλειστικά στους υπολογιστικούς πόρους που θα διαθέτει ο οικογενειακός υπολογιστής. Η εφαρμογή δεν μπορεί να χρησιμοποιήσει πόρους παρόμοιου βεληνεκούς με εφαρμογές όπως το ChatGPT της OpenAI ή το Gemini της Google. Αυτό συνεπάγεται ότι οι δυνατότητες των LLMs της εφαρμογής είναι περιορισμένες και ότι οι χρήστες της εφαρμογής για να αντλήσουν την πληροφορία που θέλουν θα πρέπει να είναι ακριβής και αναλυτικοί στον λόγο τους.

Ένας ακόμα περιορισμός είναι η διαχείριση της εφαρμογής. Η εφαρμογή έχει σχεδιαστεί για οικιακή χρήση και δεν περιλαμβάνει πολλές προηγμένες δυνατότητες για την κατανομή ρόλων ή την προσαρμογή των δυνατοτήτων της. Ο διαχειριστής της εφαρμογής έχει περιορισμένες δυνατότητες παραμετροποίησης και δεν μπορεί να περιορίσει ή να προσαρμόσει τις λειτουργίες των LLMs, σε σύγκριση με άλλες εμπορικές πλατφόρμες.

Τέλος, η εφαρμογή παρουσιάζει περιορισμούς όσον αφορά την τοπικότητά της. Επειδή η εφαρμογή λειτουργεί μόνο στο τοπικό δίκτυο του server, ο χρήστης δεν έχει τη δυνατότητα να την χρησιμοποιεί όταν βρίσκεται εκτός σπιτιού. Αυτό ίσως δημιουργεί πρόβλημα στις περιπτώσεις που ο χρήστης επιθυμεί να χρησιμοποιήσει την εφαρμογή όταν βρίσκεται σε εξωτερικού χώρους και την έχει ανάγκη. Η αναγκαιότητα για συνεχή και τοπική πρόσβαση στον server καθιστά την εφαρμογή μη διαθέσιμη σε τέτοιες περιπτώσεις.

Κεφάλαιο 2: Λειτουργίες και Χαρακτηριστικά Εφαρμογής

2.1 Επισκόπηση Εφαρμογής

Σε αυτό το κεφάλαιο αναλύονται τις βασικές λειτουργίες που θα έχει η εφαρμογή **Home AI**, εστιάζοντας στα δύο κύρια μέρη της:

- **Server:** διαχειρίζεται τη βάση δεδομένων και τις λειτουργίες αλληλεπίδρασης με τα LLMs που θα επιλέξει ο χρήστης.
- **Client:** android εφαρμογή που αναλαμβάνει την επικοινωνία με τον server και την παρουσίαση των δεδομένων στον χρήστη.

2.2 Λειτουργίες Server

Οι βασικότερες λειτουργίες του server είναι η δημιουργία και διαχείριση λογαριασμού από τον διαχειριστή και το μέλος σπιτιού, η διαχείριση των συνομιλιών και η διαχείριση των εγγράφων.

Δημιουργία και Διαχείριση Λογαριασμού

- **Δημιουργία Λογαριασμού:** ο χρήστης έχει τη δυνατότητα να δημιουργήσει ένα νέο λογαριασμό μέσω του API. Τα στοιχεία του λογαριασμού που θα αποθηκεύσει θα αποθηκευτούν σε βάση δεδομένων.
- **Επεξεργασία Στοιχείων προφίλ:** ο χρήστης θα μπορεί να επεξεργαστεί τα προσωπικά του στοιχεία και να αποθηκεύσει τις νέες αλλαγές στη βάση δεδομένων.
- **Επεξεργασία Στοιχείων προφίλ (διαχειριστής):** ο χρήστης θα μπορεί να επεξεργαστεί και να αποθηκεύσει τα προσωπικά στοιχεία άλλου χρήστη έχοντας μόνο το αναγνωριστικό του χρήστη αυτού.
- **Αλλαγή κωδικού χρήστη:** πέρα από την αλλαγή στοιχείων, για την ασφάλεια του χρήστη, ο χρήστης θα μπορεί να αλλάξει κωδικό πρόσβαση στην εφαρμογή προστατεύοντας, με αυτό τον τρόπο, τα προσωπικά του δεδομένα.
- **Αλλαγή κωδικού χρήστη (διαχειριστής):** ο χρήστης θα μπορεί να αλλάξει των κωδικό πρόσβασης ενός άλλου χρήστη έχοντας μόνο το αναγνωριστικό του χρήστη αυτού.
- **Είσοδος στην εφαρμογή:** για την εξατομικευμένη πρόσβαση του χρήστη στην εφαρμογή, ο χρήστης θα εισάγει το email και τον κωδικό πρόσβασης στην φόρμα εισόδου ώστε να λαμβάνει ένα JWT token το οποίο θα επιβεβαιώνει την ταυτότητά του με ασφάλεια .

Διαχείριση Συνομιλιών

- **Νέα συνομιλία:** ο χρήστης θα μπορεί να ξεκινήσει νέα συνομιλία με το Μεγάλο Γλωσσικό Μοντέλο της επιλογής του χωρίς το μοντέλο να έχει μνήμη από προηγούμενες συζητήσεις που έκανε με τον χρήστη.
- **Συνέχιση Προηγούμενης Συνομιλίας:** ο χρήστης θα μπορεί να συνεχίσει μια προηγούμενη συζήτηση με το Μεγάλο Γλωσσικό Μοντέλο της επιλογής του, έχοντας το μοντέλο τη γνώση των προηγούμενων μηνυμάτων που ανταλλάχθηκαν στη συγκεκριμένη συνομιλία.
- **Διαχείριση παλαιότερων συνομιλιών:** ο χρήστης μπορεί να προβάλει όλες τις συνομιλίες που είχε κάνει στο παρελθόν ή να διαγράψει συνομιλίες που δεν επιθυμεί να διατηρήσει.

Διαχείριση Εγγράφων

- **Ανέβασμα εγγράφων:** δύναται η δυνατότητα στον χρήστη να προσθέσει γνώση στο Μεγάλο Γλωσσικό Μοντέλο της επιλογής του, με χρήση RAG (Amazon Web Services, n.d.), ανεβάζοντας αρχεία που θα θέλει να αναλύσει.

- **Διαχείριση Παλαιότερων εγγράφων:** ο χρήστης μπορεί να χρησιμοποιήσει στις συνομιλίες του έγγραφα που έχει ανεβάσει στο παρελθόν στην εφαρμογή ή να διαγράψει παλιότερα έγγραφα που δεν επιθυμεί να διατηρήσει.

2.3 Χαρακτηριστικά Server

Τα βασικότερα χαρακτηριστικά του server είναι η προστασία των δεδομένων του χρήστη. Αναλυτικότερα:

- **Προστασία Δεδομένων Χρήστη από άλλους χρήστες της εφαρμογής**
Ο server δίνει πρόσβαση μόνο στους εγγεγραμμένους χρήστες της εφαρμογής. Ανάλογα με τον ρόλο που έχουν οι χρήστες θα έχουν και ανάλογη πρόσβαση στα δεδομένα της εφαρμογής. Για παράδειγμα, ένα μέλος σπιτιού δεν έχει πρόσβαση σε έγγραφα ή μηνύματα άλλων χρηστών της εφαρμογής. Γι' αυτό το λόγο το API του server εφαρμόσει τον τύπο «Ελέγχου Πρόσβασης Βάσει Ρόλων» (RBAC) (Red Hat, 2024).
- **Προστασία Δεδομένων Χρήστη από τρίτες υπηρεσίες**
Η ομαλή λειτουργία του server επιτυγχάνεται με την απουσία χρήσης τρίτων υπηρεσιών εκτός τοπικού δικτύου. Όλη επεξεργασία και αποθήκευση αρχείων και συνομιλιών γίνεται εντός τοπικού δικτύου.

2.4 Λειτουργίες Client

Ο χρήστης του client θα είναι μέλος σπιτιού. Οι λειτουργίες του client έχουν ως στόχο τη διευκόλυνση του χρήστη στην αλληλεπίδραση του με τον server. Αναλυτικά:

- **Σύνδεση με την εφαρμογή**
Ο χρήστης μπορεί να χρησιμοποιήσει την εφαρμογή μόνο αν είναι συνδεδεμένος. Επομένως αποκτά πρόσβαση εισάγοντας στην πλατφόρμα το email του και των κωδικό πρόσβασής (password) του αν είναι ήδη εγγεγραμμένος χρήστης ή κάνοντας εγγραφή στην εφαρμογή αν είναι νέος χρήστης.
- **Εισαγωγή διεύθυνσης IP server**
Ο χρήστης για να χρησιμοποιήσει την εφαρμογή θα πρέπει πρώτα ο client να συνδεθεί με τον server. Επειδή ο server θα έχει διαφορετική διεύθυνση IP ανά εγκατάσταση, ο client πρέπει να δίνει την δυνατότητα στον χρήστη να εισάγει χειροκίνητα τη διεύθυνση IP του server.
- **Αποστολή και διαχείριση μηνυμάτων**
Ο χρήστης μπορεί να ξεκινήσει μία νέα συνομιλία, να συνεχίσει ή να διαγράψει μία ήδη υπάρχουσα συνομιλία από τον λογαριασμό του.
- **Αποστολή και διαχείριση αρχείων**
Ο χρήστης μπορεί να μεταφορτώσει ένα αρχείο ή να διαγράψει ένα ήδη υπάρχων αρχείο από τον server της εφαρμογής.

2.5 Χαρακτηριστικά Client

Τα χαρακτηριστικά του client βοηθούν τον χρήστη να έχει μία πιο εύκολη και ομαλή αλληλεπίδραση με τον server. Πιο συγκεκριμένα:

- **Διατήρηση JWT token**
Με το JWT token ο χρήστης θα μπορεί να συνδέεται στην εφαρμογή χωρίς να χρειάζεται να εισάγει πολλαπλές φορές τα στοιχεία του στη φόρμα εισόδου μέχρι να λήξει το JWT token.
- **Υποστήριξη πολλαπλών τρόπων εισαγωγής νέων μηνυμάτων**
Ο χρήστης θα μπορεί να συνεχίσει μία συνομιλία είτε χρησιμοποιώντας το πληκτρολόγιο του κινητού του είτε χρησιμοποιώντας φωνητικές εντολές μέσω του μικροφώνου του κινητού του.

- **Δήλωση κατάστασης εφαρμογής**
Το σύστημα να προσφέρει οπτική ανατροφοδότηση για την κατάσταση που βρίσκεται η εφαρμογή κάθε χρονική περίοδο. Για παράδειγμα, να υπάρχει κατάσταση φόρτωσης (Loading State) όταν το Μεγάλο Γλωσσικό Μοντέλο με το οποίο έχει ξεκινήσει μία συνομιλία επεξεργάζεται το μήνυμα του χρήστη και προετοιμάζει την απάντηση.
- **Διατήρηση Διεύθυνσης IP server**
Κάθε σπίτι έχει τις δικιές του ρυθμίσεις και τοπολογίες στο τοπικό του δίκτυο. Ως αποτέλεσμα κάθε συσκευή να ανήκει σε διαφορετικό δίκτυο ή να έχει διαφορετική IP διεύθυνση από ένα άλλο δίκτυο. Η εφαρμογή πρέπει να διατηρεί την IP διεύθυνση του server όταν ο χρήστης παραμένει συνδεδεμένος στην εφαρμογή.
- **Φιλικό Περιβάλλον χρήστη**
Το περιβάλλον της εφαρμογής θα είναι φιλικό ακόμα και για αρχάριους χρήστες κινητών. Θα μπορεί να προσαρμόζεται ανάλογα με την οθόνη της συσκευής του χρήστη και να παραμετροποιείται σύμφωνα με τις επιλογές του λειτουργικού του χρήστη.

Κεφάλαιο 3: Αρχιτεκτονική Εφαρμογής

Η εφαρμογή Home AI υλοποιεί την αρχιτεκτονική client-server, ένα ευρέως διαδεδομένο αρχιτεκτονικό μοντέλο που διαχωρίζει τις λειτουργίες και τους πόρους μεταξύ του client, ο οποίος αιτείται υπηρεσίες, και του server, ο οποίος τις παρέχει. Στην περίπτωση της Home AI, ο client είναι μια εφαρμογή Android που εκτελείται στη συσκευή του χρήστη και προσφέρει μια φιλική διεπαφή για την αλληλεπίδραση με το σύστημα. Ο server, από την άλλη πλευρά, αναλαμβάνει την επεξεργασία των δεδομένων και τη διαχείριση των συνομιλιών με τα Μεγάλα Γλωσσικά Μοντέλα (LLMs), παρέχοντας την απαιτούμενη υπολογιστική ισχύ και τη λειτουργικότητα για την ανάλυση των αιτημάτων.

3.1 Λειτουργία Μοντέλου Client-Server

Η λειτουργία του μοντέλου client-server βασίζεται σε έναν κύκλο αιτημάτων (requests) και αποκρίσεων (responses) μεταξύ του client και του server.

Παράδειγμα κύκλου για το Home AI:

- 1. Αίτημα από τον Client:** Ο χρήστης αλληλεπιδρά με τον εφαρμογή Android, η οποία αποστέλλει αιτήματα στον server μέσω HTTP αιτημάτων. Αυτά τα αιτήματα μπορούν να περιλαμβάνουν αυθεντικοποίηση, αποστολή μηνυμάτων και μεταφόρτωση εγγράφων.
- 2. Επεξεργασία από τον Server:** Ο server λαμβάνει τα αιτήματα, τα επεξεργάζεται και αλληλεπιδρά με τα LLMs για να παράγει τις απαιτούμενες απαντήσεις. Επίσης, διαχειρίζεται τη βάση δεδομένων για την αποθήκευση πληροφοριών όπως το προφίλ, τις συνομιλίες και τα έγγραφα των χρηστών.
- 3. Απόκριση προς τον Client:** Ο server επιστρέφει τις απαντήσεις στον client, ο οποίος τις παρουσιάζει στον χρήστη μέσω του περιβάλλοντος εργασίας της εφαρμογής.

3.2 Αρχιτεκτονική Server

3.2.1 Δομή Server

Ο server της εφαρμογής εκτελείται μέσα σε ένα Docker container. Το docker container περιλαμβάνει ένα φάκελο "app", ο οποίος αποτελείται από τα παρακάτω αρχεία:

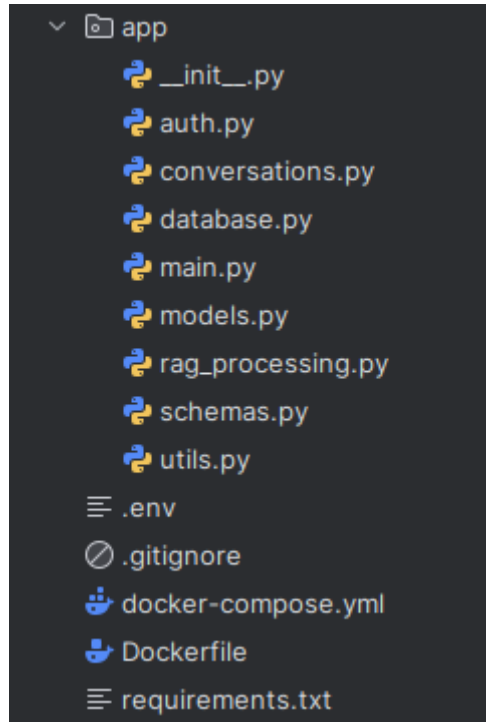
- **auth.py:** Διαχειρίζεται τις λειτουργίες αυθεντικοποίησης και τη δημιουργία JWT tokens.
- **conversation.py:** Περιλαμβάνει λειτουργίες για τη διαχείριση των συνομιλιών μεταξύ χρηστών και LLMs (LangChain) (Ollama).
- **database.py:** Συνδέει τις κλάσεις των μοντέλων της εφαρμογής με τους πίνακες της βάσης δεδομένων PostgreSQL. Εάν δεν υπάρχει ήδη βάση, την δημιουργεί κατά την εκκίνηση της εφαρμογής (PostgreSQL).
- **main.py:** Συνδέει τις διαδρομές (routes) του FastAPI με τις υπόλοιπες λειτουργίες της εφαρμογής (FastAPI).
- **models.py:** Περιλαμβάνει τις κλάσεις των μοντέλων της εφαρμογής.
- **rag_processing.py:** Υπεύθυνο για την επεξεργασία και αποθήκευση των embeddings των αρχείων που φορτώνουν οι χρήστες (Chroma).
- **schemas.py:** Περιλαμβάνει σχήματα που επαληθεύουν τη σωστή μορφοποίηση των δεδομένων κατά την εισαγωγή και εξαγωγή τους (Pydantic).
- **utils.py:** Δημιουργεί έναν χρήστη, τον Home AI, για την αποθήκευση μηνυμάτων LLMs στη βάση δεδομένων, και διαχειρίζεται το κλειδί για την κρυπτογράφηση και αποκρυπτογράφηση των JWT tokens.

Για την ομαλή λειτουργία του docker container, πριν δημιουργηθεί το docker image, ορίζονται οι μεταβλητές περιβάλλοντος της εφαρμογής στο αρχείο .env, οι βιβλιοθήκες της Python στο

requirements.txt, οι παράμετροι του docker image στο Dockerfile και μετά εκτελείται το “docker-compose.yml” αρχείο (Docker).

- **.env**: Ορίζει τις καθολικές μεταβλητές του προγράμματος.
- **requirements.txt**: Ορίζει τις βιβλιοθήκες που θα έχει ο server.
- **Dockerfile**: Καθορίζει τις οδηγίες για τη δημιουργία μίας εικόνας (image) ενός container.
- **docker-compose.yml**: Συντονίζει και διαχειρίζεται την εκτέλεση του home-ai container.

Στην Εικόνα 1 παρουσιάζεται η δομή των αρχείων του server.



Εικόνα 1: Δομή Αρχείων Server

3.2.2 Βιβλιοθήκες

Οι σημαντικότερες βιβλιοθήκες που χρησιμοποιούνται είναι:

- **FastAPI**: Επικοινωνία μεταξύ server και client.
- **SQLAlchemy**: Διαχείριση της βάσης δεδομένων.
- **Pydantic**: Επικύρωση των δεδομένων στη βάση.
- **JWT**: Διαχείριση tokens επαλήθευσης χρηστών.
- **Langchain**: Δημιουργία συζητήσεων με μνήμη για τα LLMs.
- **Ollama**: Χρήση και παραμετροποίηση των LLMs.

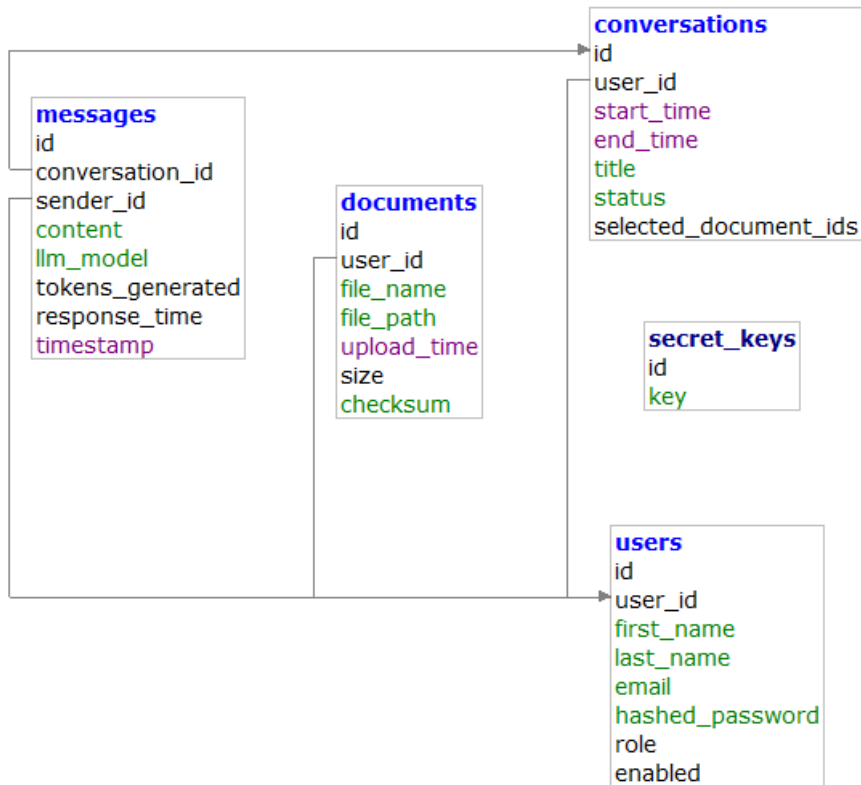
3.2.3 Τεχνολογίες

Οι βασικές τεχνολογίες που χρησιμοποιούνται στον server είναι:

- **Postgres**: Αποθήκευση δεδομένων προφίλ χρηστών, συνομιλιών και εγγράφων.
- **ChromaDB**: Επεξεργασία και αποθήκευση embeddings εγγράφων που ανεβάζουν οι χρήστες.
- **Docker**: Εύκολη εγκατάσταση και χρήση του server.

3.2.4 Διαχείριση και Βάση Δεδομένων

Η βάση δεδομένων PostgreSQL αποθηκεύει πληροφορίες χρηστών, συνομιλίες και μεταδεδομένα (metadata) εγγράφων. Αυτό επιτρέπει στον server να διαχειρίζεται αποτελεσματικά τις αλληλεπιδράσεις μεταξύ των χρηστών και LLMs. Ακολουθεί το σχήμα της βάσης δεδομένων (Διάγραμμα 1) όπως εμφανίζεται μέσα από το πρόγραμμα Adminer (Adminer, n.d.).



Διάγραμμα 1: Σχήμα Βάσης Δεδομένων PostgreSQL

3.3 Αρχιτεκτονική Client

Η υλοποίηση του client βασίζεται σε σύγχρονες τεχνολογίες και αρχιτεκτονικά πρότυπα. Είναι μια εφαρμογή Android, υλοποιημένη με Kotlin, ακολουθώντας το πρότυπο MVVM και την Clean Architecture (Dumbravan, 2022). Για τη διαχείριση των εξαρτήσεων χρησιμοποιείται το Dagger Hilt.

3.3.1 Δομή Client

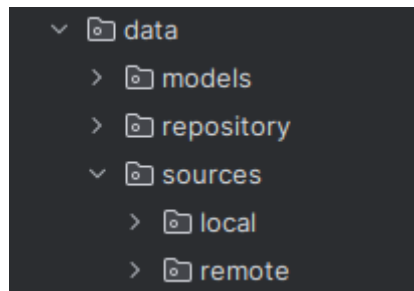
Ο κώδικας της εφαρμογής είναι χωρισμένος σε γενικές κατηγορίες, με υποκατηγορίες όπου απαιτείται. Η τρεις βασικές κατηγορίες είναι:

Επίπεδο Δεδομένων (Data Layer)

Ασχολείται με την πρόσβαση και τη διαχείριση των δεδομένων. Περιλαμβάνει τα παρακάτω πακέτα:

- **Models:** Μοντέλα για την ανταλλαγή πληροφορίας με την εφαρμογή.
- **Repository:** Υπεύθυνα για τη συλλογή δεδομένων από τοπικές βάσεις ή απομακρυσμένα APIs.
- **Sources:** Διαχειρίζεται την επικοινωνία με πηγές δεδομένων.
 - **Local:** Περιλαμβάνει το “SharedPreferencesManager.kt” το οποίο είναι υπεύθυνο για την επεξεργασία των δεδομένων στο Shared Preferences της εφαρμογής.
 - **Remote:** περιλαμβάνει αρχεία για την επικοινωνία με τον server μέσω API.

Η Εικόνα 2 παρουσιάζει την δομή των αρχείων στο επίπεδο δεδομένων.



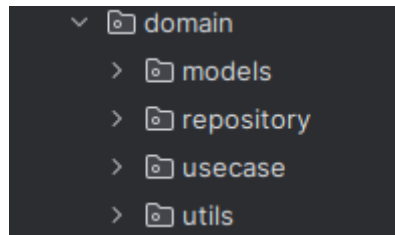
Εικόνα 2: Δομή Αρχείων Data Layer

Επίπεδο Τομέα (Domain Layer)

Υλοποιεί τις λειτουργικές δυνατότητες της εφαρμογής. Περιλαμβάνει τα παρακάτω πακέτα:

- **Models:** Μοντέλα για την εσωτερική επεξεργασία των δεδομένων.
- **Repository:** Interfaces που καθορίζουν τις λειτουργίες των repositories.
- **UseCase:** Διαχειρίζονται την επεξεργασία δεδομένων μεταξύ Data Layer και Presentation Layer.
- **Utils:** Περιέχει την κλάση “Result” που ενημερώνει την κατάσταση των δεδομένων.

Η Εικόνα 3 παρουσιάζει την δομή των αρχείων στο επίπεδο τομέα.



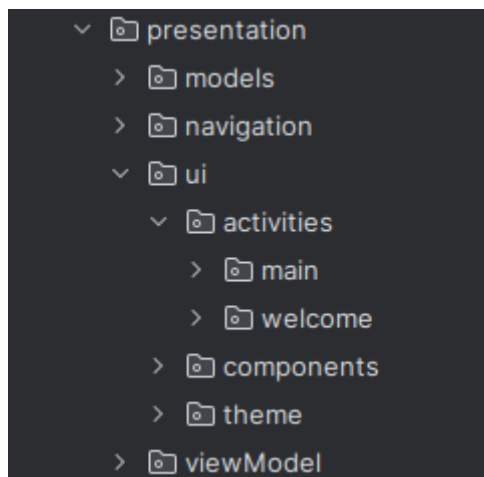
Εικόνα 3: Δομή Αρχείων Domain Layer

Επίπεδο Παρουσίασης (Presentation Layer)

Το επίπεδο αυτό επεξεργάζεται και να μετατρέπει τα αποτελέσματα της εφαρμογής σε καταστάσεις που μπορεί να κατανοήσει ο τελικός χρήστης. Περιλαμβάνει τα ακόλουθα πακέτα:

- **Models:** Μοντέλα για την εμφάνιση πληροφοριών.
- **Navigation:** Διαχείριση της πλοήγησης μεταξύ των οθονών της εφαρμογής.
- **UI** περιλαμβάνει τα:
 - **Activities:** Τα κύρια activities της εφαρμογής - WelcomeActivity και MainActivity.
 - **Components:** Γραφικά στοιχεία των οθονών (Google, n.d.).
 - **Theme:** Ορισμός χρωμάτων και στυλ (Google, n.d.).
- **ViewModels:** Καταστάσεις των οθονών για την ενημέρωση του UI.

Η Εικόνα 4 παρουσιάζει την δομή των αρχείων στο επίπεδο παρουσίασης.

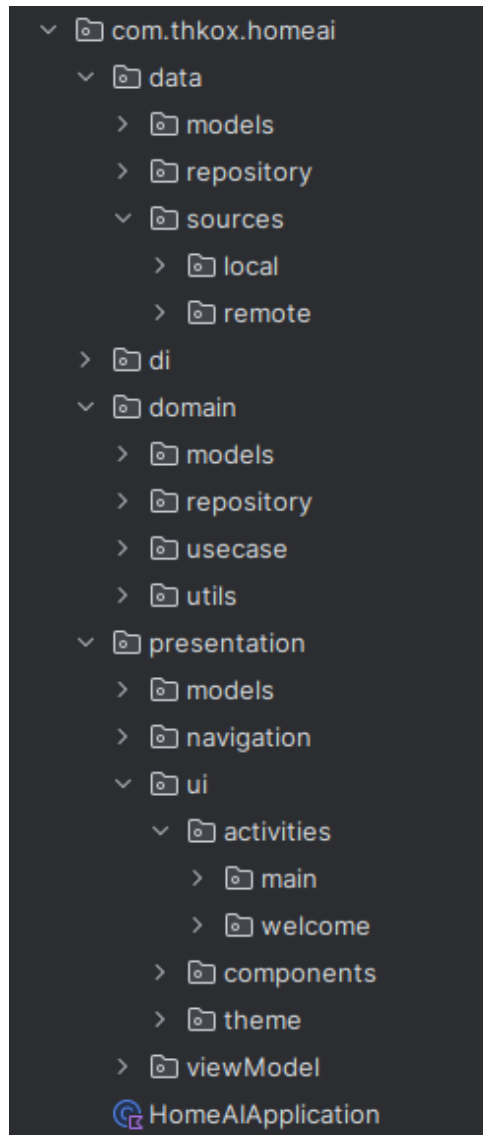


Εικόνα 4: Δομή Αρχείων Presentation Layer

Επιπλέον κατηγορία

- **Di (Dependency Injection):** Περιλαμβάνει το αρχείο AppModule.kt, όπου ορίζονται οι εξαρτήσεις μέσω Dagger Hilt.

Ακολουθεί η Εικόνα 5 που παρουσιάζει την συνολική δομή των αρχείων στην εφαρμογή android.



Εικόνα 5: Δομή Client

3.3.2 Βιβλιοθήκες

Οι βιβλιοθήκες που χρησιμοποιήθηκαν είναι:

- **Retrofit:** Εκτέλεση HTTP αιτημάτων προς τον server.
- **OkHttp:** Διαχείριση συνδέσεων HTTP.
- **Jetpack Compose:** Δημιουργία στοιχείων, οθονών του UI και άδεια χρήσης μικροφώνου (Nam, 2023).
- **Kotlin Coroutines:** Διαχείριση ασύγχρονων λειτουργιών για την ομαλή λειτουργία της εφαρμογής.

3.3.3 Τεχνολογίες

- **Kotlin:** Προτεινόμενη γλώσσα από τη Google για Android εφαρμογές.
- **Jetpack Compose:** Δημιουργία στοιχείων και οθονών του UI.
- **Material Design:** Δημιουργία σύγχρονων γραφικών στοιχείων με δυναμική παραμετροποίηση σύμφωνα με την κατάσταση του λειτουργικού συστήματος.

3.3.4 Διαχείριση Δεδομένων

Η εφαρμογή Android δεν αποθηκεύει τα δεδομένα που ανταλλάσσει με τον server. Αποθηκεύει μόνο τις κατάλληλες παραμέτρους, όπως τη διεύθυνση IP του server και το token αυθεντικοποίησης, για να μπορέσει ο χρήστης να παραμείνει συνδεδεμένος και μετά την επανεκκίνηση της εφαρμογής.

3.4 Επικοινωνία μέσω API

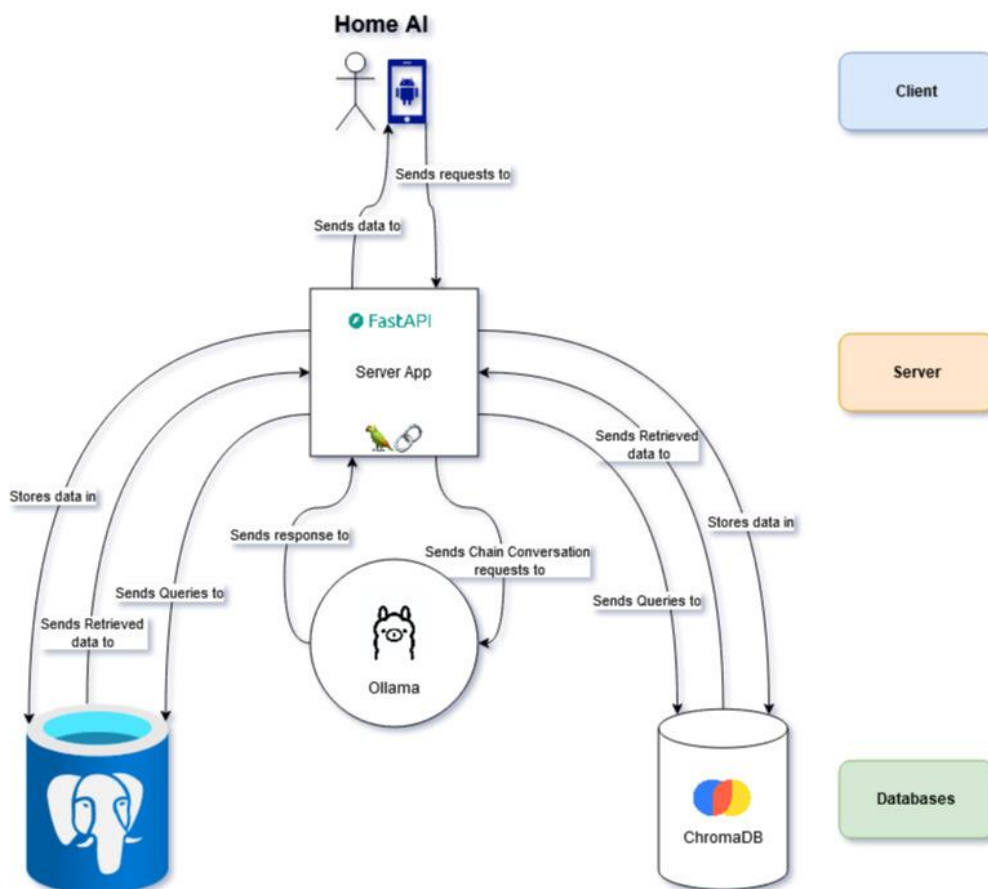
Η επικοινωνία μεταξύ client και server γίνεται μέσω HTTP αιτημάτων. Η επικοινωνία επιτυγχάνεται με τις ακόλουθες βιβλιοθήκες:

- **Για τον Client:**
 - **Retrofit:** Για την δημιουργία αιτημάτων HTTP και την εξεργασία των απαντήσεων.
 - **OkHttp:** Για την διαχείριση των συνδέσεων HTTP με τον server.
- **Για τον server:**
 - **FastAPI:** Για την ανάπτυξη του RESTful API που εξυπηρετεί τα αιτήματα του client με ασφάλεια (FastAPI).

Το RESTful API ως framework επιτρέπει την ανταλλαγή δεδομένων σε μορφή JSON, διασφαλίζοντας την αποτελεσματικότητα και την ασφάλεια των επικοινωνιών.

3.5 Αρχιτεκτονικής Συστήματος Home AI

Το Διάγραμμα 2 απεικονίζει την επικοινωνία μεταξύ των μερών της εφαρμογής.



Διάγραμμα 2: Αρχιτεκτονικής Συστήματος Home AI

Κεφάλαιο 4: Εγχειρίδιο Χρήσης

Ακολουθούν οδηγίες για την εγκατάσταση και χρήση της εφαρμογής από ένα μέλος σπιτιού.

4.1 Εγκατάσταση της εφαρμογής Home AI

Η εφαρμογή για να είναι πλήρως λειτουργική χρειάζεται να εγκατασταθεί σε έναν υπολογιστή της επιλογής σας και στις android συσκευές σας που θα θέλατε να έχετε πρόσβαση στο Home AI.

4.1.1 Εγκατάσταση Server

Απαιτήσεις συστήματος:

- **Λειτουργικό Σύστημα (OS):** Windows, Mac OS, Linux
- **Επεξεργαστής (CPU):** AMD, Intel με τουλάχιστον 4 πυρήνες
- **Μνήμη (RAM):** 16GB
- **Αποθηκευτικός Χώρος (Storage):** 50GB

Προαπαιτήσεις:

Πριν ξεκινήσετε την εγκατάσταση της εφαρμογής στον υπολογιστή θα πρέπει να έχετε μία ενεργή εγκατάσταση της Ollama, της PostgreSQL και της Docker με Docker compose. Οδηγοί εγκατάστασης θα βρείτε εδώ:

- Ollama: <https://ollama.com/download>
- PostgreSQL: <https://www.postgresql.org/download/>
- Docker: <https://docs.docker.com/engine/install/>
- Docker Compose: <https://docs.docker.com/compose/install/>

Στο Ollama θα πρέπει να έχετε κατεβάσει δύο μοντέλα: ένα μοντέλο llm που θα λαμβάνει τα μηνύματα και θα απαντάει με κείμενο (προεπιλογή είναι το "llama3.1:8b-instruct-q4_1") και ένα μοντέλο για την ανάλυση των αρχείων που θα ανεβάζει ο χρήστης (προεπιλογή είναι το "nomic-embed-text").

Στη PostgreSQL θα πρέπει να έχετε έναν ενεργό χρήστη όπου θα χρησιμοποιήσει η εφαρμογή για να αποθηκεύει και να αντλεί δεδομένα.

Στη συνέχεια, κατεβάζετε το αρχείο "home-ai-server.zip" και το αποσυμπίεζεται. Πλοηγηθείτε στον αποσυμπιεσμένο φάκελο που δημιουργήθηκε και ανοίξτε αρχείο ".env" με το πρόγραμμα επεξεργασίας κειμένων της επιλογής σας. Όταν το ανοίξετε θα δείτε τις παραμέτρους που φαίνονται στην Εικόνα 6.

```

1 # Database configuration
2 DATABASE_URL= # 192.168.4.141
3 DATABASE_USERNAME= # postgres
4 DATABASE_PASSWORD= # password1234
5 DATABASE_NAME= # homeai
6
7 # Directories
8 CHROMADB_PERSIST_DIRECTORY= # /chroma_db
9 DOCUMENTS_DIRECTORY= # /documents
10
11 # JWT configuration
12 ALGORITHM=HS256
13 ACCESS_TOKEN_EXPIRE_MINUTES= # 259200 # 6 months
14
15 # Ollama configuration
16 OLLAMA_URL= # http://192.168.4.142:11434/
17 MODEL_NAME= # llama3.1:8b-instruct-q4_1
18 EMBEDDING_MODEL_NAME= # nomic-embed-text
19
20 # App port
21 PORT= # 8000
22

```

Εικόνα 6: Παράμετροι εφαρμογής υπολογιστή

Κάθε παράμετρο έχει σε σχόλια τις προεπιλεγμένες τιμές των παραμέτρων που προτείνονται για να τρέξει η εφαρμογή στον υπολογιστή. Μπορείτε να αλλάξετε τις παραμέτρους με βάση την δική σας υλοποίηση.

Όταν αποθηκεύσετε τις παραμέτρους, κλείστε το αρχείο και ανοίξτε έναν τερματικό (terminal) από τον αποσυμπιεσμένο φάκελο.

Βεβαιωθείτε ότι στον τερματικό βρίσκεστε στο μονοπάτι που είναι ο αποσυμπιεσμένος φάκελος και γράψτε την εντολή:

‘docker-compose build’

Η διαδικασία εγκατάστασης ίσως διαρκέσει μερικά λεπτά. Όταν τελειώσει η διαδικασία θα δείτε ότι η εφαρμογή είναι έτοιμη για χρήση.

Στην Εικόνα 7 φαίνεται ένα παράδειγμα επιτυχημένης εγκατάστασης.

```

Creating home-ai ... done
Attaching to home-ai
home-ai | INFO: Started server process [1]
home-ai | INFO: Waiting for application startup.
home-ai | INFO: Application startup complete.
home-ai | INFO: Uvicorn running on http://0.0.0.0:8000 (Press CTRL+C to quit)

```

Εικόνα 7: Έναρξη εφαρμογής FastAPI

4.1.2 Εγκατάσταση Client

Απαιτήσεις συστήματος:

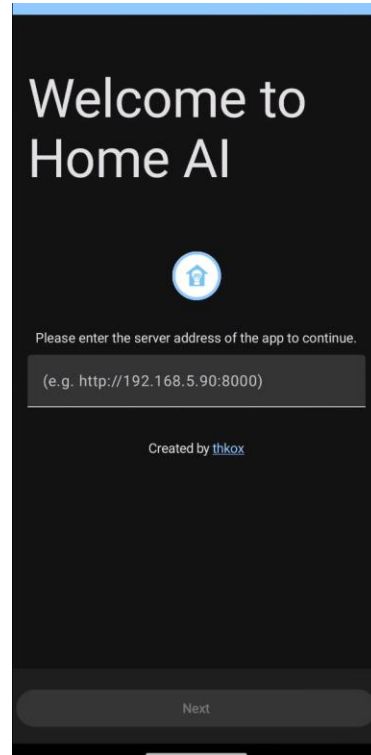
- **Λειτουργικό Σύστημα (OS):** Android SDK 33+
- **Μνήμη (RAM):** 8GB
- **Αποθηκευτικός Χώρος (Storage):** 300MB
- **Συσκευές:** Κινητά και Tablet

Εγκαταστήστε την εφαρμογή “HomeAI.apk” που θα βρείτε στο αρχείο “home-ai.zip” στο κινητό σας ακολουθώντας τις οδηγίες του κατασκευαστή του κινητού σας για εγκατάσταση εφαρμογών εκτός Google Play Store.

Αν η εγκατάσταση ολοκληρωθεί με επιτυχία, ανοίξτε την εφαρμογή και θα πρέπει να δείτε την πρώτη οθόνη της εφαρμογής για εισαγωγή της διεύθυνσης IP του υπολογιστή που τρέχει την εφαρμογή.



Εικόνα 8: Εικονίδιο εφαρμογής



Εικόνα 9: Αρχική Σελίδα Καλωσορίσματος

4.1.3 Σύνδεση της android εφαρμογής με τον server

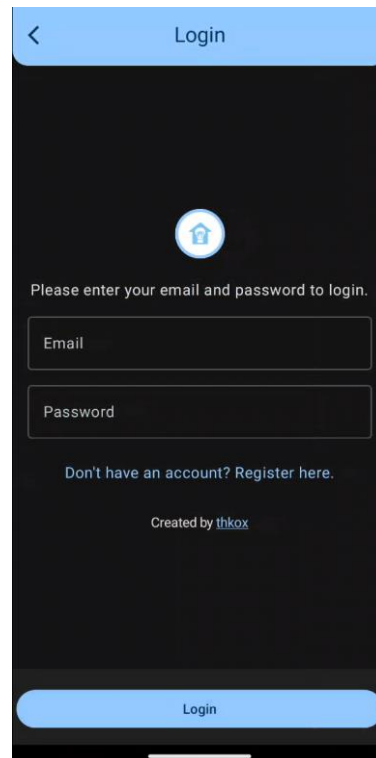
Ανοίγοντας για πρώτη φορά την εφαρμογή θα δείτε την πρώτη οθόνη καλωσορίσματος που σας ζητάει να εισάγεται την IP διεύθυνση στην οποία τρέχει η εφαρμογή. Για παράδειγμα “192.168.3.98”. Η προεπιλεγμένη θύρα (port) που ανοίγει η εφαρμογή κατά την εκτέλεσή της είναι η 8000. Σε περίπτωση που τρέχει σε άλλη θύρα συμπληρώστε τη στη διεύθυνση. Για παράδειγμα “192.168.3.98:4392”.

4.2 Λογαριασμός Χρήστη

Στην ενότητα αυτή θα βρείτε οδηγίες για την σύνδεση και την επεξεργασία στοιχείων του χρήστη για την εφαρμογή.

4.2.1 Είσοδος στην εφαρμογή

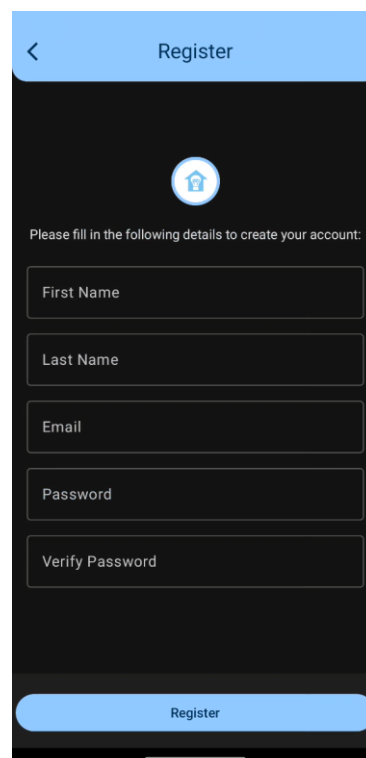
Αφού συνδέσατε την android εφαρμογή με τον server. Θα πρέπει να συνδεθείτε με τον λογαριασμό που έχετε δημιουργήσει. Αρκεί να συμπληρώσετε το email και το password που συμπληρώσατε κατά τη δημιουργία λογαριασμού.



Εικόνα 10: Οθόνη Εισόδου στην εφαρμογή

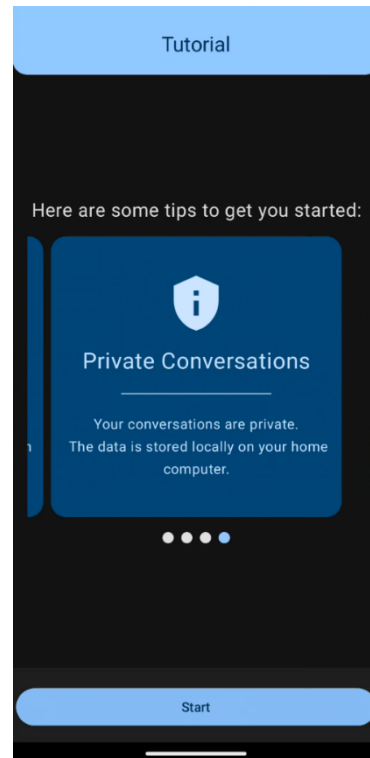
4.2.2 Δημιουργία Λογαριασμού

Αν δεν έχετε κάνει λογαριασμό και είναι η πρώτη σας φορά που χρησιμοποιείται την εφαρμογή αρκεί να πατήσετε πάνω από την επιλογή "Login" που βρίσκεται στην οθόνη εισόδου του χρήστη την επιλογή "Don't have an account? Register here." και η εφαρμογή θα σας ανακατευθύνει στην οθόνη εγγραφής όπου θα συμπληρώσετε ονοματεπώνυμο, email και password.



Εικόνα 11: Οθόνη Δημιουργία Λογαριασμού

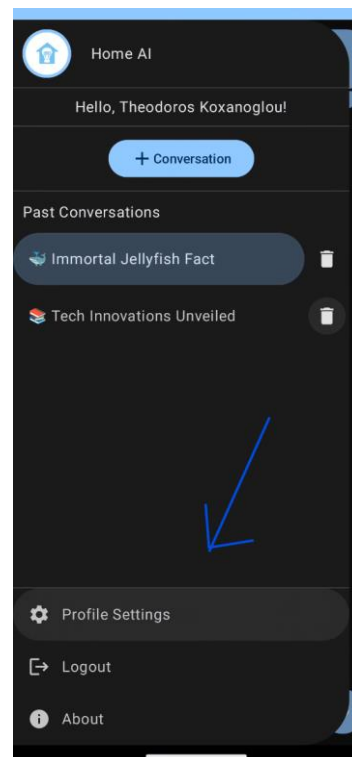
Μετά τη δημιουργία του λογαριασμού θα σας εμφανιστεί μία οθόνη καλωσορίσματος η οποία θα σας αναλύει συνοπτικά τις βασικές λειτουργίες της εφαρμογής και πως θα μπορείτε να τις αξιοποιήσετε.



Εικόνα 12: Οθόνη Γενικών Πληροφοριών

4.2.3 Επεξεργασία στοιχείων προφίλ

Σε περίπτωση που έχετε δημιουργήσει λογαριασμό στην εφαρμογή και θέλετε να επεξεργαστείτε τα στοιχεία προφίλ του λογαριασμού σας, αρκεί να πατήσετε το κουμπί χάμπουργκερ (hamburger button) που βρίσκεται στην πάνω μπάρα αριστερά και μετά να πατήσει την επιλογή με το γρανάτζι "Profile Settings".



Εικόνα 13: Μενού - Επεξεργασία Χρήστη

Όταν το πατήσετε θα σας εμφανιστεί μία οθόνη στην οποία μπορείτε να επεξεργαστείτε το ονοματεπώνυμό σας και το email σας. Όταν ολοκληρώσετε τις αλλαγές πατήστε το κουμπί "Update Profile".

Εικόνα 14: Οθόνη Επεξεργασίας Προφίλ

Εικόνα 15: Μήνυμα επιτυχούς ενημέρωσης προφίλ

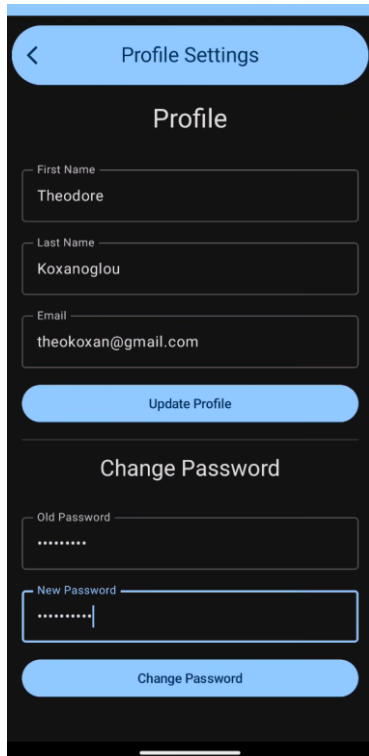
4.2.4 Αλλαγή κωδικού πρόσβασης

Για να αλλάξετε τον κωδικό πρόσβασης του χρήστη αρκεί να μεταβείτε στο μενού πατώντας το εικονίδιο στη πάνω μπάρα αριστερά και να πατήσετε την επιλογή "Profile Settings" με το γρανάζι.

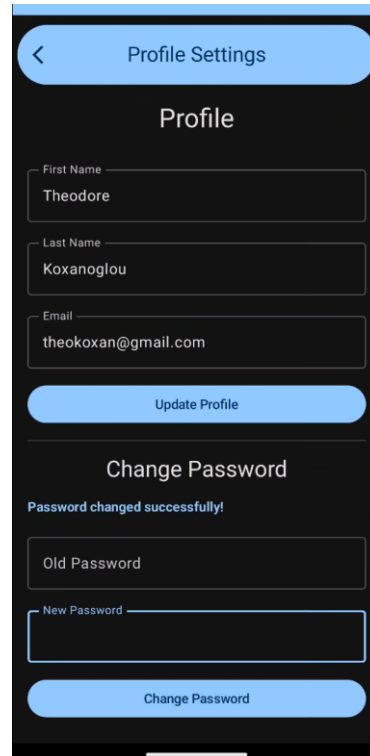
Εικόνα 16: Οθόνη Επεξεργασίας Προφίλ

Στο κάτω μέρος της οθόνης θα σας εμφανιστούν δύο πεδία. Στο πρώτο συμπληρώνετε τον παλιό κωδικό πρόσβασης που είχατε και στο δεύτερο συμπληρώνετε το νέο κωδικό πρόσβασης με τον

οποίο θα θέλατε να χρησιμοποιείται στην εφαρμογή. Όταν ολοκληρώσετε τις αλλαγές πατήστε το κουμπί “Change Password”.



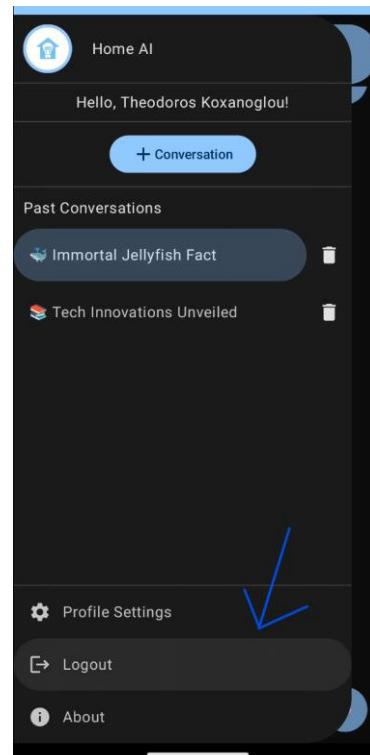
Εικόνα 17: Οθόνη Αλλαγής Κωδικού Πρόσβασης



Εικόνα 18: Επιτυχής Αλλαγή Κωδικού Πρόσβασης

4.2.5 Έξοδος από την εφαρμογή

Αν θέλετε να αποσυνδεθείτε από την εφαρμογή, το μόνο που χρειάζεται να κάνετε είναι να πατήσετε το εικονίδιο που βρίσκεται στην πάνω μπάρα αριστερά και να πατήσετε την επιλογή “Logout”.



Εικόνα 19: Μενού - Επιλογή εξόδου από τον λογαριασμό της εφαρμογής

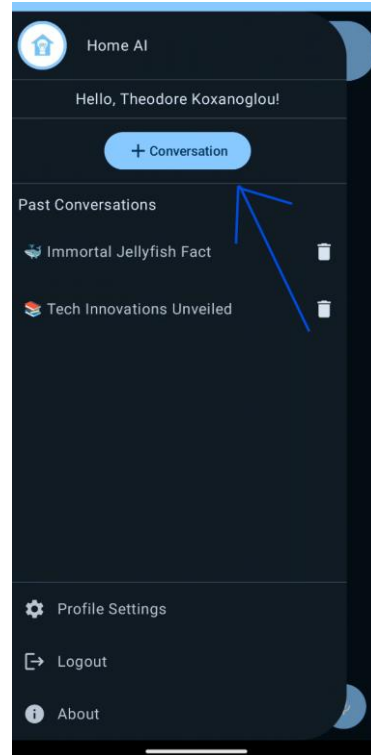
Όταν το πατήσετε η εφαρμογή θα σας ανακατευθύνει στην αρχική οθόνη καλωσορίσματος.

4.3 Συνομιλίες

Στην ενότητα αυτή υπάρχουν οδηγίες για το πως να διαχειρίζεστε τις συνομιλίες σας με το Home AI.

4.3.1 Δημιουργία νέας συνομιλίας

Για να δημιουργήσετε νέα συνομιλία αρκεί να ανοίξετε την εφαρμογή στο κινητό σας ή να πατήσετε από το μενού το κουμπί "+ Conversation". Και με τους δύο τρόπους θα σας εμφανιστεί μία καθαρή οθόνη συνομιλίας.

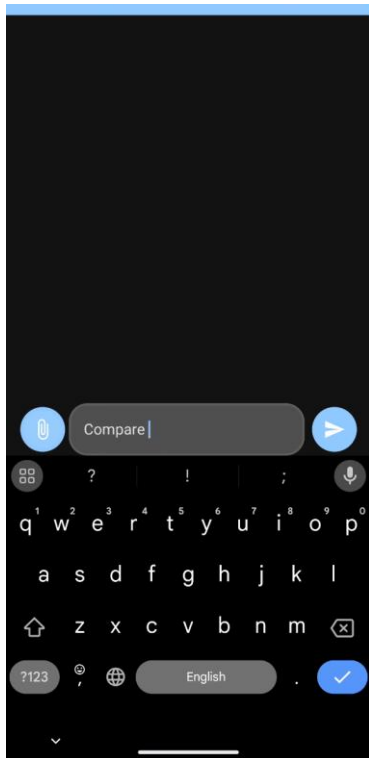


Εικόνα 20: Μενού - Δημιουργία νέας συνομιλίας

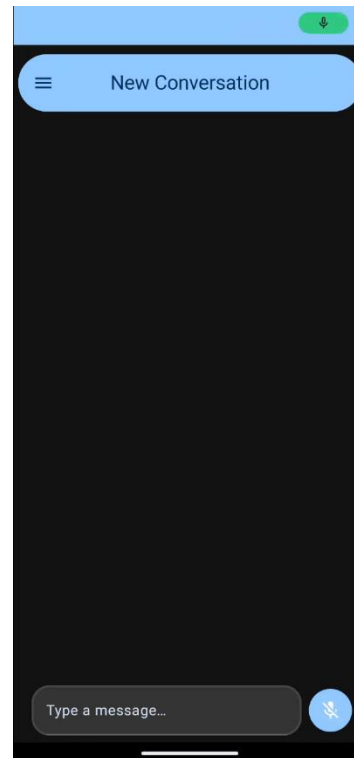
4.3.2 Τρόποι εισαγωγής μηνυμάτων στην εφαρμογή

Όταν βρίσκεστε σε κάποια συνομιλία στην οποία θα θέλατε να στείλετε κάποιο μήνυμα, το μόνο που αρκεί να κάνετε είναι να γράψετε το μήνυμα που θέλετε με το πληκτρολόγιο του κινητού σας ή να πείτε το μήνυμά σας πατώντας το εικονίδιο του μικροφώνου της εφαρμογής.

Σημείωση: Αν είναι η πρώτη φορά που θα χρησιμοποιήσετε το μικρόφωνο του κινητού στην εφαρμογή θα σας εμφανιστεί το σχετικό μήνυμα προτροπής άδειας χρήσης του μικροφώνου από την εφαρμογή.



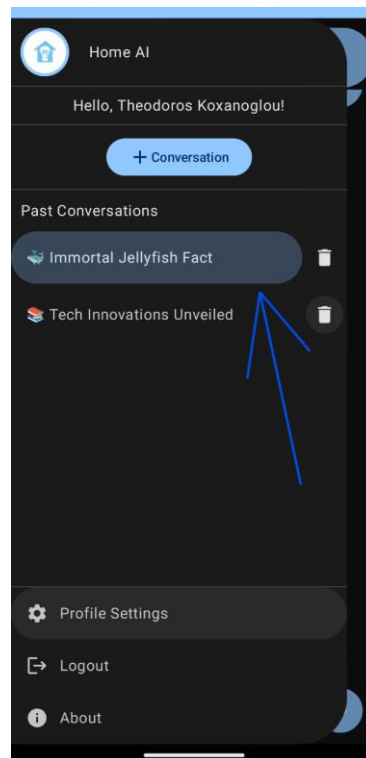
Εικόνα 21: Εισαγωγή μηνύματος μέσω πληκτρολογίου



Εικόνα 22: Εισαγωγή μηνύματος με τη χρήση μικροφώνου

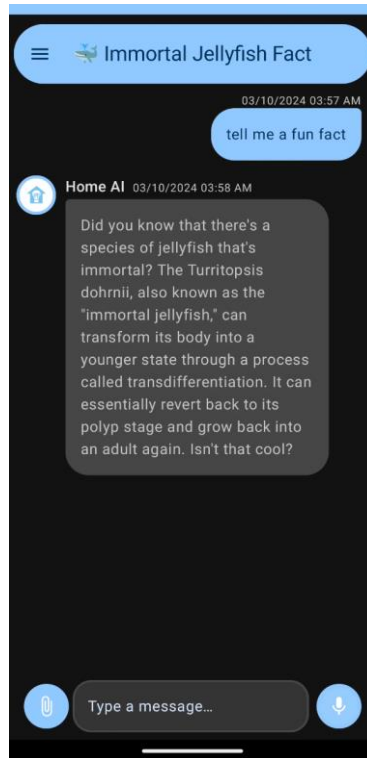
4.3.3 Συνέχιση προηγούμενης συνομιλίας

Κάθε συνομιλία που ξεκινάτε με το Home AI αποθηκεύεται στη βάση δεδομένων της εφαρμογής. Για συνεχίσετε κάποια συζήτηση με το Home AI που είχατε κάνει στο παρελθόν αρκεί να πατήσετε το μενού και επιλέξετε τον τίτλο της συζήτησης που θα θέλατε να συνεχίσετε.



Εικόνα 23: Επιλογή συνομιλίας από το μενού

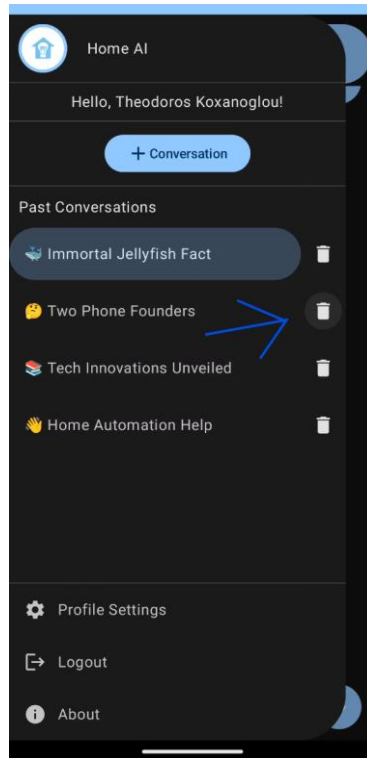
Πατώντας τον τίτλο, θα φορτώσουν στην οθόνη συνομιλιών και όλα τα μηνύματα που είχατε ανταλλάξει προηγουμένως με το Home AI στη συνομιλία αυτή.



Εικόνα 24: Εμφάνιση Ιστορικού Συνομιλίας

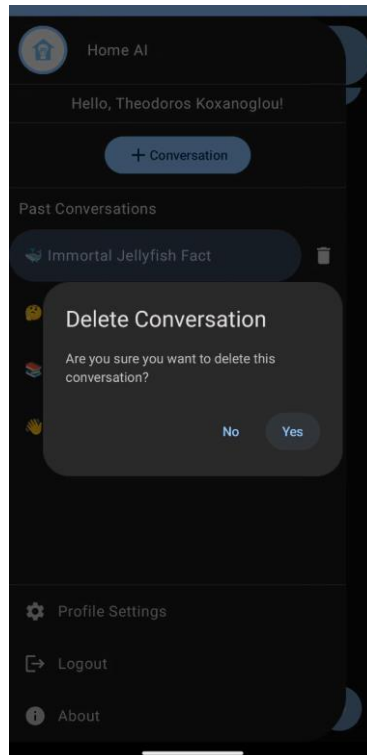
4.3.4 Διαγραφή παλαιότερων συνομιλιών

Σε περίπτωση που έχετε τελειώσει με κάποια συνομιλία και δεν θέλετε να τη συνεχίσετε, για να μην κρατάει χώρο στη βάση δεδομένων της εφαρμογής, μπορείτε να τη διαγράψετε. Για να διαγράψετε μία συνομιλία αρκεί να μεταβείτε στο μενού και να πατήσετε το εικονίδιο του κάδου δίπλα από τον τίτλο της συνομιλίας που θα θέλατε να διαγράψετε.



Εικόνα 25: Επιλογή Διαγραφή Συνομιλίας

Θα σας εμφανιστεί ένα μήνυμα επιβεβαίωσης στο οποίο θα πατήσετε ναι στη περίπτωση που θα θέλατε να συνεχίσετε τη διαγραφή.



Εικόνα 26: Μήνυμα Τελικής Επιβεβαίωσης Διαγραφής Συνομιλίας

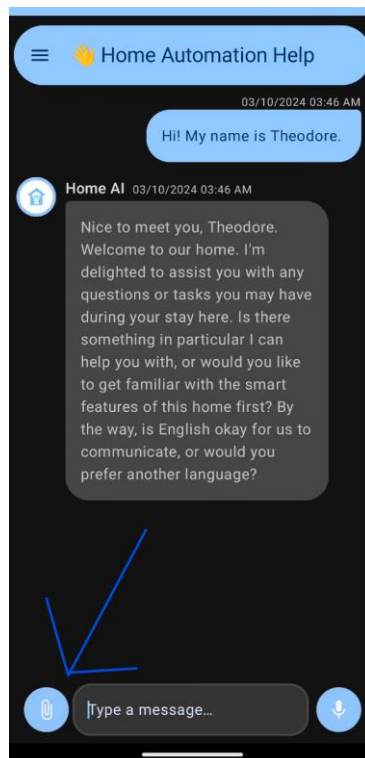
4.4 Έγγραφα

Στην ενότητα υπάρχουν οδηγίες για το πως να διαχειρίζεται τα έγγραφα που φορτώνετε στην εφαρμογή.

4.4.1 Ανέβασμα ενός έγγραφου

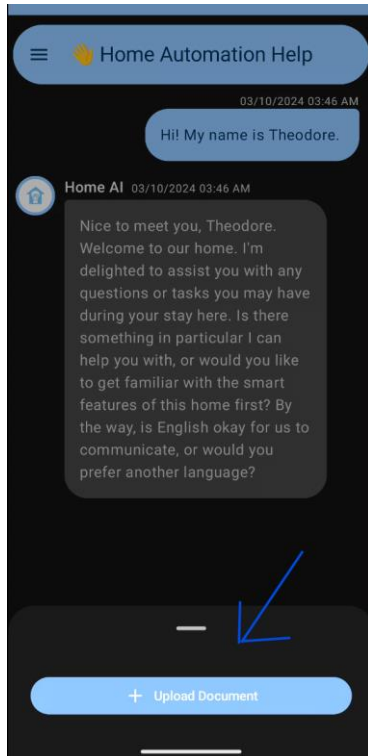
Στην αρχή ή κατά τη διάρκεια μιας συνομιλίας μπορείτε να φορτώσετε έγγραφα στην εφαρμογή τα οποία θα θέλατε να αναλύσετε περαιτέρω με το Home AI. Ουσιαστικά με την φόρτωση των αρχείων αυτών επεκτείνετε τη γνώση του Home AI πέρα από το γνωσιακό επίπεδο που έχει μέχρι στιγμής.

Για να ανεβάσετε ένα έγγραφο αρκεί να πατήσετε το κουμπί που βρίσκεται κάτω δεξιά στην οθόνη, αριστερά από το πεδίο εισαγωγής μηνύματος.

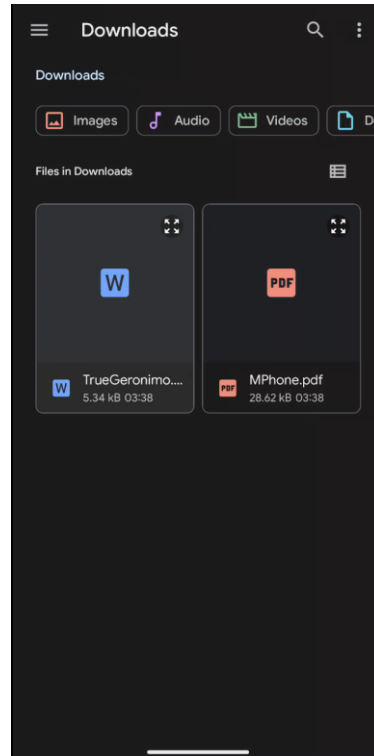


Εικόνα 27: Επισύναψη αρχείων

Πατώντας το κουμπί θα εμφανιστεί ένα αναδυόμενο παράθυρο από κάτω. Στο παράθυρο αυτό θα βρείτε ένα κουμπί που λέγεται "+ Upload Document". Πατώντας το θα σας εμφανιστεί ο διαχειριστής αρχείων του κινητού σας από τον οποίο θα επιλέξετε τα έγγραφα που θα θέλατε να ανεβάσετε.



Εικόνα 28: Φόρτωση Αρχείου

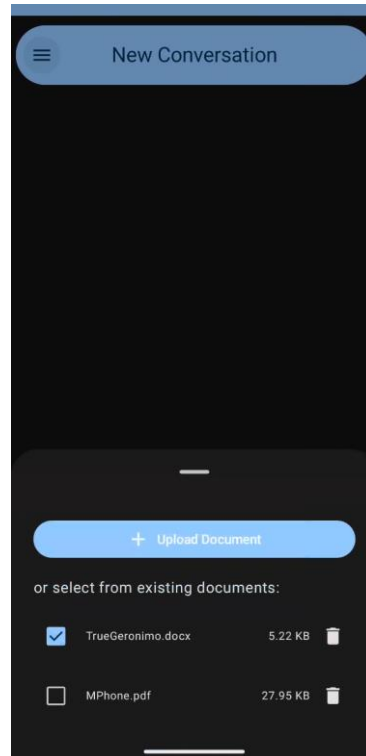


Εικόνα 29: Εξερεύνηση Αρχείων Κινητού για επιλογή εγγράφου

Σημείωση: μέχρι στιγμής η εφαρμογή μπορεί να διαχειριστεί αρχεία με κατάληξη: **.txt**, **.pdf**, **.docx** και **.csv**. Οποιοδήποτε άλλο αρχείο υπάρχει περίπτωση να μην το επεξεργαστεί σωστά.

4.4.2 Επιλογή εγγράφων σε συνομιλία

Σε περίπτωση που έχετε ήδη ανεβάσει κάποια αρχεία στην εφαρμογή τα οποία θέλετε να τα ξανά χρησιμοποιήσετε, αντί να τα ανεβάσετε ξανά, αρκεί να τα επιλέξετε από τη λίστα εγγράφων που εμφανίζεται μετά το πάτημα του κουμπιού συνδετήρα που βρίσκεται κάτω αριστερά στην οθόνη, δίπλα από το πεδίο εισαγωγής μηνύματος.

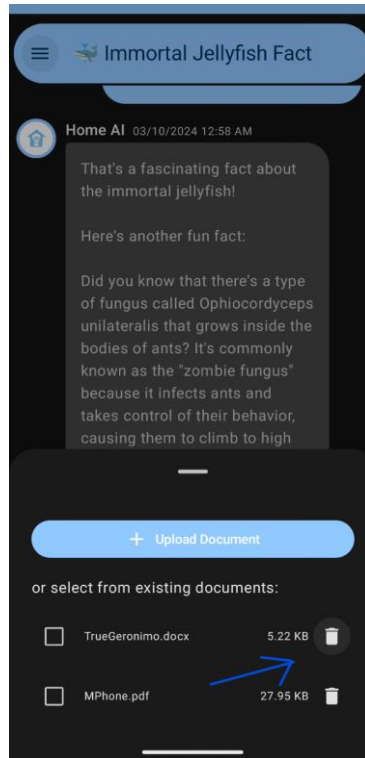


Εικόνα 30: Επιλογή Εγγράφων

Σημείωση: Σε περίπτωση που έχετε αλλάξει τοπικά στο κινητό σας την ονομασία του αρχείου και δεν θυμάστε αν το έχετε ανεβάσει εξαρχής, αρκεί να δοκιμάσετε να το ξανά ανεβάσετε. Αν υπάρχει ήδη το αρχείο στη βάση δεδομένων της εφαρμογής, δεν θα ανεβεί ξανά και θα επιλεγεί αυτόματα στη λίστα των εγγράφων.

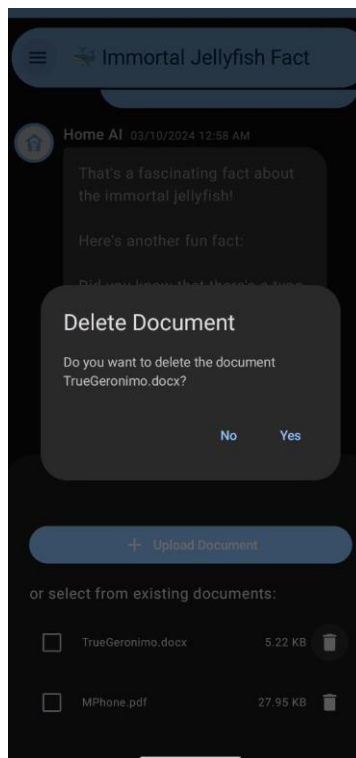
4.4.3 Διαγραφή παλαιότερων εγγράφων

Αν θέλετε να διαγράψετε κάποιο έγγραφο από τη βάση δεδομένων της εφαρμογής αρκεί να πατήσετε το εικονίδιο του συνδετήρα που βρίσκεται κάτω αριστερά στην οθόνη και στο μενού εγγράφων που θα σας εμφανιστεί να πατήσετε το εικονίδιο του κάδου δίπλα από την ονομασία του αρχείου που θα θέλατε να διαγράψετε.



Εικόνα 31: Επιλογή Διαγραφή Εγγράφου

Όταν το πατήσετε θα σας εμφανιστεί ένα μήνυμα επιβεβαίωσης για να συνεχίσετε ή να σταματήσετε την διαγραφή του εγγράφου.



Εικόνα 32: Τελικό Μήνυμα Επιβεβαίωσης διαγραφής εγγράφου

Κεφάλαιο 5: Αποτελέσματα, Οφέλη και Μελλοντικές Επεκτάσεις

5.1 Αποτελέσματα

Η εφαρμογή απέδειξε ότι είναι δυνατή η χρήση Μεγάλων Γλωσσικών Μοντέλων σε τοπικό επίπεδο χωρίς να απαιτούνται εξωτερικές υπηρεσίες νέφους. Οι χρήστες έχουν τη δυνατότητα να λαμβάνουν εξατομικευμένες απαντήσεις, βασισμένες στα προσωπικά τους δεδομένα και στα αρχεία που ανεβάζουν στην εφαρμογή. Τα δεδομένα παραμένουν εντός του τοπικού συστήματος, διασφαλίζοντας την προστασία της ιδιωτικότητάς τους, καθώς δεν κοινοποιούνται σε τρίτες υπηρεσίες. Επιπλέον, η εφαρμογή προσφέρει ένα φιλικό και εύχρηστο περιβάλλον διεπαφής χρήστη, που επιτρέπει ακόμη και σε άτομα με περιορισμένες γνώσεις πληροφορικής να αξιοποιούν τις λειτουργίες της με ευκολία και αποτελεσματικότητα.

5.2 Οφέλη

Η εφαρμογή παρέχει σημαντική αυτονομία σε σχέση με άλλες εμπορικές λύσεις που εξαρτώνται από υπηρεσίες νέφους για τη λειτουργία των Μεγάλων Γλωσσικών Μοντέλων. Αυτή η αυτονομία μειώνει το κόστος και ενισχύει την προστασία της ιδιωτικότητας των χρηστών, ιδιαίτερα σε περιβάλλοντα με υψηλές απαιτήσεις ασφάλειας δεδομένων. Η δυνατότητα προσαρμογής των απαντήσεων των LLMs βάσει των δεδομένων που παρέχουν οι χρήστες, επιτρέπει την εφαρμογή παρόμοιων τοπικών λύσεων Τεχνητής Νοημοσύνης σε διάφορους τομείς, πέραν της προσωπικής χρήσης. Η ανάπτυξη αυτής της εφαρμογής προσφέρει ουσιαστική συνεισφορά στην έρευνα για τις τοπικές λύσεις Τεχνητής Νοημοσύνης που διασφαλίζουν την ιδιωτικότητα. Παράλληλα, αποτελεί ένα πλαίσιο για μελλοντικές βελτιώσεις και επεκτάσεις, προωθώντας περαιτέρω την εφαρμογή της Τεχνητής Νοημοσύνης σε ασφαλή και αυτόνομα συστήματα.

5.3 Μελλοντικές Επεκτάσεις

Αν και η τρέχουσα υλοποίηση της εφαρμογής αποδεικνύει ότι οι τοπικές λύσεις με τη χρήση Μεγάλων Γλωσσικών Μοντέλων είναι εφικτές, υπάρχουν περιθώρια για σημαντικές βελτιώσεις. Παρακάτω παρουσιάζονται προτεινόμενες επεκτάσεις που θα μπορούσαν να ενισχύσουν τη λειτουργικότητα και την αποδοτικότητα της εφαρμογής:

- **Βελτιστοποίηση RAG:**
Η υλοποίηση RAG που έχει η εφαρμογή μπορεί να βελτιστοποιηθεί υποστηρίζοντας επιπλέον τύπους αρχείων και ενσωματώνοντας αποδοτικότερους αλγόριθμους για την εξαγωγή δεδομένων από τα έγγραφα. Αυτό θα αυξήσει το επίπεδο ευκολίας στην αναζήτηση της πληροφορίας.
- **ChromaDB σε Docker Container:**
Στην παρούσα υλοποίηση, η βάση δεδομένων ChromaDB δημιουργείται εντός του docker container της εφαρμογής. Για την καλύτερη διαχείριση και συντήρηση του κώδικα, θα ήταν δυνατό να υλοποιηθεί η ChromaDB σε ξεχωριστό docker container εντός του τοπικού δικτύου. Αυτό θα διαχωρίσει τις λειτουργίες της βάσης δεδομένων από την εφαρμογή, επιτρέποντας μεγαλύτερη ευελιξία στην αναβάθμιση και διαχείριση του κώδικα.
- **Πολυτροπική Αλληλεπίδραση (Multi-modal):**
Η ενσωμάτωση πολυτροπικών μοντέλων θα επιτρέψει στον χρήστη να λαμβάνει, πέρα από απαντήσεις σε μορφή κειμένου, εικόνες και βίντεο. Αν και αυτή η επέκταση θα αυξήσει τις απαιτήσεις σε υπολογιστική ισχύ, με την διαρκή πρόοδο στην τεχνολογία, οι πόροι αυτοί θα είναι πιθανών πιο προσβάσιμοι στο άμεσο μέλλον.
- **Προσωπική Εξατομίκευση Χρήστη:**
Αντί ο χρήστης να διατυπώνει αναλυτικά τα ερωτήματά του κάθε φορά, θα μπορούσε να παρέχει στο προφίλ του μια περιγραφή σχετικά με το ποιος είναι, ποια είναι η επαγγελματική του δραστηριότητα και ποιες πληροφορίες συνήθως αναζητά. Τα LLMs,

με βάση αυτή την πρόσθετη πληροφορία, θα μπορούν να προσφέρουν καλύτερες εξατομικευμένες απαντήσεις.

- **Δυνατότητα Ελέγχου Οικιακών Συσκευών:**

Εφόσον η εφαρμογή στοχεύει σε οικιακούς χρήστες, μια πιθανή επέκταση θα ήταν η ενσωμάτωση κάποιου framework ή API (Home Assistant, n.d.) που θα επιτρέπει στους χρήστες να ελέγχουν τις οικιακές συσκευές οι οποίες είναι συνδεδεμένες στο τοπικό δίκτυο της εφαρμογής. Αυτή η λειτουργία θα βελτίωνε την ενσωμάτωση της εφαρμογής σε έξυπνα σπίτια.

Κεφάλαιο 6: Συμπεράσματα

Η παρούσα πτυχιακή εργασία απέδειξε ότι είναι εφικτή η ανάπτυξη μιας εφαρμογής που αξιοποιεί Μεγάλα Γλωσσικά Μοντέλα σε τοπικό επίπεδο, χωρίς την ανάγκη εξάρτησης από υπηρεσίες νέφους, διασφαλίζοντας την προστασία των προσωπικών δεδομένων των χρηστών. Η εφαρμογή Home AI επιτρέπει την παροχή εξατομικευμένων απαντήσεων βάσει των δεδομένων και των αρχείων που ανεβάζουν οι χρήστες, διατηρώντας την επεξεργασία τοπικά. Αυτή η τοπική προσέγγιση εξασφαλίζει αυξημένη προστασία της ιδιωτικότητας, κάτι που αποτελεί κρίσιμο ζήτημα στο σύγχρονο τεχνολογικό περιβάλλον.

Τα κύρια οφέλη που προέκυψαν από την εφαρμογή αφορούν την αυτονομία, την ασφάλεια και την εξατομίκευση των υπηρεσιών. Η χρήση τοπικών πόρων μειώνει το κόστος και απομακρύνει την εξάρτηση από εμπορικές πλατφόρμες, διασφαλίζοντας παράλληλα την ασφάλεια των δεδομένων σε περιβάλλοντα με υψηλές απαιτήσεις προστασίας.

Η εφαρμογή αυτή αποτελεί επίσης σημαντική συνεισφορά στην έρευνα για τοπικές λύσεις Τεχνητής Νοημοσύνης, προσφέροντας ένα πλαίσιο για ασφαλή και εξατομικευμένη επεξεργασία δεδομένων, με προοπτική εφαρμογής σε πληθώρα κλάδων. Οι μελλοντικές επεκτάσεις περιλαμβάνουν βελτιώσεις στην υποστήριξη περισσότερων τύπων αρχείων, στην πολυτροπική αλληλεπίδραση και στη δυνατότητα ελέγχου οικιακών συσκευών, καθιστώντας την εφαρμογή ακόμα πιο ευέλικτη και λειτουργική για τους χρήστες.

Κατάλογος Εικόνων

ΕΙΚΟΝΑ 1: ΔΟΜΗ ΑΡΧΕΙΩΝ SERVER	9
ΕΙΚΟΝΑ 2: ΔΟΜΗ ΑΡΧΕΙΩΝ DATA LAYER	11
ΕΙΚΟΝΑ 3: ΔΟΜΗ ΑΡΧΕΙΩΝ DOMAIN LAYER	12
ΕΙΚΟΝΑ 4: ΔΟΜΗ ΑΡΧΕΙΩΝ PRESENTATION LAYER	12
ΕΙΚΟΝΑ 5: ΔΟΜΗ CLIENT	13
ΕΙΚΟΝΑ 6: ΠΑΡΑΜΕΤΡΟΙ ΕΦΑΡΜΟΓΗΣ ΥΠΟΛΟΓΙΣΤΗ	16
ΕΙΚΟΝΑ 7: ΕΝΑΡΞΗ ΕΦΑΡΜΟΓΗΣ FASTAPI	16
ΕΙΚΟΝΑ 8: ΕΙΚΟΝΙΔΙΟ ΕΦΑΡΜΟΓΗΣ	17
ΕΙΚΟΝΑ 9: ΑΡΧΙΚΗ ΣΕΛΙΔΑ ΚΑΛΩΣΟΡΙΣΜΑΤΟΣ	17
ΕΙΚΟΝΑ 10: ΟΘΟΝΗ ΕΙΣΟΔΟΥ ΣΤΗΝ ΕΦΑΡΜΟΓΗ	18
ΕΙΚΟΝΑ 11: ΟΘΟΝΗ ΔΗΜΙΟΥΡΓΙΑ ΛΟΓΑΡΙΑΣΜΟΥ	18
ΕΙΚΟΝΑ 12: ΟΘΟΝΗ ΓΕΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΩΝ	19
ΕΙΚΟΝΑ 13: ΜΕΝΟΥ - ΕΠΕΞΕΡΓΑΣΙΑ ΧΡΗΣΤΗ	19
ΕΙΚΟΝΑ 14: ΟΘΟΝΗ ΕΠΕΞΕΡΓΑΣΙΑΣ ΠΡΟΦΙΛ	20
ΕΙΚΟΝΑ 15: ΜΗΝΥΜΑ ΕΠΙΤΥΧΗΣ ΕΝΗΜΕΡΩΣΗΣ ΠΡΟΦΙΛ	20
ΕΙΚΟΝΑ 16: ΟΘΟΝΗ ΕΠΕΞΕΡΓΑΣΙΑΣ ΠΡΟΦΙΛ	20
ΕΙΚΟΝΑ 17: ΟΘΟΝΗ ΑΛΛΑΓΗΣ ΚΩΔΙΚΟΥ ΠΡΟΣΒΑΣΗΣ	21
ΕΙΚΟΝΑ 18: ΕΠΙΤΥΧΗΣ ΑΛΛΑΓΗ ΚΩΔΙΚΟΥ ΠΡΟΣΒΑΣΗΣ	21
ΕΙΚΟΝΑ 19: ΜΕΝΟΥ - ΕΠΙΛΟΓΗ ΕΞΟΔΟΥ ΑΠΟ ΤΟΝ ΛΟΓΑΡΙΑΣΜΟ ΤΗΣ ΕΦΑΡΜΟΓΗΣ	21
ΕΙΚΟΝΑ 20: ΜΕΝΟΥ - ΔΗΜΙΟΥΡΓΙΑ ΝΕΑΣ ΣΥΝΟΜΙΛΙΑΣ	22
ΕΙΚΟΝΑ 21: ΕΙΣΑΓΩΓΗ ΜΗΝΥΜΑΤΟΣ ΜΕΣΩ ΠΛΗΚΤΡΟΛΟΓΙΟΥ	23
ΕΙΚΟΝΑ 22: ΕΙΣΑΓΩΓΗ ΜΗΝΥΜΑΤΟΣ ΜΕ ΤΗ ΧΡΗΣΗ ΜΙΚΡΟΦΩΝΟΥ	23
ΕΙΚΟΝΑ 23: ΕΠΙΛΟΓΗ ΣΥΝΟΜΙΛΙΑΣ ΑΠΟ ΤΟ ΜΕΝΟΥ	23
ΕΙΚΟΝΑ 24: ΕΜΦΑΝΙΣΗ ΙΣΤΟΡΙΚΟΥ ΣΥΝΟΜΙΛΙΑΣ	24
ΕΙΚΟΝΑ 25: ΕΠΙΛΟΓΗ ΔΙΑΓΡΑΦΗ ΣΥΝΟΜΙΛΙΑΣ	25
ΕΙΚΟΝΑ 26: ΜΗΝΥΜΑ ΤΕΛΙΚΗΣ ΕΠΙΒΕΒΑΙΩΣΗΣ ΔΙΑΓΡΑΦΗΣ ΣΥΝΟΜΙΛΙΑΣ	25
ΕΙΚΟΝΑ 27: ΕΠΙΣΥΝΑΨΗ ΑΡΧΕΙΩΝ	26
ΕΙΚΟΝΑ 28: ΦΟΡΤΩΣΗ ΑΡΧΕΙΟΥ	27
ΕΙΚΟΝΑ 29: ΕΞΕΡΕΥΝΗΣΗ ΑΡΧΕΙΩΝ ΚΙΝΗΤΟΥ ΓΙΑ ΕΠΙΛΟΓΗ ΕΓΓΡΑΦΟΥ	27
ΕΙΚΟΝΑ 30: ΕΠΙΛΟΓΗ ΕΓΓΡΑΦΩΝ	28
ΕΙΚΟΝΑ 31: ΕΠΙΛΟΓΗ ΔΙΑΓΡΑΦΗ ΕΓΓΡΑΦΟΥ	29
ΕΙΚΟΝΑ 32: ΤΕΛΙΚΟ ΜΗΝΥΜΑ ΕΠΙΒΕΒΑΙΩΣΗΣ ΔΙΑΓΡΑΦΗΣ ΕΓΓΡΑΦΟΥ	29

Κατάλογος Διαγραμμάτων

ΔΙΑΓΡΑΜΜΑ 1: ΣΧΗΜΑ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ POSTGRESQL.....	10
ΔΙΑΓΡΑΜΜΑ 2: ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΣΥΣΤΗΜΑΤΟΣ HOME AI	14

Πίνακας Ορολογίας

Ξενόγλωσσος Όρος	Ελληνικός Όρος
Activity	Δραστηριότητα
Admin	Διαχειριστής
Artificial Intelligence	Τεχνητή Νοημοσύνη
Change	Αλλαγή
Clean Architecture	Καθαρή Αρχιτεκτονική
Client-Server	Πελάτης-Διακομιστής
Cloud Services	Υπηρεσίες Νέφους
Component	Συστατικό
Compose	Σύνθεση
Container	Εμπορευματοκιβώτιο (ή «Κοντέινερ»)
Conversation	Συνομιλία
Data Layer	Επίπεδο Δεδομένων
Domain Layer	Επίπεδο Τομέα
Email	Ηλεκτρονικό Ταχυδρομείο
Embeddings	Ενσωματώσεις
Framework	Πλαίσιο Εργασίας
House Member	Μέλος Οικίας
Hamburger button	Κουμπί Χάμπουργκερ
Image	Εικόνα
Large Language Models	Μεγάλα Γλωσσικά Μοντέλα
Loading State	Κατάσταση Φόρτωσης
Local	Τοπικό
Login	Σύνδεση
Logout	Αποσύνδεση
Metadata	Μεταδεδομένα
Model	Μοντέλο
Navigation	Πλοήγηση
Password	Κωδικός
Port	Θύρα
Presentation Layer	Επίπεδο Παρουσίασης
Profile Settings	Ρυθμίσεις Προφίλ
Register	Εγγραφή
Remote	Απομακρυσμένο
Repository	Αποθετήριο
Request	Αίτημα
Response	Απόκριση
Role-Based Access Control	Έλεγχος Πρόσβασης Βασισμένο σε Ρόλους
Route	Διαδρομή
Source	Πηγή
Storage	Αποθήκευση
Terminal	Τερματικό
Theme	Θέμα
Token	Διακριτικό
Update	Ενημέρωση
Use Case	Περίπτωση Χρήσης
User Interface	Διεπαφή Χρήστη
Utils	Βοηθητικά Εργαλεία
View Model	Μοντέλο Προβολής

Πίνακας Συντμήσεων-Αρκτικόλεξων

Συντομογραφία	Επεξήγηση
AI	Artificial Intelligence
API	Application Programming Interface
CPU	Central Processing Unit
DI	Dependency Injection
GB	Gigabyte
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol
JWT	JSON Web Token
LLM	Large Language Model
MB	Megabyte
MVVM	Model-View-ViewModel
OS	Operating System
RAG	Retrieval Augmented Generation
RAM	Random Access Memory
RBAC	Role-Based Access Control
SDK	Software Development Kit
UI	User Interface
MME	Μέσα Μαζικής Ενημέρωσης

Βιβλιογραφία

- Adminer. (χ.χ.). *Adminer*. Ανάκτηση από Database Management in a single PHP file: <https://www.adminer.org/>
- Amazon Web Services. (χ.χ.). *What is RAG (Retrieval-Augmented Generation)?* Ανάκτηση από Amazon Web Services: <https://aws.amazon.com/what-is/retrieval-augmented-generation/>
- AT&T. (2024, March 28). *How Artificial Intelligence (AI) Helps Improve Accessibility*. Ανάκτηση από AT&T Accessibility: <https://about.att.com/sites/accessibility/stories/how-ai-helps-accessibility>
- Chroma. (χ.χ.). *Getting Started*. Ανάκτηση από Chroma: <https://docs.trychroma.com/>
- Council, E. P. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on Artificial Intelligence (The AI Act). *Official Journal of the European Union*. Ανάκτηση από https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689
- Dell'Acqua, F., McFowland III, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K. C., Rajendran, S., . . . Lakhani, K. R. (2023). *Navigating the Jagged Technological*. Cambridge, MA: MIT Sloan School of Management.
- Docker. (χ.χ.). *How Compose works*. Ανάκτηση από dockerdocs: <https://docs.docker.com/compose/intro/compose-application-model/>
- Dumbravan, A. (2022). *Clean Android Architecture: Take a layered approach to writing clean, testable, and decoupled Android applications*. Birmingham: Packt Publishing Ltd. Ανάκτηση από <https://www.packtpub.com/en-ca/product/clean-android-architecture-9781803234588>
- FastAPI. (χ.χ.). *APIRouter class*. Ανάκτηση από FastAPI: <https://fastapi.tiangolo.com/reference/apirouter/>
- FastAPI. (χ.χ.). *Security*. Ανάκτηση από FastAPI: <https://fastapi.tiangolo.com/tutorial/security/>
- Fei Liu, Z. K. (2024). Optimizing RAG Techniques for Automotive Industry PDF Chatbots: A Case Study with Locally Deployed Ollama Models. *arXiv*. doi:10.48550/arXiv.2408.05933
- Google. (χ.χ.). *Get started with Jetpack Compose*. Ανάκτηση από Android Developers: <https://developer.android.com/develop/ui/compose/documentation>
- Google. (χ.χ.). *Material Design for Android*. Ανάκτηση από Android Developers: <https://developer.android.com/develop/ui/views/theming/look-and-feel>
- Gruber, J. B. (2024). rollama: An R package for using generative large language models through Ollama. *arXiv*. doi:10.48550/arXiv.2404.07654
- Home Assistant. (χ.χ.). *Installation*. Ανάκτηση από Home Assistant: <https://www.home-assistant.io/installation/>
- LangChain. (χ.χ.). *How to add memory to chatbots*. Ανάκτηση από LangChain: https://python.langchain.com/docs/how_to/chatbots_memory/
- Michael Haenlein, A. K. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*. doi:10.1177/0008125619864925

Nam, L. (2023, May 14). *Jetpack Compose Permissions using Accompanist library, ModalBottomSheet*. Ανάκτηση από Medium: <https://medium.com/@lukohnam/jetpack-compose-permissions-using-accompanist-library-b1c0fbbe8831>

Ollama. (χ.χ.). *API*. Ανάκτηση από Github: <https://github.com/ollama/ollama/blob/main/docs/api.md>

PostgreSQL. (χ.χ.). *CREATE DATABASE*. Ανάκτηση από PostgreSQL: <https://www.postgresql.org/docs/current/sql-createdatabase.html>

Pydantic. (χ.χ.). *Overview*. Ανάκτηση από Pydantic: <https://docs.pydantic.dev/1.10/>

Red Hat. (2024, May 14). <https://www.redhat.com/en/topics/security/what-is-role-based-access-control>. Ανάκτηση από Red Hat: <https://www.redhat.com/en/topics/security/what-is-role-based-access-control>