



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ
ΤΜΗΜΑ ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ
ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΙΔΙΚΕΥΣΗ: “ΠΡΟΗΓΜΕΝΑ ΠΛΗΡΟΦΟΡΙΑΚΑ
ΣΥΣΤΗΜΑΤΑ”

Estia Group: Εφαρμογή για Ανάκτηση και Προβολή Αγγελιών Ακινήτων

Η ΓΛΩΣΣΑ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ ΡΥΤΗΘΩΝ

Επιβλέπων Καθηγητής: Ν. Σγούρος

Ιωάννα Κανδή, Κωνσταντίνος Μαυρογιώργος

ΑΜ: ME2136, ME2144

Email: {ioannakandi,kostismvg}@gmail.com

Περιεχόμενα

Περιεχόμενα.....	1
Κατάλογος Πινάκων.....	1
Κατάλογος Εικόνων	1
Περίληψη.....	2
1. Χρησιμοποιούμενες Τεχνολογίες	3
2. Συλλεχθέντα Δεδομένα	4
3. Προεπεξεργασία Δεδομένων	4
4. Παρεχόμενες Υπηρεσίες Ιστού	4
5. Διεπαφή Χρήστη.....	6
6. Συμπεράσματα	8

Κατάλογος Πινάκων

Πίνακας 1: Πρότυπο πίνακα για υλοποιημένες υπηρεσίες.....	4
Πίνακας 2: Υπηρεσία ανάκτησης αγγελιών ακινήτων.....	5
Πίνακας 3: Υπηρεσία ανάκτησης στατιστικών στοιχείων	5

Κατάλογος Εικόνων

Εικόνα 1: Ενδεικτικά στιγμιότυπα οθόνης από τη διεπαφή χρήστη.....	7
--	---

Περίληψη

Η παρούσα εργασία αποτελεί την απαλλακτική εργασία για το μάθημα «Η Γλώσσα Προγραμματισμού Python» του 2^{ου} εξαμήνου της μεταπτυχιακής ειδίκευσης «Προηγμένα Πληροφοριακά Συστήματα» του Τμήματος Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς. Ειδικότερα, σκοπός της εργασίας είναι ο σχεδιασμός και η υλοποίηση μιας εφαρμογής για την προβολή αγγελιών ακινήτων προς αγορά. Πιο συγκεκριμένα, στη συγκεκριμένη εφαρμογή ανακτώνται δεδομένα από την ιστοσελίδα της Χρυσής Ευκαιρίας αναφορικά με ακίνητα προς πώληση στις περιοχές Τρίκαλα, Μεσολόγγι και Κρήτη. Στο παρόν έγγραφο περιγράφονται συνοπτικά οι χρησιμοποιούμενες τεχνολογίες υλοποίησης, τα συλλεχθέντα δεδομένα, η προεπεξεργασία που υπέστησαν, οι παρεχόμενες υπηρεσίες και η αντίστοιχη διεπαφή χρήστη που αναπτύχθηκε.

1. Χρησιμοποιούμενες Τεχνολογίες

Αναφορικά με τις τεχνολογίες που αξιοποιήθηκαν για την υλοποίηση της εφαρμογής, αυτές είναι οι ακόλουθες:

- Python: Αξιοποιήθηκε για την υλοποίηση του back-end, δηλαδή όλων των παρεχόμενων υπηρεσιών, συμπεριλαμβανομένων της συλλογής των δεδομένων από την Χρυσή Ευκαιρία, της επεξεργασίας αυτών και της ανάκτησης και αυτών από τους χρήστες της διεπαφής. Ειδικότερα, τα modules της Python που χρησιμοποιήθηκαν είναι τα ακόλουθα:
 - BeautifulSoup (bs4): Χρησιμοποιήθηκε για την πραγματοποίηση web scraping στην ιστοσελίδα της Χρυσής Ευκαιρίας και την ανάκτηση των αντίστοιχων δεδομένων.
 - Pymongo: Χρησιμοποιήθηκε για την αποθήκευση των δεδομένων στην αντίστοιχη μη σχεσιακή βάση δεδομένων MongoDB.
 - Flask: Για την ανάπτυξη της εφαρμογής με χρήση υπηρεσιών ιστού, έτσι ώστε αυτές να είναι διαθέσιμες στους χρήστες μέσω της διεπαφής η οποία αναπτύχθηκε με σχετικές web τεχνολογίες.
 - Flask Cors: Χρησιμοποιήθηκε για την αποφυγή εμφάνισης του σφάλματος με ασφαλούς σύνδεσης (Cross Origin).
 - Pandas: Απαιτούμενη για την διαχείριση των δεδομένων που συλλέγονται, καθώς και για την επεξεργασία και ανάλυσή τους (π.χ. εξαγωγή στατιστικών στοιχείων).
 - JSON: Αξιοποιήθηκε για επιστροφή δεδομένων στο front-end σε JSON μορφή.
 - Requests: Απαραίτητη για την πραγματοποίηση κλήσεων HTTP σε APIs τρίτων.
- MongoDB: Αξιοποιήθηκε ως βάση δεδομένων της εφαρμογής.
- Bootstrap 5: για την μορφοποίηση των παραπάνω τμημάτων ούτως ώστε να είναι ευδιάκριτα και ελκυστικά προς τον χρήστη. Το Bootstrap είναι ένα framework το οποίο αποτελείται από HTML, CSS και JavaScript. Ουσιαστικά, παρέχει στους προγραμματιστές έτοιμες κλάσεις τις οποίες μπορούν να χρησιμοποιήσουν έτσι ώστε να υλοποιήσουν γρήγορα μία πολύ ελκυστική διεπαφή χρήστη.
- jQuery: αποτελεί μία βιβλιοθήκη JavaScript η οποία παίρνει πολλές κοινές εργασίες που απαιτούν πολλές γραμμές κώδικα JavaScript για να ολοκληρωθούν και τις «τυλίγει» σε μεθόδους που μπορούν να χρησιμοποιηθούν με μία γραμμή κώδικα. Μέσω της jQuery γίνονται και οι κλήσεις Ajax που έχουν υλοποιηθεί στο back-end και προβάλλονται τα αντίστοιχα αποτελέσματα στους χρήστες.

2. Συλλεχθέντα Δεδομένα

Τα συλλεχθέντα δεδομένα ανακτήθηκαν από την [Χρυσή Ευκαιρία](#) και αφορούσαν αγγελίες πώλησης ακινήτων -Μονοκατοικίες και Διαμερίσματα- στις περιοχές της Κρήτης, του Μεσολογίου και των Τρικάλων. Η συλλογή των δεδομένων δεν παρέχεται από την διεπαφή του χρήστη καθώς η ιστοσελίδα της Χρυσής Ευκαιρίας ανανεώνεται διαρκώς, ούτως ώστε να μην είναι δυνατή η αυτόματη ανάκτηση των δεδομένων με την τεχνική του web scraping. Γι' αυτό το λόγο, τα δεδομένα ανακτήθηκαν μία φορά και εν συνεχεία οι αντίστοιχες HTML σελίδες αποθηκεύτηκαν τοπικά, έτσι ώστε να τεθούν υπό επεξεργασία, και οι χρήσιμες πληροφορίες να αποθηκευτούν, τελικά, στη βάση δεδομένων.

3. Προεπεξεργασία Δεδομένων

Προτού αποθηκευτούν τα δεδομένα που ανακτήθηκαν από τη Χρυσή Ευκαιρία στη βάση δεδομένων, ήταν απαραίτητη η προεπεξεργασία τους. Ειδικότερα, για κάθε ανακτηθείσα HTML σελίδα έπρεπε, αρχικά να πραγματοποιηθεί web scraping για τον εντοπισμό της απαραίτητης πληροφορίας. Για κάθε ακίνητο ανακτήθηκαν οι ακόλουθες πληροφορίες:

- Τίτλος Αγγελίας
- Τιμή (€)
- Τιμή / Τετραγωνικό Μέτρο
- Όροφος
- Αριθμός Υπνοδωματίων
- Αριθμός Μπάνιων
- Έτος Κατασκευής

Έπειτα, ήταν αναγκαία η αφαίρεση σημείων στίξης και συμβόλων από τα δεδομένα, προκειμένου αυτά να είναι επεξεργάσιμα. Επιπρόσθετα, κάποιες από τις προαναφερθείσες πληροφορίες ήταν σε μορφή κειμένου, η οποία δεν προσφερόταν για την ζητούμενη επεξεργασία. Για παράδειγμα, η μεταβλητή «Όροφος» έπαιρνε τις τιμές «Υπόγειο», «Ισόγειο», «1^{ος}» κ.ο.κ. Σε αυτή τη περίπτωση, ανατέθηκαν αριθμοί στις αντίστοιχες τιμές: δηλαδή, στη τιμή «Υπόγειο» ανατέθηκε η τιμή -1, στη τιμή «Ισόγειο» ανατέθηκε ο αριθμός 0, στη τιμή «1^{ος}» ανατέθηκε η τιμή 1 κ.ο.κ.

4. Παρεχόμενες Υπηρεσίες Ιστού

Στη συγκεκριμένη ενότητα περιγράφονται οι υλοποιημένες υπηρεσίες ιστού σε μορφή πινάκων. Οι πίνακες έχουν την ακόλουθη μορφή:

Όνομα Υπηρεσίας	
Περιγραφή	Περιγραφή της λειτουργίας της υπηρεσίας
Endpoint URL	http://localhost:5000/...
HTTP μέθοδος	Δυνατές τιμές: GET, POST, PUT, DELETE
Παράμετροι	Οι παράμετροι που χρειάζεται η υπηρεσία (αν υπάρχουν)
Επιστροφή	Τι επιστρέφει η υπηρεσία στον χρήστη

Πίνακας 1: Πρότυπο πίνακα για υλοποιημένες υπηρεσίες.

getSelectedHouses	
Περιγραφή	Χρησιμοποιείται για την ανάκτηση αγγελιών ακινήτων
Endpoint URL	http://localhost:5000/getSelectedHouses/{location}/{level}
HTTP μέθοδος	GET
Παράμετροι	location: η τοποθεσία για την οποία ενδιαφέρεται ο χρήστης level: ο τύπος ακινήτου (αν πρόκειται για μονοκατοικία ή διαμέρισμα)
Επιστροφή	Επιστρέφει τα δεδομένα που ζητήθηκαν σε JSON μορφή

Πίνακας 2: Υπηρεσία ανάκτησης αγγελιών ακινήτων

getStatistics	
Περιγραφή	Χρησιμοποιείται για την ταυτοποίηση και την είσοδο των χρηστών στο σύστημα
Endpoint URL	http://localhost:5000/getStatistics/{location}/{level}
HTTP μέθοδος	GET
Παράμετροι	location: η τοποθεσία για την οποία ενδιαφέρεται ο χρήστης level: ο τύπος ακινήτου (αν πρόκειται για μονοκατοικία ή διαμέρισμα)
Επιστροφή	Επιστρέφει τα δεδομένα που ζητήθηκαν σε JSON μορφή

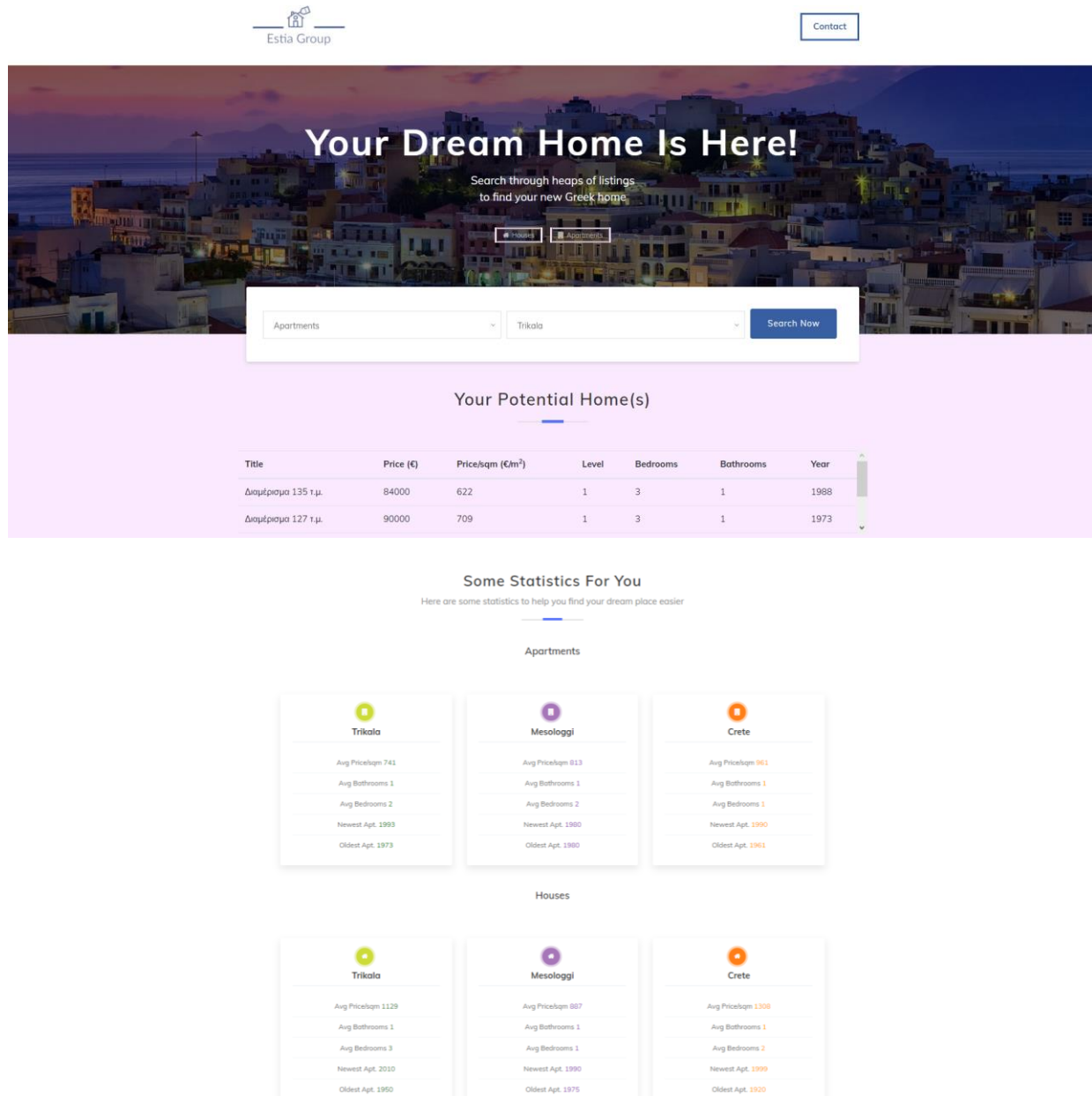
Πίνακας 3: Υπηρεσία ανάκτησης στατιστικών στοιχείων

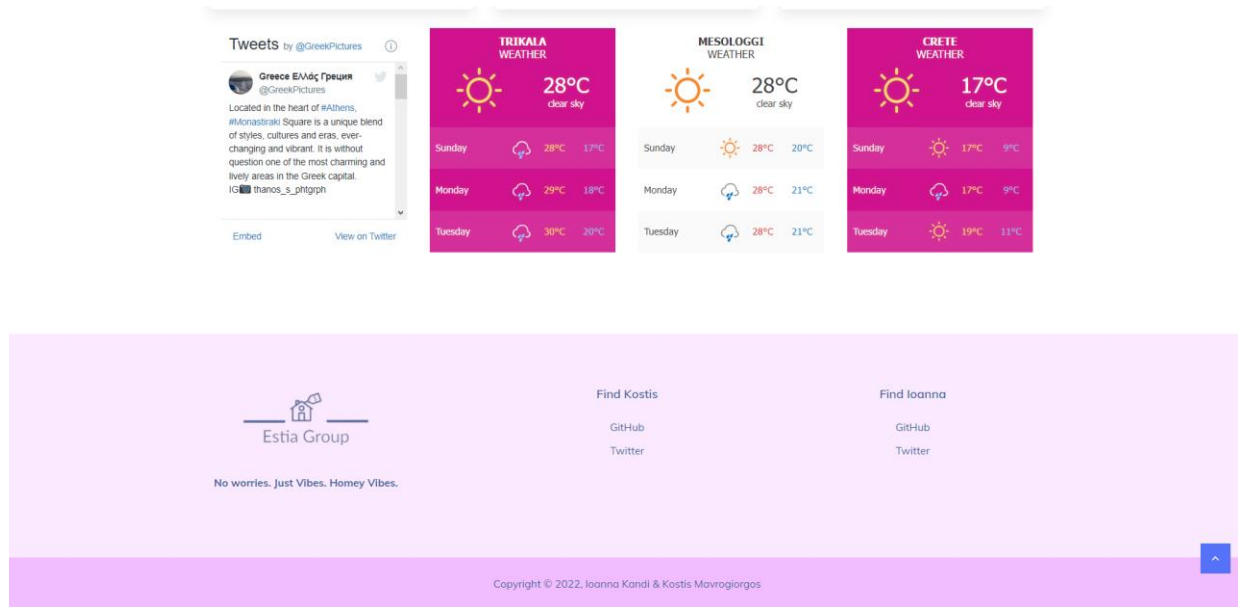
Εκτός από τις υλοποιημένες υπηρεσίες ιστού οι οποίες αναλύθηκαν παραπάνω, η εφαρμογή επικοινωνεί και με δύο (2) APIs τρίτων, όπως φαίνεται παρακάτω:

- Twitter: Χρησιμοποιείται για την ανάκτηση tweets που περιέχουν φωτογραφικό υλικό από την Ελλάδα και το οποίο ίσως είναι χρήσιμο σε κάποιον χρήστη που αναζητά την ιδανική περιοχή για να αγοράσει σπίτι. Το συγκεκριμένο περιεχόμενο είναι διαθέσιμο στη διεπαφή χρήστη, όπως φαίνεται στην ενότητα 5.
- WeatherWidget.io: Αξιοποιείται για τη πρόγνωση του καιρού σε διάστημα τριών (3) ημερών στις περιοχές της Κρήτης, του Μεσολογίου και των Τρικάλων. Το συγκεκριμένο περιεχόμενο είναι διαθέσιμο στη διεπαφή χρήστη, όπως φαίνεται στην ενότητα 5.

5. Διεπαφή Χρήστη

Ενδεικτικά στιγμιότυπα από τη διεπαφή χρήστη παρουσιάζονται στην ακόλουθη εικόνα:





Εικόνα 1: Ενδεικτικά στιγμιότυπα οθόνης από τη διεπαφή χρήστη

Επίσης, αξίζει να σημειωθεί ότι ο κώδικας της εφαρμογής είναι ανεβασμένος και σε ιδιωτικό repository στο GitHub, στη διεύθυνση <https://github.com/ioannakandi/estiaGroup>. Για να έχει κανείς πρόσβαση θα πρέπει να επικοινωνήσει σε ένα από τα emails που αναγράφονται στο εξώφυλλο της εργασίας, έτσι ώστε να προστεθεί ως contributor.

6. Συμπεράσματα

Η συγκεκριμένη εργασία είχε ως στόχο την ανάπτυξη εφαρμογής για ανάκτηση και προβολή αγγελιών ακινήτων. Ειδικότερα, η εφαρμογή που αναπτύχθηκε έχει τη δυνατότητα ανάκτησης δεδομένων σχετικά με αγγελίες ακινήτων από τη Χρυσή Ευκαιρία αναφορικά με τις περιοχές των Τρικάλων, του Μεσολογγίου και της Κρήτης, υπολογισμού σχετικών στατιστικών και προβολής αυτών και αναζήτησης αγγελιών ακινήτων βάσει κριτηρίων. Ως μελλοντικές βελτιώσεις της παρούσας εργασίας θα μπορούσαν να θεωρηθούν η προσθήκη περισσότερων περιοχών και η ανάκτηση αγγελιών από περισσότερες ιστοσελίδες. Ωστόσο, οι παραπάνω βελτιώσεις δεν είναι εύκολα υλοποιήσιμες με την τεχνική του web scraping, καθώς η πλειοψηφία των διαθέσιμων ιστοσελίδων αλλάζουν ανά τακτά χρονικά διαστήματα τη μορφή τους, έτσι ώστε να αποθαρρύνουν την αυτοματοποιημένη λήψη του περιεχομένου τους.