



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Τίτλος Πτυχιακής Εργασίας	Μεγάλα Γλωσσικά Μοντέλα με χρήση RAG για την Αντιστοίχιση Υποψηφίων με Θέσεις Εργασίας RAG-Based LLMs for Job Matching Tasks
Ονοματεπώνυμο Φοιτητή	Συρίγου Στυλιανή
Πατρώνυμο	Λοΐζος
Αριθμός Μητρώου	Π22167
Επιβλέπων Καθηγητής	Διονύσιος Σωτηρόπουλος, Αναπληρωτής Καθηγητής

Ημερομηνία Παράδοσης / Σεπτέμβριος 2026

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

Στο σημείο αυτό οφείλω ένα μεγάλο ευχαριστώ σε όλους όσοι στάθηκαν δίπλα μου κατά τη διάρκεια της εκπόνησης της παρούσας πτυχιακής εργασίας. Ξεχωριστές ευχαριστίες οφείλω στον επιβλέποντα καθηγητή μου, Διονύσιο Σωτηρόπουλο, για τη συνεχή καθοδήγηση, τις εύστοχες παρατηρήσεις και την πολύτιμη υποστήριξη που μου προσέφερε σε κάθε στάδιο της εργασίας. Ακόμη, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στην οικογένειά μου, για την υπομονή και τη συμπαράσταση που έδειξαν καθ' όλη τη διάρκεια των φοιτητικών μου χρόνων. Κλείνοντας, θα ήθελα να ευχαριστήσω το σύνολο των καθηγητών του τμήματος, για τις γνώσεις και τα εφόδια που μου παρείχαν κατά τη διάρκεια των σπουδών μου.

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία πραγματεύεται τον σχεδιασμό και την υλοποίηση ενός καινοτόμου, αυτοματοποιημένου συστήματος αξιολόγησης και κατάταξης βιογραφικών σημειωμάτων με τη χρήση τεχνολογιών Τεχνητής Νοημοσύνης (AI). Βασικός στόχος της εφαρμογής είναι η αυτοματοποίηση της διαδικασίας επιλογής προσωπικού (Recruitment), μέσω της μέτρησης του βαθμού συνάφειας μεταξύ των προσόντων των υποψηφίων και των απαιτήσεων μιας συγκεκριμένης θέσης εργασίας (Job Description).

Το σύστημα αναπτύχθηκε σε περιβάλλον Python (Streamlit UI) και υποστηρίζει την πολυτροπική εισαγωγή δεδομένων, επεξεργαζόμενο αρχεία μορφής PDF, εικόνων (JPG/PNG), καθώς και μαζικών αρχείων ZIP. Για την εξαγωγή κειμένου από οπτικά έγγραφα ενσωματώθηκε η τεχνολογία Οπτικής Αναγνώρισης Χαρακτήρων **EasyOCR**.

Η αρχιτεκτονική του συστήματος βασίζεται στην προηγμένη τεχνική **RAG (Retrieval-Augmented Generation)**, η οποία επιτρέπει στο Μεγάλο Γλωσσικό Μοντέλο **Phi-3 (της Microsoft)** να αναλύει δυναμικά και σε πραγματικό χρόνο τα δεδομένα. Η ανάκτηση των πληροφοριών πραγματοποιείται μέσω της διανυσματικής βάσης δεδομένων **ChromaDB**, όπου τα βιογραφικά έχουν μετατραπεί σε μαθηματικά διανύσματα (**Embeddings**) για την επίτευξη σημασιολογικής αναζήτησης (Semantic Search). Ο συντονισμός και η διασύνδεση των υποσυστημάτων πραγματοποιείται μέσω του πλαισίου **LangChain**.

Το τελικό περιβάλλον εργασίας προσφέρει στον διαχειριστή HR μια διπλή οπτική σύγκριση: στην αριστερή πλευρά κατατάσσονται οι νεοεισαχθέντες υποψήφιοι, ενώ στη δεξιά πλευρά ανακτώνται οι βέλτιστοι υποψήφιοι από τη διανυσματική βάση, η οποία δημιουργείται μέσω της διαδικασίας ingestion από δημόσιο σύνολο δεδομένων 962 βιογραφικών [40]. Η κατάταξη γίνεται αυτόματα σε φθίνουσα σειρά βάσει ενός ποσοστιαίου δείκτη καταλληλότητας (Score 0-100%). Παράλληλα, το σύστημα παράγει μια ποιοτική σύνοψη για κάθε υποψήφιο και προτείνει τρεις εξειδικευμένες ερωτήσεις συνέντευξης, εστιάζοντας στα εντοπισμένα κενά ή τις αδυναμίες του βιογραφικού. Λόγω της τοπικής εκτέλεσης του μοντέλου μέσω της πλατφόρμας **Ollama**, διασφαλίζεται πλήρως η ιδιωτικότητα και η ασφάλεια των ευαίσθητων προσωπικών δεδομένων.

Λέξεις – κλειδιά: Αξιολόγηση Βιογραφικών, Κατάταξη Υποψηφίων, Περιγραφή Θέσης Εργασίας, Διανυσματική Βάση Δεδομένων, Τοπικό Γλωσσικό Μοντέλο, Αναγνώριση Κειμένου.

ABSTRACT

This thesis presents the design and implementation of an innovative automated system for curriculum vitae (CV) evaluation and candidate ranking using Artificial Intelligence (AI) technologies. The primary objective of the application is to automate the recruitment process by measuring the degree of relevance between candidates' qualifications and the requirements of a specific job position, as defined in a Job Description.

The system was developed in Python using the Streamlit framework and supports multimodal data input, enabling the processing of PDF documents, image files (JPG/PNG), and bulk ZIP archives. To extract textual information from visual documents, the EasyOCR Optical Character Recognition (OCR) technology was integrated into the platform.

The system architecture is based on the advanced Retrieval-Augmented Generation (RAG) paradigm, which enables the Microsoft Phi-3 Large Language Model (LLM) to dynamically analyze and interpret data in real time. Information retrieval is performed through the

ChromaDB vector database, where CVs are transformed into mathematical vector representations (embeddings) to facilitate semantic search and similarity matching. The orchestration and integration of all system components are achieved through the LangChain framework.

The final user interface provides HR professionals with a dual-view comparison environment. On the left side, newly submitted candidates are ranked according to their suitability for the target position, while on the right side, the most relevant candidates are retrieved from the pre-existing vector database, which is built through the ingestion process from a public dataset of 962 CVs. Candidate ranking is performed automatically in descending order based on a suitability score ranging from 0 to 100%.

In addition, the system generates a qualitative summary for each candidate and proposes three tailored interview questions, focusing on identified skill gaps, missing qualifications, or weaknesses within the candidate's profile. By leveraging local model execution through the Ollama platform, the proposed solution ensures full privacy, security, and confidentiality of sensitive personal data, eliminating the need for external cloud-based processing.

Keywords: CV Evaluation, Candidate Ranking, Recruitment, Job Description, Vector Database, Large Language Model, Semantic Search, Optical Character Recognition, Retrieval-Augmented Generation (RAG), Artificial Intelligence.

Πίνακας περιεχομένων

Περιεχόμενα

Copyright ©	2
Ευχαριστίες	3
ΠΕΡΙΛΗΨΗ	4
ABSTRACT	4
Περιεχόμενα	6
Κατάλογος Εικόνων	7
1. Εισαγωγή	9
2. Στόχοι	9
3. Θεωρητικό Πλαίσιο και Βιβλιογραφική Ανασκόπηση	11
3.1 Η Ψηφιοποίηση των Διαδικασιών HR και η Ανάγκη για Αυτοματοποιημένη Διαλογή....	11
3.2 Αρχιτεκτονική RAG, Σημασιολογική Αναζήτηση και Μεγάλα Γλωσσικά Μοντέλα	11
3.3 Τοπική Υπολογιστική Ισχύς, Ιδιωτικότητα Δεδομένων και Τοπική Εκτέλεση (Local Deployment).....	12
3.4 Σχεδιασμός Εμπειρίας Χρήστη (UX) και Διασθητική Διεπαφή HR	13
3.5 Προηγμένες Τεχνολογίες και Μελλοντικές Προοπτικές.....	14
3.5.1 Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence – XAI) στη Διοίκηση Ανθρωπίνου Δυναμικού	14
3.5.2 Παραγωγική Τεχνητή Νοημοσύνη και Ευφυείς Πράκτορες (Agentic HR Systems)	14
3.6 Επιχειρησιακή Βιωσιμότητα, Επαγγελματική Συμπεριφορά και Κοινωνικός Αντίκτυπος15	
3.7 Πρακτικές Εφαρμογές και Μελέτες Περίπτωσης στον Τομέα του Ανθρωπίνου Δυναμικού	15
4. Περιγραφή Προβλήματος	16
5. Αναφορά σε παρόμοιες εφαρμογές	17
5.1. Eightfold.ai	18
5.2. Ideal	18
5.3. Συγκριτική αξιολόγηση με προτεινόμενη λύση	19
6. Περιγραφή της λύσης	19
7. Ανάλυση απαιτήσεων	20
7.1. Ρόλοι χρηστών και λειτουργίες	21
7.2. Απαιτήσεις εγκατάστασης και λειτουργίας	21
7.3. Απαιτήσεις χρήσης	22
7.4. Περιορισμοί.....	22
8. Σχεδιασμός της εφαρμογής	23
8.1. Αρχιτεκτονική Συστήματος.....	23
8.2. Τεχνολογίες που Χρησιμοποιήθηκαν.....	24
8.2.1. Python & Streamlit (Web UI)	24
8.2.2. Ollama & Phi-3 (Local Inference)	24
8.2.3. ChromaDB (Vector Database)	24
8.2.4. EasyOCR (Optical Character Recognition).....	25

8.3. Περιγραφή Βάσης Δεδομένων	25
9. Παραδείγματα από τη χρήση του συστήματος	26
10. Εγχειρίδιο Σύντομης Αναφοράς	45
10.1. Εισαγωγή Αρχείων	45
10.1.1. Σελίδα Εισαγωγής Βιογραφικών (Upload CV)	46
10.1.2. Επεξεργασία Αρχείων (File Processing)	46
10.2. Εισαγωγή Περιγραφής Θέσης Εργασίας (Job Description)	46
10.3. Αποτελέσματα Αντιστοίχισης Υποψηφίων από Εισαγωγή (Candidate Matching)	47
10.3.1. Λίστα Υποψηφίων από Εισαγωγή	47
10.3.2. Αναλυτική Αξιολόγηση Υποψηφίου από Εισαγωγή	48
10.4. Δημιουργία Ερωτήσεων Συνέντευξης	49
10.5. Αποτελέσματα Αντιστοίχισης Υποψηφίων (Candidate Matching)	50
10.5.1. Λίστα Υποψηφίων	50
10.5.2. Αναλυτική Αξιολόγηση Υποψηφίου	51
10.6. Δημιουργία Ερωτήσεων Συνέντευξης	52
11. Αναλυτικό εγχειρίδιο χρήστη	53
11.1. Εισαγωγή Περιγραφής Θέσης Εργασίας	53
11.2. Προαιρετική Εισαγωγή Νέων Βιογραφικών	54
11.3. Αξιολόγηση των Νέων Βιογραφικών	55
11.4. Αναζήτηση Υποψηφίων από τη Βάση Δεδομένων	56
11.5. Παρουσίαση Αποτελεσμάτων	57
12. Αξιολόγηση της αποτελεσματικότητας του συστήματος	58
12.1 Σκοπός της αξιολόγησης	58
12.2 Μεθοδολογία αξιολόγησης	58
12.3 Μετρικές αξιολόγησης	59
12.4 Αποτελέσματα αξιολόγησης	59
12.5 Συνολική αξιολόγηση	59
13. Οφέλη	60
14. Συμπεράσματα	61
15. Πίνακας Ορολογίας	62
16. Πίνακας συντηρήσεων-αρτικόλεξων-ακρονυμίων	64
17. Βιβλιογραφία	65

Κατάλογος Εικόνων

Εικόνα 1: Αρχική διεπαφή της εφαρμογής Smart HR Ranking & Evaluation	25
Εικόνα 2: Εισαγωγή της περιγραφής της θέσης εργασίας	26
Εικόνα 3: Έναρξη της διαδικασίας αξιολόγησης των βιογραφικών	27
Εικόνα 4: Ανάκτηση των σχετικών βιογραφικών από τη βάση δεδομένων	28
Εικόνα 5: Ολοκλήρωση της διαδικασίας ανάκτησης και προετοιμασία της αξιολόγησης	29
Εικόνα 6: Παρουσίαση των αποτελεσμάτων αξιολόγησης των υποψηφίων	30

Εικόνα 7: Παρουσίαση των αποτελεσμάτων αξιολόγησης των υποψηφίων	31
Εικόνα 8: Αναλυτική προβολή των πληροφοριών ενός υποψηφίου	32
Εικόνα 9: Δημιουργία ερωτήσεων συνέντευξης για τον υποψήφιο	32
Εικόνα 10: Επιλογή νέων βιογραφικών προς αξιολόγηση	33
Εικόνα 11: Εκκίνηση της επεξεργασίας των νέων βιογραφικών	34
Εικόνα 12: Ολοκλήρωση της μεταφόρτωσης των νέων βιογραφικών	35
Εικόνα 13: Έναρξη αξιολόγησης των νέων βιογραφικών	36
Εικόνα 14: Πρόσδος της διαδικασίας αξιολόγησης	37
Εικόνα 15: Ολοκλήρωση της αξιολόγησης και κατάταξη των νέων υποψηφίων	37
Εικόνα 16: Αναλυτική αξιολόγηση του καταλληλότερου υποψηφίου	38
Εικόνα 17: Αυτόματη δημιουργία ερωτήσεων συνέντευξης για τον υποψήφιο	39
Εικόνα 18: Αναλυτική αξιολόγηση του δεύτερου υποψηφίου	40
Εικόνα 19: Αναλυτική αξιολόγηση του τρίτου υποψηφίου	41
Εικόνα 20: Αξιολόγηση υποψηφίου με χαμηλότερο ποσοστό συνάφειας	42
Εικόνα 21: Αναλυτική αξιολόγηση υποψηφίου από τη βάση δεδομένων	42
Εικόνα 22: Αναλυτική αξιολόγηση υποψηφίου με χαμηλότερο ποσοστό συνάφειας από τη βάση δεδομένων	43
Εικόνα 23: Σελίδα εισαγωγής αρχείων βιογραφικών	44
Εικόνα 24: Σελίδα επεξεργασίας και ανάλυσης αρχείων	45
Εικόνα 25: Σελίδα εισαγωγής απαιτήσεων θέσης εργασίας	45
Εικόνα 26: Προβολή αποτελεσμάτων αντιστοίχισης υποψηφίων	46
Εικόνα 27: Αναλυτική παρουσίαση αποτελέσματος αξιολόγησης	47
Εικόνα 28: Παραγωγή ερωτήσεων συνέντευξης από το σύστημα	48
Εικόνα 29: Προβολή αποτελεσμάτων αντιστοίχισης υποψηφίων	49
Εικόνα 30: Αναλυτική παρουσίαση αποτελέσματος αξιολόγησης	49
Εικόνα 31: Παραγωγή ερωτήσεων συνέντευξης από το σύστημα	50
Εικόνα 32: Εισαγωγή Job Description	51
Εικόνα 33: Εισαγωγή νέων βιογραφικών	52
Εικόνα 34: Αξιολόγηση των νέων βιογραφικών	53
Εικόνα 35: Αναζήτηση υποψηφίων από τη βάση δεδομένων	54
Εικόνα 36: Τελική παρουσίαση αποτελεσμάτων	55

1. Εισαγωγή

Η ραγδαία αύξηση του όγκου των διαθέσιμων δεδομένων και η συνεχώς αυξανόμενη ζήτηση για αποτελεσματική διαχείριση ανθρώπινου δυναμικού έχουν καταστήσει τη διαδικασία επιλογής προσωπικού μία από τις πλέον απαιτητικές λειτουργίες των σύγχρονων οργανισμών. Τα τμήματα Ανθρώπινου Δυναμικού (Human Resources - HR) καλούνται συχνά να αξιολογήσουν μεγάλο αριθμό βιογραφικών σημειωμάτων για την κάλυψη μίας ή περισσότερων θέσεων εργασίας, γεγονός που συνεπάγεται σημαντική κατανάλωση χρόνου και πόρων. Η χειροκίνητη αξιολόγηση των υποψηφίων ενδέχεται να οδηγήσει σε καθυστερήσεις, αυξημένο λειτουργικό κόστος και περιορισμένη αντικειμενικότητα κατά τη διαδικασία λήψης αποφάσεων. Για τον λόγο αυτό, η αξιοποίηση τεχνολογιών Τεχνητής Νοημοσύνης και προηγμένων συστημάτων επεξεργασίας δεδομένων αναδεικνύεται ως μία ιδιαίτερα ελκυστική προσέγγιση για την αυτοματοποίηση και τη βελτιστοποίηση της διαδικασίας προσλήψεων.

Η παρούσα πτυχιακή εργασία επικεντρώνεται στον σχεδιασμό και την ανάπτυξη μιας διαδικτυακής εφαρμογής για την αυτοματοποιημένη αξιολόγηση και κατάταξη βιογραφικών σημειωμάτων, με βάση τον βαθμό συνάφειας των προσόντων των υποψηφίων προς τις απαιτήσεις μιας συγκεκριμένης θέσης εργασίας. Ο χρήστης του συστήματος, ο οποίος αντιστοιχεί στον υπεύθυνο προσλήψεων ή διαχειριστή ανθρώπινου δυναμικού, έχει τη δυνατότητα να εισάγει την περιγραφή μιας θέσης εργασίας στην αγγλική γλώσσα, να μεταφορτώνει νέα βιογραφικά σημειώματα σε διαφορετικές μορφές αρχείων και να λαμβάνει αυτοματοποιημένα αποτελέσματα αξιολόγησης και κατάταξης.

Το σύστημα παρέχει μία ολοκληρωμένη εικόνα της καταλληλότητας κάθε υποψηφίου, παρουσιάζοντας ποσοτικό δείκτη συνάφειας, περιγραφική σύνοψη των βασικών προσόντων και εξατομικευμένες προτεινόμενες ερωτήσεις συνέντευξης, οι οποίες εστιάζουν σε πιθανά κενά γνώσεων ή εμπειρίας. Παράλληλα, προσφέρει τη δυνατότητα σύγκρισης μεταξύ των νέων υποψηφίων και των πλέον κατάλληλων βιογραφικών που ανακτώνται από μία προϋπάρχουσα βάση δεδομένων.

Για την υλοποίηση της εφαρμογής χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python, ενώ το γραφικό περιβάλλον αναπτύχθηκε με τη χρήση του πλαισίου Streamlit. Η αρχιτεκτονική του συστήματος βασίζεται στην τεχνική Retrieval-Augmented Generation (RAG), αξιοποιώντας το μεγάλο γλωσσικό μοντέλο Phi-3 της Microsoft μέσω της πλατφόρμας Ollama. Επιπλέον, για τη διαχείριση, αποθήκευση και σημασιολογική αναζήτηση των δεδομένων ενσωματώθηκε η διανυσματική βάση δεδομένων ChromaDB, η οποία επιτρέπει την αναπαράσταση των βιογραφικών σε μορφή embeddings και τη διεξαγωγή αναζητήσεων με βάση τη σημασιολογική ομοιότητα. Μέσω του συνδυασμού αυτών των τεχνολογιών, το προτεινόμενο σύστημα επιτυγχάνει την αποτελεσματική υποστήριξη της διαδικασίας επιλογής προσωπικού, συμβάλλοντας στη μείωση του χρόνου αξιολόγησης και στη βελτίωση της ποιότητας των τελικών αποφάσεων πρόσληψης.

2. Στόχοι

Η παρούσα εργασία έχει ως κύριο στόχο τον σχεδιασμό και την ανάπτυξη μιας καινοτόμου εφαρμογής αξιολόγησης και κατάταξης βιογραφικών σημειωμάτων, η οποία μετασχηματίζει την παραδοσιακή και χρονοβόρα διαδικασία επιλογής προσωπικού σε μια αυτοματοποιημένη και ευφυή ψηφιακή διαδικασία. Η εφαρμογή αξιοποιεί σύγχρονες τεχνολογίες Τεχνητής Νοημοσύνης με σκοπό τη βελτίωση της αποδοτικότητας, της ακρίβειας και της ευελιξίας των διαδικασιών προσλήψεων που πραγματοποιούνται από τα τμήματα Ανθρώπινου Δυναμικού (Human Resources – HR).

Οι επιμέρους στόχοι της εργασίας συνοψίζονται ως εξής:

- **Βελτιστοποίηση της Αναζήτησης και Κατάταξης Υποψηφίων**

Η παροχή δυνατότητας αυτοματοποιημένης αξιολόγησης και κατάταξης των υποψηφίων με βάση τον βαθμό συνάφειας των προσόντων τους προς τις απαιτήσεις μιας συγκεκριμένης θέσης εργασίας. Η κατάταξη πραγματοποιείται σε φθίνουσα σειρά βάσει ενός ποσοτικού δείκτη καταλληλότητας (Score), επιτρέποντας την ταχεία ανάδειξη των πλέον κατάλληλων υποψηφίων.

- **Υποστήριξη της Διαδικασίας Συνέντευξης**

Η λειτουργία του συστήματος ως εργαλείου υποστήριξης των υπευθύνων προσλήψεων μέσω της ανάλυσης των δεξιοτήτων και της επαγγελματικής εμπειρίας κάθε υποψηφίου. Για τον σκοπό αυτό παράγονται εξατομικευμένες προτεινόμενες ερωτήσεις συνέντευξης, οι οποίες εστιάζουν σε πιθανά κενά γνώσεων, δεξιοτήτων ή εμπειρίας.

- **Εξοικονόμηση Χρόνου και Πόρων**

Η μείωση του απαιτούμενου χρόνου κατά το αρχικό στάδιο αξιολόγησης των αιτήσεων μέσω της αυτόματης δημιουργίας συνοπτικών περιγραφών των βιογραφικών σημειωμάτων, περιορίζοντας την ανάγκη χειροκίνητης εξέτασης μεγάλου όγκου εγγράφων.

- **Σημασιολογική Αντιστοίχιση Υποψηφίων**

Η αξιοποίηση τεχνικών σημασιολογικής αναζήτησης (Semantic Search) για τον εντοπισμό των πλέον κατάλληλων υποψηφίων βάσει του πραγματικού νοήματος των δεξιοτήτων και των προσόντων τους και όχι αποκλειστικά βάσει προκαθορισμένων λέξεων-κλειδίων.

- **Υποστήριξη Πολυτροπικών Δεδομένων**

Η δυνατότητα επεξεργασίας διαφορετικών μορφών αρχείων, όπως PDF, JPG, PNG και ZIP, επιτρέποντας την αποτελεσματική διαχείριση τόσο ψηφιακών όσο και σαρωμένων βιογραφικών σημειωμάτων μέσω τεχνολογιών αναγνώρισης κειμένου.

- **Διασφάλιση Ιδιωτικότητας και Ασφάλειας Δεδομένων**

Η προστασία των προσωπικών δεδομένων των υποψηφίων μέσω της τοπικής εκτέλεσης του γλωσσικού μοντέλου και της αποφυγής αποστολής ευαίσθητων πληροφοριών σε εξωτερικές υπολογιστικές υποδομές ή υπηρεσίες νέφους (cloud services), συμβάλλοντας στη συμμόρφωση με τις απαιτήσεις του Γενικού Κανονισμού Προστασίας Δεδομένων (GDPR) [31].

- **Αξιοποίηση Σύγχρονων Τεχνολογιών Τεχνητής Νοημοσύνης**

Η ενσωμάτωση προηγμένων τεχνολογιών και εργαλείων, όπως οι τεχνικές Retrieval-Augmented Generation (RAG), η διανυσματική βάση δεδομένων ChromaDB, το γλωσσικό μοντέλο Phi-3, το EasyOCR και το πλαίσιο LangChain, με στόχο τη βελτίωση της ακρίβειας και της αξιοπιστίας των αποτελεσμάτων.

- **Βελτίωση της Εμπειρίας Χρήστη**

Η ανάπτυξη ενός σύγχρονου, φιλικού και διαδραστικού περιβάλλοντος χρήσης μέσω του πλαισίου Streamlit, το οποίο επιτρέπει στον υπεύθυνο προσλήψεων να πραγματοποιεί άμεση οπτική σύγκριση μεταξύ νέων υποψηφίων και υποψηφίων που ανακτώνται από την υφιστάμενη βάση δεδομένων, διευκολύνοντας τη διαδικασία λήψης αποφάσεων.

3. Θεωρητικό Πλαίσιο και Βιβλιογραφική Ανασκόπηση

3.1 Η Ψηφιοποίηση των Διαδικασιών HR και η Ανάγκη για Αυτοματοποιημένη Διαλογή

Η ραγδαία εξέλιξη των τεχνολογιών πληροφορικής και η εκτεταμένη ψηφιοποίηση των επιχειρησιακών διαδικασιών έχουν μετασχηματίσει σημαντικά τον τομέα της Διοίκησης Ανθρωπίνου Δυναμικού (Human Resource Management - HRM). Η αξιολόγηση και διαλογή βιογραφικών σημειωμάτων (resume screening) αποτελεί σήμερα μία από τις πλέον απαιτητικές διαδικασίες κατά την επιλογή προσωπικού, καθώς οι υπεύθυνοι προσλήψεων καλούνται να διαχειριστούν μεγάλο αριθμό αιτήσεων που υποβάλλονται σε ψηφιακή μορφή.

Στο πλαίσιο αυτό, τα Συστήματα Παρακολούθησης Υποψηφίων (Applicant Tracking Systems - ATS) έχουν εξελιχθεί από απλά εργαλεία αποθήκευσης και διαχείρισης βιογραφικών σε ολοκληρωμένες ψηφιακές πλατφόρμες που υποστηρίζουν την οργάνωση, αξιολόγηση και παρακολούθηση των υποψηφίων μέσω αυτοματοποιημένων διαδικασιών και αλγοριθμικών μεθόδων.

Η διαδικασία αξιολόγησης ξεκινά με την εξαγωγή και επεξεργασία του περιεχομένου των βιογραφικών. Η χρήση τεχνολογιών Οπτικής Αναγνώρισης Χαρακτήρων (Optical Character Recognition - OCR) επιτρέπει τη μετατροπή αδόμητων δεδομένων, όπως σαρωμένα έγγραφα και εικόνες, σε επεξεργάσιμο ψηφιακό κείμενο. Με τον τρόπο αυτό καθίσταται δυνατή η ανάλυση πληροφοριών που διαφορετικά δεν θα μπορούσαν να αξιοποιηθούν από αυτοματοποιημένα συστήματα.

Η ενσωμάτωση τεχνολογιών Τεχνητής Νοημοσύνης στη διαδικασία επιλογής προσωπικού συμβάλλει σημαντικά στη μείωση του χρόνου αξιολόγησης, στον περιορισμό της κόπωσης κατά τη λήψη αποφάσεων και στη βελτίωση της συνέπειας των αποτελεσμάτων [27]. Παράλληλα, η αυτοματοποίηση μειώνει την επίδραση της ανθρώπινης υποκειμενικότητας και των ασυνείδητων προκαταλήψεων κατά το αρχικό στάδιο αξιολόγησης [2]. Ως αποτέλεσμα, οι οργανισμοί αποκτούν μια πιο ολοκληρωμένη και άμεσα προσβάσιμη εικόνα της διαθέσιμης δεξαμενής ταλέντων, γεγονός που επιταχύνει τη διαδικασία πρόσληψης και συμβάλλει στη βελτιστοποίηση της αξιοποίησης των διαθέσιμων πόρων.

3.2 Αρχιτεκτονική RAG, Σημασιολογική Αναζήτηση και Μεγάλα Γλωσσικά Μοντέλα

Τα Μεγάλα Γλωσσικά Μοντέλα (Large Language Models - LLMs) και οι τεχνικές Σημασιολογικής Αναζήτησης (Semantic Search) αποτελούν βασικά συστατικά των σύγχρονων συστημάτων επεξεργασίας φυσικής γλώσσας [4, 17, 19]. Σε αντίθεση με τα παραδοσιακά συστήματα διαλογής βιογραφικών, τα οποία βασίζονται κυρίως στην αντιστοίχιση λέξεων-κλειδιών, οι σύγχρονες προσεγγίσεις αξιοποιούν διανυσματικές αναπαραστάσεις κειμένου, γνωστές ως embeddings [5, 26].

Μέσω των embeddings, το περιεχόμενο ενός βιογραφικού σημειώματος και μιας περιγραφής θέσης εργασίας μετατρέπεται σε μαθηματικά διανύσματα υψηλής διάστασης [18, 22]. Η σύγκριση μεταξύ τους πραγματοποιείται μέσω μετρικών σημασιολογικής ομοιότητας, όπως η ομοιότητα συνημιτόνου (Cosine Similarity), επιτρέποντας τον εντοπισμό υποψηφίων που

παρουσιάζουν ουσιαστική συνάφεια με τις απαιτήσεις της θέσης, ακόμη και όταν χρησιμοποιούν διαφορετική ορολογία ή συνώνυμες εκφράσεις [28].

Μία από τις σημαντικότερες εξελίξεις στον χώρο των γλωσσικών μοντέλων είναι η αρχιτεκτονική Retrieval-Augmented Generation (RAG) [7]. Η προσέγγιση αυτή συνδυάζει τις δυνατότητες ανάκτησης πληροφορίας με τις δυνατότητες παραγωγής φυσικής γλώσσας, επιτρέποντας σε ένα γλωσσικό μοντέλο να αξιοποιεί εξωτερικές πηγές γνώσης κατά τη διαδικασία λήψης αποφάσεων [14].

Στην αρχιτεκτονική RAG, το σύστημα ανακτά αρχικά τα πλέον σχετικά έγγραφα από μία διανυσματική βάση δεδομένων, όπως η ChromaDB. Στη συνέχεια, τα ανακτημένα δεδομένα παρέχονται ως πρόσθετο πλαίσιο στο γλωσσικό μοντέλο, το οποίο παράγει αναλυτικές συνόψεις, αξιολογήσεις και εξατομικευμένες προτάσεις. Η προσέγγιση αυτή βελτιώνει σημαντικά την ποιότητα των αποτελεσμάτων, καθώς το μοντέλο βασίζεται σε πραγματικά δεδομένα του οργανισμού και όχι αποκλειστικά στη γνώση που απέκτησε κατά την εκπαίδευσή του.

Παράλληλα, ιδιαίτερο ενδιαφέρον παρουσιάζουν τα Μικρά Γλωσσικά Μοντέλα (Small Language Models - SLMs), τα οποία μπορούν να εκτελούνται τοπικά σε υπολογιστικές υποδομές οργανισμών χωρίς να απαιτείται η χρήση εξωτερικών υπηρεσιών νέφους. Η τοπική εκτέλεση μοντέλων μέσω πλατφορμών όπως το Ollama προσφέρει σημαντικά πλεονεκτήματα ως προς την προστασία της ιδιωτικότητας και την ασφάλεια των δεδομένων, καθώς τα προσωπικά στοιχεία των υποψηφίων παραμένουν εντός του οργανισμού καθ' όλη τη διάρκεια της επεξεργασίας τους.

Για τη διασύνδεση όλων των επιμέρους τεχνολογιών χρησιμοποιούνται σύγχρονα πλαίσια εντοπισμού, όπως το LangChain, το οποίο επιτρέπει την αποτελεσματική συνεργασία μεταξύ γλωσσικών μοντέλων, διανυσματικών βάσεων δεδομένων και μηχανισμών ανάκτησης πληροφορίας. Με τον τρόπο αυτό δημιουργούνται ολοκληρωμένα συστήματα τεχνητής νοημοσύνης, ικανά να υποστηρίξουν πολύπλοκες διαδικασίες λήψης αποφάσεων στον τομέα της διαχείρισης ανθρώπινου δυναμικού.

3.3 Τοπική Υπολογιστική Ισχύς, Ιδιωτικότητα Δεδομένων και Τοπική Εκτέλεση (Local Deployment)

Η επεξεργασία μεγάλου όγκου δεδομένων κειμένου, η εφαρμογή τεχνολογιών Οπτικής Αναγνώρισης Χαρακτήρων (Optical Character Recognition - OCR) και η αξιοποίηση προηγμένων μοντέλων Τεχνητής Νοημοσύνης σε πραγματικό χρόνο απαιτούν σημαντικούς υπολογιστικούς πόρους. Στη σύγχρονη βιομηχανία λογισμικού, η κάλυψη αυτών των απαιτήσεων επιτυγχάνεται συχνά μέσω υπηρεσιών Υπολογιστικού Νέφους (Cloud Computing), οι οποίες παρέχουν κλιμακούμενη υπολογιστική ισχύ και αποθηκευτικούς πόρους.

Ωστόσο, στον τομέα της διαχείρισης ανθρώπινου δυναμικού, η επεξεργασία βιογραφικών σημειωμάτων συνεπάγεται τη διαχείριση ιδιαίτερα ευαίσθητων προσωπικών δεδομένων, όπως στοιχεία ταυτότητας, επαγγελματική εμπειρία, ακαδημαϊκά προσόντα και πληροφορίες επικοινωνίας. Η αποστολή και αποθήκευση των δεδομένων αυτών σε εξωτερικές υποδομές ενδέχεται να δημιουργήσει ζητήματα ασφάλειας, ιδιωτικότητας και κανονιστικής συμμόρφωσης, ιδιαίτερα σε περιβάλλοντα που διέπονται από αυστηρά νομοθετικά πλαίσια, όπως ο Γενικός Κανονισμός Προστασίας Δεδομένων (General Data Protection Regulation - GDPR) [30].

Για την αντιμετώπιση των παραπάνω προκλήσεων, τα τελευταία χρόνια παρατηρείται αυξανόμενο ενδιαφέρον για την τοπική εκτέλεση (Local Deployment) μοντέλων τεχνητής

νοημοσύνης. Μέσω πλατφορμών όπως το Ollama, καθίσταται δυνατή η εκτέλεση γλωσσικών μοντέλων απευθείας στις υπολογιστικές υποδομές του οργανισμού, χωρίς την ανάγκη αποστολής δεδομένων σε εξωτερικές υπηρεσίες νέφους. Με τον τρόπο αυτό, διασφαλίζεται ότι όλες οι διαδικασίες επεξεργασίας πραγματοποιούνται εντός του οργανισμού, μειώνοντας σημαντικά τους κινδύνους διαρροής ή μη εξουσιοδοτημένης πρόσβασης στα δεδομένα.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η αξιοποίηση Μικρών Γλωσσικών Μοντέλων (Small Language Models - SLMs), όπως το Phi-3, τα οποία προσφέρουν υψηλή απόδοση με σχετικά περιορισμένες απαιτήσεις σε υπολογιστικούς πόρους. Η δυνατότητα αυτή επιτρέπει την ανάπτυξη αποτελεσματικών εφαρμογών τεχνητής νοημοσύνης ακόμη και σε συστήματα χωρίς εξειδικευμένο εξοπλισμό υψηλών επιδόσεων.

Επιπλέον, η τοπική εκτέλεση συμβάλλει στη μείωση του λειτουργικού κόστους, καθώς εξαλείφεται η ανάγκη χρήσης εξωτερικών υπηρεσιών μέσω προγραμματιστικών διεπαφών (APIs), οι οποίες συχνά συνοδεύονται από οικονομικές χρεώσεις ανά χρήση. Παράλληλα, ενισχύεται η επιχειρησιακή ανεξαρτησία του οργανισμού και περιορίζεται η εξάρτηση από τρίτους παρόχους υπηρεσιών.

Ως εκ τούτου, η τοπική υπολογιστική ισχύς και η τοπική εκτέλεση γλωσσικών μοντέλων αποτελούν μια ιδιαίτερα ελκυστική προσέγγιση για εφαρμογές Διοίκησης Ανθρωπίνου Δυναμικού, συνδυάζοντας υψηλό επίπεδο προστασίας δεδομένων, μειωμένο λειτουργικό κόστος και αυξημένο έλεγχο επί των διαδικασιών επεξεργασίας πληροφοριών.

3.4 Σχεδιασμός Εμπειρίας Χρήστη (UX) και Διαισθητική Διεπαφή HR

Η αποτελεσματικότητα ενός συστήματος υποστήριξης αποφάσεων δεν εξαρτάται αποκλειστικά από την ακρίβεια των αλγορίθμων που χρησιμοποιεί, αλλά και από την ποιότητα της αλληλεπίδρασης που προσφέρει στον τελικό χρήστη. Στην περίπτωση συστημάτων αξιολόγησης και κατάταξης βιογραφικών σημειωμάτων, οι κύριοι χρήστες είναι οι επαγγελματίες του τομέα Ανθρωπίνου Δυναμικού, οι οποίοι καλούνται να λαμβάνουν κρίσιμες αποφάσεις σε περιορισμένο χρονικό διάστημα.

Οι σύγχρονες αρχές σχεδιασμού εμπειρίας χρήστη (User Experience - UX) δίνουν ιδιαίτερη έμφαση στην απλότητα, τη σαφήνεια και τη λειτουργικότητα των διεπαφών. Ένα επιτυχημένο σύστημα οφείλει να επιτρέπει την εύκολη πλοήγηση, την ταχεία εισαγωγή δεδομένων και την άμεση κατανόηση των αποτελεσμάτων, χωρίς να απαιτεί εξειδικευμένες τεχνικές γνώσεις από τον χρήστη. Η προσέγγιση του σχεδιασμού με επίκεντρο τον χρήστη (User-Centered Design) επιδιώκει ακριβώς τη γεφύρωση του χάσματος μεταξύ των σύνθετων τεχνολογικών μηχανισμών και των πραγματικών επιχειρησιακών αναγκών.

Σύμφωνα με τη σχετική βιβλιογραφία στον τομέα της αλληλεπίδρασης ανθρώπου-υπολογιστή (Human-Computer Interaction - HCI), η οπτική παρουσίαση της πληροφορίας επηρεάζει σημαντικά την αποτελεσματικότητα της λήψης αποφάσεων. Για τον λόγο αυτό, η παράλληλη παρουσίαση πληροφοριών (side-by-side comparison) αποτελεί μια ιδιαίτερα αποδοτική σχεδιαστική πρακτική, καθώς επιτρέπει στον χρήστη να συγκρίνει άμεσα διαφορετικές κατηγορίες δεδομένων χωρίς να απαιτείται συνεχής εναλλαγή οθονών ή προβολών.

Στο πλαίσιο της παρούσας εργασίας, η διεπαφή σχεδιάστηκε ώστε να προσφέρει ταυτόχρονη προβολή των νεοεισαχθέντων υποψηφίων και των πλέον σχετικών βιογραφικών που ανακτώνται από την υπάρχουσα βάση δεδομένων. Η προσέγγιση αυτή ενισχύει τη συγκριτική αξιολόγηση και διευκολύνει την ταχύτερη αναγνώριση των πλέον κατάλληλων υποψηφίων.

Παράλληλα, ιδιαίτερη έμφαση δόθηκε στις αρχές της Επεξηγήσιμης Τεχνητής Νοημοσύνης (Explainable Artificial Intelligence - XAI). Αντί της απλής εμφάνισης ενός αριθμητικού δείκτη καταλληλότητας, το σύστημα παρέχει επιπρόσθετη αιτιολόγηση των αποτελεσμάτων μέσω ποιοτικών συνόψεων και εξατομικευμένων ερωτήσεων συνέντευξης. Η πρακτική αυτή αυξάνει τη διαφάνεια του συστήματος και ενισχύει την εμπιστοσύνη των χρηστών προς τις προτεινόμενες αξιολογήσεις.

Η ανάπτυξη της διεπαφής πραγματοποιήθηκε με τη χρήση του πλαισίου Streamlit, το οποίο επιτρέπει τη δημιουργία σύγχρονων και διαδραστικών διαδικτυακών εφαρμογών με έμφαση στην ευχρηστία [32]. Μέσω της αξιοποίησης του συγκεκριμένου εργαλείου, δημιουργήθηκε ένα καθαρό και επαγγελματικό περιβάλλον εργασίας που υποστηρίζει αποτελεσματικά τις καθημερινές ανάγκες των υπευθύνων προσλήψεων και συμβάλλει στη βελτίωση της συνολικής εμπειρίας χρήσης.

3.5 Προηγμένες Τεχνολογίες και Μελλοντικές Προοπτικές

3.5.1 Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence – XAI) στη Διοίκηση Ανθρώπινου Δυναμικού

Ένα από τα σημαντικότερα ζητήματα που συνοδεύουν την αξιοποίηση συστημάτων Τεχνητής Νοημοσύνης στη λήψη αποφάσεων είναι το φαινόμενο της «μαύρης κούτας» (black box) [24]. Στην περίπτωση αυτή, ένα σύστημα είναι σε θέση να παράγει αποτελέσματα χωρίς να καθίσταται σαφής ο τρόπος με τον οποίο οδηγήθηκε σε αυτά. Η έλλειψη διαφάνειας αποτελεί σημαντικό εμπόδιο για την υιοθέτηση τεχνολογιών τεχνητής νοημοσύνης σε κρίσιμους τομείς, όπως η επιλογή προσωπικού και η διαχείριση ανθρώπινου δυναμικού.

Η Εξηγήσιμη Τεχνητή Νοημοσύνη (Explainable Artificial Intelligence – XAI) αναπτύχθηκε με στόχο την ενίσχυση της διαφάνειας και της ερμηνευσιμότητας των αλγοριθμικών αποφάσεων [3]. Στο πλαίσιο της αξιολόγησης βιογραφικών σημειωμάτων, η XAI δεν περιορίζεται στην παραγωγή ενός αριθμητικού δείκτη καταλληλότητας, αλλά παρέχει επεξηγήσεις σχετικά με τα χαρακτηριστικά, τις δεξιότητες και την επαγγελματική εμπειρία που επηρέασαν την τελική αξιολόγηση του υποψηφίου [15].

Η δυνατότητα παροχής δομημένων και κατανοητών εξηγήσεων ενισχύει την εμπιστοσύνη των χρηστών προς το σύστημα και διευκολύνει τη διαδικασία λήψης αποφάσεων. Επιπλέον, η σύνδεση των αποτελεσμάτων με εξατομικευμένες ερωτήσεις συνέντευξης συμβάλλει στη δημιουργία ενός περιβάλλοντος μεγαλύτερης διαφάνειας και αντικειμενικότητας, μειώνοντας τον κίνδυνο αυθαίρετων ή μη τεκμηριωμένων αποφάσεων κατά την επιλογή προσωπικού.

3.5.2 Παραγωγική Τεχνητή Νοημοσύνη και Ευφυείς Πράκτορες (Agentic HR Systems)

Η ανάπτυξη των Μεγάλων Γλωσσικών Μοντέλων (Large Language Models – LLMs) έχει οδηγήσει στη δημιουργία νέων μορφών ευφυών πληροφοριακών συστημάτων, τα οποία υπερβαίνουν τις δυνατότητες των παραδοσιακών εφαρμογών αυτοματοποίησης. Μία από τις σημαντικότερες εξελίξεις αποτελεί η εμφάνιση των ευφυών πρακτόρων (Agentic AI Systems), οι οποίοι είναι σε θέση να αντιλαμβάνονται το πλαίσιο μιας εργασίας, να ανακτούν πληροφορίες από πολλαπλές πηγές και να εκτελούν σύνθετες διαδικασίες με περιορισμένη ανθρώπινη παρέμβαση. [6, 25]

Στον τομέα της επιλογής προσωπικού, οι ευφυείς πράκτορες μπορούν να λειτουργήσουν ως ψηφιακοί συνεργάτες των επαγγελματιών Ανθρώπινου Δυναμικού. Μέσω της επεξεργασίας φυσικής γλώσσας, έχουν τη δυνατότητα να απαντούν σε εξειδικευμένα ερωτήματα σχετικά με τους υποψηφίους, να συγκρίνουν προφίλ, να αναζητούν κατάλληλους υποψηφίους για συγκεκριμένες θέσεις εργασίας και να προτείνουν εναλλακτικές στρατηγικές στελέχωσης.

Η αλληλεπίδραση με φυσική γλώσσα καθιστά τα συστήματα αυτά περισσότερο προσιτά και λειτουργικά, μετατρέποντάς τα από απλά εργαλεία ανάλυσης δεδομένων σε ενεργούς υποστηρικτές της διαδικασίας λήψης αποφάσεων. Η εξέλιξη αυτή αναμένεται να διαδραματίσει καθοριστικό ρόλο στο μέλλον της ψηφιακής διαχείρισης ανθρώπινου δυναμικού.

3.6 Επιχειρησιακή Βιωσιμότητα, Επαγγελματική Συμπεριφορά και Κοινωνικός Αντίκτυπος

Η αξιοποίηση συστημάτων τεχνητής νοημοσύνης στη διαχείριση ανθρώπινου δυναμικού δεν συνδέεται αποκλειστικά με τη βελτίωση της αποδοτικότητας των διαδικασιών προσλήψεων, αλλά και με τη δημιουργία πιο βιώσιμων και κοινωνικά υπεύθυνων επιχειρησιακών πρακτικών. Η αυτοματοποίηση της αξιολόγησης βιογραφικών συμβάλλει στη μείωση του χρόνου επεξεργασίας των αιτήσεων, στη βελτιστοποίηση της αξιοποίησης των διαθέσιμων πόρων και στη μείωση του κόστους που σχετίζεται με λανθασμένες ή μη αποδοτικές προσλήψεις.

Παράλληλα, η διατήρηση και αξιοποίηση ιστορικών δεδομένων υποψηφίων επιτρέπει στους οργανισμούς να επαναχρησιμοποιούν αποτελεσματικά τη διαθέσιμη δεξαμενή ταλέντων, περιορίζοντας την ανάγκη για επαναλαμβανόμενες διαδικασίες αναζήτησης προσωπικού. Η δυνατότητα αυτή ενισχύει τη βιωσιμότητα των διαδικασιών στελέχωσης και συμβάλλει στη δημιουργία μακροχρόνιας αξίας για τον οργανισμό.

Επιπλέον, η ψηφιακή διαχείριση των δεδομένων διευκολύνει την εσωτερική κινητικότητα ταλέντων (talent mobility), επιτρέποντας την επαναξιολόγηση παλαιότερων υποψηφίων για νέες θέσεις εργασίας. Η πρακτική αυτή προάγει την αποτελεσματικότερη αξιοποίηση του ανθρώπινου δυναμικού και ενισχύει τη στρατηγική διαχείριση ταλέντων.

Από κοινωνική σκοπιά, η χρήση τεχνολογιών τεχνητής νοημοσύνης μπορεί να συμβάλει στη βελτίωση της διαφάνειας και της αντικειμενικότητας κατά τη διαδικασία επιλογής προσωπικού, εφόσον εφαρμόζονται κατάλληλοι μηχανισμοί ελέγχου και επεξηγησιμότητας [11]. Η αξιολόγηση όλων των υποψηφίων με βάση κοινά και προκαθορισμένα κριτήρια συμβάλλει στον περιορισμό των προκαταλήψεων και στη διαμόρφωση ενός δικαιότερου εργασιακού περιβάλλοντος [13].

Παρότι η ανάπτυξη και συντήρηση τέτοιων συστημάτων απαιτεί αρχικές επενδύσεις σε τεχνολογική υποδομή και εξειδικευμένο προσωπικό, τα μακροπρόθεσμα οφέλη σε επιχειρησιακό, κοινωνικό και οργανωσιακό επίπεδο καθιστούν την υιοθέτησή τους ιδιαίτερα ελκυστική. Ως εκ τούτου, οι τεχνολογίες τεχνητής νοημοσύνης αναμένεται να διαδραματίσουν καθοριστικό ρόλο στη διαμόρφωση ενός πιο αποδοτικού, διαφανούς και βιώσιμου περιβάλλοντος εργασίας στο μέλλον.

3.7 Πρακτικές Εφαρμογές και Μελέτες Περίπτωσης στον Τομέα του Ανθρώπινου Δυναμικού

Η αξιοποίηση τεχνολογιών Τεχνητής Νοημοσύνης στη διαχείριση ανθρώπινου δυναμικού έχει γνωρίσει σημαντική ανάπτυξη τα τελευταία χρόνια, με αποτέλεσμα την εμφάνιση προηγμένων συστημάτων που υποστηρίζουν τη διαδικασία αναζήτησης, αξιολόγησης και επιλογής

προσωπικού [8, 29]. Πολλές σύγχρονες πλατφόρμες διαχείρισης ταλέντων ενσωματώνουν τεχνικές μηχανικής μάθησης, επεξεργασίας φυσικής γλώσσας και σημασιολογικής αναζήτησης, με σκοπό τη βελτίωση της αποδοτικότητας των διαδικασιών προσλήψεων [9].

Ενδεικτικά παραδείγματα αποτελούν πλατφόρμες όπως το Ideal, το Eightfold.ai και το Fetcher, οι οποίες αξιοποιούν αλγορίθμους τεχνητής νοημοσύνης για την ανάλυση μεγάλου όγκου βιογραφικών σημειωμάτων και την αυτόματη αντιστοίχιση υποψηφίων με διαθέσιμες θέσεις εργασίας [10]. Τα συστήματα αυτά είναι σε θέση να δημιουργούν προφίλ δεξιοτήτων, να αξιολογούν την επαγγελματική εμπειρία των υποψηφίων και να εντοπίζουν συσχετίσεις μεταξύ των απαιτήσεων μιας θέσης εργασίας και των προσόντων των ενδιαφερομένων.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η χρήση σημασιολογικών τεχνικών αναζήτησης, οι οποίες επιτρέπουν τον εντοπισμό κατάλληλων υποψηφίων ακόμη και όταν χρησιμοποιείται διαφορετική ορολογία για την περιγραφή παρόμοιων γνώσεων ή επαγγελματικών εμπειριών. Μέσω των διανυσματικών αναπαραστάσεων κειμένου (embeddings), τα συστήματα αυτά αξιολογούν το πραγματικό νόημα των πληροφοριών και όχι αποκλειστικά τη λεξική ταύτιση λέξεων-κλειδίων [21].

Παράλληλα, η ενσωμάτωση τεχνολογιών Οπτικής Αναγνώρισης Χαρακτήρων (Optical Character Recognition – OCR) επιτρέπει την αυτοματοποιημένη εξαγωγή κειμένου από σαρωμένα έγγραφα και αρχεία εικόνων, διευρύνοντας το εύρος των μορφών δεδομένων που μπορούν να υποστηριχθούν από τα σύγχρονα συστήματα διαχείρισης υποψηφίων. Με τον τρόπο αυτό καθίσταται δυνατή η αποτελεσματική επεξεργασία μεγάλου αριθμού εγγράφων ανεξαρτήτως της μορφής τους.

Επιπλέον, η εμφάνιση συστημάτων που βασίζονται σε Μεγάλα Γλωσσικά Μοντέλα (Large Language Models – LLMs), όπως το ChatGPT Enterprise και άλλες εξειδικευμένες εφαρμογές τεχνητής νοημοσύνης, έχει επεκτείνει σημαντικά τις δυνατότητες των πληροφοριακών συστημάτων ανθρώπινου δυναμικού. Τα συστήματα αυτά επιτρέπουν την αλληλεπίδραση μέσω φυσικής γλώσσας, την παραγωγή συνόψεων, την υποστήριξη της διαδικασίας λήψης αποφάσεων και τη δημιουργία εξατομικευμένων προτάσεων για τη διεξαγωγή συνεντεύξεων.

Οι παραπάνω εφαρμογές καταδεικνύουν ότι τεχνολογίες όπως η σημασιολογική αναζήτηση, οι διανυσματικές βάσεις δεδομένων και η αρχιτεκτονική Retrieval-Augmented Generation (RAG) δεν αποτελούν πλέον αποκλειστικά αντικείμενο ακαδημαϊκής έρευνας, αλλά έχουν ήδη ενσωματωθεί σε πραγματικά επιχειρησιακά περιβάλλοντα. Η συνεχής εξέλιξή τους αναμένεται να μετασχηματίσει περαιτέρω τον τρόπο με τον οποίο οι οργανισμοί εντοπίζουν, αξιολογούν και επιλέγουν ανθρώπινο δυναμικό.

4. Περιγραφή Προβλήματος

Η διαδικασία επιλογής προσωπικού αποτελεί μία από τις σημαντικότερες λειτουργίες των τμημάτων Ανθρώπινου Δυναμικού (Human Resources – HR), καθώς επηρεάζει άμεσα τη στελέχωση, την παραγωγικότητα και τη συνολική αποτελεσματικότητα ενός οργανισμού. Παρά τη σημαντική πρόοδο που έχει σημειωθεί στον τομέα της ψηφιοποίησης των επιχειρησιακών διαδικασιών, η αρχική αξιολόγηση και διαλογή βιογραφικών σημειωμάτων εξακολουθεί σε πολλές περιπτώσεις να βασίζεται σε χειροκίνητες διαδικασίες, οι οποίες απαιτούν σημαντικό χρόνο και ανθρώπινη προσπάθεια.

Οι σύγχρονοι οργανισμοί καλούνται συχνά να διαχειριστούν μεγάλο αριθμό αιτήσεων για την κάλυψη μίας ή περισσότερων θέσεων εργασίας. Η εξέταση κάθε βιογραφικού ξεχωριστά, η σύγκριση των προσόντων των υποψηφίων με τις απαιτήσεις της θέσης και η τελική επιλογή των

επικρατέστερων υποψηφίων αποτελούν διαδικασίες που επιβαρύνουν σημαντικά τον φόρτο εργασίας των υπευθύνων προσλήψεων. Η συνεχής λήψη αποφάσεων κατά τη διάρκεια αυτής της διαδικασίας μπορεί να οδηγήσει στο φαινόμενο της κόπωσης λήψης αποφάσεων (decision fatigue), το οποίο επηρεάζει αρνητικά την ποιότητα της κρίσης και την αποτελεσματικότητα των αποφάσεων που λαμβάνονται.

Παράλληλα, η ανάγκη για ταχεία κάλυψη κενών θέσεων εργασίας δημιουργεί πρόσθετη χρονική πίεση στα τμήματα Ανθρωπίνου Δυναμικού. Ως αποτέλεσμα, οι υπεύθυνοι προσλήψεων συχνά καλούνται να αξιολογήσουν μεγάλο αριθμό υποψηφίων σε περιορισμένο χρονικό διάστημα, γεγονός που αυξάνει την πιθανότητα παράβλεψης κατάλληλων υποψηφίων ή λήψης λιγότερο τεκμηριωμένων αποφάσεων.

Ένα επιπλέον πρόβλημα αφορά τη διαχείριση και αξιοποίηση των ήδη διαθέσιμων δεδομένων υποψηφίων. Πολλοί οργανισμοί διαθέτουν ιστορικά αρχεία βιογραφικών σημειωμάτων τα οποία έχουν συλλεχθεί κατά τη διάρκεια προηγούμενων διαδικασιών πρόσληψης. Ωστόσο, η αναζήτηση και επαναξιοποίηση αυτών των δεδομένων είναι συχνά δύσκολη, ιδιαίτερα όταν τα βιογραφικά είναι αποθηκευμένα σε διαφορετικές μορφές αρχείων ή όταν δεν υπάρχουν κατάλληλοι μηχανισμοί ταξινόμησης και αναζήτησης. Ως συνέπεια, οι επιχειρήσεις ενδέχεται να προχωρούν σε νέες διαδικασίες αναζήτησης προσωπικού, παρόλο που κατάλληλοι υποψήφιοι υπάρχουν ήδη στο διαθέσιμο ιστορικό δεδομένων.

Η πολυμορφία των αρχείων που υποβάλλονται από τους υποψηφίους αποτελεί επίσης σημαντική πρόκληση. Τα βιογραφικά σημειώματα μπορεί να βρίσκονται σε μορφή PDF, εικόνας ή σαρωμένων εγγράφων, γεγονός που δυσχεραίνει την αυτοματοποιημένη επεξεργασία και αξιολόγησή τους. Η απουσία μηχανισμών εξαγωγής και ομογενοποίησης της πληροφορίας περιορίζει σημαντικά τη δυνατότητα αποτελεσματικής διαχείρισης των δεδομένων αυτών.

Επιπλέον, η αξιολόγηση της καταλληλότητας ενός υποψηφίου δεν μπορεί να βασίζεται αποκλειστικά στην ύπαρξη συγκεκριμένων λέξεων-κλειδιών μέσα στο βιογραφικό σημείωμα. Η αποτελεσματική αντιστοίχιση μεταξύ υποψηφίων και θέσεων εργασίας απαιτεί κατανόηση του πραγματικού νοήματος των δεξιοτήτων, της επαγγελματικής εμπειρίας και των προσόντων που περιγράφονται στα έγγραφα. Τα παραδοσιακά συστήματα αναζήτησης αδυνατούν συχνά να εντοπίσουν τέτοιες σημασιολογικές συσχετίσεις, με αποτέλεσμα να περιορίζεται η ποιότητα των προτεινόμενων υποψηφίων.

Λαμβάνοντας υπόψη τις παραπάνω προκλήσεις, καθίσταται αναγκαία η ανάπτυξη ενός σύγχρονου συστήματος που θα υποστηρίζει την αυτοματοποιημένη εξαγωγή, επεξεργασία, ανάλυση και αξιολόγηση βιογραφικών σημειωμάτων. Ένα τέτοιο σύστημα οφείλει να αξιοποιεί τεχνολογίες τεχνητής νοημοσύνης, σημασιολογικής αναζήτησης και επεξεργασίας φυσικής γλώσσας, με στόχο τη βελτίωση της ταχύτητας, της ακρίβειας και της αξιοπιστίας της διαδικασίας επιλογής προσωπικού.

Η παρούσα εργασία επιχειρεί να αντιμετωπίσει τα προβλήματα αυτά μέσω της ανάπτυξης μιας ολοκληρωμένης εφαρμογής αξιολόγησης και κατάταξης βιογραφικών σημειωμάτων. Η εφαρμογή αξιοποιεί τεχνολογίες OCR, διανυσματικές βάσεις δεδομένων, αρχιτεκτονική Retrieval-Augmented Generation (RAG) και γλωσσικά μοντέλα τεχνητής νοημοσύνης, με σκοπό την υποστήριξη των υπευθύνων προσλήψεων στη λήψη τεκμηριωμένων και αποδοτικών αποφάσεων.

5. Αναφορά σε παρόμοιες εφαρμογές

Στο χώρο της ψηφιακής διαχείρισης και αξιολόγησης ανθρώπινου δυναμικού έχουν αναπτυχθεί αρκετές εφαρμογές που έχουν ως στόχο την καλύτερη οργάνωση των βιογραφικών σημειωμάτων αλλά και την παροχή έξυπνων συστάσεων με γνώμονα τις απαιτήσεις των επιχειρήσεων. Η ανάλυση των εφαρμογών αυτών είναι ιδιαίτερα σημαντική προκειμένου να μπορέσουν να γίνουν κατανοητά τα πλεονεκτήματα και οι ελλείψεις τους, καθώς και για την ανάδειξη του τρόπου με τον οποίο η προτεινόμενη λύση διαφοροποιείται όχι ως προς το ίδιο το αντικείμενο, αλλά σε σχέση με την εμπειρία, την ευελιξία και τις ανάγκες του χρήστη (HR Manager).

5.1. Eightfold.ai

Η εφαρμογή Eightfold.ai αποτελεί ένα χαρακτηριστικό παράδειγμα πλατφόρμας διαχείρισης ταλέντων που απευθύνεται σε μεγάλους οργανισμούς και επιχειρήσεις. Βασικές λειτουργίες της εφαρμογής, που διαθέτει όλες τις δυνατότητες ενός πλήρους συστήματος ATS (Applicant Tracking System), είναι η συγκέντρωση των βιογραφικών των υποψηφίων, καθώς και η δημιουργία σημασιολογικών συσχετίσεων ανάμεσα στις δεξιότητές τους και τις επιθυμητές θέσεις εργασίας. Σε αυτές τις δυνατότητες προστίθεται ακόμα και η χαρτογράφηση της μελλοντικής εξέλιξης των εργαζομένων σε επιχειρησιακό μοντέλο, προκειμένου να υπάρχει μια πιο αποδοτική οργάνωση για επερχόμενες ανάγκες στελέχωσης.

Το Eightfold.ai συλλέγει επίσης στατιστικά για τον κάθε υποψήφιο σε σχέση με την επαγγελματική του πορεία, ενώ για κάθε μεμονωμένο προφίλ που βρίσκεται στην ηλεκτρονική βάση καταγράφει σημαντικές πληροφορίες που σχετίζονται κυρίως με τη συνάφεια των δεξιοτήτων και τις πιθανότητες επιτυχίας του στη θέση. Ακόμα μια χρήσιμη λειτουργία της εφαρμογής είναι η ύπαρξη ενότητας για την πρόβλεψη μελλοντικών αναγκών της αγοράς, προκειμένου να μπορεί ο οργανισμός να οραματιστεί σε μεγαλύτερο εύρος τη στελέχωσή του, ενισχύοντας έτσι τη στρατηγική ευελιξία του τμήματος HR.

Τέλος, δίνει τη δυνατότητα διασύνδεσης με εξωτερικές πλατφόρμες εύρεσης εργασίας (όπως το LinkedIn), προσδίδοντας μια ευρύτερη δικτυακή διάσταση στην εφαρμογή που συνοδεύει τον γενικά οργανωτικό της χαρακτήρα, παρουσιάζοντας βέβαια έλλειψη στον τομέα της προστασίας δεδομένων, καθώς βασίζεται αποκλειστικά σε εξωτερικά συστήματα Cloud.

5.2. Ideal

Η εφαρμογή Ideal, όπως και η προαναφερόμενη, αποτελεί επίσης μια ολοκληρωμένη πλατφόρμα αξιολόγησης υποψηφίων, παρουσιάζοντας παρόμοια θεμελιώδη χαρακτηριστικά, ιδιαίτερα στις κεντρικές λειτουργίες. Στις λειτουργίες αυτές συγκαταλέγονται η μαζική εισαγωγή βιογραφικών από τον χρήστη και στη συνέχεια η αυτόματη βαθμολόγηση (screening) των υποψηφίων που έχει στη διάθεσή του, αξιοποιώντας το υπάρχον pool της εταιρείας.

Το Ideal αποτελεί μια εφαρμογή βασιζόμενη στην τεχνητή νοημοσύνη, καθώς τροφοδοτείται κατά τη σύνδεση του χρήστη με βασικές πληροφορίες σχετικά με τις απαιτήσεις της θέσης εργασίας και τα επιθυμητά χαρακτηριστικά των ιδανικών υπαλλήλων. Ενισχύει τη χρηστικότητα ενσωματώνοντας την τεχνητή νοημοσύνη στη λειτουργία της προβολής της κατάταξης και επίσης με την ύπαρξη ψηφιακού βοηθού για την παροχή απαντήσεων σε τυχόν διευκρινιστικά ερωτήματα γύρω από τα προσόντα των υποψηφίων.

Στο πεδίο της δημιουργίας μικρών λιστών (shortlists), η εφαρμογή επιτρέπει στον χρήστη να επιλέξει μια συγκεκριμένη κατηγορία προσόντων και με βάση αυτή να παραχθούν διαφορετικές

επιλογές κατάταξης, αξιοποιώντας τις δυνατότητες της βάσης δεδομένων του. Επιπλέον, μπορεί να αναθέσει κάθε υποψήφιο σε συγκεκριμένα στάδια της διαδικασίας προσλήψεων κάνοντας χρήση ενός ψηφιακού πίνακα ροής εργασιών.

Ένα ακόμα κοινό στοιχείο με το Eightfold.ai είναι η συλλογή στατιστικών στοιχείων για την αποδοτικότητα των προσλήψεων, ενισχύοντας τον εταιρικό χαρακτήρα της πλατφόρμας. Ωστόσο, η υπερβολική έμφαση στα πολύπλοκα εταιρικά στατιστικά και στην εξάρτηση από τρίτα συστήματα Cloud μετατοπίζει την προσοχή του χρήστη από τον βασικό του στόχο, ο οποίος είναι η άμεση, απλή και ασφαλής τοπική διαχείριση και αξιολόγηση των βιογραφικών του.

5.3. Συγκριτική αξιολόγηση με προτεινόμενη λύση

Σε σύγκριση με τις παραπάνω λύσεις που εξετάστηκαν, η συγκεκριμένη εφαρμογή που πραγματεύεται η παρούσα πτυχιακή εργασία διαφοροποιείται με συγκεκριμένες καινοτομίες και τεχνολογίες. Οι τεχνολογίες αυτές συνδυάζουν τα προτερήματα ήδη υπάρχοντων εφαρμογών, προσδίδοντας ταυτόχρονα ευελιξία, απόλυτη ασφάλεια και ενισχυμένη αλληλεπίδραση με τον χρήστη.

Η προτεινόμενη εφαρμογή υποστηρίζει βασικές λειτουργίες όπως και οι υπόλοιπες εναλλακτικές λύσεις στον τομέα των συστημάτων ATS και της αξιολόγησης προσωπικού. Ωστόσο, ενσωματώνει λειτουργίες προσθήκης, διαχείρισης και προβολής αναλυτικών πληροφοριών για κάθε υποψήφιο, προσφέροντας στον HR Manager μια ολοκληρωμένη side-by-side εικόνα της δεξαμενής ταλέντων του. Για τη δημιουργία των προτάσεων και της κατάταξης, η εφαρμογή λαμβάνει υπόψη ως κριτήρια την περιγραφή της θέσης εργασίας και τα πραγματικά προσόντα των υποψηφίων μέσω σημασιολογικής αναζήτησης με τη χρήση της διανυσματικής βάσης ChromaDB. Με αυτόν τον τρόπο παρέχει ρεαλιστικές, αντικειμενικές και πρακτικές προτάσεις βασισμένες στις απαιτήσεις που ο ίδιος ο διαχειριστής όρισε.

Επιπλέον, η υλοποίηση των προσωποποιημένων προτάσεων (σύνοψη προσόντων και παραγωγή ερωτήσεων συνέντευξης μέσω του μοντέλου Phi-3) εστιάζει κυρίως στην ευχρηστία και στην επεξηγησιμότητα. Αυτό έρχεται σε αντίθεση με παρεμφερείς εμπορικές εφαρμογές που βασικός τους στόχος είναι τα περίπλοκα εταιρικά στατιστικά και οι κοινωνικές αλληλεπιδράσεις με εξωτερικά δίκτυα, μετατοπίζοντας το ενδιαφέρον από τις ουσιαστικές λειτουργίες της διαλογής. Τέλος, η αποκλειστικά τοπική εκτέλεση του συστήματος (local deployment μέσω Ollama) επιτρέπει στον χρήστη να επεξεργάζεται ευαίσθητα δεδομένα με πλήρη συμμόρφωση προς τον GDPR, δίνοντάς του έτσι μεγαλύτερο έλεγχο και ασφάλεια στην καθημερινή επιχειρησιακή του εμπειρία.

Συνοψίζοντας, υπάρχει μεγάλη ανάγκη στον τομέα των εφαρμογών διαχείρισης ανθρώπινου δυναμικού προκειμένου να εξοικονομούν χρόνο οι υπεύθυνοι προσλήψεων από τις καθημερινές τους δραστηριότητες ρουτίνας. Στο μεγάλο εύρος των εφαρμογών είναι αναμενόμενο κάποιες λειτουργίες να είναι παρόμοιες, όμως σημασία έχει πώς ξεχωρίζει η κάθε μία με τα δικά της κριτήρια, εστιάζοντας στις τεχνολογίες που αυτή επιθυμεί. Η προτεινόμενη εφαρμογή δίνει προτεραιότητα στην απλή, καθαρή λειτουργικότητα (Streamlit UI) προκειμένου να υπάρχει πρόσβαση από όλους τους χρήστες, αλλά και στην εξατομικευμένη ευελιξία και ασφάλεια που αναμένουν με τη χρήση της.

6. Περιγραφή της λύσης

Η προτεινόμενη λύση που προβάλλεται στη συγκεκριμένη πτυχιακή εργασία είναι μια διαδικτυακή εφαρμογή (Web Application) που χρησιμοποιείται προκειμένου να διευκολύνει και να αποδεσμεύσει τους χρήστες (HR Managers) από το πρόβλημα της διαχείρισης και αξιολόγησης βιογραφικών σημειωμάτων. Τα ζητήματα της καθημερινότητας ενός οργανισμού, όπως είναι και αυτό της επιλογής προσωπικού, συχνά απαιτούν περισσότερη από την αναμενόμενη ενέργεια, οδηγώντας τα στελέχη σε πνευματική κόπωση (*decision fatigue*). Η εφαρμογή υλοποιήθηκε με γνώμονα τις πραγματικές ανάγκες του χρήστη για την εξάλειψη των προβλημάτων αυτών και διαφοροποιείται από παρεμφερείς εφαρμογές, όπου η εστίαση συχνά μετατοπίζεται σε δευτερεύουσες λειτουργίες και περίπλοκα εταιρικά στατιστικά.

Η υλοποίηση της εφαρμογής έγινε εξ ολοκλήρου στη γλώσσα προγραμματισμού Python, τόσο για τη δομή και τη διαχείριση των δεδομένων όσο και για την ενσωμάτωση των μοντέλων τεχνητής νοημοσύνης, ενώ για τη σχεδίαση του γραφικού περιβάλλοντος χρησιμοποιήθηκε το πλαίσιο Streamlit. Η αρχιτεκτονική του συστήματος αξιοποιεί τις τεχνολογίες ChromaDB για τη διανυσματική αποθήκευση, Ollama για την τοπική εκτέλεση του μοντέλου Phi-3, και LangChain για τον συντονισμό της πληροφορίας, διασφαλίζοντας την ταχύτητα, τη σημασιολογική ακρίβεια και κυρίως την ασφάλεια των ευαίσθητων προσωπικών δεδομένων. Με την ενσωμάτωση αυτών των τεχνολογιών διασφαλίζεται η άμεση αλληλεπίδραση με τον χρήστη, η πλήρης συμμόρφωση με τον GDPR, καθώς και η μελλοντική επεκτασιμότητα σε ένα σταθερό περιβάλλον.

Πιο συγκεκριμένα, ο χρήστης έχει τη δυνατότητα για προσθήκη των βιογραφικών σημειωμάτων σε πολλαπλές μορφές αρχείων (PDF, JPG, PNG, ZIP) προκειμένου να έχει πλήρη εικόνα για τη δεξαμενή των υποψηφίων και τις επιλογές του. Προσφέρει, επίσης, εύκολη αναζήτηση, επεξεργασία μέσω OCR (EasyOCR) και επισκόπηση των στοιχείων με βάση τις απαιτήσεις της θέσης εργασίας. Στις λειτουργίες που προσφέρει η εφαρμογή είναι και η δημιουργία προσωποποιημένων προτάσεων και αξιολογήσεων, όπου με τη χρήση δεδομένων όπως η περιγραφή της θέσης εργασίας στην αγγλική γλώσσα, προβάλλονται ως αποτέλεσμα πρακτικές και ρεαλιστικές ιδέες στον χρήστη.

Οι προτάσεις αυτές δεν προκύπτουν αυθαίρετα, αλλά προέρχονται από την ταξινόμηση και φθίνουσα κατάταξη των καλύτερων υποψηφίων βάσει σημασιολογικής συνάφειας (Score) και στηρίζονται στις προσωπικές επιλογές και προτιμήσεις του χρήστη, εξατομικεύοντας την εμπειρία του. Μια ακόμα λειτουργία της εφαρμογής είναι η δυνατότητα αυτόματης παραγωγής μιας ποιοτικής σύνοψης για κάθε βιογραφικό, καθώς και τριών στοχευμένων ερωτήσεων συνέντευξης που εστιάζουν στα κενά του υποψηφίου, έτσι ώστε να είναι άμεσα προσβάσιμα για μελλοντική χρήση στη συνέντευξη. Όλα τα παραπάνω, σε συνδυασμό και με τον βασικό στόχο της εφαρμογής, την έμφαση στην ευχρηστία, την επεξηγησιμότητα και την τοπική ασφάλεια των δεδομένων, καταδεικνύουν τα ιδιαίτερα χαρακτηριστικά της λύσης.

Με αυτό τον τρόπο, η εφαρμογή συμβάλλει στην καλύτερη οργάνωση της καθημερινότητας του τμήματος HR, μειώνοντας δραστικά τον χρόνο που δαπνάζεται για τη χειροκίνητη διαλογή βιογραφικών. Παρά τις ομοιότητες που υπάρχουν ανάμεσα στις διάφορες εφαρμογές της αγοράς, η προτεινόμενη λύση ξεχωρίζει γιατί δίνει προτεραιότητα στην απλή λειτουργικότητα, στην τοπική ιδιωτικότητα και στην ευελιξία που περιμένουν οι χρήστες.

7. Ανάλυση απαιτήσεων

Στην ενότητα αυτή παρουσιάζονται οι βασικές λειτουργίες της εφαρμογής, οι απαιτήσεις για την εγκατάσταση και τη χρήση της, καθώς και οι περιορισμοί που προκύπτουν από τον σχεδιασμό και την υλοποίησή της.

7.1. Ρόλοι χρηστών και λειτουργίες

Η εφαρμογή απευθύνεται σε χρήστες που ασχολούνται με τη διαδικασία αξιολόγησης και επιλογής προσωπικού, όπως υπεύθυνοι ανθρώπινου δυναμικού (HR Managers) ή υπεύθυνοι προσλήψεων. Δεν υποστηρίζεται σύστημα αυθεντικοποίησης ή διαχωρισμός ρόλων χρηστών, καθώς όλοι οι χρήστες έχουν πρόσβαση στις ίδιες λειτουργίες της εφαρμογής. Η αλληλεπίδραση πραγματοποιείται μέσω ενός ενιαίου περιβάλλοντος εργασίας, το οποίο επιτρέπει την εισαγωγή δεδομένων και την προβολή των αποτελεσμάτων της αξιολόγησης.

Οι βασικές λειτουργίες που παρέχει η εφαρμογή είναι οι ακόλουθες:

- **Εισαγωγή περιγραφής θέσης εργασίας:** Ο χρήστης εισάγει την περιγραφή της διαθέσιμης θέσης εργασίας (Job Description) στο αντίστοιχο πεδίο της εφαρμογής.
- **Μεταφόρτωση βιογραφικών σημειωμάτων:** Υποστηρίζεται η εισαγωγή βιογραφικών σε μορφές PDF, JPG, PNG και ZIP.
- **Αυτόματη εξαγωγή κειμένου:** Τα αρχεία εικόνας επεξεργάζονται μέσω της τεχνολογίας EasyOCR ώστε το περιεχόμενό τους να μετατραπεί σε επεξεργάσιμο κείμενο.
- **Αξιολόγηση υποψηφίων:** Με την επιλογή της αντίστοιχης λειτουργίας, το σύστημα αξιολογεί τα εισαγόμενα βιογραφικά σε σχέση με την περιγραφή της θέσης εργασίας αξιοποιώντας την αρχιτεκτονική Retrieval-Augmented Generation (RAG).
- **Σημασιολογική αναζήτηση:** Η εφαρμογή αναζητά τα πιο σχετικά βιογραφικά που υπάρχουν ήδη αποθηκευμένα στη διανυσματική βάση δεδομένων ChromaDB και, αφού αφαιρέσει τυχόν διπλότυπες εγγραφές, προωθεί τους τρεις επικρατέστερους υποψηφίους προς αξιολόγηση.
- **Παράλληλη παρουσίαση αποτελεσμάτων:** Τα αποτελέσματα εμφανίζονται σε δύο στήλες. Στην αριστερή πλευρά παρουσιάζονται οι αξιολογήσεις των νέων βιογραφικών που εισήγαγε ο χρήστης, ενώ στη δεξιά πλευρά εμφανίζονται οι καταλληλότεροι υποψήφιοι που ανακτήθηκαν από τη βάση δεδομένων.
- **Παρουσίαση αποτελεσμάτων αξιολόγησης:** Για κάθε υποψήφιο εμφανίζεται το ποσοστό αντιστοίχισης με τη θέση εργασίας, συνοδευόμενο από σύντομη περιγραφή των βασικών χαρακτηριστικών του.
- **Αυτόματη δημιουργία ερωτήσεων συνέντευξης:** Για κάθε αξιολογημένο υποψήφιο δημιουργούνται τρεις ενδεικτικές ερωτήσεις συνέντευξης, οι οποίες βασίζονται στο περιεχόμενο του βιογραφικού και στις απαιτήσεις της θέσης.

7.2. Απαιτήσεις εγκατάστασης και λειτουργίας

Για την ορθή εγκατάσταση και λειτουργία της εφαρμογής απαιτούνται τα ακόλουθα:

Περιβάλλον Python: Η εφαρμογή έχει αναπτυχθεί στη γλώσσα Python και απαιτεί εγκατεστημένη έκδοση Python συμβατή με τις χρησιμοποιούμενες βιβλιοθήκες.

Απαραίτητες βιβλιοθήκες: Για την εκτέλεση της εφαρμογής απαιτείται η εγκατάσταση των βιβλιοθηκών Streamlit, LangChain, ChromaDB, EasyOCR, Ollama και των υπόλοιπων εξαρτήσεων του έργου.

Τοπική εγκατάσταση Ollama: Η εφαρμογή απαιτεί εγκατεστημένη την πλατφόρμα Ollama, μέσω της οποίας εκτελείται το γλωσσικό μοντέλο Phi-3 τοπικά στον υπολογιστή.

Τοπική διανυσματική βάση δεδομένων: Η ChromaDB πρέπει να έχει δημιουργηθεί και να περιέχει τα απαραίτητα embeddings των βιογραφικών που χρησιμοποιούνται για τη σημασιολογική αναζήτηση.

Δικαιώματα πρόσβασης στα αρχεία: Ο χρήστης πρέπει να διαθέτει δικαιώματα ανάγνωσης των αρχείων που εισάγει, καθώς και πρόσβαση στους φακέλους όπου αποθηκεύονται τα δεδομένα της εφαρμογής.

7.3. Απαιτήσεις χρήσης

Για την αποτελεσματική χρήση της εφαρμογής ισχύουν οι παρακάτω προϋποθέσεις:

Περιγραφή θέσης εργασίας: Η περιγραφή της θέσης εργασίας πρέπει να εισάγεται στην αγγλική γλώσσα, ώστε να επιτυγχάνεται η βέλτιστη απόδοση του γλωσσικού μοντέλου κατά τη διαδικασία αξιολόγησης.

Διαθέσιμα βιογραφικά: Τα βιογραφικά πρέπει να βρίσκονται αποθηκευμένα τοπικά στον υπολογιστή του χρήστη σε μορφή PDF, JPG, PNG ή ZIP πριν από τη μεταφόρτωσή τους στην εφαρμογή.

Βάση δεδομένων βιογραφικών: Για τη λειτουργία της σημασιολογικής αναζήτησης απαιτείται να υπάρχει ήδη διανυσματική βάση δεδομένων ChromaDB, η οποία περιλαμβάνει τα διαθέσιμα βιογραφικά που χρησιμοποιούνται για την ανάκτηση των καταλληλότερων υποψηφίων.

7.4. Περιορισμοί

Η εφαρμογή παρουσιάζει ορισμένους περιορισμούς που σχετίζονται με την υλοποίηση και τις απαιτήσεις εκτέλεσής της.

Υπολογιστικοί πόροι: Η απόδοση της εφαρμογής εξαρτάται από τους διαθέσιμους πόρους του υπολογιστή (κυρίως CPU και RAM), καθώς όλες οι διαδικασίες επεξεργασίας πραγματοποιούνται τοπικά.

Μέγεθος υποσυνόλου βιογραφικών: Για λόγους υπολογιστικής αποδοτικότητας και μείωσης του χρόνου απόκρισης της εφαρμογής, στην τρέχουσα υλοποίηση χρησιμοποιείται υποσύνολο 50 βιογραφικών από το συνολικό dataset των 962 βιογραφικών. Η επιλογή αυτή αφορά την εκτέλεση της εφαρμογής σε πραγματικό χρόνο και μπορεί να τροποποιηθεί ανάλογα με τους διαθέσιμους υπολογιστικούς πόρους και τις απαιτήσεις της εκάστοτε χρήσης.

Τοπική εκτέλεση μοντέλου: Για την αξιολόγηση των υποψηφίων απαιτείται η λειτουργία της πλατφόρμας Ollama και η διαθεσιμότητα του μοντέλου Phi-3 στον υπολογιστή όπου εκτελείται η εφαρμογή.

Τοπική επεξεργασία δεδομένων: Η εφαρμογή έχει σχεδιαστεί ώστε η επεξεργασία των βιογραφικών να πραγματοποιείται τοπικά, χωρίς χρήση εξωτερικών υπηρεσιών τεχνητής νοημοσύνης, γεγονός που αυξάνει την ιδιωτικότητα των δεδομένων αλλά καθιστά την απόδοση εξαρτώμενη από τις δυνατότητες του τοπικού συστήματος.

Υποστηριζόμενες μορφές αρχείων: Η εφαρμογή υποστηρίζει μόνο αρχεία PDF, JPG, PNG και ZIP. Άλλοι τύποι αρχείων δεν μπορούν να υποβληθούν προς επεξεργασία.

8. Σχεδιασμός της εφαρμογής

8.1. Αρχιτεκτονική Συστήματος

Η προτεινόμενη εφαρμογή ακολουθεί μια local-first (τοπικής εκτέλεσης) αρχιτεκτονική, αξιοποιώντας πλήρως ένα προηγμένο οικοσύστημα εργαλείων Τεχνητής Νοημοσύνης και διαχείρισης δεδομένων για την επεξεργασία και την αξιολόγηση βιογραφικών σημειωμάτων. Η δομή του συστήματος έχει σχεδιαστεί έτσι ώστε να προσφέρει μια απρόσκοπτη εμπειρία χρήστη, διασφαλίζοντας παράλληλα ότι τα ευαίσθητα προσωπικά δεδομένα των υποψηφίων παραμένουν ιδιωτικά και προστατευμένα εντός των υπολογιστικών υποδομών του οργανισμού σε πραγματικό χρόνο.

Η αρχιτεκτονική της εφαρμογής αποτελείται από τα εξής βασικά μέρη:

- **Web Client (Streamlit UI / Python):** Αποτελεί το περιβάλλον αλληλεπίδρασης του χρήστη (HR Manager). Διαχειρίζεται τη μεταφόρτωση ενός αρχείου (PDF, εικόνα) ή ενός συμπιεσμένου αρχείου ZIP για μαζική εισαγωγή, την εισαγωγή της περιγραφής της θέσης εργασίας και την side-by-side (παράλληλη) προβολή της κατάταξης των υποψηφίων, εξασφαλίζοντας μια δομημένη πλοήγηση και μια διεπαφή όσο το δυνατόν φιλικότερη προς τον χρήστη.
- **Πλαίσιο Ενvorχήστρωσης (LangChain):** Λειτουργεί ως η κεντρική υποδομή που συνδέει και συντονίζει όλα τα επιμέρους υποσυστήματα. Αναλαμβάνει τη ροή της πληροφορίας, τροφοδοτώντας τα δεδομένα των βιογραφικών από τη φάση της ανάκτησης στη φάση της παραγωγής αποτελεσμάτων από το γλωσσικό μοντέλο.
- **Μηχανή Εξαγωγής Κειμένου (EasyOCR):** Υπεύθυνη για την επεξεργασία των οπτικών εγγράφων και των αρχείων εικόνας (JPG/PNG). Μετατρέπει αυτόματα τα σκαναρισμένα βιογραφικά σε επεξεργάσιμο ψηφιακό κείμενο, ώστε να μπορούν να αναλυθούν σημασιολογικά από το σύστημα.
- **Διανυσματική Βάση Δεδομένων (ChromaDB):** Χρησιμοποιείται ως η κύρια βάση δεδομένων για τη σημασιολογική αναζήτηση (Semantic Search). Εδώ αποθηκεύονται τα βιογραφικά σημειώματα αφού πρώτα μετατραπούν σε μαθηματικά διανύσματα (Embeddings), επιτρέποντας την άμεση ανάκτηση των βέλτιστων υποψηφίων με βάση το πραγματικό νόημα των προσόντων τους και όχι απλές λέξεις-κλειδιά.
- **Τοπικό Γλωσσικό Μοντέλο (Phi-3 μέσω Ollama):** Αποτελεί τον πυρήνα της τεχνητής νοημοσύνης του συστήματος. Εκτελείται αποκλειστικά τοπικά μέσω της πλατφόρμας

Ollama και αναλαμβάνει να αναλύσει δυναμικά τα δεδομένα, να παράγει το ποσοστό συνάφειας (Score), την ποιοτική σύνοψη για κάθε υποψήφιο, καθώς και τις τρεις εξειδικευμένες ερωτήσεις συνέντευξης.

Η επιλογή αυτής της αρχιτεκτονικής διασφαλίζει ότι η εφαρμογή προσφέρει κορυφαία τεχνική υπεροχή και ακρίβεια, χωρίς να εξαρτάται από εξωτερικές υπηρεσίες Cloud, διατηρώντας τα δεδομένα του τμήματος Ανθρωπίνου Δυναμικού απόλυτα ασφαλή και συγχρονισμένα στην τοπική υποδομή.

8.2. Τεχνολογίες που Χρησιμοποιήθηκαν

Για την υλοποίηση της εφαρμογής επιλέχθηκαν τεχνολογίες αιχμής που προσφέρουν σταθερότητα, υψηλή απόδοση και απόλυτη ασφάλεια στην επεξεργασία κειμένου, εγγράφων και δεδομένων.

8.2.1. Python & Streamlit (Web UI)

Για την ανάπτυξη του συστήματος χρησιμοποιήθηκε αποκλειστικά η γλώσσα προγραμματισμού Python [42], τόσο για τη δομή και τη διαχείριση της επιχειρησιακής λογικής των δεδομένων όσο και για την ενσωμάτωση μοντέλων Τεχνητής Νοημοσύνης (AI), αξιοποιώντας το πλούσιο οικοσύστημα βιβλιοθηκών της (LangChain, ChromaDB client, Sentence-Transformers) [35, 37, 38].

Ως περιβάλλον ανάπτυξης και σχεδίασης της διεπαφής χρησιμοποιήθηκε το πλαίσιο Streamlit [44]. Μέσω του Streamlit σχεδιάστηκε ένα σύγχρονο και καθαρό γραφικό περιβάλλον (UI), το οποίο επιτρέπει την οπτικοποίηση των αποτελεσμάτων σε διπλή παράλληλη διάταξη (Side-by-side View). Η επιλογή αυτή διευκολύνει τον χρήστη (HR Manager) στην άμεση σύγκριση και αναζήτηση των κατάλληλων υποψηφίων, προσφέροντας μια εξαιρετικά φιλική και διαδραστική εμπειρία χρήσης χωρίς την ανάγκη περίπλοκων δομών ιστότοπου.

8.2.2. Ollama & Phi-3 (Local Inference)

Η πλατφόρμα Ollama χρησιμοποιήθηκε για την υλοποίηση του συστήματος τοπικής εκτέλεσης (Local Inference) του Μεγάλου Γλωσσικού Μοντέλου Phi-3 της Microsoft. [1, 34] Στην εφαρμογή μας, η Ollama διαχειρίζεται αυτόματα τη φιλοξενία του μοντέλου, την κατανομή των υπολογιστικών πόρων (CPU/RAM) και την επεξεργασία των αιτημάτων (prompts) σε πραγματικό χρόνο [39].

Η επιλογή αυτή απαλλάσσει την εφαρμογή από την ανάγκη διαβίβασης ευαίσθητων προσωπικών δεδομένων σε εξωτερικές υπηρεσίες Cloud ή τρίτα APIs. Με αυτόν τον τρόπο, διασφαλίζεται η απόλυτη ιδιωτικότητα των βιογραφικών σημειωμάτων, καθώς όλες οι διεργασίες ανάλυσης, βαθμολόγησης και παραγωγής ερωτήσεων εκτελούνται περιορισμένα και με ασφάλεια σε τοπικό επίπεδο, εντός των υποδομών του οργανισμού.

8.2.3. ChromaDB (Vector Database)

Σε αντίθεση με τις παραδοσιακές σχεσιακές βάσεις δεδομένων, η ChromaDB αποτελεί μια διανυσματική βάση δεδομένων (Vector Database), η οποία επιτρέπει την αποθήκευση και αναζήτηση δεδομένων με βάση τη σημασιολογική τους ομοιότητα [33]. Η λειτουργία αυτή είναι ιδιαίτερα χρήσιμη στην προτεινόμενη εφαρμογή, καθώς επιτρέπει την εύρεση των βιογραφικών που σχετίζονται περισσότερο με την περιγραφή μιας θέσης εργασίας.

Στην εφαρμογή, τα βιογραφικά σημειώματα αποθηκεύονται στη ChromaDB ως έγγραφα (documents), συνοδευόμενα από τις αντίστοιχες διανυσματικές αναπαραστάσεις τους (embeddings). Τα embeddings δημιουργούνται με τη χρήση κατάλληλου μοντέλου μετασχηματισμού κειμένου και χρησιμοποιούνται για τη σημασιολογική σύγκριση μεταξύ της περιγραφής της θέσης εργασίας και των διαθέσιμων βιογραφικών.

Κατά τη διαδικασία αξιολόγησης, η περιγραφή της θέσης εργασίας μετατρέπεται επίσης σε embedding και χρησιμοποιείται ως ερώτημα αναζήτησης στη ChromaDB. Η βάση δεδομένων επιστρέφει τα βιογραφικά με τη μεγαλύτερη σημασιολογική συνάφεια, τα οποία στη συνέχεια αποστέλλονται στο γλωσσικό μοντέλο Phi-3 για την παραγωγή της τελικής αξιολόγησης, του ποσοστού αντιστοίχισης και των ερωτήσεων συνέντευξης.

Η χρήση της ChromaDB επιτρέπει την ταχεία ανάκτηση των καταλληλότερων υποψηφίων, μειώνοντας τον χρόνο αναζήτησης και βελτιώνοντας τη συνολική απόδοση της εφαρμογής, χωρίς να απαιτούνται σύνθετα ερωτήματα όπως στις συμβατικές βάσεις δεδομένων.

8.2.4. EasyOCR (Optical Character Recognition)

Λόγω της φύσης της εφαρμογής, η επεξεργασία και η εξαγωγή πληροφοριών από αρχεία εικόνων (JPG/PNG) ή σκαναρισμένων εγγράφων υψηλής ποιότητας είναι κρίσιμη. Η βιβλιοθήκη EasyOCR προσφέρει την απαραίτητη υπολογιστική υποδομή προκειμένου να αναγνωρίζει και να μετατρέπει το οπτικό περιεχόμενο των βιογραφικών σε ψηφιακό κείμενο [36, 41].

Η αναγνώριση κειμένου εκτελείται μία φορά, κατά τη μεταφόρτωση κάθε αρχείου, και το αποτέλεσμα αξιοποιείται στη συνέχεια σε μορφή απλού κειμένου. Με τον τρόπο αυτό αποφεύγεται η επαναλαμβανόμενη επεξεργασία των εικόνων και δεν απαιτείται μόνιμη αποθήκευση των αρχικών αρχείων πολυμέσων.

8.3. Περιγραφή Βάσης Δεδομένων

Η εφαρμογή χρησιμοποιεί τη διανυσματική βάση δεδομένων **ChromaDB** για την αποθήκευση και τη σημασιολογική αναζήτηση των βιογραφικών σημειωμάτων. Σε αντίθεση με τις παραδοσιακές σχεσιακές βάσεις δεδομένων, η ChromaDB οργανώνει τα δεδομένα με βάση τις διανυσματικές αναπαραστάσεις (embeddings) των εγγράφων, επιτρέποντας την ανάκτηση πληροφοριών σύμφωνα με τη σημασιολογική τους ομοιότητα και όχι μόνο με λέξεις-κλειδιά [16].

Τα βιογραφικά που χρησιμοποιούνται από την εφαρμογή εισάγονται στη βάση δεδομένων μέσω της διαδικασίας δημιουργίας της βάσης (ingestion process). Κατά τη διαδικασία αυτή, το περιεχόμενο κάθε βιογραφικού μετατρέπεται σε διανυσματική αναπαράσταση (embedding) με τη χρήση του μοντέλου **sentence-transformers/all-MiniLM-L6-v2** και αποθηκεύεται στη ChromaDB μαζί με το αντίστοιχο κείμενο και τα απαραίτητα μεταδεδομένα [43].

Κατά τη διαδικασία αξιολόγησης, η περιγραφή της θέσης εργασίας μετατρέπεται επίσης σε embedding και χρησιμοποιείται ως ερώτημα αναζήτησης στη βάση δεδομένων. Η ChromaDB εντοπίζει τα βιογραφικά με τη μεγαλύτερη σημασιολογική συνάφεια και επιστρέφει τα πιο σχετικά αποτελέσματα, τα οποία προωθούνται στο γλωσσικό μοντέλο **Phi-3** για την τελική αξιολόγηση των υποψηφίων.

Η χρήση της ChromaDB επιτρέπει γρήγορη και αποδοτική ανάκτηση των καταλληλότερων υποψηφίων, ακόμη και όταν η περιγραφή της θέσης εργασίας και τα βιογραφικά δεν περιέχουν ακριβώς τις ίδιες λέξεις. Με τον τρόπο αυτό, η εφαρμογή αξιοποιεί τη σημασιολογική πληροφορία του κειμένου και όχι μόνο τη λεκτική ομοιότητα [12, 20].

Δομή των εγγράφων στη ChromaDB

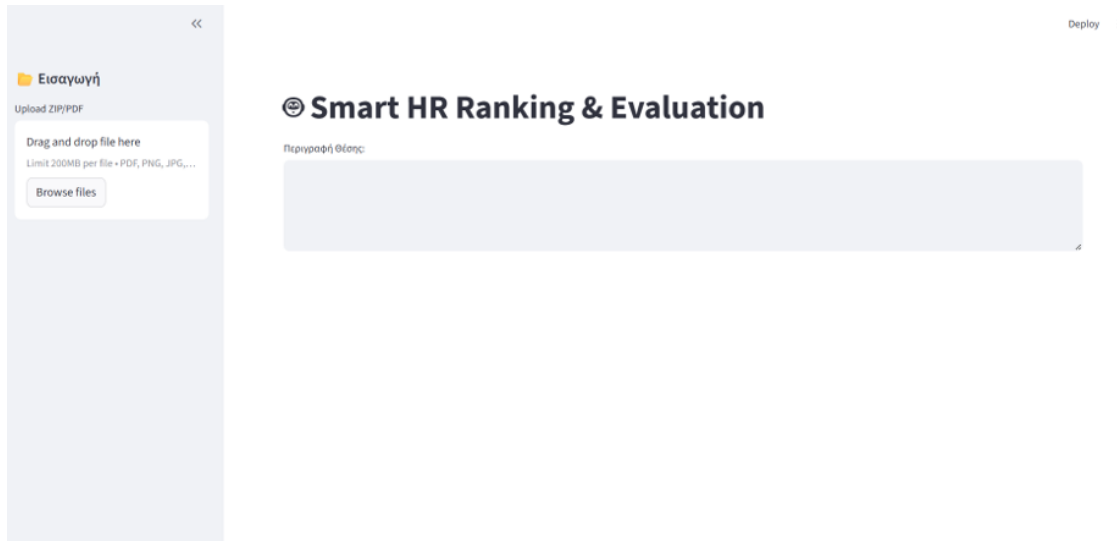
Για κάθε βιογραφικό που αποθηκεύεται στη βάση δεδομένων διατηρούνται τα παρακάτω στοιχεία:

- **Document:** Το πλήρες κείμενο του βιογραφικού σημειώματος.
- **Embedding:** Η διανυσματική αναπαράσταση του βιογραφικού, η οποία δημιουργείται από το μοντέλο **sentence-transformers/all-MiniLM-L6-v2** και χρησιμοποιείται για τη σημασιολογική αναζήτηση [23].
- **Metadata:** Συνοδευτικές πληροφορίες που σχετίζονται με το βιογραφικό και χρησιμοποιούνται για την οργάνωση και διαχείριση των δεδομένων στη βάση.

Η δομή αυτή επιτρέπει στη ChromaDB να ανακτά αποτελεσματικά τα βιογραφικά που παρουσιάζουν τη μεγαλύτερη σημασιολογική ομοιότητα με την περιγραφή της θέσης εργασίας, συμβάλλοντας στην ταχύτερη και ακριβέστερη αξιολόγηση των υποψηφίων.

9. Παραδείγματα από τη χρήση του συστήματος

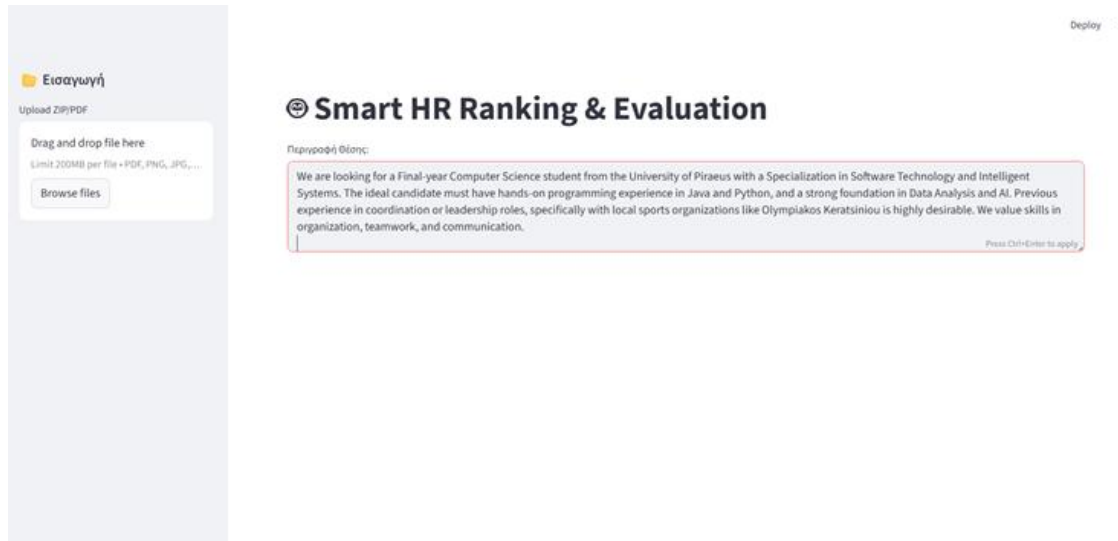
Για την πλήρη κατανόηση της λειτουργίας της πλατφόρμας, η παρούσα ενότητα παρουσιάζει τον τρόπο με τον οποίο το σύστημα αποκρίνεται ανάλογα με τις επιλογές, τα φορτωμένα δεδομένα και τις αναζητήσεις του εκάστοτε υπεύθυνου προσλήψεων (HR Manager).



Εικόνα 1: Αρχική διεπαφή της εφαρμογής Smart HR Ranking & Evaluation

Στην Εικόνα 1 παρουσιάζεται η αρχική διεπαφή της εφαρμογής **Smart HR Ranking & Evaluation**, η οποία αναπτύχθηκε για την αυτοματοποιημένη αξιολόγηση και κατάταξη βιογραφικών σημειωμάτων με τη χρήση Μεγάλων Γλωσσικών Μοντέλων (Large Language Models - LLMs) και της αρχιτεκτονικής **Retrieval Augmented Generation (RAG)**. Η εφαρμογή έχει σχεδιαστεί με στόχο να παρέχει ένα απλό και εύχρηστο περιβάλλον, μέσω του οποίου ο υπεύθυνος ανθρώπινου δυναμικού μπορεί να πραγματοποιήσει ολόκληρη τη διαδικασία αξιολόγησης των υποψηφίων χωρίς να απαιτούνται εξειδικευμένες τεχνικές γνώσεις. Η κύρια λειτουργία της βασίζεται στη σύγκριση των απαιτήσεων μιας θέσης εργασίας με τα προσόντα που περιγράφονται στα διαθέσιμα βιογραφικά, ώστε να υπολογιστεί ο βαθμός συνάφειας κάθε υποψηφίου.

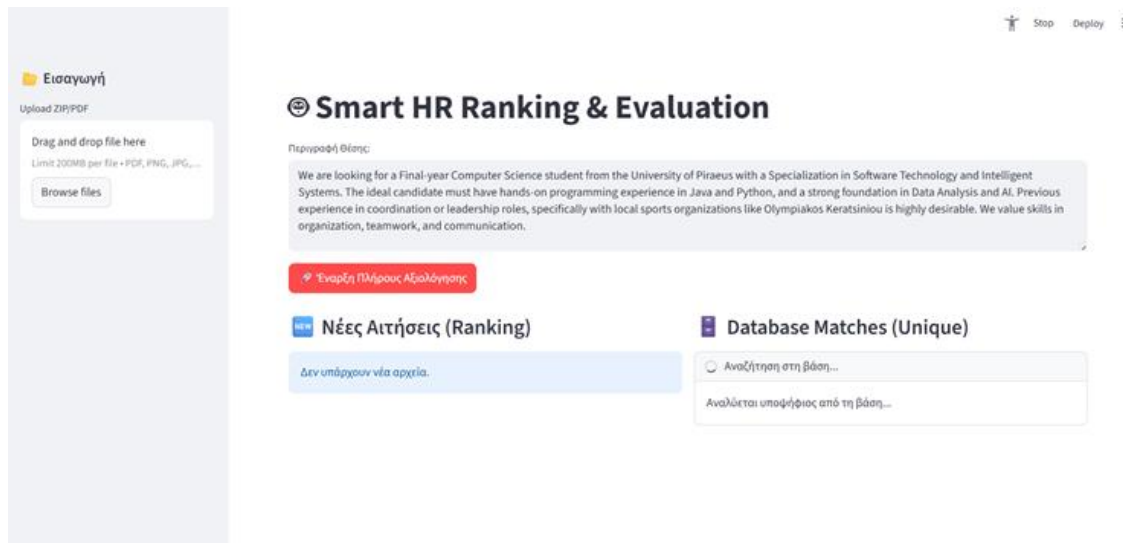
Στο αριστερό τμήμα της οθόνης βρίσκεται η περιοχή μεταφόρτωσης αρχείων, μέσω της οποίας ο χρήστης μπορεί να εισαγάγει νέα βιογραφικά σε μορφή **PDF**, **JPG**, **PNG** ή συμπιεσμένου αρχείου **ZIP**, επιτρέποντας την εισαγωγή ενός αρχείου κάθε φορά ή, μέσω του τελευταίου, πολλών βιογραφικών ταυτόχρονα. Στο κεντρικό τμήμα της εφαρμογής εμφανίζεται το πεδίο εισαγωγής της περιγραφής της θέσης εργασίας (**Job Description**), όπου ο χρήστης καταχωρίζει τις απαιτήσεις της θέσης στα αγγλικά. Η περιγραφή αυτή αποτελεί το βασικό ερώτημα του συστήματος και χρησιμοποιείται για την αναζήτηση των πιο σχετικών βιογραφικών στη διανυσματική βάση δεδομένων **ChromaDB**. Στη συνέχεια, το γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω του **Ollama**, αξιολογεί κάθε υποψήφιο, υπολογίζει το ποσοστό συνάφειας με τη θέση εργασίας και δημιουργεί μια σύντομη αιτιολόγηση καθώς και προτεινόμενες ερωτήσεις συνέντευξης, υποστηρίζοντας αποτελεσματικά τη διαδικασία επιλογής προσωπικού.



Εικόνα 2: Εισαγωγή της περιγραφής της θέσης εργασίας

Στην Εικόνα 2 παρουσιάζεται το στάδιο κατά το οποίο ο υπεύθυνος ανθρώπινου δυναμικού συμπληρώνει την περιγραφή της διαθέσιμης θέσης εργασίας (**Job Description**). Στο συγκεκριμένο παράδειγμα η θέση αφορά έναν **Software Engineer**, όπου αναφέρονται αναλυτικά τα απαραίτητα προσόντα, όπως η γνώση της γλώσσας προγραμματισμού **Python**, η εμπειρία σε τεχνολογίες υπολογιστικού νέφους (**AWS**), η εξοικείωση με βάσεις δεδομένων **SQL**, καθώς και δεξιότητες συνεργασίας και επίλυσης προβλημάτων. Η περιγραφή αυτή αποτελεί το βασικό σημείο αναφοράς της διαδικασίας αξιολόγησης, καθώς καθορίζει τα κριτήρια με τα οποία θα συγκριθούν όλα τα διαθέσιμα βιογραφικά.

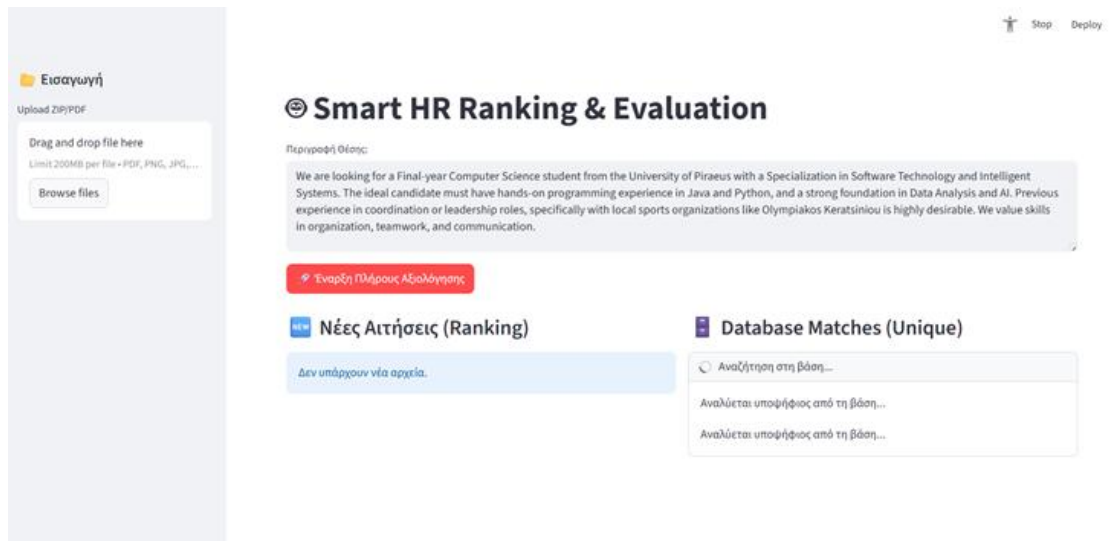
Μετά την εισαγωγή της περιγραφής, το κείμενο χρησιμοποιείται ως είσοδος του συστήματος RAG για την αναζήτηση των πιο σχετικών βιογραφικών στη διανυσματική βάση δεδομένων **ChromaDB**. Κάθε βιογραφικό έχει ήδη μετατραπεί σε διανυσματική αναπαράσταση (embedding), γεγονός που επιτρέπει την αναζήτηση με βάση τη σημασιολογική ομοιότητα και όχι αποκλειστικά την ταύτιση λέξεων-κλειδιών. Με τον τρόπο αυτό, το σύστημα μπορεί να εντοπίσει υποψηφίους που διαθέτουν τις απαιτούμενες γνώσεις και εμπειρία ακόμη και όταν χρησιμοποιείται διαφορετική ορολογία στο βιογραφικό τους. Η σωστή και αναλυτική περιγραφή της θέσης εργασίας συμβάλλει σημαντικά στη βελτίωση της ακρίβειας των αποτελεσμάτων και στην αποτελεσματικότερη κατάταξη των υποψηφίων.



Εικόνα 3: Έναρξη της διαδικασίας αξιολόγησης των βιογραφικών

Στην Εικόνα 3 παρουσιάζεται το στάδιο κατά το οποίο ο χρήστης επιλέγει την έναρξη της διαδικασίας αξιολόγησης, πατώντας το κουμπί «Έναρξη Πλήρους Αξιολόγησης». Με την ενέργεια αυτή, η εφαρμογή ξεκινά την εκτέλεση όλων των απαραίτητων διαδικασιών που απαιτούνται για την ανάλυση των βιογραφικών και την αντιστοίχισή τους με την περιγραφή της θέσης εργασίας που έχει εισαχθεί. Παράλληλα, εμφανίζονται ενημερωτικά μηνύματα («Αξιολόγηση νέων βιογραφικών...», «Αναζήτηση στη βάση...»), τα οποία πληροφορούν τον χρήστη ότι η εφαρμογή βρίσκεται σε κατάσταση επεξεργασίας και ότι η διαδικασία εκτελείται κανονικά.

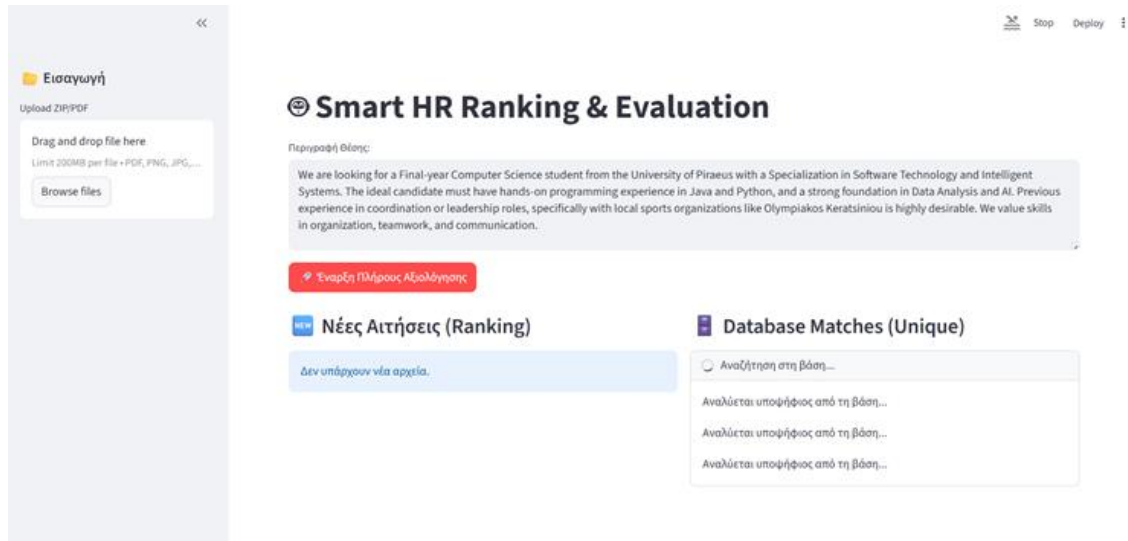
Κατά τη διάρκεια αυτού του σταδίου ενεργοποιείται η αρχιτεκτονική **RAG (Retrieval Augmented Generation)**, η οποία αρχικά πραγματοποιεί αναζήτηση των πιο σχετικών βιογραφικών στη βάση δεδομένων **ChromaDB** και στη συνέχεια προωθεί τα αποτελέσματα στο γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω του **Ollama**. Το μοντέλο αναλύει το περιεχόμενο κάθε βιογραφικού σε συνδυασμό με τις απαιτήσεις της θέσης εργασίας και υπολογίζει το ποσοστό συνάφειας κάθε υποψηφίου. Με τον τρόπο αυτό, η αξιολόγηση πραγματοποιείται αυτόματα, χωρίς την ανάγκη χειροκίνητης εξέτασης όλων των βιογραφικών, μειώνοντας σημαντικά τον απαιτούμενο χρόνο και διευκολύνοντας τη διαδικασία επιλογής προσωπικού.



Εικόνα 4: Ανάκτηση των σχετικών βιογραφικών από τη βάση δεδομένων

Στην Εικόνα 4 παρουσιάζεται το στάδιο κατά το οποίο η εφαρμογή πραγματοποιεί την ανάκτηση των πιο σχετικών βιογραφικών από τη διανυσματική βάση δεδομένων **ChromaDB**. Μετά την εισαγωγή της περιγραφής της θέσης εργασίας και την έναρξη της αξιολόγησης, το σύστημα συγκρίνει το περιεχόμενο της περιγραφής με τα διανυσματικά δεδομένα (embeddings) που έχουν δημιουργηθεί για κάθε βιογραφικό. Η διαδικασία αυτή πραγματοποιείται μέσω της αρχιτεκτονικής **Retrieval Augmented Generation (RAG)**, η οποία επιτρέπει την αναζήτηση των υποψηφίων με τη μεγαλύτερη σημασιολογική συνάφεια, ανεξάρτητα από τη χρήση ακριβώς των ίδιων λέξεων ή φράσεων.

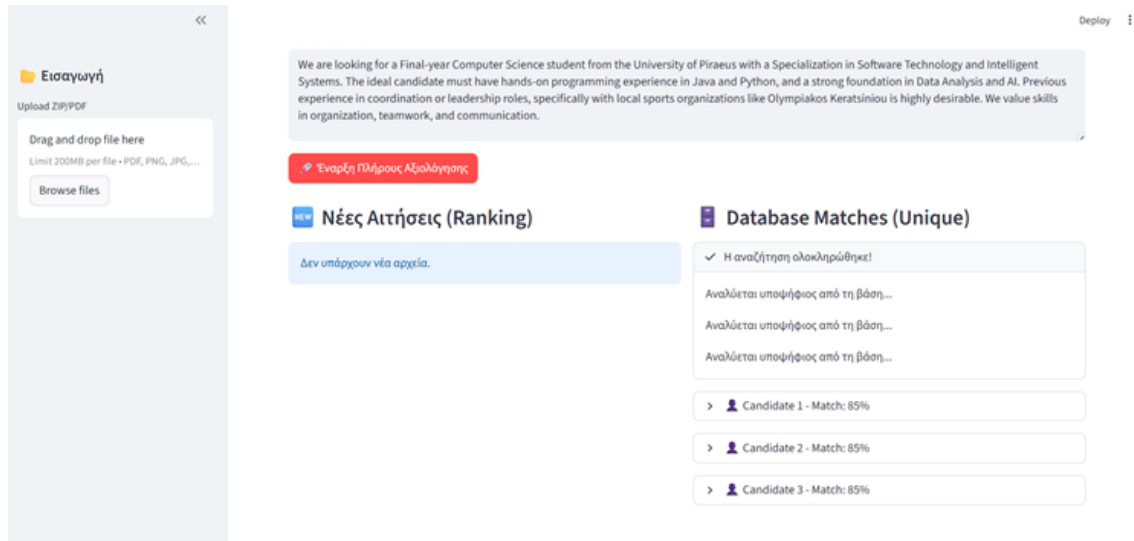
Παράλληλα, η εφαρμογή εμφανίζει στον χρήστη την εξέλιξη της διαδικασίας μέσω διαδοχικών ενημερωτικών μηνυμάτων, παρέχοντας άμεση εικόνα για την κατάσταση της αξιολόγησης. Με την ολοκλήρωση της ανάκτησης, τα βιογραφικά που παρουσιάζουν τη μεγαλύτερη συνάφεια προωθούνται στο γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω του **Ollama**. Το μοντέλο αναλαμβάνει να πραγματοποιήσει την τελική αξιολόγηση κάθε υποψηφίου, υπολογίζοντας το ποσοστό συνάφειας, δημιουργώντας μια σύντομη αιτιολόγηση της βαθμολογίας και προτείνοντας κατάλληλες ερωτήσεις που μπορούν να χρησιμοποιηθούν σε μια ενδεχόμενη συνέντευξη. Έτσι, η εφαρμογή συνδυάζει την ταχύτητα της διανυσματικής αναζήτησης με τις δυνατότητες κατανόησης και παραγωγής κειμένου που προσφέρει ένα σύγχρονο γλωσσικό μοντέλο.



Εικόνα 5: Ολοκλήρωση της διαδικασίας ανάκτησης και προετοιμασία της αξιολόγησης

Στην Εικόνα 5 παρουσιάζεται το στάδιο κατά το οποίο ολοκληρώνεται η διαδικασία ανάκτησης των σχετικών βιογραφικών και το σύστημα προχωρά στην τελική φάση της αξιολόγησης. Αφού έχουν εντοπιστεί οι υποψήφιοι με τη μεγαλύτερη συνάφεια προς την περιγραφή της θέσης εργασίας, η εφαρμογή εμφανίζει την πρόοδο της επεξεργασίας, ενημερώνοντας τον χρήστη ότι η διαδικασία εξελίσσεται κανονικά. Η λειτουργία αυτή προσφέρει άμεση πληροφόρηση για την κατάσταση του συστήματος, ιδιαίτερα όταν αξιολογείται μεγάλος αριθμός βιογραφικών, βελτιώνοντας τη συνολική εμπειρία χρήσης.

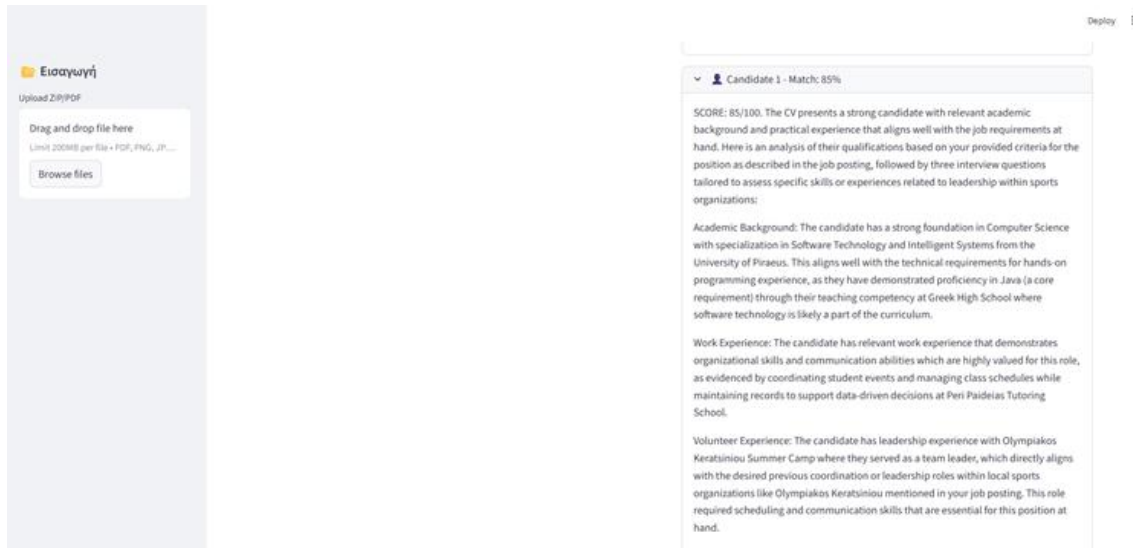
Στο συγκεκριμένο στάδιο, τα βιογραφικά που ανακτήθηκαν από τη βάση **ChromaDB** αποστέλλονται στο γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω του **Ollama**. Το μοντέλο αναλύει το περιεχόμενο κάθε βιογραφικού σε συνδυασμό με τις απαιτήσεις της θέσης εργασίας και υπολογίζει το ποσοστό συνάφειας κάθε υποψηφίου. Παράλληλα, δημιουργεί μια σύντομη περιγραφή που αιτιολογεί τη βαθμολογία, καθώς και προτεινόμενες ερωτήσεις συνέντευξης προσαρμοσμένες στο προφίλ του κάθε υποψηφίου. Με την ολοκλήρωση αυτού του σταδίου, η εφαρμογή είναι έτοιμη να παρουσιάσει τα αποτελέσματα στον χρήστη, ταξινομημένα από τον καταλληλότερο προς τον λιγότερο κατάλληλο υποψήφιο, διευκολύνοντας σημαντικά τη διαδικασία επιλογής προσωπικού.



Εικόνα 6: Παρουσίαση των αποτελεσμάτων αξιολόγησης των υποψηφίων

Στην Εικόνα 6 παρουσιάζονται τα πρώτα αποτελέσματα της διαδικασίας αξιολόγησης, μετά την ολοκλήρωση της επεξεργασίας των βιογραφικών από το σύστημα. Στη δεξιά στήλη παρουσιάζονται τα βιογραφικά που ανακτήθηκαν από τη βάση δεδομένων ChromaDB, ενώ η αριστερή στήλη παραμένει κενή, καθώς στο συγκεκριμένο παράδειγμα δεν έχουν μεταφορτωθεί νέα βιογραφικά. Κάθε υποψήφιος συνοδεύεται από το αντίστοιχο ποσοστό συνάφειας, το οποίο εκφράζει τον βαθμό στον οποίο τα προσόντα και οι δεξιότητές του ανταποκρίνονται στις απαιτήσεις της θέσης εργασίας. Η κατάταξη πραγματοποιείται αυτόματα σε φθίνουσα σειρά, επιτρέποντας στον υπεύθυνο ανθρώπινου δυναμικού να εντοπίσει άμεσα τους καταλληλότερους υποψηφίους.

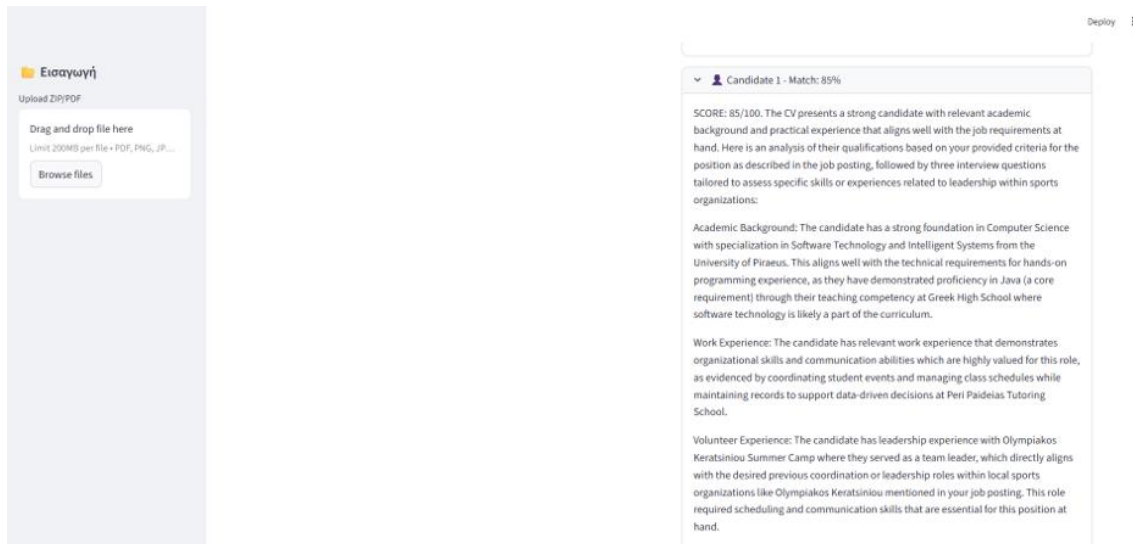
Εκτός από το ποσοστό συνάφειας, για κάθε υποψήφιο εμφανίζεται μια σύντομη περιγραφή των βασικών χαρακτηριστικών του, η οποία δημιουργείται από το γλωσσικό μοντέλο **Phi-3**. Η περιγραφή αυτή συνοψίζει τα σημαντικότερα στοιχεία του βιογραφικού και αιτιολογεί συνοπτικά τη βαθμολογία που έλαβε. Με αυτόν τον τρόπο, ο χρήστης δεν χρειάζεται να εξετάσει αναλυτικά κάθε βιογραφικό, αλλά μπορεί να αποκτήσει μια γρήγορη και ολοκληρωμένη εικόνα για όλους τους υποψηφίους, διευκολύνοντας σημαντικά τη διαδικασία επιλογής προσωπικού.



Εικόνα 7: Παρουσίαση των αποτελεσμάτων αξιολόγησης των υποψηφίων

Στην Εικόνα 7 παρουσιάζεται η ανάπτυξη μιας εγγραφής της λίστας αποτελεσμάτων. Με την επιλογή ενός υποψηφίου εμφανίζεται ολόκληρο το κείμενο της αξιολόγησης που παρήγαγε το γλωσσικό μοντέλο Phi-3, το οποίο ξεκινά με τη βαθμολογία συνάφειας και συνεχίζει με ανάλυση του ακαδημαϊκού υποβάθρου, της επαγγελματικής εμπειρίας και των υπόλοιπων στοιχείων του βιογραφικού σε σχέση με τις απαιτήσεις της θέσης εργασίας.

Η αναλυτική αυτή προβολή αποτελεί εφαρμογή των αρχών της Επεξηγήσιμης Τεχνητής Νοημοσύνης (ΧΑΙ): ο χρήστης δεν βλέπει μόνο τον αριθμητικό δείκτη, αλλά και την τεκμηρίωσή του. Με τον τρόπο αυτό μπορεί να κρίνει κατά πόσον η βαθμολογία δικαιολογείται, χωρίς να χρειάζεται να διαβάσει ολόκληρο το βιογραφικό.

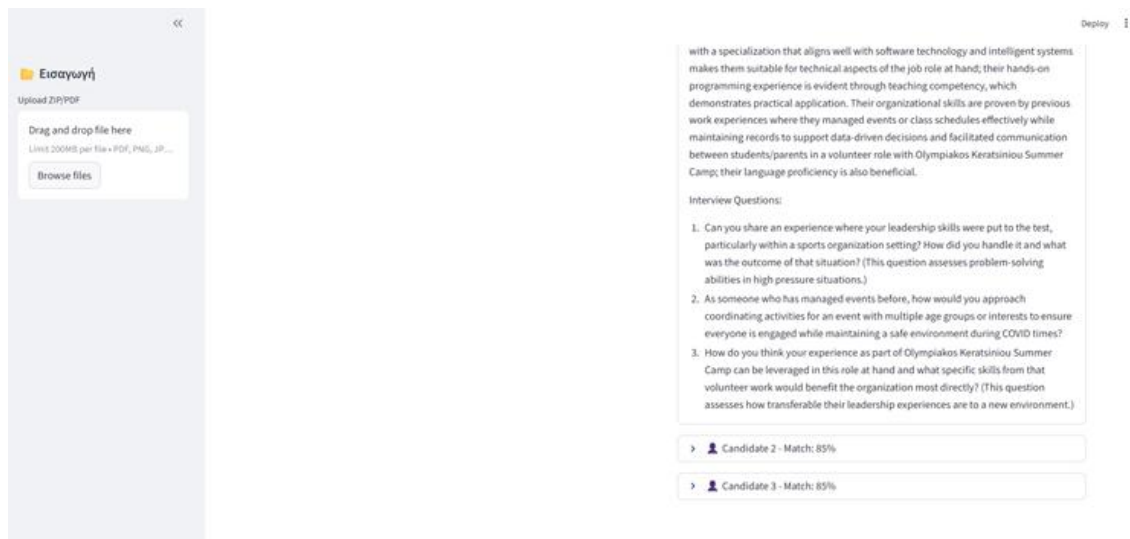


Εικόνα 8: Αναλυτική προβολή των πληροφοριών ενός υποψηφίου

Στην Εικόνα 8 παρουσιάζεται η αναλυτική προβολή ενός υποψηφίου, η οποία εμφανίζεται μετά την επιλογή του αντίστοιχου αποτελέσματος από τη λίστα αξιολόγησης. Εκτός από το συνολικό ποσοστό συνάφειας, εμφανίζεται αναλυτικότερη περιγραφή των γνώσεων, της επαγγελματικής εμπειρίας και των δεξιοτήτων που εντοπίστηκαν στο βιογραφικό. Η πληροφορία αυτή

δημιουργείται δυναμικά από το γλωσσικό μοντέλο **Phi-3**, το οποίο αναλύει το περιεχόμενο του βιογραφικού σε συνδυασμό με τις απαιτήσεις της θέσης εργασίας και παρέχει μια σύντομη αλλά ουσιαστική αιτιολόγηση της αξιολόγησης.

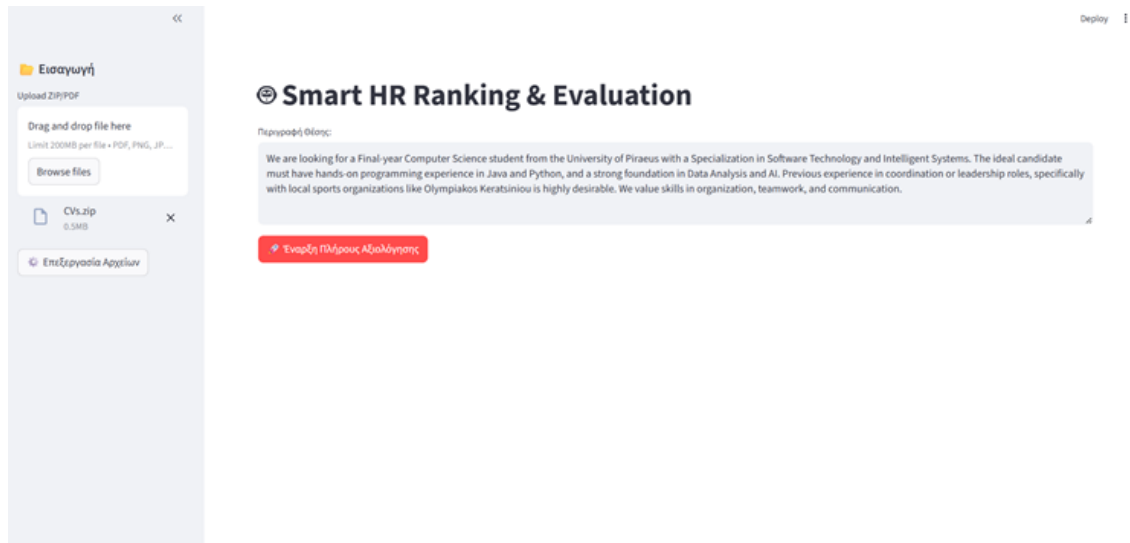
Η δυνατότητα αυτή επιτρέπει στον υπεύθυνο ανθρώπινου δυναμικού να κατανοήσει εύκολα τους λόγους για τους οποίους ένας υποψήφιος θεωρείται κατάλληλος για τη συγκεκριμένη θέση, χωρίς να χρειάζεται να διαβάσει ολόκληρο το βιογραφικό. Παράλληλα, η εφαρμογή διατηρεί τη σειρά κατάταξης όλων των υποψηφίων, δίνοντας τη δυνατότητα άμεσης σύγκρισης μεταξύ διαφορετικών βιογραφικών. Με τον τρόπο αυτό, η διαδικασία επιλογής γίνεται πιο γρήγορη, αντικειμενική και αποτελεσματική, καθώς ο χρήστης μπορεί να εστιάσει στους υποψηφίους με τη μεγαλύτερη συνάφεια και να προχωρήσει ευκολότερα στο επόμενο στάδιο της συνέντευξης.



Εικόνα 9: Δημιουργία ερωτήσεων συνέντευξης για τον υποψήφιο

Στην Εικόνα 9 παρουσιάζεται ένα από τα σημαντικότερα χαρακτηριστικά της εφαρμογής, δηλαδή η αυτόματη δημιουργία ερωτήσεων συνέντευξης για κάθε υποψήφιο. Μετά την ολοκλήρωση της αξιολόγησης, το γλωσσικό μοντέλο **Phi-3** αναλύει το περιεχόμενο του βιογραφικού και δημιουργεί τρεις εξατομικευμένες ερωτήσεις, οι οποίες σχετίζονται με τις γνώσεις, την επαγγελματική εμπειρία και τις δεξιότητες που αναφέρονται σε αυτό. Οι ερωτήσεις αυτές δεν είναι προκαθορισμένες, αλλά παράγονται δυναμικά κάθε φορά που πραγματοποιείται νέα αξιολόγηση.

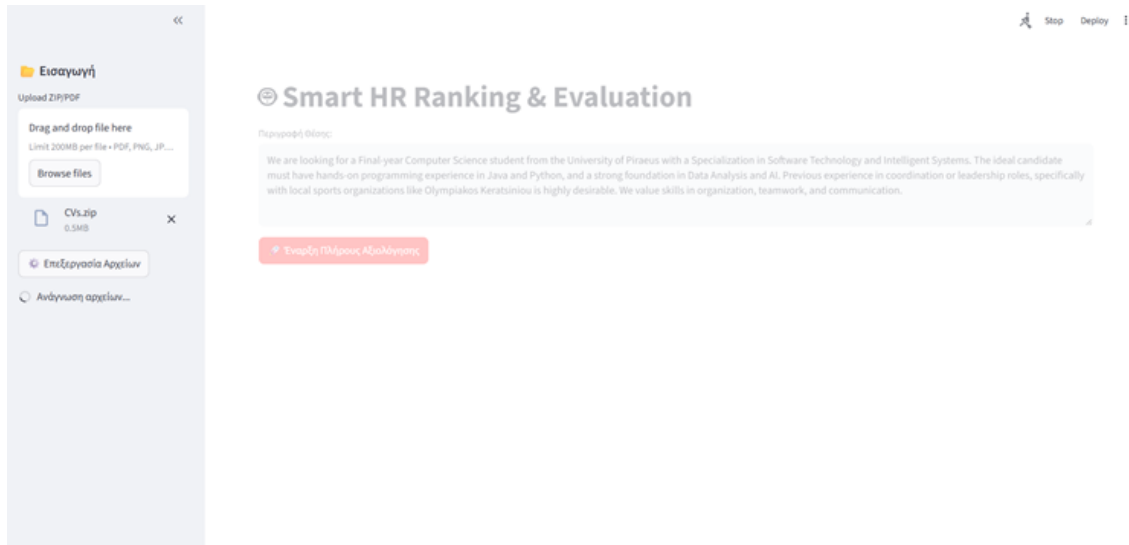
Η λειτουργία αυτή προσφέρει σημαντική υποστήριξη στον υπεύθυνο ανθρώπινου δυναμικού, καθώς μειώνει τον χρόνο προετοιμασίας της συνέντευξης και συμβάλλει στη διερεύνηση των σημαντικότερων στοιχείων του βιογραφικού. Με τον τρόπο αυτό, ο αξιολογητής μπορεί να εστιάσει σε συγκεκριμένες τεχνικές γνώσεις ή επαγγελματικές εμπειρίες του υποψηφίου, πραγματοποιώντας μια πιο στοχευμένη και αποτελεσματική διαδικασία επιλογής προσωπικού.



Εικόνα 10: Επιλογή νέων βιογραφικών προς αξιολόγηση

Στην Εικόνα 10 παρουσιάζεται η διαδικασία επιλογής νέων βιογραφικών από τον υπολογιστή του χρήστη. Η εφαρμογή επιτρέπει την εισαγωγή αρχείων σε μορφή **ZIP**, δίνοντας τη δυνατότητα μεταφόρτωσης πολλών βιογραφικών ταυτόχρονα. Η λειτουργία αυτή διευκολύνει σημαντικά τη χρήση της εφαρμογής σε περιπτώσεις όπου πρέπει να αξιολογηθεί μεγάλος αριθμός υποψηφίων, καθώς αποφεύγεται η ανάγκη επιλογής κάθε αρχείου ξεχωριστά.

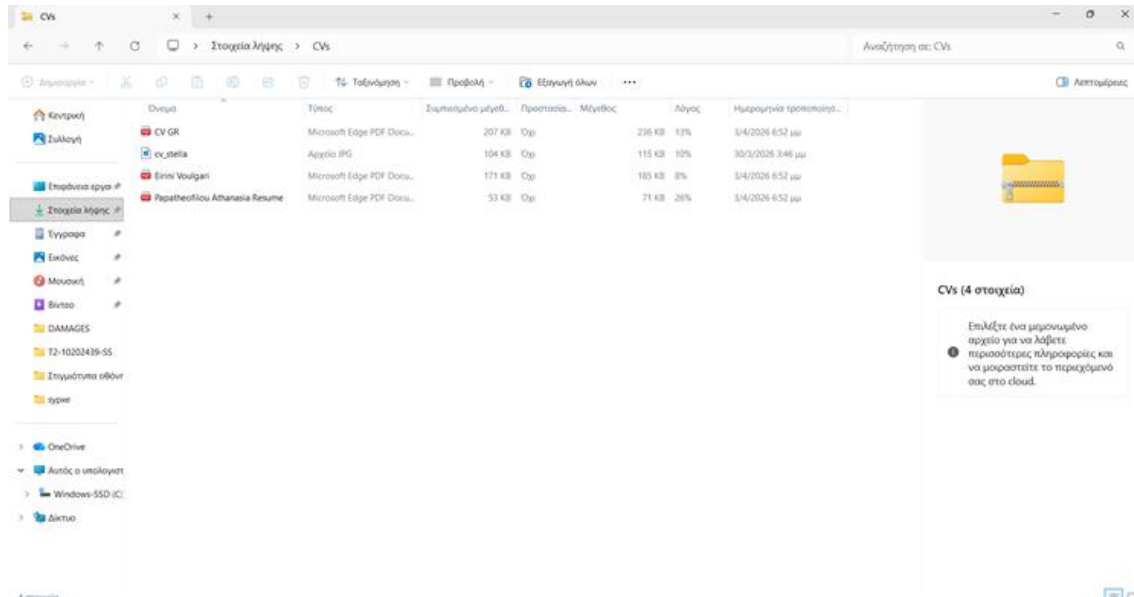
Μετά την επιλογή του συμπιεσμένου αρχείου, η εφαρμογή αναλαμβάνει αυτόματα την αποσυμπίεση και την επεξεργασία του περιεχομένου του. Στη συνέχεια, κάθε βιογραφικό προετοιμάζεται για ανάλυση, ανεξάρτητα από τη μορφή στην οποία βρίσκεται. Εφόσον υπάρχουν αρχεία εικόνας, χρησιμοποιείται η βιβλιοθήκη **EasyOCR** για την εξαγωγή του κειμένου, ενώ τα αρχεία PDF διαβάζονται απευθείας. Με τον τρόπο αυτό, όλα τα βιογραφικά μετατρέπονται σε ενιαία μορφή κειμένου, ώστε να μπορούν να αξιολογηθούν από το σύστημα.



Εικόνα 11: Εκκίνηση της επεξεργασίας των νέων βιογραφικών

Στην Εικόνα 11 παρουσιάζεται το στάδιο κατά το οποίο ξεκινά η επεξεργασία των νέων βιογραφικών που έχουν μεταφορτωθεί από τον χρήστη. Μετά την επιλογή του αρχείου, η εφαρμογή εμφανίζει σχετικό μήνυμα προόδου, ενημερώνοντας ότι η διαδικασία βρίσκεται σε εξέλιξη. Κατά τη διάρκεια αυτού του σταδίου πραγματοποιείται ο έλεγχος των αρχείων, η ανάγνωση του περιεχομένου τους και η προετοιμασία τους για την επόμενη φάση της αξιολόγησης.

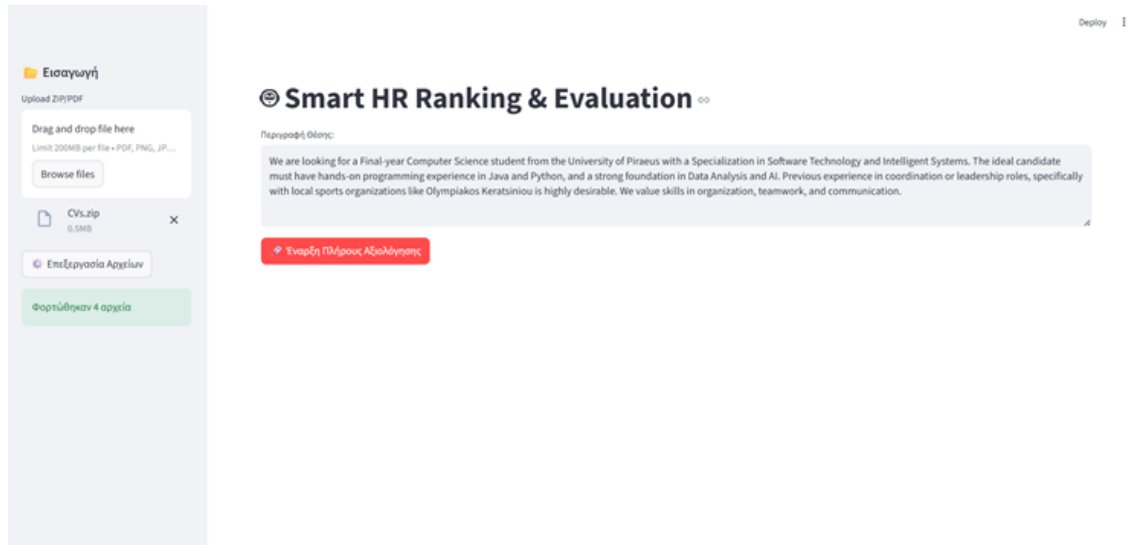
Αφού ολοκληρωθεί η επεξεργασία, το εξαγόμενο κείμενο κάθε βιογραφικού διατηρείται στην τρέχουσα συνεδρία εργασίας της εφαρμογής και προωθείται στο γλωσσικό μοντέλο **Phi-3** για αξιολόγηση ως προς την περιγραφή της θέσης εργασίας. Τα νέα βιογραφικά αξιολογούνται ανεξάρτητα από εκείνα που είναι ήδη αποθηκευμένα στη **ChromaDB** και παρουσιάζονται σε ξεχωριστή στήλη, ώστε ο υπεύθυνος προσλήψεων να τα συγκρίνει με τους υποψηφίους της βάσης. Με τον τρόπο αυτό επιτρέπεται η άμεση αξιολόγηση νέων υποψηφίων χωρίς να απαιτείται επανεκκίνηση της εφαρμογής ή προηγούμενη καταχώρισή τους στη βάση δεδομένων, η οποία πραγματοποιείται μέσω της ξεχωριστής διαδικασίας δημιουργίας της βάσης (ingestion process) που περιγράφεται στην ενότητα 8.3.



Εικόνα 12: Ολοκλήρωση της μεταφόρτωσης των νέων βιογραφικών

Στην Εικόνα 12 παρουσιάζεται η επιτυχής ολοκλήρωση της μεταφόρτωσης των νέων βιογραφικών στην εφαρμογή. Μετά την επιλογή του συμπιεσμένου αρχείου, το σύστημα ολοκληρώνει την αποσυμπίεση και την επεξεργασία όλων των εγγράφων, εμφανίζοντας σχετικό μήνυμα επιβεβαίωσης προς τον χρήστη. Με τον τρόπο αυτό διασφαλίζεται ότι όλα τα αρχεία έχουν διαβαστεί επιτυχώς και είναι πλέον διαθέσιμα για αξιολόγηση, χωρίς να απαιτείται κάποια επιπλέον ενέργεια από τον χρήστη.

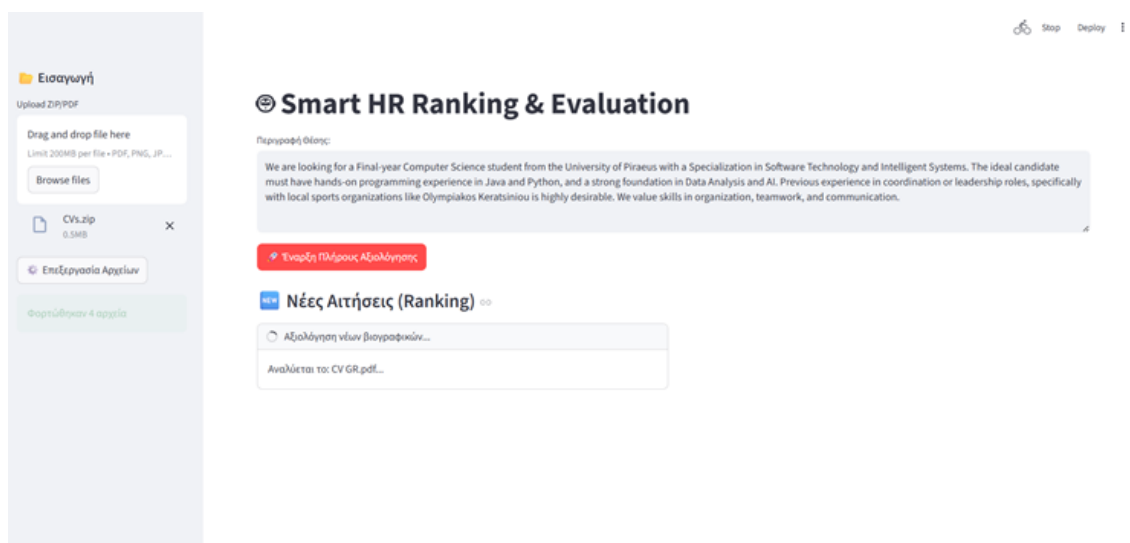
Κατά τη διάρκεια της διαδικασίας αυτής, η εφαρμογή αναγνωρίζει αυτόματα τον τύπο κάθε αρχείου και εξάγει το περιεχόμενό του. Για τα αρχεία εικόνας χρησιμοποιείται η βιβλιοθήκη **EasyOCR**, ενώ τα αρχεία PDF επεξεργάζονται απευθείας, ώστε όλα τα βιογραφικά να μετατραπούν σε ενιαία μορφή κειμένου. Με την ολοκλήρωση της μεταφόρτωσης, τα νέα δεδομένα είναι έτοιμα να συμμετάσχουν στην επόμενη φάση της αξιολόγησης μαζί με τα βιογραφικά που υπάρχουν ήδη στη βάση δεδομένων.



Εικόνα 13: Έναρξη αξιολόγησης των νέων βιογραφικών

Στην Εικόνα 13 παρουσιάζεται η έναρξη της διαδικασίας αξιολόγησης των νέων βιογραφικών που προστέθηκαν από τον χρήστη. Με την επιλογή της αντίστοιχης λειτουργίας, η εφαρμογή ξεκινά την ανάλυση των εγγράφων και τη σύγκρισή τους με την περιγραφή της θέσης εργασίας που έχει ήδη εισαχθεί. Παράλληλα, εμφανίζεται μήνυμα ενημέρωσης σχετικά με την πρόοδο της διαδικασίας, ώστε ο χρήστης να γνωρίζει ότι η αξιολόγηση εκτελείται κανονικά.

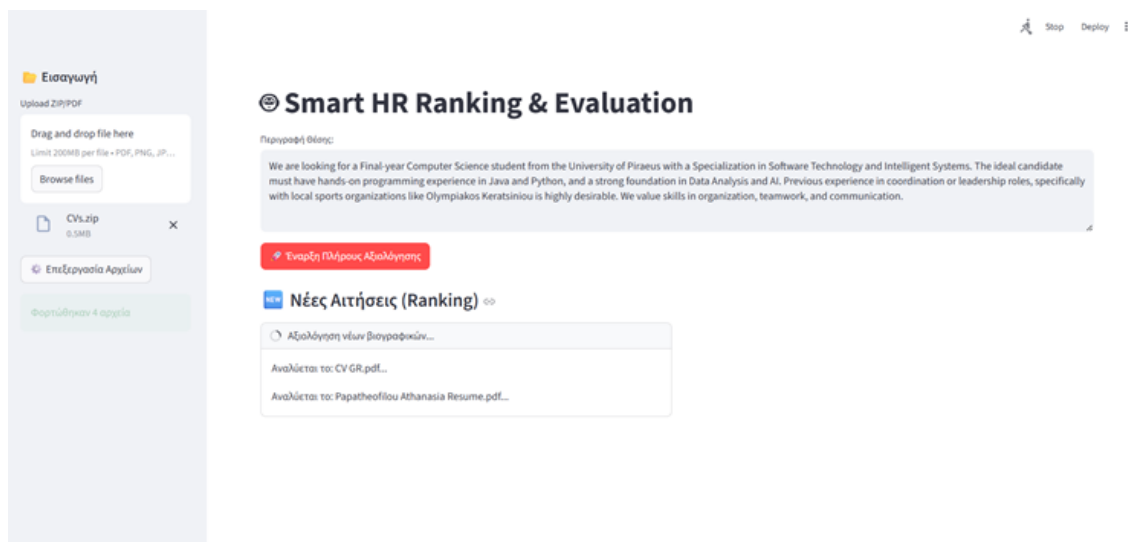
Σε αυτό το στάδιο το κείμενο κάθε νέου βιογραφικού προωθείται απευθείας, μαζί με την περιγραφή της θέσης εργασίας, στο γλωσσικό μοντέλο Phi-3, το οποίο εκτελείται τοπικά μέσω του Ollama. Το μοντέλο αναλύει κάθε νέο βιογραφικό, εντοπίζει τις σημαντικότερες δεξιότητες και εμπειρίες και υπολογίζει τον βαθμό συνάφειας κάθε υποψηφίου με τις απαιτήσεις της θέσης εργασίας.



Εικόνα 14: Πρόοδος της διαδικασίας αξιολόγησης

Στην Εικόνα 14 παρουσιάζεται η εξέλιξη της διαδικασίας αξιολόγησης των νέων βιογραφικών. Η εφαρμογή ενημερώνει συνεχώς τον χρήστη για την πρόοδο της επεξεργασίας, εμφανίζοντας σχετικό μήνυμα καθώς αναλύονται διαδοχικά τα διαθέσιμα αρχεία. Η δυνατότητα αυτή είναι ιδιαίτερα χρήσιμη όταν αξιολογείται μεγάλος αριθμός βιογραφικών, καθώς παρέχει άμεση εικόνα της κατάστασης του συστήματος και του χρόνου που απαιτείται μέχρι την ολοκλήρωση της διαδικασίας.

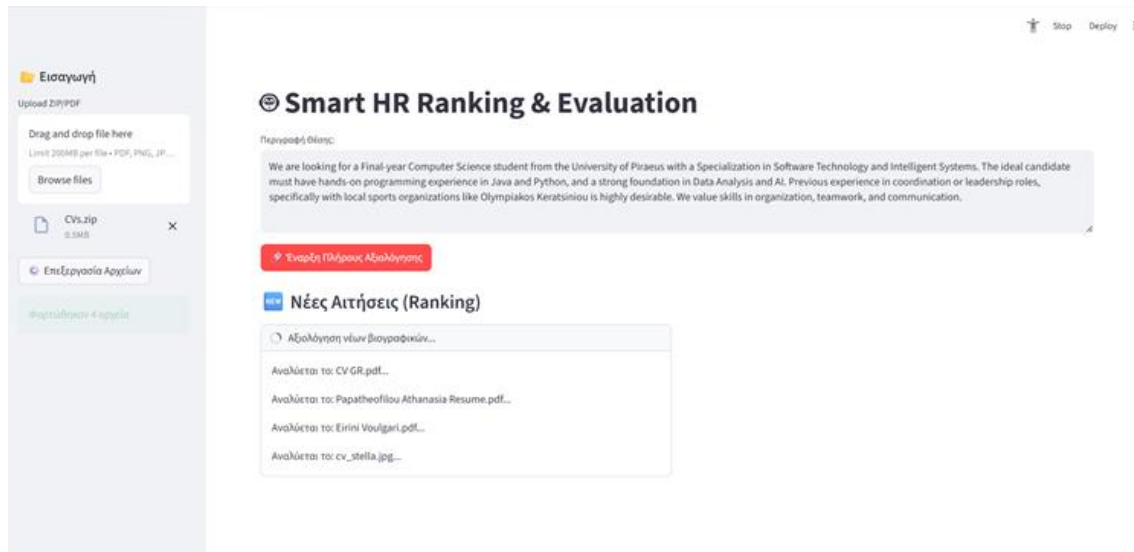
Κατά την επεξεργασία κάθε βιογραφικού, το γλωσσικό μοντέλο **Phi-3** εξετάζει το περιεχόμενο του σε συνδυασμό με την περιγραφή της θέσης εργασίας και δημιουργεί την αντίστοιχη αξιολόγηση. Παράλληλα, υπολογίζεται το ποσοστό συνάφειας, παράγεται μια σύντομη αιτιολόγηση της βαθμολογίας και προετοιμάζονται οι προτεινόμενες ερωτήσεις συνέντευξης. Με την ολοκλήρωση της διαδικασίας, όλα τα αποτελέσματα συγκεντρώνονται και παρουσιάζονται στον χρήστη ταξινομημένα από τον καταλληλότερο προς τον λιγότερο κατάλληλο υποψήφιο.



Εικόνα 15: Ολοκλήρωση της αξιολόγησης και κατάταξη των νέων υποψηφίων

Στην Εικόνα 15 παρουσιάζονται τα αποτελέσματα μετά την ολοκλήρωση της αξιολόγησης των νέων βιογραφικών που μεταφορτώθηκαν από τον χρήστη. Η εφαρμογή έχει πλέον επεξεργαστεί όλα τα διαθέσιμα αρχεία και εμφανίζει τους υποψηφίους ταξινομημένους σε φθίνουσα σειρά, με βάση το ποσοστό συνάφειας που υπολογίστηκε για κάθε έναν από αυτούς. Η κατάταξη αυτή επιτρέπει στον υπεύθυνο ανθρώπινου δυναμικού να εντοπίσει άμεσα τους καταλληλότερους υποψηφίους, χωρίς να απαιτείται η χειροκίνητη εξέταση όλων των βιογραφικών.

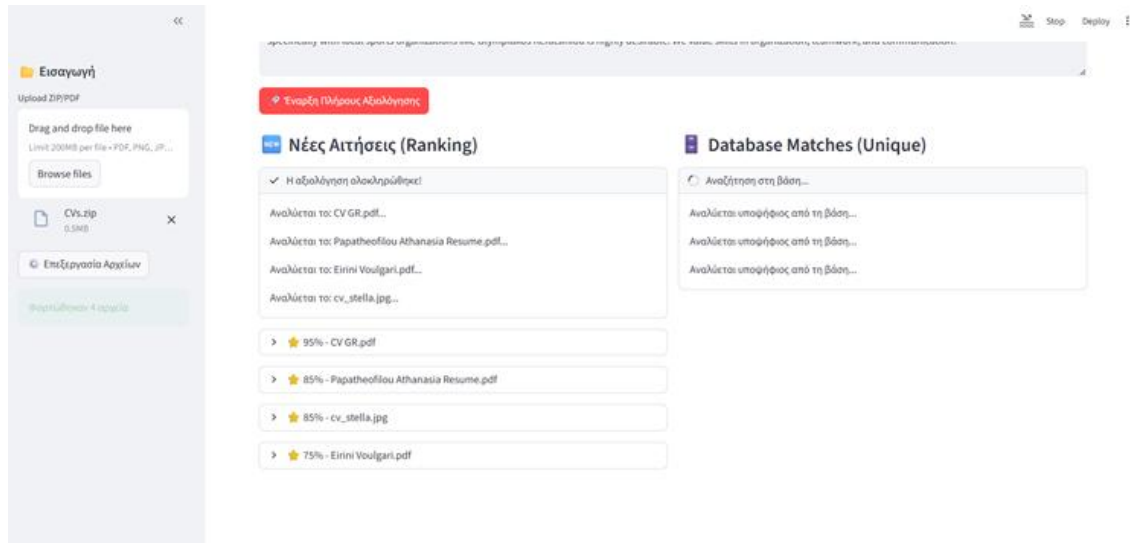
Για κάθε υποψήφιο εμφανίζεται το ποσοστό συνάφειας, καθώς και μια σύντομη περιγραφή που συνοψίζει τα βασικά χαρακτηριστικά του βιογραφικού. Η πληροφορία αυτή δημιουργείται αυτόματα από το γλωσσικό μοντέλο **Phi-3**, το οποίο συνδυάζει τα δεδομένα του βιογραφικού με τις απαιτήσεις της θέσης εργασίας. Με τον τρόπο αυτό, η εφαρμογή προσφέρει μια ολοκληρωμένη εικόνα των αποτελεσμάτων και διευκολύνει τη διαδικασία επιλογής των υποψηφίων που θα προχωρήσουν στο επόμενο στάδιο της συνέντευξης.



Εικόνα 16: Αναλυτική αξιολόγηση του καταλληλότερου υποψηφίου

Στην Εικόνα 16 παρουσιάζεται η αναλυτική αξιολόγηση του υποψηφίου που κατέλαβε την πρώτη θέση στην κατάταξη. Εκτός από το συνολικό ποσοστό συνάφειας, η εφαρμογή εμφανίζει μια αναλυτική περιγραφή των γνώσεων, της επαγγελματικής εμπειρίας και των τεχνικών δεξιοτήτων που εντοπίστηκαν στο βιογραφικό και συνέβαλαν στη συγκεκριμένη βαθμολογία. Η πληροφορία αυτή παρέχει στον υπεύθυνο ανθρώπινου δυναμικού μια συνοπτική αλλά ολοκληρωμένη εικόνα του προφίλ του υποψηφίου, χωρίς να απαιτείται η ανάγνωση ολόκληρου του βιογραφικού.

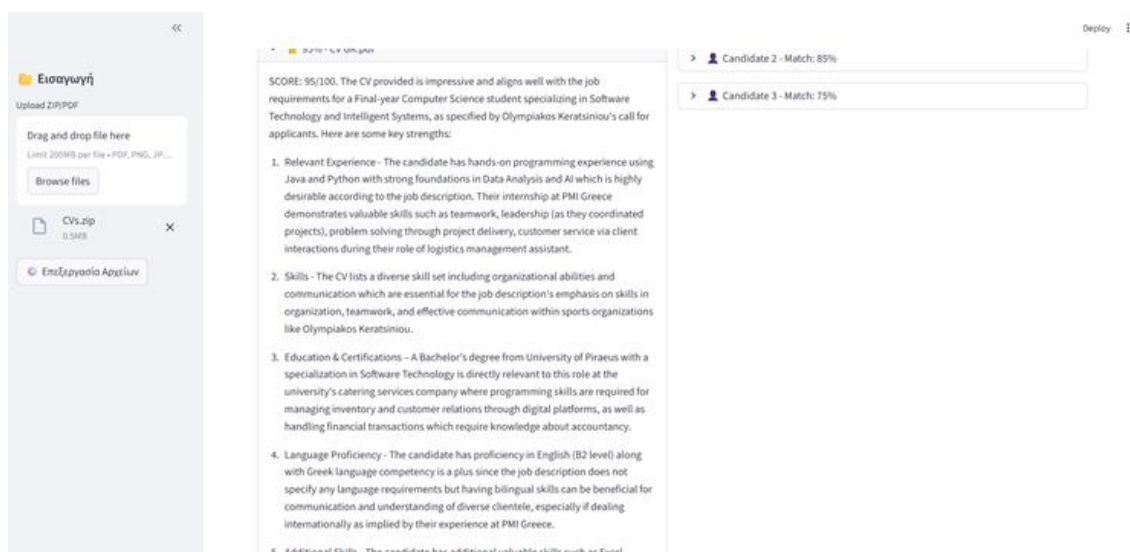
Η αξιολόγηση πραγματοποιείται από το γλωσσικό μοντέλο **Phi-3**, το οποίο λαμβάνει υπόψη τόσο τις απαιτήσεις της θέσης εργασίας όσο και το περιεχόμενο του βιογραφικού. Με βάση τη σημασιολογική αντιστοίχιση των πληροφοριών, το μοντέλο υπολογίζει το ποσοστό συνάφειας και δημιουργεί την αντίστοιχη αιτιολόγηση. Η δυνατότητα αυτή συμβάλλει στη λήψη πιο αντικειμενικών αποφάσεων, καθώς η κατάταξη βασίζεται στα πραγματικά προσόντα του υποψηφίου και όχι αποκλειστικά στη χειροκίνητη αξιολόγηση από τον χρήστη.



Εικόνα 17: Αυτόματη δημιουργία ερωτήσεων συνέντευξης για τον υποψήφιο

Στην Εικόνα 17 παρουσιάζονται οι ερωτήσεις συνέντευξης που δημιουργούνται αυτόματα για τον υποψήφιο που αξιολογήθηκε. Μετά την ολοκλήρωση της ανάλυσης του βιογραφικού, το γλωσσικό μοντέλο **Phi-3** δημιουργεί τρεις εξατομικευμένες ερωτήσεις, οι οποίες βασίζονται στις γνώσεις, την επαγγελματική εμπειρία και τις δεξιότητες που αναφέρονται στο βιογραφικό. Οι ερωτήσεις αυτές δεν είναι προκαθορισμένες, αλλά παράγονται δυναμικά κάθε φορά, ανάλογα με το περιεχόμενο του εκάστοτε υποψηφίου.

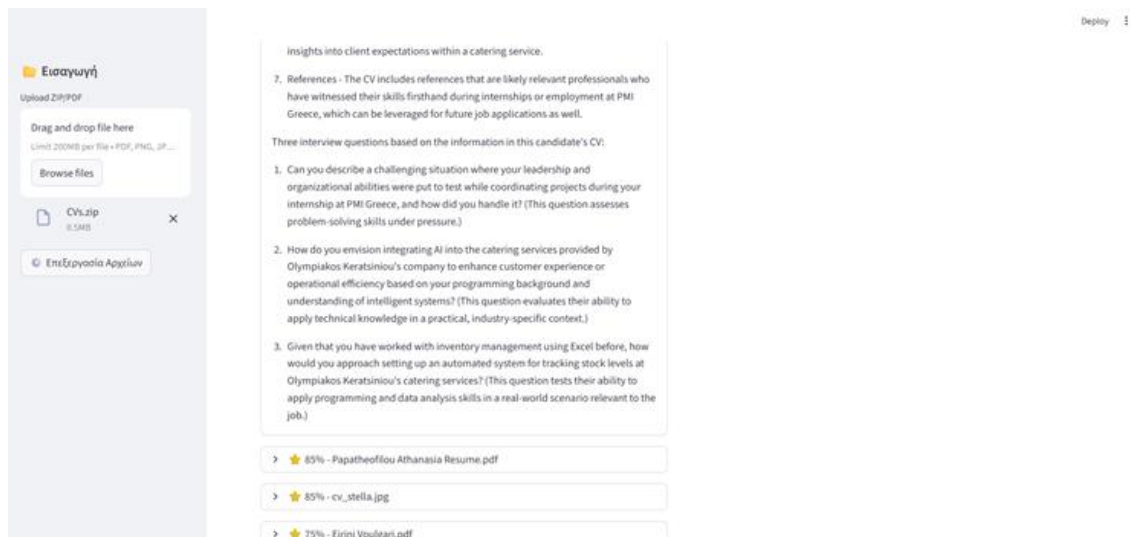
Η λειτουργία αυτή αποτελεί ένα επιπλέον εργαλείο υποστήριξης για τον υπεύθυνο ανθρώπινου δυναμικού, καθώς συμβάλλει στην καλύτερη προετοιμασία της συνέντευξης και επιτρέπει την εστίαση στα σημαντικότερα σημεία του βιογραφικού. Με τον τρόπο αυτό, η εφαρμογή δεν περιορίζεται μόνο στην κατάταξη των υποψηφίων, αλλά υποστηρίζει και το επόμενο στάδιο της διαδικασίας επιλογής προσωπικού, παρέχοντας χρήσιμες και στοχευμένες ερωτήσεις που μπορούν να αξιοποιηθούν κατά τη διάρκεια της συνέντευξης.



Εικόνα 18: Αναλυτική αξιολόγηση του δεύτερου υποψηφίου

Στην Εικόνα 18 παρουσιάζεται η αναλυτική αξιολόγηση του δεύτερου υποψηφίου που προέκυψε από τη διαδικασία κατάταξης. Όπως και στην περίπτωση του πρώτου υποψηφίου, εμφανίζεται το συνολικό ποσοστό συνάφειας, καθώς και μια σύντομη περιγραφή των σημαντικότερων χαρακτηριστικών του βιογραφικού. Η περιγραφή αυτή δημιουργείται αυτόματα από το γλωσσικό μοντέλο **Phi-3** και συνοψίζει τις τεχνικές γνώσεις, την επαγγελματική εμπειρία και τις δεξιότητες που θεωρήθηκαν περισσότερο σχετικές με τις απαιτήσεις της συγκεκριμένης θέσης εργασίας.

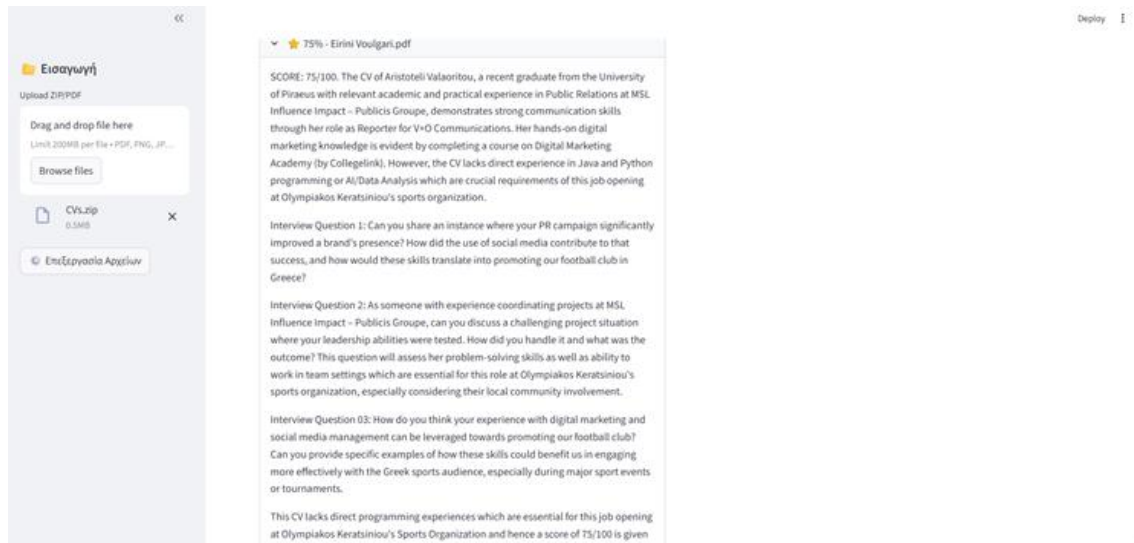
Η παρουσίαση των αποτελεσμάτων με αυτόν τον τρόπο επιτρέπει στον υπεύθυνο ανθρώπινου δυναμικού να συγκρίνει εύκολα διαφορετικούς υποψηφίους χωρίς να χρειάζεται να εξετάσει αναλυτικά κάθε βιογραφικό. Παράλληλα, η εφαρμογή διατηρεί την κατάταξη όλων των υποψηφίων σύμφωνα με το ποσοστό συνάφειας, διευκολύνοντας τη διαδικασία επιλογής και επιτρέποντας την άμεση αναγνώριση των επικρατέστερων υποψηφίων για το επόμενο στάδιο της συνέντευξης.



Εικόνα 19: Αναλυτική αξιολόγηση του τρίτου υποψηφίου

Στην Εικόνα 19 παρουσιάζεται η αξιολόγηση ενός ακόμη υποψηφίου, ο οποίος κατατάσσεται χαμηλότερα σε σχέση με τους προηγούμενους. Η εφαρμογή εμφανίζει το ποσοστό συνάφειας μαζί με μια συνοπτική περιγραφή των γνώσεων και της επαγγελματικής εμπειρίας που εντοπίστηκαν στο βιογραφικό. Η αξιολόγηση πραγματοποιείται με βάση τη σημασιολογική αντιστοίχιση των απαιτήσεων της θέσης εργασίας με τα στοιχεία του βιογραφικού, προσφέροντας μια αντικειμενική εκτίμηση της καταλληλότητας του υποψηφίου.

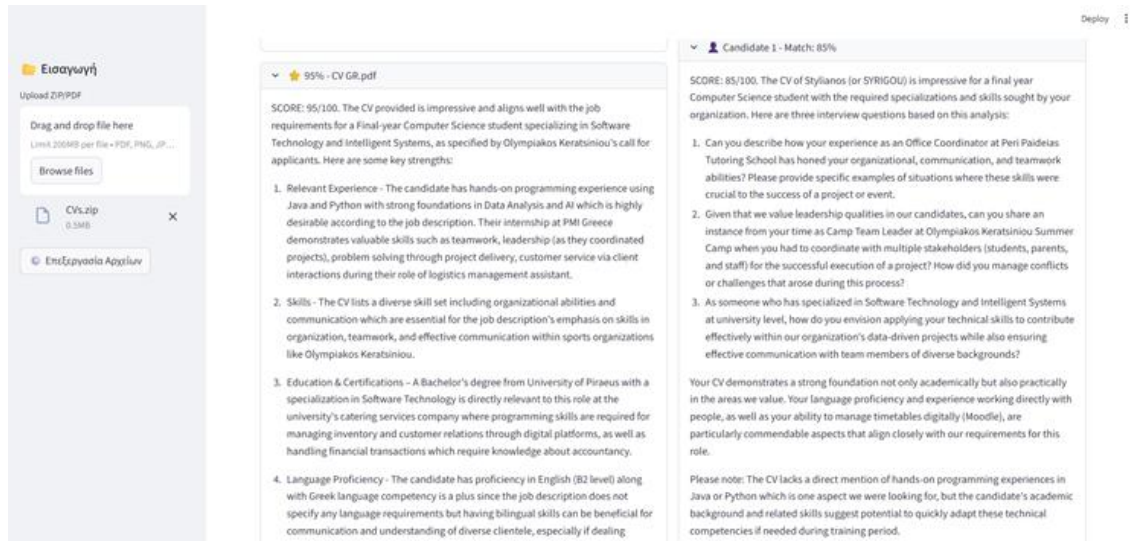
Η δυνατότητα προβολής αναλυτικών πληροφοριών για κάθε υποψήφιο επιτρέπει στον χρήστη να κατανοήσει τους λόγους για τους οποίους ένας υποψήφιος έλαβε συγκεκριμένη βαθμολογία. Με αυτόν τον τρόπο, η εφαρμογή δεν περιορίζεται μόνο στην παρουσίαση ενός ποσοστού, αλλά παρέχει και μια σύντομη τεκμηρίωση της αξιολόγησης, διευκολύνοντας τη λήψη αποφάσεων και τη σύγκριση μεταξύ διαφορετικών υποψηφίων.



Εικόνα 20: Αξιολόγηση υποψηφίου με χαμηλότερο ποσοστό συνάφειας

Στην Εικόνα 20 παρουσιάζεται η αξιολόγηση ενός υποψηφίου με χαμηλότερο ποσοστό συνάφειας σε σχέση με τους υπόλοιπους. Η εφαρμογή εμφανίζει τόσο τη βαθμολογία όσο και μια συνοπτική περιγραφή των χαρακτηριστικών του βιογραφικού, επισημαίνοντας τα στοιχεία που θεωρήθηκαν σχετικά αλλά και εκείνα που δεν καλύπτουν πλήρως τις απαιτήσεις της θέσης εργασίας. Με τον τρόπο αυτό, ο χρήστης μπορεί να κατανοήσει γιατί ο συγκεκριμένος υποψήφιος κατατάχθηκε χαμηλότερα από τους υπόλοιπους.

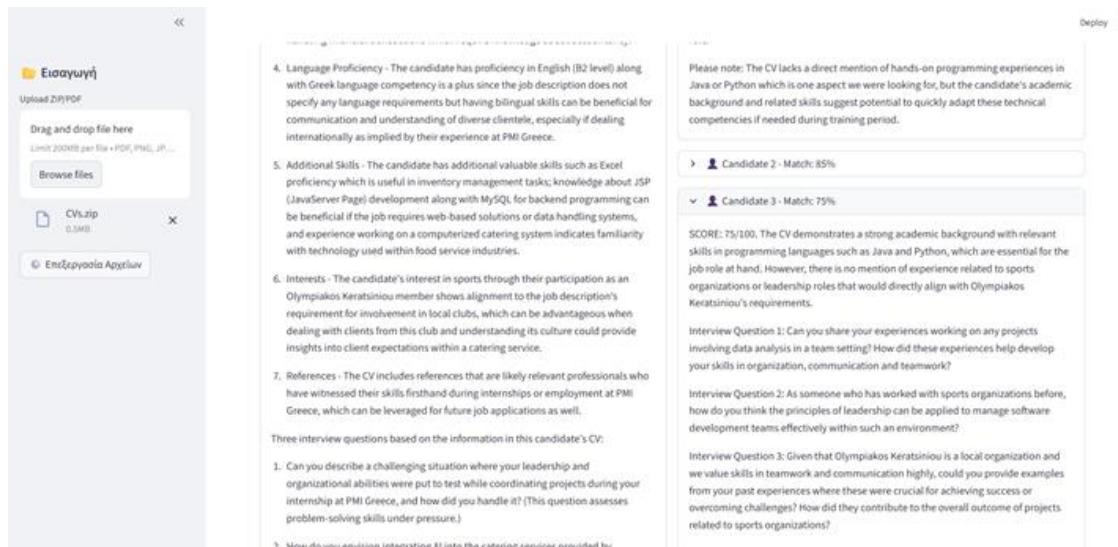
Η συγκεκριμένη λειτουργία ενισχύει τη διαφάνεια της διαδικασίας αξιολόγησης, καθώς δεν παρουσιάζεται μόνο το τελικό αποτέλεσμα αλλά και η βασική αιτιολόγησή του. Έτσι, ο υπεύθυνος ανθρώπινου δυναμικού μπορεί να εξετάσει τόσο τους υποψηφίους με υψηλή βαθμολογία όσο και εκείνους που εμφανίζουν μικρότερη συνάφεια, διατηρώντας μια ολοκληρωμένη εικόνα όλων των αποτελεσμάτων που παράγαγε το σύστημα.



Εικόνα 21: Αναλυτική αξιολόγηση υποψηφίου από τη βάση δεδομένων

Στην Εικόνα 21 παρουσιάζεται η αναλυτική αξιολόγηση ενός από τους υποψηφίους που ανακτήθηκαν από τη βάση δεδομένων μέσω της διαδικασίας αναζήτησης του συστήματος RAG. Ο υποψήφιος εμφανίζεται με το αντίστοιχο ποσοστό συνάφειας, ενώ ακολουθεί αναλυτική περιγραφή των προσόντων, των τεχνικών δεξιοτήτων, της εκπαίδευσης και της επαγγελματικής εμπειρίας που συνέβαλαν στη συγκεκριμένη βαθμολογία. Η ανάλυση αυτή δημιουργείται αυτόματα από το γλωσσικό μοντέλο **Phi-3**, το οποίο συγκρίνει το περιεχόμενο του βιογραφικού με τις απαιτήσεις της θέσης εργασίας και αιτιολογεί τη συνολική αξιολόγηση.

Παράλληλα, η εφαρμογή δημιουργεί τρεις εξατομικευμένες ερωτήσεις συνέντευξης, οι οποίες βασίζονται αποκλειστικά στα χαρακτηριστικά του συγκεκριμένου βιογραφικού. Οι ερωτήσεις αυτές μπορούν να χρησιμοποιηθούν από τον υπεύθυνο ανθρώπινου δυναμικού κατά τη διάρκεια της συνέντευξης, επιτρέποντας μια περισσότερο στοχευμένη αξιολόγηση του υποψηφίου. Με τον τρόπο αυτό, το σύστημα υποστηρίζει όχι μόνο την επιλογή των καταλληλότερων βιογραφικών, αλλά και την αποτελεσματικότερη προετοιμασία της διαδικασίας συνέντευξης.



Εικόνα 22: Αναλυτική αξιολόγηση υποψηφίου με χαμηλότερο ποσοστό συνάφειας από τη βάση δεδομένων

Στην Εικόνα 22 παρουσιάζεται η αναλυτική αξιολόγηση ενός υποψηφίου που προέρχεται επίσης από τη βάση δεδομένων, αλλά συγκέντρωσε χαμηλότερο ποσοστό συνάφειας σε σχέση με τους υπόλοιπους. Η εφαρμογή εμφανίζει τη συνολική βαθμολογία, καθώς και την αιτιολόγηση που δημιουργήσε το γλωσσικό μοντέλο **Phi-3**, επισημαίνοντας τόσο τα στοιχεία που ικανοποιούν τις απαιτήσεις της θέσης όσο και τις ελλείψεις που οδήγησαν στη χαμηλότερη κατάταξη. Με αυτόν τον τρόπο, ο χρήστης μπορεί να κατανοήσει με σαφήνεια τους λόγους της συγκεκριμένης αξιολόγησης.

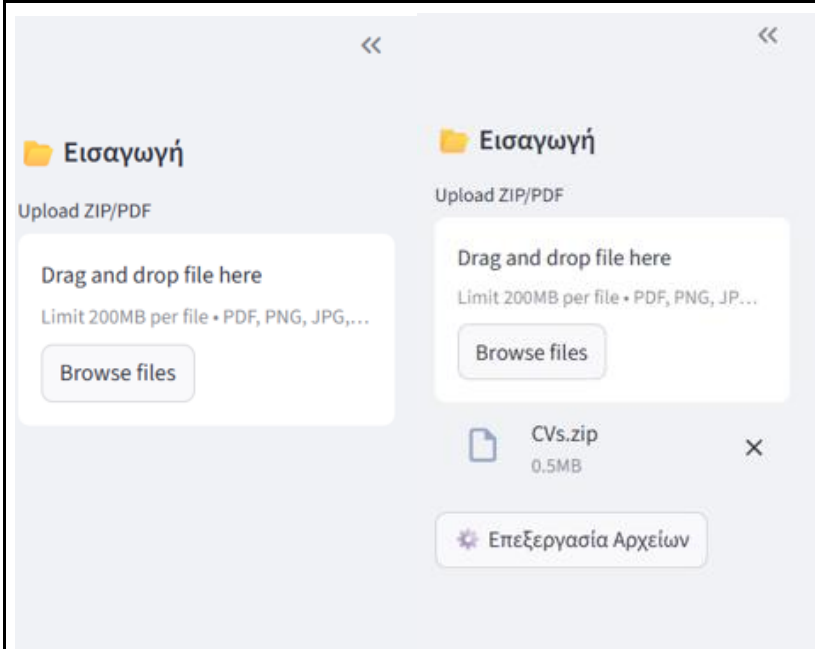
Επιπλέον, δημιουργούνται αυτόματα τρεις προτεινόμενες ερωτήσεις συνέντευξης, οι οποίες εστιάζουν στις εμπειρίες και στις δεξιότητες που αναφέρονται στο συγκεκριμένο βιογραφικό. Η δυνατότητα αυτή επιτρέπει στον υπεύθυνο ανθρώπινου δυναμικού να διερευνήσει περαιτέρω τα δυνατά σημεία αλλά και τις πιθανές αδυναμίες του υποψηφίου, διατηρώντας μια ενιαία και αντικειμενική διαδικασία αξιολόγησης για όλα τα βιογραφικά που ανακτώνται από τη βάση δεδομένων.

10.Εγχειρίδιο Σύντομης Αναφοράς

Στην ενότητα αυτή θα παρουσιαστούν συνοπτικά οι βασικές λειτουργίες της εφαρμογής Smart HR Ranking & Evaluation, τις οποίες μπορεί να αξιοποιήσει ο χρήστης για την αυτόματη αξιολόγηση βιογραφικών και την αντιστοίχιση υποψηφίων με διαθέσιμες θέσεις εργασίας.

10.1. Εισαγωγή Αρχείων

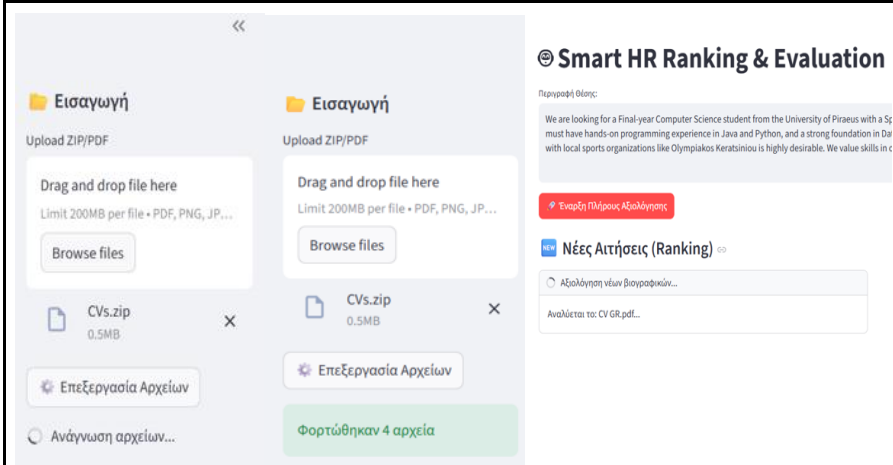
10.1.1. Σελίδα Εισαγωγής Βιογραφικών (Upload CV)



Εικόνα 23: Σελίδα εισαγωγής αρχείων βιογραφικών

1. Επιλογή και μεταφόρτωση αρχείου βιογραφικού υποψηφίου.
2. Υποστήριξη εισαγωγής αρχείων διαφορετικών μορφών.
3. Προβολή των αρχείων που έχουν εισαχθεί στο σύστημα.
4. Εκκίνηση διαδικασίας επεξεργασίας των βιογραφικών.

10.1.2. Επεξεργασία Αρχείων (File Processing)



Εικόνα 24: Σελίδα επεξεργασίας και ανάλυσης αρχείων

1. Έλεγχος των αρχείων που έχουν εισαχθεί.
2. Εξαγωγή κειμένου από τα βιογραφικά μέσω PyPDF και EasyOCR.
3. Μετατροπή όλων των αρχείων σε ενιαία μορφή κειμένου.
4. Προετοιμασία των δεδομένων για την αξιολόγηση.

10.2. Εισαγωγή Περιγραφής Θέσης Εργασίας (Job Description)

 Smart HR Ranking & Evaluation Περιγραφή Θέσης: We are looking for a Final-year Computer Science student from the University of Piraeus with a Specialization in Software Technology and Intelligent Systems. The ideal candidate must have hands-on programming experience in Java and Python, and a strong foundation in Data Analysis and AI. Previous experience in coordination or leadership roles, specifically with local sports organizations like Olympiakos Keratsiniou is highly desirable. We value skills in organization, teamwork, and communication. Press Ctrl+Enter to apply Εικόνα 25: Σελίδα εισαγωγής απαιτήσεων θέσης εργασίας	1. Εισαγωγή περιγραφής της διαθέσιμης θέσης εργασίας. 2. Καθορισμός απαιτούμενων δεξιοτήτων και προσόντων. 3. Ανάλυση των απαιτήσεων μέσω του συστήματος RAG. 4. Δημιουργία βάσης σύγκρισης μεταξύ υποψηφίων και θέσης εργασίας.
--	---

10.3. Αποτελέσματα Αντιστοίχισης Υποψηφίων από Εισαγωγή (Candidate Matching)

10.3.1. Λίστα Υποψηφίων από Εισαγωγή

Ενάρξη Πλήρους Αξιολόγησης

NEW

Νέες Αιτήσεις (Ranking)

✓ Η αξιολόγηση ολοκληρώθηκε!

Αναλύεται το: CV GR.pdf...

Αναλύεται το: Papatheofilou Athanasia Resume.pdf...

Αναλύεται το: Eirini Voulgari.pdf...

Αναλύεται το: cv_stella.jpg...

> ⭐ 95% - CV GR.pdf

> ⭐ 85% - Papatheofilou Athanasia Resume.pdf

> ⭐ 85% - cv_stella.jpg

> ⭐ 75% - Eirini Voulgari.pdf

1. Παρουσίαση των υποψηφίων που αξιολογήθηκαν.
2. Εμφάνιση ποσοστού αντιστοίχισης (Match Score).
3. Ταξινόμηση υποψηφίων βάσει καταλληλότητας.
4. Επιλογή υποψηφίου για αναλυτική προβολή.

Εικόνα 26: Προβολή αποτελεσμάτων αντιστοίχισης υποψηφίων	
---	--

10.3.2. Αναλυτική Αξιολόγηση Υποψηφίου από Εισαγωγή

ΣΥΝΟΛΟ ΑΞΙΟΛΟΓΗΣΗΣ SCORE: 95/100. The CV provided is impressive and aligns well with the job requirements for a Final-year Computer Science student specializing in Software Technology and Intelligent Systems, as specified by Olympiakos Keratsiniou's call for applicants. Here are some key strengths: 1. Relevant Experience - The candidate has hands-on programming experience using Java and Python with strong foundations in Data Analysis and AI which is highly desirable according to the job description. Their internship at PMI Greece demonstrates valuable skills such as teamwork, leadership (as they coordinated projects), problem solving through project delivery, customer service via client interactions during their role of logistics management assistant. 2. Skills - The CV lists a diverse skill set including organizational abilities and communication which are essential for the job description's emphasis on skills in organization, teamwork, and effective communication within sports organizations like Olympiakos Keratsiniou. 3. Education & Certifications - A Bachelor's degree from University of Piraeus with a specialization in Software Technology is directly relevant to this role at the university's catering services company where programming skills are required for managing inventory and customer relations through digital platforms, as well as handling financial transactions which require knowledge about accountancy. 4. Language Proficiency - The candidate has proficiency in English (B2 level) along with Greek language competency is a plus since the job description does not specify any language requirements but having bilingual skills can be beneficial for communication and understanding of diverse clientele, especially if dealing internationally as implied by their experience at PMI Greece. 5. Additional Skills - The candidate has additional valuable skills such as Excel Εικόνα 27: Αναλυτική παρουσίαση αποτελέσματος αξιολόγησης	1. Συνολική βαθμολογία αντιστοίχισης. 2. Παρουσίαση θετικών χαρακτηριστικών του υποψηφίου. 3. Εντοπισμός ελλείψεων σε σχέση με τη θέση εργασίας. 4. Παροχή αιτιολόγησης μέσω του γλωσσικού μοντέλου.
--	--

10.4. Δημιουργία Ερωτήσεων Συνέντευξης

insights into client expectations within a catering service. 7. References - The CV includes references that are likely relevant professionals who have witnessed their skills firsthand during internships or employment at PMI Greece, which can be leveraged for future job applications as well. Three interview questions based on the information in this candidate's CV: 1. Can you describe a challenging situation where your leadership and organizational abilities were put to test while coordinating projects during your internship at PMI Greece, and how did you handle it? (This question assesses problem-solving skills under pressure.) 2. How do you envision integrating AI into the catering services provided by Olympiakos Keratsiniou's company to enhance customer experience or operational efficiency based on your programming background and understanding of intelligent systems? (This question evaluates their ability to apply technical knowledge in a practical, industry-specific context.) 3. Given that you have worked with inventory management using Excel before, how would you approach setting up an automated system for tracking stock levels at Olympiakos Keratsiniou's catering services? (This question tests their ability to apply programming and data analysis skills in a real-world scenario relevant to the job.) > 🌟 85% - Papatheofilou Athanasia Resume.pdf > 🌟 85% - cv_stella.jpg > 🌟 75% - Eirini Voulgari.pdf Εικόνα 28: Παραγωγή ερωτήσεων συνέντευξης από το σύστημα	1. Δημιουργία εξατομικευμένων ερωτήσεων με βάση το βιογραφικό. 2. Σύνδεση ερωτήσεων με τις απαιτήσεις της θέσης εργασίας. 3. Αξιολόγηση πιθανών σημείων προς διερεύνηση. 4. Υποστήριξη της διαδικασίας επιλογής προσωπικού.
---	--

10.5. Αποτελέσματα Αντιστοίχισης Υποψηφίων (Candidate Matching)

10.5.1. Λίστα Υποψηφίων

Deploy ersity of Piraeus with a Specialization in Software Technology and Intelligent nce in Java and Python, and a strong foundation in Data Analysis and AI. Previous rts organizations like Olympiakos Keratsiniou is highly desirable. We value skills Database Matches (Unique) ✓ Η αναζήτηση ολοκληρώθηκε! Αναλύεται υποψήφιος από τη βάση... Αναλύεται υποψήφιος από τη βάση... Αναλύεται υποψήφιος από τη βάση... > Candidate 1 - Match: 85% > Candidate 2 - Match: 85% > Candidate 3 - Match: 85% Εικόνα 29: Προβολή αποτελεσμάτων αντιστοίχισης υποψηφίων	5. Παρουσίαση των υποψηφίων που αξιολογήθηκαν. 6. Εμφάνιση ποσοστού αντιστοίχισης (Match Score). 7. Ταξινόμηση υποψηφίων βάσει καταλληλότητας. 8. Επιλογή υποψηφίου για αναλυτική προβολή.
--	--

10.5.2. Αναλυτική Αξιολόγηση Υποψηφίου

Deploy Candidate 1 - Match: 85% SCORE: 85/100. The CV presents a strong candidate with relevant academic background and practical experience that aligns well with the job requirements at hand. Here is an analysis of their qualifications based on your provided criteria for the position as described in the job posting, followed by three interview questions tailored to assess specific skills or experiences related to leadership within sports organizations: Academic Background: The candidate has a strong foundation in Computer Science with specialization in Software Technology and Intelligent Systems from the University of Piraeus. This aligns well with the technical requirements for hands-on programming experience, as they have demonstrated proficiency in Java (a core requirement) through their teaching competency at Greek High School where software technology is likely a part of the curriculum. Work Experience: The candidate has relevant work experience that demonstrates organizational skills and communication abilities which are highly valued for this role, as evidenced by coordinating student events and managing class schedules while maintaining records to support data-driven decisions at Peri Paideias Tutoring School. Volunteer Experience: The candidate has leadership experience with Olympiakos Keratsiniou Summer Camp where they served as a team leader, which directly aligns with the desired previous coordination or leadership roles within local sports organizations like Olympiakos Keratsiniou mentioned in your job posting. This role required scheduling and communication skills that are essential for this position at hand. Εικόνα 30: Αναλυτική παρουσίαση αποτελέσματος αξιολόγησης	5. Συνολική βαθμολογία αντιστοίχισης. 6. Παρουσίαση θετικών χαρακτηριστικών του υποψηφίου. 7. Εντοπισμός ελλείψεων σε σχέση με τη θέση εργασίας. 8. Παροχή αιτιολόγησης μέσω του γλωσσικού μοντέλου.
---	--

10.6. Δημιουργία Ερωτήσεων Συνέντευξης

Deploy ⓘ with a specialization that aligns well with software technology and intelligent systems makes them suitable for technical aspects of the job role at hand; their hands-on programming experience is evident through teaching competency, which demonstrates practical application. Their organizational skills are proven by previous work experiences where they managed events or class schedules effectively while maintaining records to support data-driven decisions and facilitated communication between students/parents in a volunteer role with Olympiakos Keratsiniou Summer Camp; their language proficiency is also beneficial. Interview Questions: 1. Can you share an experience where your leadership skills were put to the test, particularly within a sports organization setting? How did you handle it and what was the outcome of that situation? (This question assesses problem-solving abilities in high pressure situations.) 2. As someone who has managed events before, how would you approach coordinating activities for an event with multiple age groups or interests to ensure everyone is engaged while maintaining a safe environment during COVID times? 3. How do you think your experience as part of Olympiakos Keratsiniou Summer Camp can be leveraged in this role at hand and what specific skills from that volunteer work would benefit the organization most directly? (This question assesses how transferable their leadership experiences are to a new environment.) >  Candidate 2 - Match: 85% >  Candidate 3 - Match: 85% Εικόνα 31: Παραγωγή ερωτήσεων συνέντευξης από το σύστημα <tr><td> 5. Δημιουργία εξατομικευμένων ερωτήσεων με βάση το βιογραφικό. 6. Σύνδεση ερωτήσεων με τις απαιτήσεις της θέσης εργασίας. 7. Αξιολόγηση πιθανών σημείων προς διερεύνηση. 8. Υποστήριξη της διαδικασίας επιλογής προσωπικού. </td></tr>	5. Δημιουργία εξατομικευμένων ερωτήσεων με βάση το βιογραφικό. 6. Σύνδεση ερωτήσεων με τις απαιτήσεις της θέσης εργασίας. 7. Αξιολόγηση πιθανών σημείων προς διερεύνηση. 8. Υποστήριξη της διαδικασίας επιλογής προσωπικού.
5. Δημιουργία εξατομικευμένων ερωτήσεων με βάση το βιογραφικό. 6. Σύνδεση ερωτήσεων με τις απαιτήσεις της θέσης εργασίας. 7. Αξιολόγηση πιθανών σημείων προς διερεύνηση. 8. Υποστήριξη της διαδικασίας επιλογής προσωπικού.	

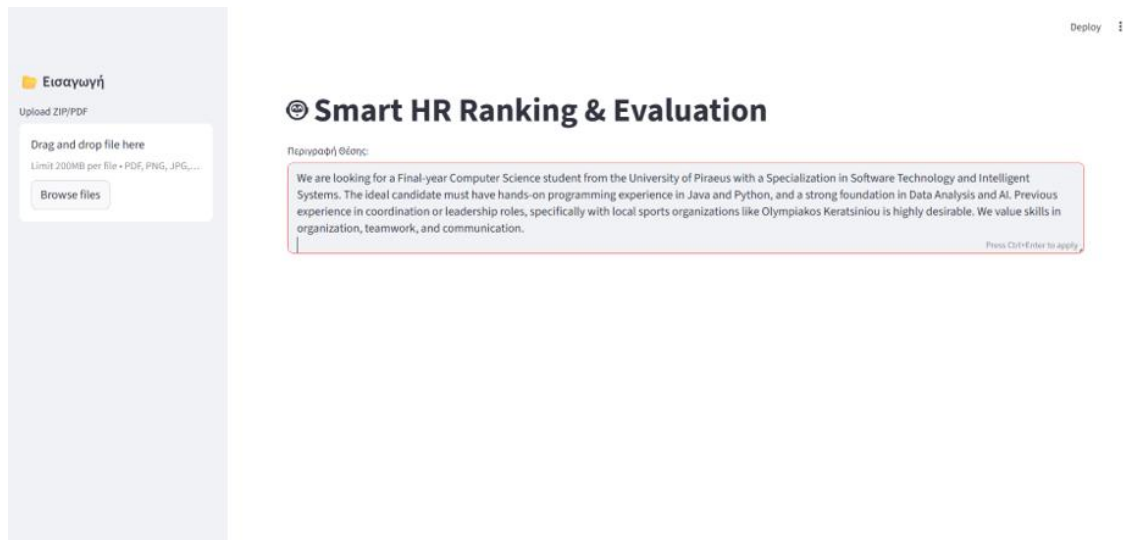
11. Αναλυτικό εγχειρίδιο χρήστη

Το αναλυτικό εγχειρίδιο χρήστη παρουσιάζει τη διαδικασία λειτουργίας της εφαρμογής Smart HR Ranking & Evaluation από τη σκοπιά του τελικού χρήστη. Περιγράφονται αναλυτικά όλα τα βήματα που απαιτούνται για την εισαγωγή της περιγραφής μιας θέσης εργασίας, την προαιρετική προσθήκη νέων βιογραφικών, την αξιολόγηση των υποψηφίων και την παρουσίαση των αποτελεσμάτων. Με τον τρόπο αυτό, ο χρήστης μπορεί να αξιοποιήσει αποτελεσματικά όλες τις δυνατότητες της εφαρμογής, είτε επιθυμεί να αξιολογήσει νέα βιογραφικά που διαθέτει είτε να αναζητήσει τους καταλληλότερους υποψηφίους από τη βάση δεδομένων.

11.1. Εισαγωγή Περιγραφής Θέσης Εργασίας

Η διαδικασία αξιολόγησης ξεκινά με την εισαγωγή της περιγραφής της διαθέσιμης θέσης εργασίας. Στο αντίστοιχο πεδίο ο χρήστης καταχωρίζει τις απαιτήσεις της θέσης, όπως τις απαραίτητες τεχνικές γνώσεις, την απαιτούμενη επαγγελματική εμπειρία, τις επιθυμητές δεξιότητες και οποιοδήποτε άλλο χαρακτηριστικό θεωρεί σημαντικό. Η περιγραφή αυτή αποτελεί το βασικό σημείο αναφοράς της εφαρμογής, καθώς πάνω σε αυτή βασίζεται ολόκληρη η διαδικασία αντιστοίχισης των υποψηφίων.

Μετά την εισαγωγή της, το σύστημα χρησιμοποιεί την περιγραφή ως ερώτημα αναζήτησης (query) στην αρχιτεκτονική **Retrieval Augmented Generation (RAG)**. Το κείμενο μετατρέπεται σε διανυσματική αναπαράσταση για την αναζήτηση των σχετικών βιογραφικών στη βάση, ενώ η ίδια η περιγραφή προωθείται σε μορφή κειμένου στο γλωσσικό μοντέλο Phi-3 για την τελική αξιολόγηση.



Εικόνα 32: Εισαγωγή Job Description

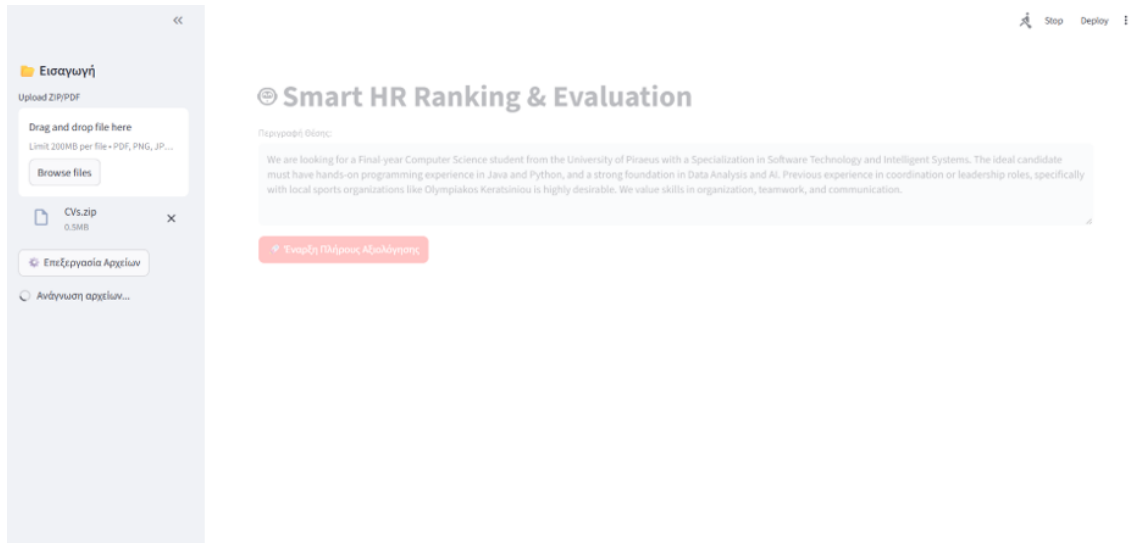
Βήματα χρήσης

1. Ο χρήστης εισάγει την περιγραφή της θέσης εργασίας.
2. Καθορίζει τα απαιτούμενα προσόντα και τις δεξιότητες.
3. Επιβεβαιώνει το κείμενο που θα χρησιμοποιηθεί στην αξιολόγηση.
4. Η περιγραφή χρησιμοποιείται ως βάση για την αναζήτηση και σύγκριση των υποψηφίων.

11.2. Προαιρετική Εισαγωγή Νέων Βιογραφικών

Μετά την εισαγωγή της περιγραφής της θέσης, ο χρήστης έχει τη δυνατότητα να προσθέσει νέα βιογραφικά τα οποία δεν υπάρχουν ήδη στη βάση δεδομένων. Η διαδικασία αυτή είναι προαιρετική και επιτρέπει την αξιολόγηση επιπλέον υποψηφίων χωρίς να απαιτείται προηγούμενη αποθήκευσή τους στη βάση.

Η εφαρμογή υποστηρίζει αρχεία μορφής **PDF**, **JPG**, **PNG** και **ZIP**, επιτρέποντας την επεξεργασία τόσο εγγράφων όσο και εικόνων. Για τα αρχεία εικόνας χρησιμοποιείται η βιβλιοθήκη **EasyOCR**, η οποία μετατρέπει το περιεχόμενο σε κείμενο, ενώ τα αρχεία PDF επεξεργάζονται απευθείας. Στη συνέχεια, όλα τα βιογραφικά προετοιμάζονται για την αξιολόγηση.



Εικόνα 33: Εισαγωγή νέων βιογραφικών

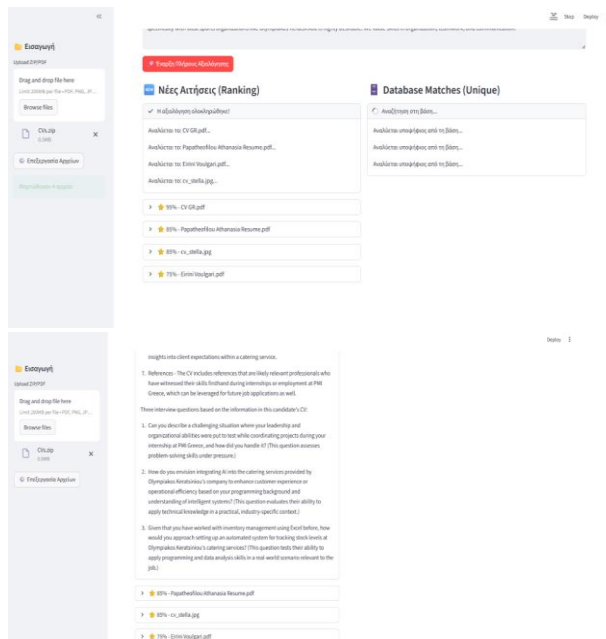
Βήματα χρήσης

1. Ο χρήστης επιλέγει την περιοχή μεταφόρτωσης αρχείων.
2. Επιλέγει ένα βιογραφικό ή ένα συμπιεσμένο αρχείο ZIP με πολλά βιογραφικά από τον υπολογιστή του.
3. Η εφαρμογή αναγνωρίζει αυτόματα τον τύπο κάθε αρχείου.
4. Τα δεδομένα προετοιμάζονται για αξιολόγηση.

11.3. Αξιολόγηση των Νέων Βιογραφικών

Εφόσον ο χρήστης έχει προσθέσει νέα βιογραφικά, μπορεί να ξεκινήσει τη διαδικασία αξιολόγησής τους. Η εφαρμογή εξάγει το κείμενο από κάθε αρχείο και το προωθεί, μαζί με την περιγραφή της θέσης εργασίας, στο γλωσσικό μοντέλο **Phi-3**, το οποίο αναλαμβάνει τη σύγκριση του περιεχομένου με τις απαιτήσεις της θέσης.

Στη συνέχεια, το μοντέλο **Phi-3** αναλύει κάθε υποψήφιο, υπολογίζει το ποσοστό συνάφειας, δημιουργεί σύντομη αιτιολόγηση της αξιολόγησης και παράγει προτεινόμενες ερωτήσεις συνέντευξης.



Εικόνα 34: Αξιολόγηση των νέων βιογραφικών

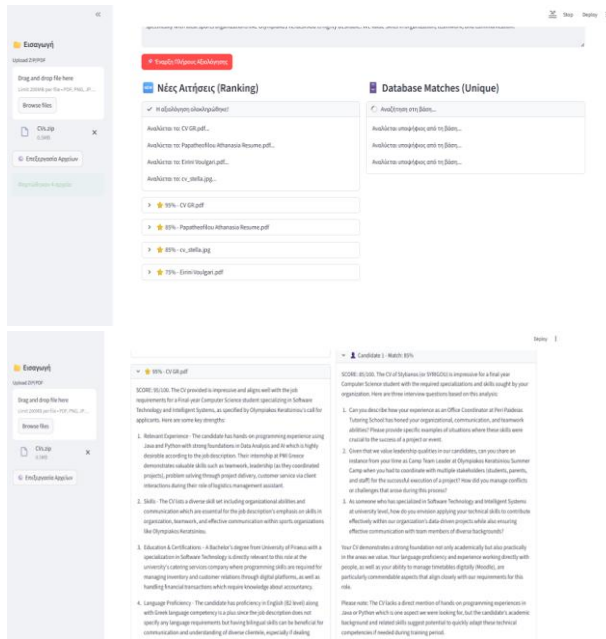
Βήματα χρήσης

1. Ο χρήστης επιλέγει την έναρξη αξιολόγησης.
2. Η εφαρμογή επεξεργάζεται όλα τα νέα βιογραφικά.
3. Υπολογίζεται το ποσοστό συνάφειας κάθε υποψηφίου.
4. Παρουσιάζονται τα αποτελέσματα αξιολόγησης.

11.4. Αναζήτηση Υποψηφίων από τη Βάση Δεδομένων

Σε περίπτωση που ο χρήστης δεν επιθυμεί να προσθέσει νέα βιογραφικά, η εφαρμογή μπορεί να αξιοποιήσει τα βιογραφικά που υπάρχουν ήδη αποθηκευμένα στη βάση **ChromaDB**. Η περιγραφή της θέσης εργασίας χρησιμοποιείται ως ερώτημα αναζήτησης και η αρχιτεκτονική **RAG** ανακτά τους υποψηφίους που παρουσιάζουν τη μεγαλύτερη σημασιολογική ομοιότητα με τις απαιτήσεις της θέσης.

Τα ανακτημένα βιογραφικά αποστέλλονται στο μοντέλο **Phi-3**, το οποίο πραγματοποιεί την τελική αξιολόγηση και δημιουργεί τη βαθμολογία, την αιτιολόγηση και τις προτεινόμενες ερωτήσεις συνέντευξης για κάθε υποψήφιο.



Εικόνα 35: Αναζήτηση υποψηφίων από τη βάση δεδομένων

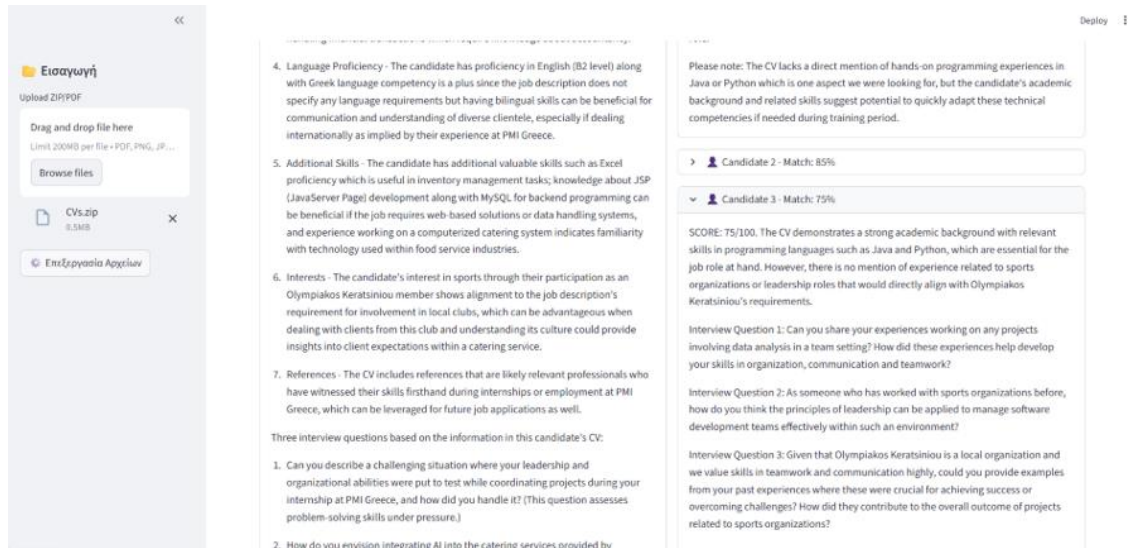
Βήματα χρήσης

1. Ο χρήστης επιλέγει την αξιολόγηση των βιογραφικών της βάσης.
2. Η εφαρμογή πραγματοποιεί σημασιολογική αναζήτηση στη βάση ChromaDB.
3. Ανακτώνται οι τρεις επικρατέστεροι υποψήφιοι, μετά την αφαίρεση τυχόν διπλότυπων εγγράφων.
4. Τα αποτελέσματα προωθούνται στο γλωσσικό μοντέλο για αξιολόγηση.

11.5. Παρουσίαση Αποτελεσμάτων

Μετά την ολοκλήρωση της αξιολόγησης, είτε αυτή αφορά νέα βιογραφικά είτε βιογραφικά της βάσης δεδομένων, η εφαρμογή παρουσιάζει τα αποτελέσματα ταξινομημένα κατά φθίνουσα σειρά, σύμφωνα με το ποσοστό συνάφειας κάθε υποψηφίου με τη θέση εργασίας.

Για κάθε υποψήφιο εμφανίζεται η συνολική βαθμολογία, σύντομη περιγραφή των προσόντων του, αναλυτική αιτιολόγηση της αξιολόγησης και τρεις εξατομικευμένες ερωτήσεις συνέντευξης που δημιουργήθηκαν αυτόματα από το γλωσσικό μοντέλο.



Εικόνα 36: Τελική παρουσίαση αποτελεσμάτων

Βήματα χρήσης

1. Ο χρήστης εξετάζει την τελική κατάταξη των υποψηφίων.
2. Επιλέγει έναν υποψήφιο για αναλυτική προβολή.
3. Μελετά την αιτιολόγηση και τις προτεινόμενες ερωτήσεις συνέντευξης.
4. Χρησιμοποιεί τα αποτελέσματα για τη λήψη της τελικής απόφασης.

12. Αξιολόγηση της αποτελεσματικότητας του συστήματος

12.1 Σκοπός της αξιολόγησης

Για την αποτίμηση της αποτελεσματικότητας του προτεινόμενου συστήματος πραγματοποιήθηκε αξιολόγηση της διαδικασίας αντιστοίχισης βιογραφικών σημειωμάτων με περιγραφές θέσεων εργασίας. Στόχος της αξιολόγησης είναι να εξεταστεί κατά πόσο ο μηχανισμός σημασιολογικής αναζήτησης που βασίζεται σε διανυσματικές αναπαραστάσεις είναι σε θέση να ανακτά τα βιογραφικά που παρουσιάζουν τη μεγαλύτερη σημασιολογική συνάφεια με μία δεδομένη θέση εργασίας.

Η αξιολόγηση βασίζεται στο δημόσιο dataset βιογραφικών που χρησιμοποιήθηκε για τη δημιουργία της διανυσματικής βάσης ChromaDB. Στο dataset περιλαμβάνονται 962 βιογραφικά, τα οποία είναι οργανωμένα σε διαφορετικές επαγγελματικές κατηγορίες. Η περιγραφή της θέσης εργασίας χρησιμοποιείται ως ερώτημα αναζήτησης, μετατρέπεται σε embedding και συγκρίνεται σημασιολογικά με τις διανυσματικές αναπαραστάσεις των βιογραφικών.

12.2 Μεθοδολογία αξιολόγησης

Για την αξιολόγηση χρησιμοποιήθηκε το σύνολο των διαθέσιμων βιογραφικών του dataset, το οποίο περιλαμβάνει 962 βιογραφικά. Για καθεμία από τις 25 επαγγελματικές κατηγορίες δημιουργήθηκαν δύο διαφορετικές περιγραφές θέσης εργασίας, με αποτέλεσμα να πραγματοποιηθούν συνολικά 50 αναζητήσεις. Κάθε περιγραφή θέσης εργασίας

χρησιμοποιήθηκε ως ερώτημα αναζήτησης στη διανυσματική βάση δεδομένων, με στόχο την ανάκτηση των βιογραφικών που παρουσιάζουν τη μεγαλύτερη σημασιολογική ομοιότητα.

12.3 Μετρικές αξιολόγησης

Για την ποσοτική αξιολόγηση της απόδοσης χρησιμοποιήθηκαν μετρικές που βασίζονται στη θέση στην οποία εμφανίζεται ένα σχετικό βιογραφικό στα αποτελέσματα αναζήτησης. Εξετάστηκε ειδικότερα η επιτυχία ανάκτησης σχετικού υποψηφίου στα πρώτα αποτελέσματα (Top-K), ώστε να αποτυπωθεί κατά πόσο το σύστημα είναι σε θέση να περιορίζει αποτελεσματικά το σύνολο των υποψηφίων και να προτείνει τους καταλληλότερους υποψηφίους στις υψηλότερες θέσεις της κατάταξης.

Μετρική Αποτέλεσμα	
Top-1 Accuracy	82.00%
Top-3 Accuracy	84.00%
Top-5 Accuracy	88.00%
MRR	0.8347

12.4 Αποτελέσματα αξιολόγησης

Για την αξιολόγηση της αποτελεσματικότητας του συστήματος πραγματοποιήθηκαν συνολικά 50 αναζητήσεις, οι οποίες αντιστοιχούν σε δύο διαφορετικές περιγραφές θέσης εργασίας για καθεμία από τις 25 επαγγελματικές κατηγορίες του dataset. Τα ερωτήματα μετατράπηκαν σε διανυσματικές αναπαραστάσεις και χρησιμοποιήθηκαν για την ανάκτηση των πλέον σημασιολογικά σχετικών βιογραφικών από τη διανυσματική βάση δεδομένων.

Τα αποτελέσματα έδειξαν ότι το σύστημα πέτυχε Top-1 Accuracy 82,00%, δηλαδή σε 82% των περιπτώσεων το πρώτο αποτέλεσμα της αναζήτησης ανήκε στην ίδια επαγγελματική κατηγορία με την αναμενόμενη. Όταν εξετάστηκαν τα τρία πρώτα αποτελέσματα, η απόδοση αυξήθηκε σε 84,00%, ενώ στα πέντε πρώτα αποτελέσματα έφτασε το 88,00%.

Η τιμή MRR (Mean Reciprocal Rank) υπολογίστηκε ίση με 0,8347, υποδεικνύοντας ότι το σχετικό βιογραφικό εμφανίζεται, κατά μέσο όρο, σε υψηλή θέση της κατάταξης των αποτελεσμάτων.

Τα αποτελέσματα καταδεικνύουν την ικανότητα του συστήματος να εντοπίζει βιογραφικά που παρουσιάζουν σημασιολογική συνάφεια με τις απαιτήσεις μιας θέσης εργασίας. Η αύξηση της απόδοσης από το Top-1 στο Top-5 δείχνει επίσης ότι, σε ορισμένες περιπτώσεις, το καταλληλότερο αποτέλεσμα δεν εμφανίζεται απαραίτητα στην πρώτη θέση, αλλά περιλαμβάνεται στις πρώτες θέσεις της λίστας προτεινόμενων υποψηφίων.

12.5 Συνολική αξιολόγηση

Συνολικά, τα αποτελέσματα της αξιολόγησης καταδεικνύουν ότι το προτεινόμενο σύστημα είναι σε θέση να πραγματοποιεί αποτελεσματική σημασιολογική αντιστοίχιση μεταξύ περιγραφών θέσεων εργασίας και βιογραφικών σημειωμάτων. Η επίτευξη Top-1 Accuracy 82,00% δείχνει ότι στην πλειονότητα των περιπτώσεων το σύστημα κατατάσσει ένα βιογραφικό της αναμενόμενης επαγγελματικής κατηγορίας στην πρώτη θέση. Παράλληλα, η επίδοση 88,00% στο Top-5

δείχνει ότι το σχετικό βιογραφικό περιλαμβάνεται στις πέντε υψηλότερες θέσεις στην πλειονότητα των αναζητήσεων.

Οι περιπτώσεις στις οποίες δεν ανακτήθηκε βιογραφικό της αναμενόμενης κατηγορίας στα πρώτα αποτελέσματα παρουσιάζουν ιδιαίτερο ενδιαφέρον. Για παράδειγμα, στις αναζητήσεις που αφορούσαν τις κατηγορίες HR και Operations Manager, τα αποτελέσματα παρουσίασαν κυρίως βιογραφικά της κατηγορίας PMO. Αντίστοιχα, στην κατηγορία Testing εμφανίστηκαν κυρίως βιογραφικά της κατηγορίας Automation Testing. Οι αποκλίσεις αυτές μπορούν να αποδοθούν, μεταξύ άλλων, στις εννοιολογικές και δεξιολογικές ομοιότητες που παρουσιάζουν ορισμένες επαγγελματικές κατηγορίες, με αποτέλεσμα τα βιογραφικά των συγκεκριμένων κατηγοριών να παρουσιάζουν υψηλή σημασιολογική ομοιότητα με τα αντίστοιχα ερωτήματα αναζήτησης.

Η τιμή MRR ίση με 0,8347 ενισχύει την παραπάνω παρατήρηση, καθώς υποδεικνύει ότι τα σχετικά αποτελέσματα εμφανίζονται γενικά σε υψηλές θέσεις της κατάταξης. Η διαφορά μεταξύ της Top-1 και της Top-5 Accuracy δείχνει επίσης ότι σε ορισμένες περιπτώσεις το σχετικό βιογραφικό δεν κατατάσσεται στην πρώτη θέση, αλλά εξακολουθεί να βρίσκεται μεταξύ των πέντε κορυφαίων αποτελεσμάτων.

Συνεπώς, η συνολική αξιολόγηση καταδεικνύει ότι η χρήση διανυσματικών αναπαραστάσεων και σημασιολογικής αναζήτησης μπορεί να αποτελέσει αποτελεσματικό μηχανισμό για τον περιορισμό του αρχικού συνόλου υποψηφίων και την ανάδειξη των περισσότερο σχετικών βιογραφικών. Τα αποτελέσματα δείχνουν ότι το σύστημα μπορεί να λειτουργήσει ως εργαλείο υποστήριξης της διαδικασίας αντιστοίχισης υποψηφίων με θέσεις εργασίας, παρέχοντας μία αρχική κατάταξη υποψηφίων με βάση τη σημασιολογική συνάφεια των βιογραφικών τους με τις απαιτήσεις της εκάστοτε θέσης.

13. Οφέλη

Η εφαρμογή Smart HR Ranking & Evaluation προσφέρει σημαντικά οφέλη κατά τη διαδικασία αξιολόγησης και επιλογής υποψηφίων, συμβάλλοντας στην αποτελεσματικότερη υποστήριξη των υπευθύνων ανθρώπινου δυναμικού (HR Managers). Μέσω της αξιοποίησης τεχνολογιών Τεχνητής Νοημοσύνης και σημασιολογικής αναζήτησης, επιτρέπει την ταχύτερη και περισσότερο οργανωμένη επεξεργασία μεγάλου αριθμού βιογραφικών σημειωμάτων. Ένα από τα σημαντικότερα πλεονεκτήματα της εφαρμογής είναι η **εξοικονόμηση χρόνου** κατά το αρχικό στάδιο επιλογής υποψηφίων. Αντί της χειροκίνητης ανάγνωσης και σύγκρισης πολλών βιογραφικών, ο χρήστης εισάγει την περιγραφή της θέσης εργασίας και το σύστημα αναλαμβάνει την αναζήτηση και αξιολόγηση των καταλληλότερων υποψηφίων. Με τον τρόπο αυτό μειώνεται σημαντικά ο χρόνος που απαιτείται για τη διαδικασία του αρχικού screening.

Παράλληλα, η εφαρμογή συμβάλλει στη **βελτίωση της οργάνωσης των βιογραφικών**. Τα βιογραφικά που έχουν αποθηκευτεί στη διανυσματική βάση δεδομένων ChromaDB μπορούν να αναζητηθούν σημασιολογικά με βάση τις απαιτήσεις κάθε θέσης εργασίας, επιτρέποντας την αξιοποίηση ήδη διαθέσιμων υποψηφίων χωρίς να απαιτείται νέα χειροκίνητη αναζήτηση. Ένα ακόμη σημαντικό πλεονέκτημα είναι η **υποστήριξη της διαδικασίας λήψης αποφάσεων**. Για κάθε υποψήφιο παρουσιάζεται ποσοστό αντιστοίχισης με την περιγραφή της θέσης εργασίας, συνοπτική αιτιολόγηση των αποτελεσμάτων, καθώς και αυτόματα παραγόμενες ερωτήσεις συνέντευξης. Οι πληροφορίες αυτές λειτουργούν υποστηρικτικά για τον υπεύθυνο προσλήψεων, χωρίς να αντικαθιστούν την τελική ανθρώπινη απόφαση.

Η εφαρμογή αξιοποιεί το γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω της πλατφόρμας **Ollama**, ενώ η διαχείριση των βιογραφικών πραγματοποιείται μέσω της **ChromaDB**. Η τοπική εκτέλεση του μοντέλου επιτρέπει την επεξεργασία των δεδομένων χωρίς να απαιτείται αποστολή των βιογραφικών σε εξωτερικές υπηρεσίες, ενισχύοντας έτσι την προστασία των προσωπικών δεδομένων των υποψηφίων. Επιπλέον, η εφαρμογή διαθέτει ένα απλό και εύχρηστο περιβάλλον χρήσης, το οποίο επιτρέπει την εύκολη εισαγωγή της περιγραφής θέσης εργασίας, τη μεταφόρτωση νέων βιογραφικών και την προβολή των αποτελεσμάτων αξιολόγησης. Η διεπαφή έχει σχεδιαστεί έτσι ώστε να διευκολύνει τον χρήστη, ακόμη και χωρίς ιδιαίτερες τεχνικές γνώσεις.

Συνολικά, η εφαρμογή συμβάλλει στην επιτάχυνση της διαδικασίας αξιολόγησης, στη βελτίωση της οργάνωσης των διαθέσιμων βιογραφικών και στην αποτελεσματικότερη υποστήριξη της διαδικασίας επιλογής προσωπικού. Ο συνδυασμός της σημασιολογικής αναζήτησης με τις δυνατότητες ενός μεγάλου γλωσσικού μοντέλου προσφέρει ένα σύγχρονο εργαλείο υποστήριξης των διαδικασιών ανθρώπινου δυναμικού, διατηρώντας παράλληλα τον άνθρωπο ως τον τελικό υπεύθυνο για τη λήψη των αποφάσεων.

14. Συμπεράσματα

Η παρούσα πτυχιακή εργασία είχε ως αντικείμενο την ανάπτυξη μιας εφαρμογής υποστήριξης της διαδικασίας αξιολόγησης βιογραφικών σημειωμάτων, αξιοποιώντας τεχνολογίες Retrieval-Augmented Generation (RAG) και Μεγάλων Γλωσσικών Μοντέλων (Large Language Models - LLMs). Στόχος της εφαρμογής ήταν η διευκόλυνση της διαδικασίας αντιστοίχισης υποψηφίων με μία θέση εργασίας μέσω της σημασιολογικής αναζήτησης και της αυτόματης ανάλυσης των βιογραφικών.

Κατά την υλοποίηση της εφαρμογής χρησιμοποιήθηκαν σύγχρονες τεχνολογίες όπως η **Streamlit** για την ανάπτυξη της διεπαφής χρήστη, η **ChromaDB** για την αποθήκευση και αναζήτηση διανυσματικών αναπαραστάσεων των βιογραφικών, η βιβλιοθήκη **EasyOCR** για την εξαγωγή κειμένου από εικόνες και το γλωσσικό μοντέλο **Phi-3**, το οποίο εκτελείται τοπικά μέσω της πλατφόρμας **Ollama**. Ο συνδυασμός των τεχνολογιών αυτών επέτρεψε την ανάπτυξη ενός ολοκληρωμένου συστήματος που μπορεί να αναλύει βιογραφικά διαφορετικών μορφών και να παρουσιάζει αποτελέσματα αξιολόγησης με τρόπο εύκολα κατανοητό από τον χρήστη.

Η εφαρμογή επιτυγχάνει τον βασικό σκοπό της, καθώς δίνει τη δυνατότητα εισαγωγής περιγραφής θέσης εργασίας, αξιολόγησης νέων ή ήδη αποθηκευμένων βιογραφικών, παρουσίασης ποσοστού αντιστοίχισης κάθε υποψηφίου, δημιουργίας σύντομης αιτιολόγησης της αξιολόγησης και παραγωγής ενδεικτικών ερωτήσεων συνέντευξης. Οι λειτουργίες αυτές μπορούν να υποστηρίξουν τον υπεύθυνο ανθρώπινου δυναμικού κατά το αρχικό στάδιο επιλογής υποψηφίων, χωρίς όμως να αντικαθιστούν την τελική ανθρώπινη κρίση.

Παράλληλα, η χρήση της αρχιτεκτονικής RAG σε συνδυασμό με τη διανυσματική βάση δεδομένων ChromaDB αποδεικνύει ότι είναι δυνατή η αποτελεσματική αξιοποίηση σημασιολογικής αναζήτησης για εφαρμογές επιλογής προσωπικού. Η δυνατότητα ανάκτησης των πιο σχετικών βιογραφικών και η αξιοποίησή τους από το γλωσσικό μοντέλο οδηγούν σε περισσότερο στοχευμένες και τεκμηριωμένες απαντήσεις σε σχέση με μια απλή αναζήτηση λέξεων-κλειδιών.

Κατά την ανάπτυξη της εφαρμογής εντοπίστηκαν και ορισμένοι περιορισμοί. Η ποιότητα των αποτελεσμάτων εξαρτάται σε σημαντικό βαθμό από την ποιότητα των βιογραφικών και της περιγραφής της θέσης εργασίας που εισάγει ο χρήστης, καθώς και από τις δυνατότητες του χρησιμοποιούμενου γλωσσικού μοντέλου. Επιπλέον, η εφαρμογή αποτελεί πρωτότυπο ερευνητικό σύστημα και δεν έχει αξιολογηθεί σε πραγματικό περιβάλλον λειτουργίας οργανισμού ή τμήματος ανθρώπινου δυναμικού.

Ως μελλοντικές επεκτάσεις μπορούν να εξεταστούν η υποστήριξη περισσότερων γλωσσών, η αξιοποίηση νεότερων ή μεγαλύτερων γλωσσικών μοντέλων, η ενσωμάτωση πρόσθετων πηγών δεδομένων για την αναζήτηση υποψηφίων, καθώς και η ανάπτυξη μηχανισμών που θα επιτρέπουν την ευκολότερη διαχείριση και ενημέρωση της βάσης βιογραφικών. Επιπλέον, ιδιαίτερο ενδιαφέρον παρουσιάζει η αξιολόγηση της εφαρμογής από επαγγελματίες του χώρου του ανθρώπινου δυναμικού, ώστε να εξαχθούν συμπεράσματα σχετικά με τη χρηστικότητα και την αποτελεσματικότητά της σε πραγματικές διαδικασίες επιλογής προσωπικού.

Συνολικά, η εργασία απέδειξε ότι η αξιοποίηση τεχνικών Retrieval-Augmented Generation σε συνδυασμό με τοπικά γλωσσικά μοντέλα μπορεί να αποτελέσει μία αξιόπιστη προσέγγιση για την υποστήριξη της διαδικασίας αξιολόγησης υποψηφίων. Η εφαρμογή που αναπτύχθηκε καλύπτει τους βασικούς στόχους της εργασίας και μπορεί να αποτελέσει τη βάση για περαιτέρω έρευνα και ανάπτυξη στον τομέα των ευφύων συστημάτων υποστήριξης αποφάσεων για τη διοίκηση ανθρώπινου δυναμικού.

15. Πίνακας Ορολογίας

Όρος	Ορισμός
Applicant (Υποψήφιος)	Το άτομο που υποβάλλει αίτηση/βιογραφικό σημείωμα για μια συγκεκριμένη θέση εργασίας.
Candidate Ranking (Κατάταξη Υποψηφίων)	Η διαδικασία ταξινόμησης των υποψηφίων βάσει του βαθμού συνάφειάς τους με τις απαιτήσεις μιας θέσης εργασίας.
ChromaDB	Ανοιχτού κώδικα διανυσματική βάση δεδομένων, η οποία χρησιμοποιήθηκε στην παρούσα εργασία για την αποθήκευση και αναζήτηση των embeddings των βιογραφικών.
CV (Curriculum Vitae)	Βιογραφικό σημείωμα· έγγραφο που συνοψίζει την εκπαίδευση, την εργασιακή εμπειρία και τις δεξιότητες ενός υποψηφίου.
Embedding	Αριθμητική (διανυσματική) αναπαράσταση κειμένου υψηλής διάστασης, η οποία αποτυπώνει τη σημασιολογική του σημασία.
HR Manager (Υπεύθυνος Προσλήψεων)	Το στέλεχος του τμήματος Ανθρώπινου Δυναμικού που είναι υπεύθυνο για τη διαδικασία επιλογής προσωπικού.
Job Description (Περιγραφή Θέσης Εργασίας)	Κείμενο που περιγράφει τις απαιτήσεις, τα καθήκοντα και τα προσόντα που ζητούνται για μια συγκεκριμένη θέση εργασίας.

Keyword Matching (Αντιστοίχιση Λέξεων-Κλειδιών)	Παραδοσιακή μέθοδος αναζήτησης/αξιολόγησης βιογραφικών βάσει ακριβούς αντιστοίχισης συγκεκριμένων λέξεων, χωρίς κατανόηση του εννοιολογικού περιεχομένου.
Large Language Model – LLM (Μεγάλο Γλωσσικό Μοντέλο)	Μοντέλο τεχνητής νοημοσύνης εκπαιδευμένο σε τεράστιο όγκο κειμενικών δεδομένων, ικανό να κατανοεί και να παράγει φυσική γλώσσα.
Local Inference (Τοπική Εξαγωγή Συμπερασμάτων)	Η εκτέλεση των υπολογισμών ενός μοντέλου τεχνητής νοημοσύνης απευθείας στον τοπικό υπολογιστή/server, χωρίς αποστολή δεδομένων σε εξωτερικό cloud API.
OCR (Optical Character Recognition)	Τεχνολογία μετατροπής εικόνας κειμένου (π.χ. σκαναρισμένο βιογραφικό) σε επεξεργάσιμο ψηφιακό κείμενο.
Phi-3	Μικρό γλωσσικό μοντέλο (SLM) της Microsoft, σχεδιασμένο ώστε να εκτελείται αποδοτικά τοπικά με περιορισμένους υπολογιστικούς πόρους.
Prompt (Προτροπή)	Το κείμενο εισόδου που δίνεται σε ένα γλωσσικό μοντέλο προκειμένου να καθοδηγήσει την απόκρισή του.
Retrieval-Augmented Generation – RAG	Αρχιτεκτονική τεχνητής νοημοσύνης που συνδυάζει ανάκτηση πληροφορίας από εξωτερική βάση γνώσης με την παραγωγή φυσικής γλώσσας από ένα LLM.
Resume Screening (Διαλογή Βιογραφικών)	Η διαδικασία αρχικής αξιολόγησης και φιλτραρίσματος βιογραφικών σημειωμάτων με σκοπό τον εντοπισμό των πλέον κατάλληλων υποψηφίων.
Semantic Search (Σημασιολογική Αναζήτηση)	Μέθοδος αναζήτησης που εντοπίζει περιεχόμενο βάσει εννοιολογικής συνάφειας και όχι απλής αντιστοίχισης λέξεων-κλειδιών.
Similarity Score (Βαθμός Ομοιότητας)	Αριθμητική τιμή που εκφράζει τον βαθμό εννοιολογικής συνάφειας μεταξύ δύο κειμένων (π.χ. βιογραφικού και περιγραφής θέσης).
Streamlit	Πλαίσιο ανάπτυξης (framework) της Python για τη δημιουργία διαδραστικών web εφαρμογών με ελάχιστο κώδικα front-end.
Vector Database (Διανυσματική Βάση Δεδομένων)	Βάση δεδομένων εξειδικευμένη στην αποθήκευση και αναζήτηση embeddings, με χρήση μετρικών ομοιότητας αντί για ακριβή αντιστοίχιση τιμών.
Small Language Model – SLM (Μικρό Γλωσσικό Μοντέλο)	Γλωσσικό μοντέλο μειωμένου μεγέθους σε σχέση με τα LLMs, σχεδιασμένο για τοπική/offline εκτέλεση με χαμηλότερες απαιτήσεις πόρων.
Explainable AI – XAI (Εξηγήσιμη Τεχνητή Νοημοσύνη)	Σύνολο τεχνικών που καθιστούν τις αποφάσεις ενός μοντέλου τεχνητής νοημοσύνης κατανοητές και ερμηνεύσιμες από τον άνθρωπο.
Cosine Similarity (Ομοιότητα Συνημιτόνου)	Μετρική σημασιολογικής ομοιότητας μεταξύ δύο διανυσμάτων (embeddings), η οποία χρησιμοποιείται για τον εντοπισμό εννοιολογικά συναφών κειμένων.

Agentic AI (Ευφυείς Πράκτορες)	Συστήματα τεχνητής νοημοσύνης ικανά να εκτελούν πολύβημες εργασίες αυτόνομα, λαμβάνοντας αποφάσεις χωρίς συνεχή ανθρώπινη καθοδήγηση.
Ollama	Πλατφόρμα που επιτρέπει την τοπική εκτέλεση γλωσσικών μοντέλων (LLMs/SLMs) σε προσωπικό υπολογιστή ή server, χωρίς σύνδεση σε εξωτερικό cloud API.

16. Πίνακας συντμήσεων-αρτικόλεξων-ακρωνυμίων

Ακρωνύμιο	Επεξήγηση
AI	Artificial Intelligence (Τεχνητή Νοημοσύνη)
API	Application Programming Interface
ATS	Applicant Tracking System
AWS	Amazon Web Services
CPU	Central Processing Unit
CSS	Cascading Style Sheets
CV	Curriculum Vitae (Βιογραφικό Σημείωμα)
GDPR	General Data Protection Regulation
GPU	Graphics Processing Unit
HCI	Human-Computer Interaction
HR	Human Resources (Ανθρώπινο Δυναμικό)
HRM	Human Resource Management
HTML	HyperText Markup Language
ID	Identifier
MRR	Mean Reciprocal Rank
JPG	Joint Photographic Experts Group (μορφή αρχείου εικόνας)
LLM	Large Language Model
OCR	Optical Character Recognition
PDF	Portable Document Format

PNG	Portable Network Graphics
RAG	Retrieval-Augmented Generation
RAM	Random Access Memory
SLM	Small Language Model
SQL	Structured Query Language
TXT	Plain Text (μορφή αρχείου)
UI	User Interface
UID	User Identifier
UX	User Experience
XAI	Explainable Artificial Intelligence
ZIP	Αρχείο συμπιεσμένης μορφής (ZIP archive)

17.Βιβλιογραφία

1. Abdin, M., Aneja, J., Awadalla, H., et al. (2024). Phi-3 Technical Report: A Highly Capable Language Model Locally on Your Phone. arXiv preprint arXiv:2404.14219.
2. Albaroudi, E., Mansouri, T., & Alameer, A. (2024). A Comprehensive Review of AI Techniques for Addressing Algorithmic Bias in Job Hiring. AI (MDPI), 5(1), 383–404.
3. Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. Information Fusion, 58, 82–115.
4. Chowdhery, A., Narang, S., Devlin, J., et al. (2022). PaLM: Scaling Language Modeling with Pathways. arXiv preprint arXiv:2204.02311.
5. Liu, Y., Ott, M., Goyal, N., et al. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv preprint arXiv:1907.11692.
6. Gan, C., Zhang, Q., & Mori, T. (2024). Application of LLM Agents in Recruitment: A Novel Framework for Resume Screening. arXiv preprint arXiv:2401.08315.
7. Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., & Wang, H. (2023). Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv preprint arXiv:2312.10997.

8. González-Morales, A., Martín-López, M. M., & Góngora-Larequi, F. (2026). A systematic review and descriptive analysis of artificial intelligence applied to recruitment and personnel selection in the present and possible future. *Discover Artificial Intelligence*, 6, 762.
9. González-Morales, A., Martín-López, M. M., & Góngora-Larequi, F. (2026). A systematic review and descriptive analysis of artificial intelligence applied to recruitment and personnel selection in the present and possible future. *Discover Artificial Intelligence*, 6, 762.
10. Maheshwary, S., & Misra, H. (2018). Matching Resumes to Jobs via Deep Siamese Network. In *Companion Proceedings of the The Web Conference 2018 (WWW '18)*, 87–88.
11. Jobin, A., Ienca, M., & Vayena, E. (2019). The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*, 1(9), 389–399.
12. Douze, M., Guzhva, A., Deng, C., et al. (2024). The Faiss library. *arXiv preprint arXiv:2401.08281*.
13. Köchling, A., & Wehner, M. C. (2020). Discriminated by an Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of HR Recruitment and HR Development. *Business Research*, 13, 795–848.
14. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 9459–9474.
15. Doshi-Velez, F., & Kim, B. (2017). Towards A Rigorous Science of Interpretable Machine Learning. *arXiv preprint arXiv:1702.08608*.
16. Jégou, H., Douze, M., & Schmid, C. (2011). Product Quantization for Nearest Neighbor Search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(1), 117–128.
17. Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.
18. Muennighoff, N., Tazi, N., Magne, L., & Reimers, N. (2022). MTEB: Massive Text Embedding Benchmark. *arXiv preprint arXiv:2210.07316*.
19. Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). Large Language Models: A Survey. *arXiv preprint arXiv:2402.06196*.
20. Khattab, O., & Zaharia, M. (2020). ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT. *Proceedings of SIGIR 2020*.
21. Gugnani, A., & Misra, H. (2020). Implicit Skills Extraction Using Document Embedding and Its Use in Job Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(08), 13283–13290.

22. Conneau, A., Khandelwal, K., Goyal, N., et al. (2020). Unsupervised Cross-lingual Representation Learning at Scale. *Proceedings of ACL 2020*, 8440–8451.
23. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3982–3992.
24. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Pedreschi, D., & Giannotti, F. (2018). A Survey of Methods for Explaining Black Box Models. *ACM Computing Surveys*, 51(5), Article 93.
25. Poole, D. L., & Mackworth, A. K. (2023). *Artificial Intelligence: Foundations of Computational Agents* (3rd ed.). Cambridge University Press.
26. Salton, G., & Buckley, C. (1988). Term-Weighting Approaches in Automatic Text Retrieval. *Information Processing & Management*, 24(5), 513–523.
27. Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial Intelligence in Human Resources Management: Challenges and a Path Forward. *California Management Review*, 61(4), 15–42.
28. Nexus: A Multi-Modal Framework for Semantic Job Matching and AI-Driven Resume Analysis. (2026). *Procedia Computer Science*, 283, 922–935.
29. Mori, M., Sasseti, S., & Cavaliere, V. (2024). A systematic literature review on artificial intelligence in recruiting and selection: a matter of ethics. *Personnel Review*.
30. European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data (General Data Protection Regulation – GDPR). *Official Journal of the European Union*.
31. European Parliament and Council of the EU (2024). Regulation (EU) 2024/1689 — Artificial Intelligence Act. *OJ L*, 2024/1689, 12.7.2024.
32. ISO/IEC (2011). ISO/IEC 25010:2011 — SQuaRE.
33. Chroma. (2024). Chroma: The Open-Source Embedding Database Documentation. Ανακτήθηκε από: <https://docs.trychroma.com/>
34. Hugging Face. (2024). Microsoft Phi-3 Model Card. Ανακτήθηκε από: <https://huggingface.co/microsoft>
35. Wolf, T. et al. (2020). *Transformers: State-of-the-Art NLP*. *EMNLP 2020: System Demonstrations*, 38–45.
36. JaidedAI. (2023). EasyOCR: Ready-to-use OCR library with 80+ supported languages. *GitHub Repository*. Ανακτήθηκε από: <https://github.com/JaidedAI/EasyOCR>

37. LangChain. (2024). LangChain Documentation. Ανακτήθηκε από:
<https://python.langchain.com/>
38. Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., & Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585, 357–362.
39. Ollama. (2024). Ollama: Get up and running with Llama 3, Mistral, Phi-3 and other large language models locally. Ανακτήθηκε από: <https://ollama.com/>
40. McKinney, W. (2010). Data Structures for Statistical Computing in Python. *SciPy 2010*, 56–61.
41. Clark, A., & Contributors. (2026). python-pillow/Pillow (v12.3.0) [Computer software]. Zenodo.
42. Python Software Foundation. (2024). Python 3 Documentation. Ανακτήθηκε από:
<https://docs.python.org/3/>
43. Reimers, N. (2024). Sentence Transformers Documentation. Ανακτήθηκε από:
<https://www.sbert.net/>
44. Streamlit Inc. (2024). Streamlit Documentation. Ανακτήθηκε από:
<https://docs.streamlit.io/>